Course Code: EEE 325

Course Title: Physical Electronics

Units: 3

TEXTBOOKS

1. "Physics of Semiconductor Devices" S. M. Sze and K. K. Ng.

2. "Fundamentals of Microelectronics" B. Razavi.

3. "Electronic Principles" A. Malvino & D. Bates.

4. "Engineering Electronics" G. E. Happel & W. M. Hesselberth.

5. "Microelectronic Circuits" Sedra/Smith

6. "Solid State Physical Electronics" Aldert Van der Ziel

7. "CMOS Circuit Design, Layout, and Simulation" R. J. Baker.

8. "Electronic Devices and Circuit Theory" R. Boylestad & L. Nashelsky.

9. "Principles of Electronics" V. K. Mehta & R. Mehta.

ABOUT THE GRADE

- Examination: 60%
- Attendance: 10%
- Assignment: 15%
- Test:[15%]
- Hand written assignments will not be accepted.

FREE ELECTRONS MOTIONING STATIC & MAGNETIC FIELDS.

The term **static** refers to a situation where the fields do not vary with time. A **static electric field** (also referred to as electrostatic field) is created by charges that are fixed in space.

A **static electric field** is just the effect of a rest charge(i.e. a charge that is not moving, say Q) and does not involve any magnetic field and also is inversely proportional to the distance(**r)** of the point(of interest) from the charge.

==Electric charges are source of electric fields.== An electric field exerts force on an electric charge, whether the charge happens to be moving or at rest.

ELECTRIC FIELD

In 1875 Coulomb studied the magnitude of the forces between two electrically charged bodies. He formulated the law governing the electrostatic forces of attraction, called Coulomb's law.

The law states that:

In a given medium the force of attraction or repulsion F_e between two bodies with charges of Q and q is directly proportional to the charges and inversely proportional to the square of their separation r.

$$F_e = \frac{1}{4\pi\epsilon_o}\frac{Qq}{r^2}$$

The constant of proportionality $\frac{1}{4\pi\epsilon_o}$ has been chosen to have the value $9.05\text{x}10^9 m/F$. The force described by Coulomb's law is an action at a distance. The charge Q acts through a distance to exert an electric force F_e on the test charge q.

ϵ_o =is called the permittivity constant of vacuum

We imagine that existence of Q changes the state of the space by producing an electric field in the space surrounding itself. If a test charge q is placed at some point distance r from the charge Q, the electric field E at that point is given by:

$$E = \frac{F_e}{q}$$

The unit of E is newton/coulomb (N/C)

Since $F_e = \dfrac{1}{4\pi\epsilon_o}\dfrac{Qq}{r^2}$, the magnitude of E is given by

$$E = \dfrac{Q}{4\pi\epsilon_o r^2}$$

Since F_e is a vector and q is a scalar, the direction of E is the same as that of F_e, i.e. the direction in which a test positive charge placed at the point would tend to move.

The electric field is the component of electromagnetic field. It is a vector field and is generated by electric charges or time varying magnetic field. Since the magnetic field is described as a function of electric field, the equations of both fields are coupled and together form Maxwell's equations that describe both fields as a function of charges and currents.

A uniform electric field is constant at every point. It can be approximated by placing two conducting plates parallel to each other and maintaining a voltage (potential difference) between them.

THE CONCEPT OF ELECTRIC POTENTIAL

Electric potential is a scalar field which is comparable to gravitational potential. It is the work done in taking a unit positive charge from infinity to the point in question.

Like the gravitational potential, it is inversely proportional to the distance r from the charge Q which creates the field

$$V = \frac{Q}{4\pi\epsilon_o r}$$

It should also be noted that if a charge Q is moved through a distance d (in metres) and potential difference V (in volts) the work W done, in joules,

W=QV

and the electric field and potential difference are related by

$$E = \frac{V}{d}$$

Calculate the magnitude of the force of repulsion between two equal charges of 2.0 μC, separated by a distance of 1 m. in a vacuum. [Take $\frac{1}{4\pi\epsilon_o} = 9 \times 10^9$].

Solution

$$F = \frac{Qq}{4\pi\epsilon_o r^2} = \frac{9 \times 10^9 \times (2.0 \times 10^{-6})^2}{1^2}$$
$$= 0.036 N$$

MAGNETIC FIELD

Electric current is the source of magnetic field. A magnetic field is the magnetic effect of electric currents and magnetic materials. The magnetic field at any given point is specified by both a direction and a magnitude (or strength); as such it is a vector field. The term is used for two distinct but closely related fields denoted by the symbols B and H,

where H is measured in units of amperes per meter (symbol: $A \cdot m^{-1}$ or A/m) in the SI. B is measured in tesla (symbol: T) and newton per meter per ampere (symbol: $N \cdot m^{-1} \cdot A^{-1}$ or N/(m·A) in the SI. B is most commonly defined in terms of the Lorentz force it exerts on moving electric charges.

When a charged particle is placed in an electromagnetic field, it experiences a force, called Lorentz force.

Magnetic fields can be produced by moving electric charges and the intrinsic magnetic moments of elementary particles associated with a fundamental quantum property, their spin. In special relativity, electric and magnetic fields are two interrelated aspects of a single object, called the electromagnetic tensor;

The split of this tensor into electric and magnetic fields depends on the relative velocity of the observer and charge. In quantum physics, the electromagnetic field is quantized and electromagnetic interactions result from the exchange of photons.
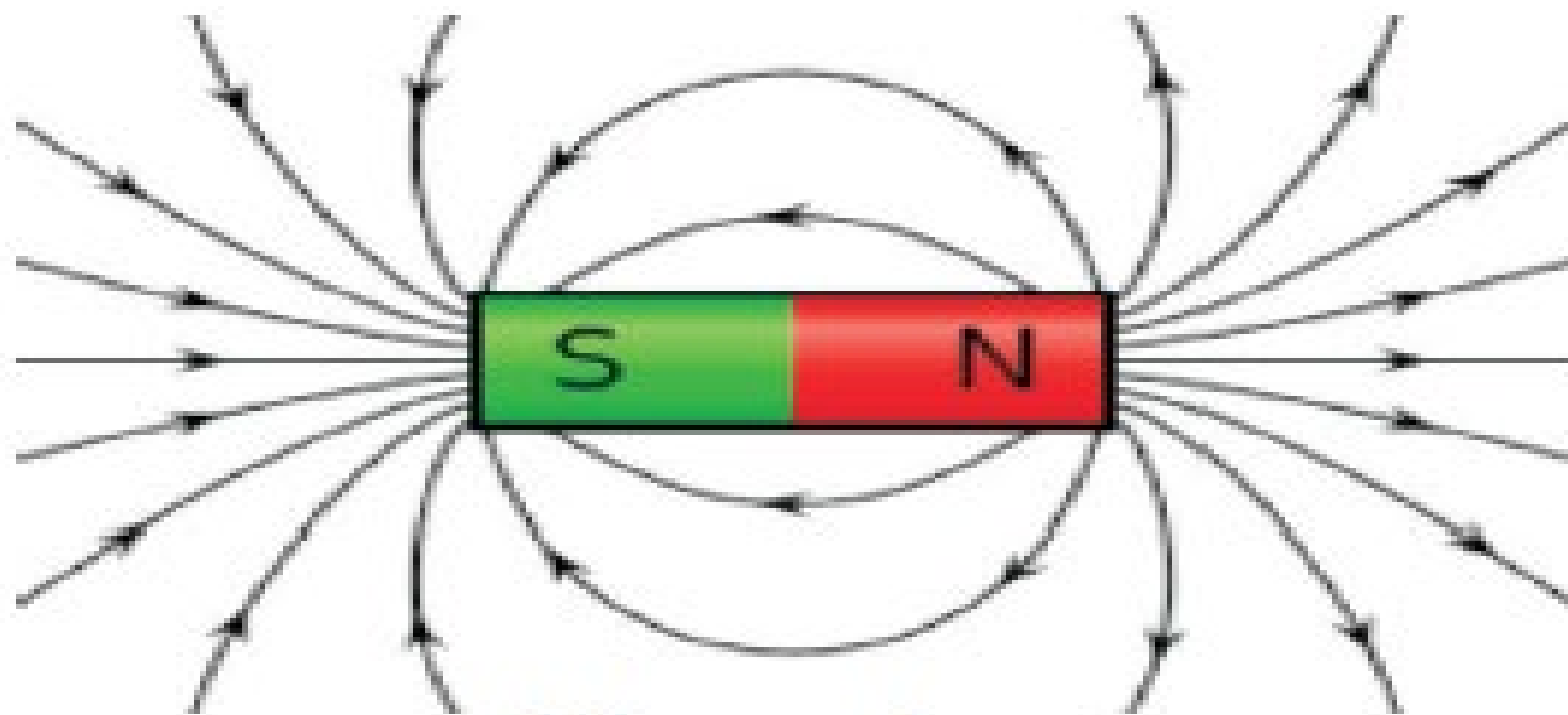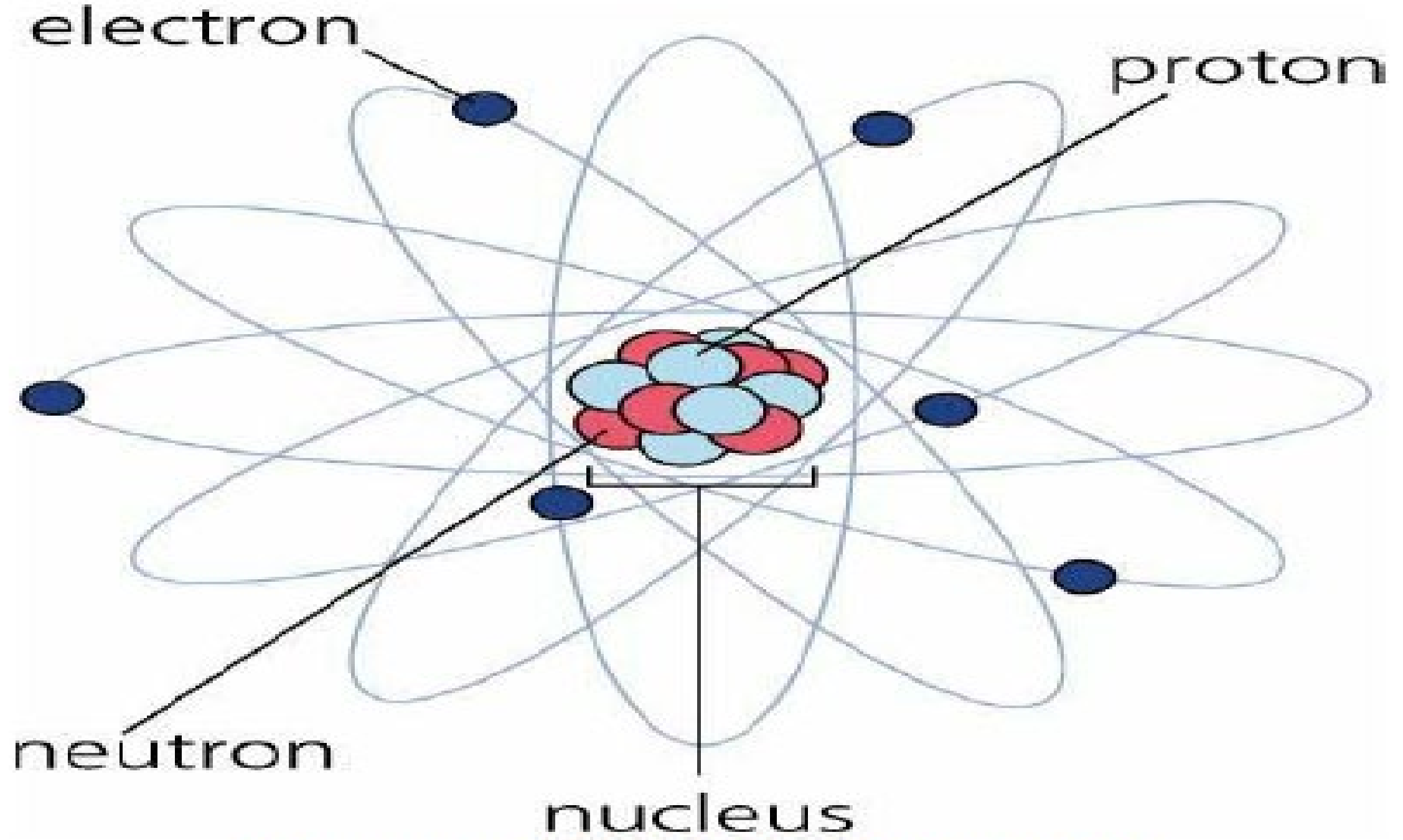
Figure-1

CHARGE AND ELECTRIC FORCE

There are two types of charges:

1 Positive charge (+)

2 Negative charge (-)

Conservation law of charges:

• Charge can be neither created nor destroyed. It can be only transferred.

• The unit of charge is Coulomb. It is denoted by C.

Fig. 1.1: Structure of the atom

Each atoms consist of three types of particles.

Protons

Protons are positive charge (+) and it has charge of + 1.6 X 10^{-19} coulombs (C).

Electrons
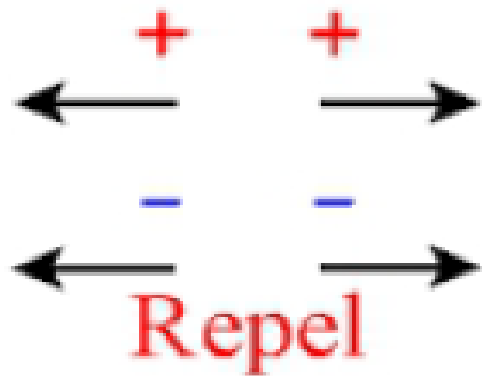
Electrons are negative charge (-) and it has charge of – 1.6 X 10^{-19} coulombs (C).

Neutron

The neutron has zero charge.

COULOMB'S FIRST LAW:

"Same kind of charges repel each other and different kind of charges attract each other".



Fig. 1.2: (a) Repulsion (b) Attraction

COULOMB'S SECOND LAW (INVERSE SQUARE LAW):

"The force between two charges Q1 and Q2 separated by a distance R is proportional to the product between the charges and inversely proportional to the square of distance between them".

Coulomb's law has given an expression to calculate the electric force in Newton (N) on one point charge by the other.
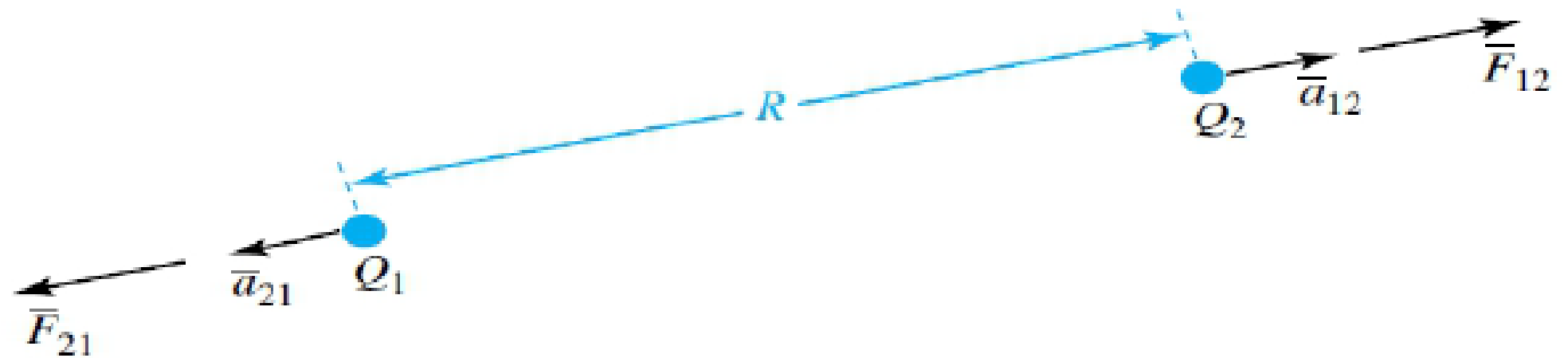
Fig. 1.3: Illustration of coulomb's law.

Force on Q_1 due to Q_2 → $\bar{F}_{21} = [Q_1Q_2/4\pi\varepsilon_0\varepsilon_r R^2]\,\bar{a}_{21} = [K\,Q_1Q_2/R^2]\,\bar{a}_{21}$

Force on Q_2 due to Q_1 → $\bar{F}_{12} = [Q_1Q_2/4\,\pi\varepsilon_0\varepsilon_r R^2]\,\bar{a}_{12} = [K\,Q_1Q_2/R^2]\,\bar{a}_{12}$

Where,

F = Force between two charges in N.

K = 1 / $4\pi\varepsilon_0\varepsilon_r$

Q_1, Q_2 = Point charges in C.

ε_0 = Permittivity of the free space medium (8.854×10^{-12} F/m).

ε_r = Relative permittivity of the medium (for air $\varepsilon_r = 1$).

a_{12} and a_{21} are unit vectors along the line joining Q1 and Q2.

R = Distance between two charges in meter.

- Forces $\bar{F}_{21}$ and $\bar{F}_{12}$ are due to Q_2 and Q_1. They are equal in magnitude and opposite of each other in direction.
- The magnitude of the force is proportional to the product of the charge magnitudes.
- The magnitude of the force is inversely proportional to the square of the distance between the charges.
- The magnitude of the force depends on the medium.
- The direction of the force is along the line joining the charges.

Ex 1: (A) A small region of an impure silicon crystal with dimensions 1.25×10^{-6} m $\times 10^{-3}$ m $\times 10^{-3}$ m has only the ions (with charge $+1.6 \times 10^{-19}$ C) present with a volume density of $10^{25}/m^3$. The rest of the crystal volume contains equal densities of electrons (with charge -1.6×10^{-19} C) and positive ions. Find the net total charge of the crystal.

(B) Consider the charge of part (A) as a point charge Q_1. Determine the force exerted by this on a charge $Q_2 = 3\mu C$ when the charges are separated by a distance of 2 m in free space as shown in figure.

(C) If the another charge $Q_3 = -2\mu C$ is added to the system 1 m above Q_2, as shown in figure, calculate the force excreted on Q_2.
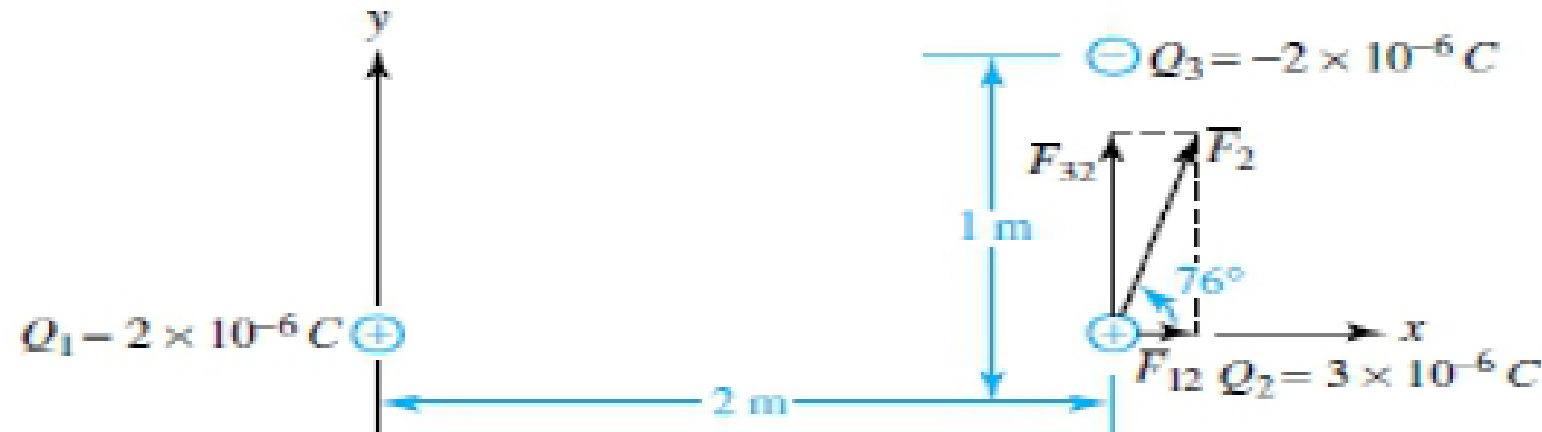
Solution:

- In the region where both ions and free electrons exist, their opposite charge cancel. So the net charge density is zero.
- The volume charge density = volume × charge

$$\rho = (10^{25})(1.6 \times 10^{-19}) = 1.6 \times 10^{6} \ C/m^3$$

- The net total charge is then calculated as:

$$Q = \rho \times v = (1.6 \times 10^{6})(1.25 \times 10^{-6} \times 10^{-3} \times 10^{-3}) = 2 \times 10^{-6} \ C$$

- The force that Q_1 exerts on Q_2 is in the positive direction of x.
- Now as we know that;

$$F = Q_1 Q_2 / 4\pi \varepsilon_0 \varepsilon_r R^2$$

Where; $\varepsilon_0 = 8.854 \ X \ ^{10}\text{-}12 \ F/m = 10^{-9} / 36\pi$ and $\varepsilon_r = 1$ (Relative permittivity of air)

$$\bar{F}_{12} = \frac{(3 \times 10^{-6})(2 \times 10^{-6})}{4\pi(10^{-9}/36\pi)2^2} \bar{a}_x = \bar{a}_x \, 13.5 \times 10^{-3} \, \text{N}$$

- The force that Q_3 exerts on Q_2 is in the positive direction of y.

$$\bar{F}_{32} = \frac{(3 \times 10^{-6})(-2 \times 10^{-6})}{4\pi(10^{-9}/36\pi)1^2}(-\bar{a}_y) = \bar{a}_y \, 54 \times 10^{-3} \, \text{N}$$

- The resultant force F_2 acting on Q_2 is in the superposition of F_{12} and F_{32} due to Q_1 and Q_3 respectively.

$$\bar{F}_2 = \sqrt{F_{12}^2 + F_{32}^2} \, \angle \tan^{-1} \frac{F_{32}}{\bar{F}_{12}}$$

$$= \sqrt{13.5^2 + 54^2} \times 10^{-3} \, \angle \tan^{-1} \frac{54}{13.5}$$

$$= 55.7 \times 10^{-3} \, \angle 76° \, \text{N}$$

CURRENT AND MAGNETIC FORCE:

Current:

Rate of flow of electrons is known as current. I = Q / t Where,
I = Current, Q = Charge, t = Time
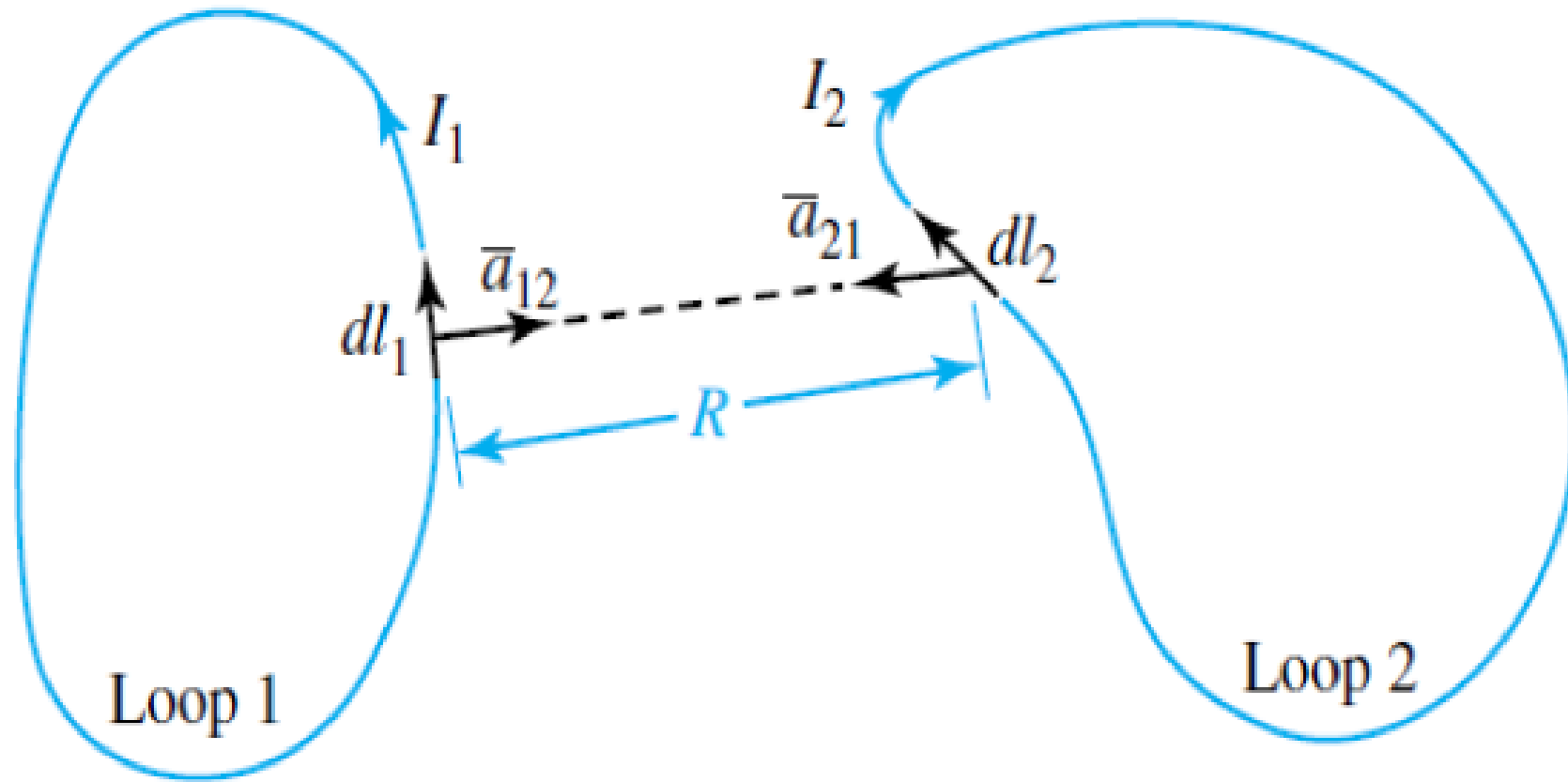
- If charge varies with time, so

$$i = dq / dt$$

Where,

  dq = change in charge

  dt = change in time

- Unit of current is Ampere. It is denoted as A.
- 1 A = 1 C/sec
- 1 Coulomb of charge is the charge carried by 6.25×10^{18} electrons.
- 1 mA = 10^{-3} A and 1 μA = 10^{-6} A

➤ Magnetic Force:

- *Ampere's law of force:*
- It is concerned with magnetic forces associated with two loops of wire carrying currents.

Fig. 1.4 Illustration of ampere's law of force

- It shows two loops of wire in free space carrying currents I_1 and I_2. Considering a differential element dl_1 of loop 1 and a differential element dl_2 of loop 2.
- The differential magnetic force by the differential current element is

$$d\bar{F}_{21} = I_1 \, d\bar{l}_1 \times \frac{\mu_0}{4\pi R^2} \left[I_2 \, d\bar{l}_2 \times \bar{a}_{21} \right]$$

$$d\bar{F}_{12} = I_2 \, d\bar{l}_2 \times \frac{\mu_0}{4\pi R^2} \left[I_2 \, d\bar{l}_2 \times \bar{a}_{12} \right]$$

- Where $\bar{a}_{21}$ and $\bar{a}_{12}$ are unit vectors along the line joining the two current elements, R is the distance between the centers of the elements, μ_0 is the permeability of free space.
- The magnitude of the force is proportional to the product of the two currents and the product of the lengths of the two current elements.
- The magnitude of the force is inversely proportional to the square of the distance between the current elements.
- Each current element is acted upon by a magnetic field due to the other current element,

$$d\bar{F}_{21} = I_1 \, d\bar{l}_1 \times B_2$$

$$d\bar{F}_{12} = I_2 \, d\bar{l}_2 \times B_1$$

- Where B is known as the magnetic flux density with unit of N/A·m, Webers per square meter (Wb/m^2) or tesla (T).
- Current distribution is the source of magnetic field, just as charge distribution is the source of electric field.

$$\bar{B}_2 = \frac{\mu_0}{4\pi R^2} (I_2 \, d\bar{l}_2 \times \bar{a}_{21})$$
$$\bar{B}_1 = \frac{\mu_0}{4\pi R^2} (I_1 \, d\bar{l}_1 \times \bar{a}_{12})$$

- This equation is known as the **Biot–Savart law.**

- *Lorentz force equation:*
- Current is due to the flow of charges.
- $I = dq/dt$ and $d\bar{l} = \bar{v} \, dt$, where v is the velocity.

$$d\bar{F} = \frac{dq}{dt} (\bar{v} \, dt) \times \bar{B} = dq \, (\bar{v} \times \bar{B})$$

- Thus the force F experienced by a charge q moving with a velocity v in a magnetic field of flux density B is given by

$$\bar{F} = q \, (\bar{v} \times \bar{B})$$

- Total force acting on a charge q moving with velocity v in a region characterized by electric field intensity E and a magnetic field of flux density B is

$$\bar{F} = \bar{F}_E + \bar{F}_M = q \, (\bar{E} + \bar{v} \times \bar{B})$$

ELECTRONIC STRUCTURE OF MATTER & ATOMIC MODELS

THE ATOM

All matter is composed of atoms; all atoms consist of electrons, protons, and neutrons except normal hydrogen, which does not have a neutron.

Each element in the periodic table has a unique atomic structure. Since hydrogen is simplest of the elements known, we shall consider its spectrum in detail in this course. At first, the atom was thought to be a tiny indivisible sphere.

Later it was shown that the atom was not a single particle but was made up of a small, dense nucleus around which electrons orbit at great distances from the nucleus, similar to the way planets orbit the sun.

Niels Bohr proposed that the electrons in an atom circle the nucleus in different obits, similar to the way planets orbit the sun in our solar system.

The Bohr model is often referred to as the planetary model.

Another view of the atom called the ***quantum model*** is considered a more accurate representation, but it is difficult to visualize. For most practical purposes in electronics, the Bohr model suffices and is commonly used because it is easy to visualize.

Atoms are electrically neutral. Therefore, an atom must also contain some positive charge to neutralise the negative charge of the electrons.

THE BOHR MODEL

An **atom** is the smallest particle of an element that retains the characteristics of that element. Each of the known 118 elements has atoms that are different from the atoms of all other elements. This gives each element a unique atomic structure.

According to the classical Bohr model, atoms have a planetary type of structure that consists of a central nucleus surrounded by orbiting electrons, as illustrated in Figure 1–1.
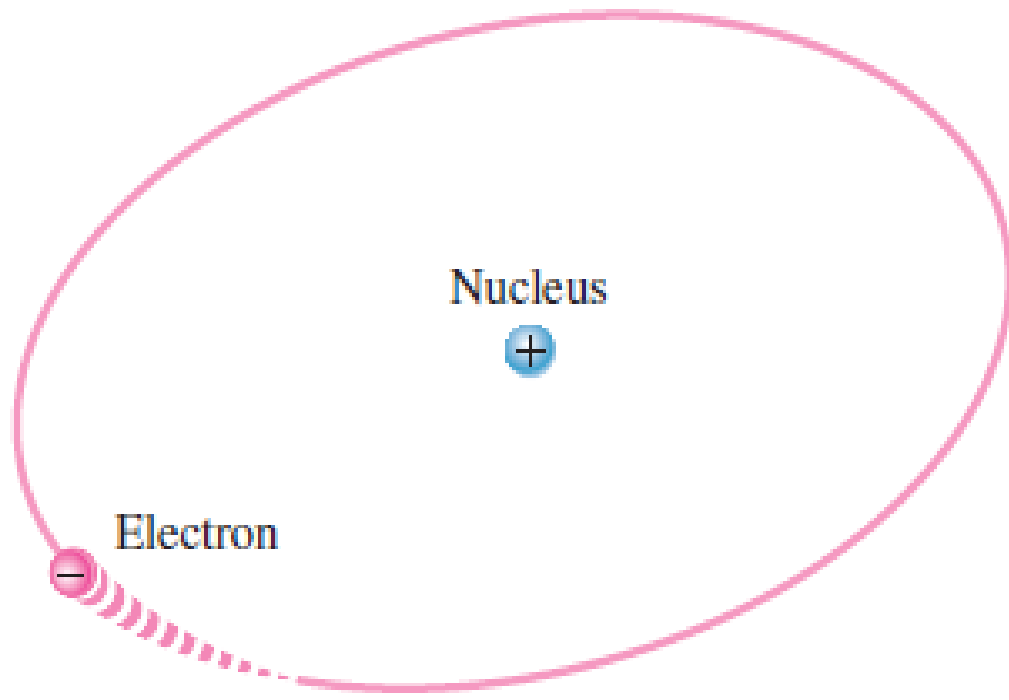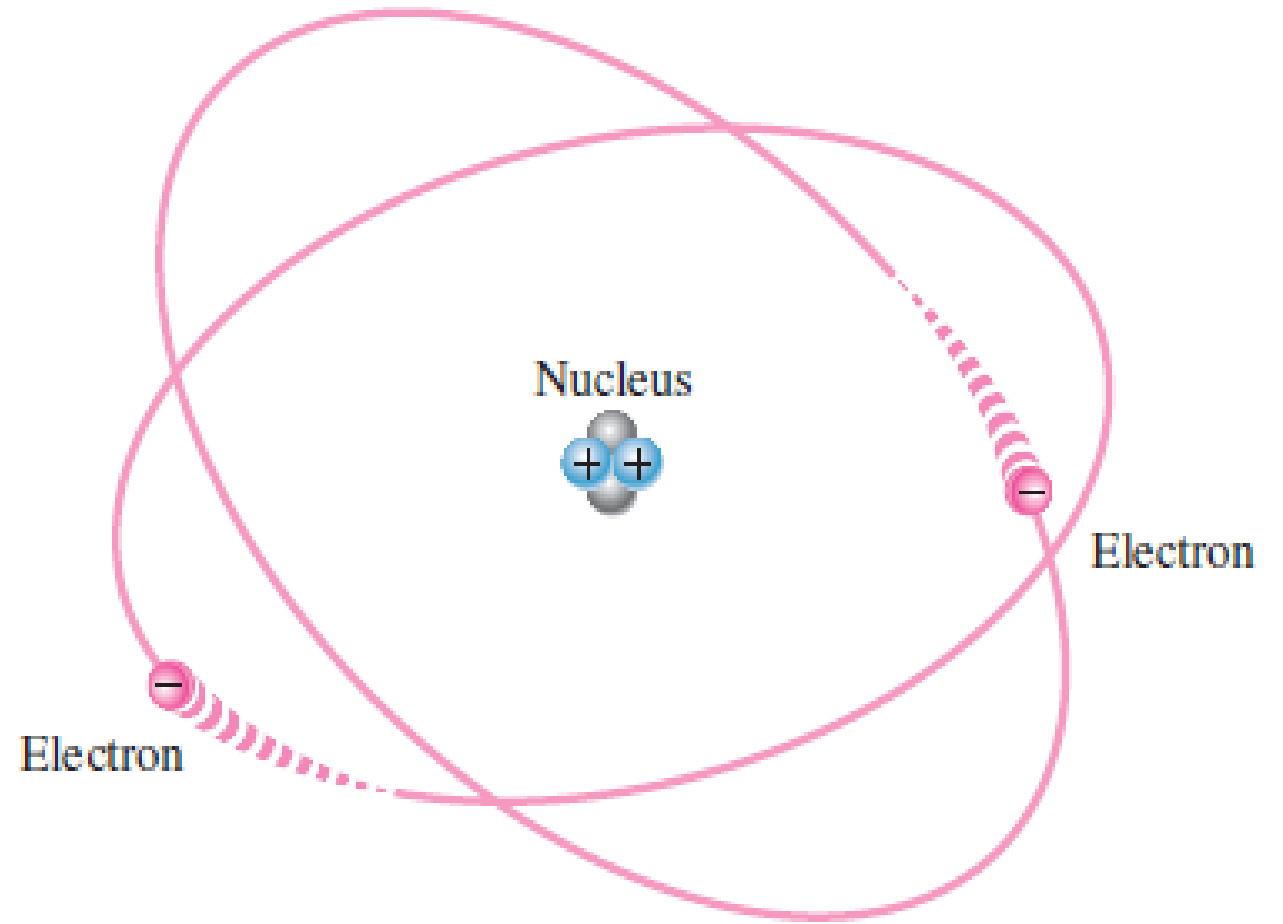
Electron + Proton Neutron
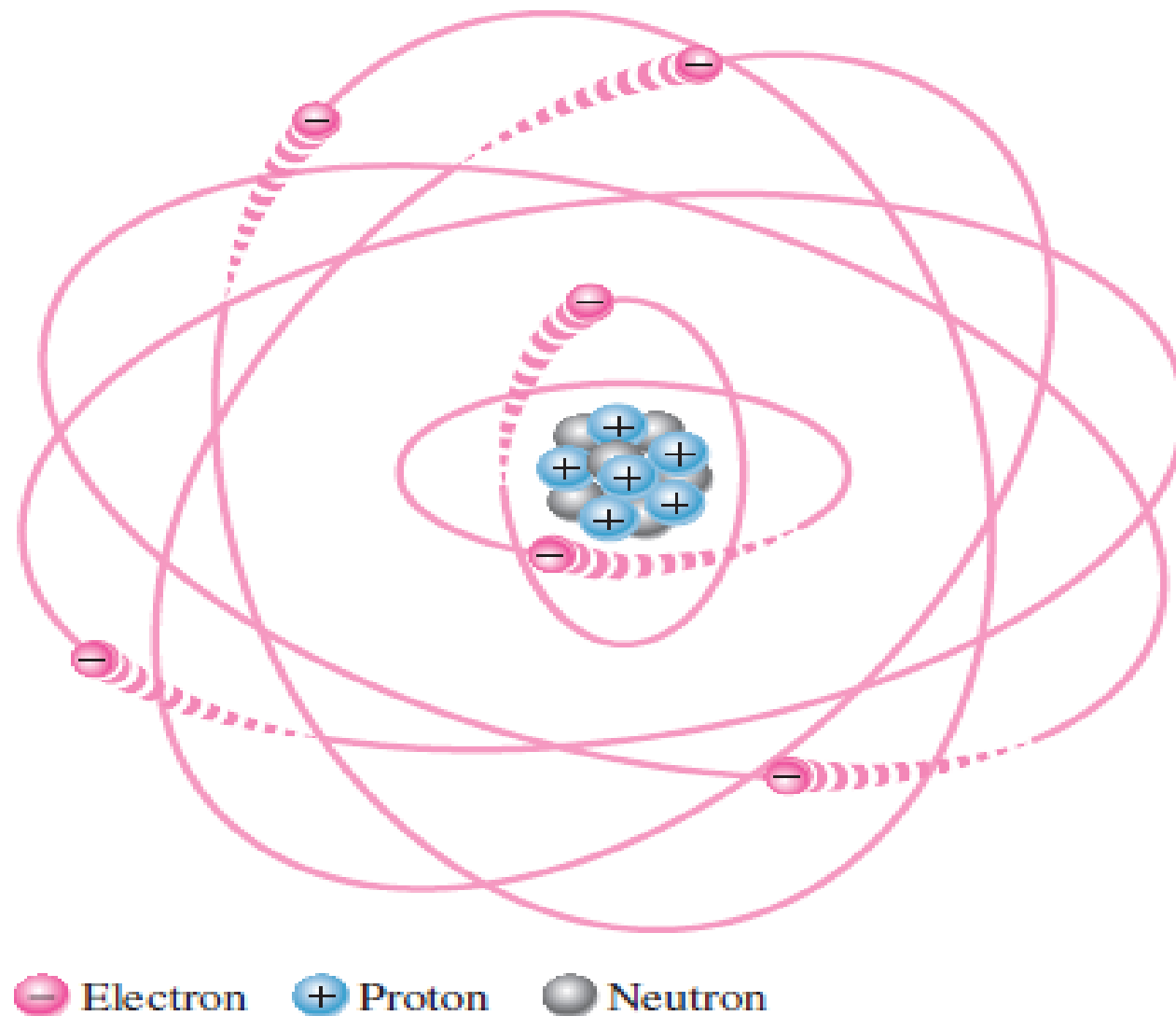
▲ FIGURE 1–1

The Bohr model of an atom showing electrons in orbits around the nucleus, which consists of protons and neutrons. The "tails" on the electrons indicate motion.

The **nucleus** consists of positively charged particles called **protons** and uncharged particles called **neutrons**. The basic particles of negative charge are called **electrons**. Each type of atom has a certain number of electrons and protons that distinguishes it from the atoms of all other elements.

For example, the simplest atom is that of hydrogen, which has one proton and one electron, as shown in Figure 1–2(a). As another example, the helium atom, shown in Figure 1–2(b), has two protons and two neutrons in the nucleus and two electrons orbiting the nucleus.

(a) Hydrogen atom

(b) Helium atom

▲ FIGURE 1–2

Two simple atoms, hydrogen and helium.

ATOMIC NUMBER

The **atomic number** equals the number of protons in the nucleus, which is the same as the number of electrons in an electrically balanced (neutral) atom. For example, hydrogen has an atomic number of 1 and helium has an atomic number of 2.

In their normal (or neutral) state, all atoms of a given element have the same number of electrons as protons; the positive charges cancel the negative charges, and the atom has a net charge of zero. Atomic numbers of all the elements are shown on the periodic table of the elements in Figure 1–3.

ELECTRONS AND SHELLS

Energy Levels: Electrons orbit the nucleus of an atom at certain distances from the nucleus. Electrons near the nucleus have less energy than those in more distant orbits.

Only discrete (separate and distinct) values of electron energies exist within atomic structures. Therefore, electrons must orbit only at discrete distances from the nucleus.

The Bohr model of the silicon atom is shown in Figure 1–4. Notice that there are 14 electrons surrounding the nucleus with exactly 14 protons, and usually 14 neutrons in the nucleus.

Illustration of the Bohr model of the silicon atom.

Shell 3 / Shell 2 / Shell 1
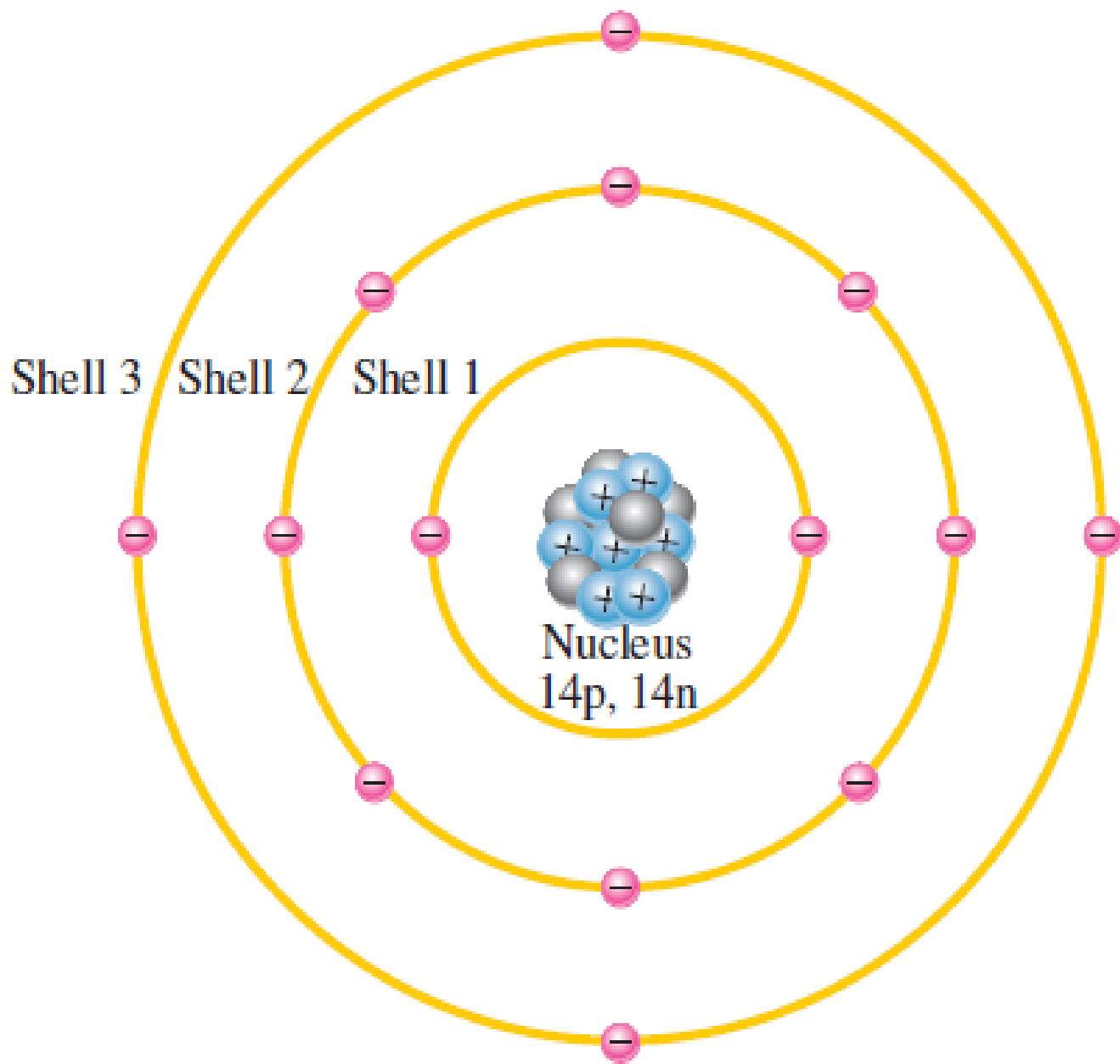
Nucleus
14p, 14n

Each discrete distance (**orbit**) from the nucleus corresponds to a certain energy level. In an atom, the orbits are grouped into energy levels known as **shells**. A given atom has a fixed number of shells. Each shell has a fixed maximum number of electrons. The shells (energy levels) are designated 1, 2, 3, and so on, with 1 being closest to the nucleus.

The maximum number of electrons (N_e) that can exist in each shell of an atom is a fact of nature and can be calculated by the formula,

$$\boldsymbol{N_e = 2n^2} \qquad\qquad Eq.\,1-1$$

Where n is the number of the shell. The maximum number of electrons that can exist in the innermost shell (shell 1) is

$$N_e = 2n^2 = 2(1)^2 = 2$$

The maximum number of electrons that can exist in shell 2 is

$$N_e = 2n^2 = 2(2)^2 = 8$$

The maximum number of electrons that can exist in shell 3 is

$$N_e = 2n^2 = 2(3)^2 = 18$$

The maximum number of electrons that can exist in shell 4 is

$$N_e = 2n^2 = 2(4)^2 = 32$$

VALENCE ELECTRONS

Electrons that are in orbits farther from the nucleus have higher energy and are less tightly bound to the atom than those closer to the nucleus. **This is because the force of attraction between the positively charged nucleus and the negatively charged electron decreases with increasing distance from the nucleus**.

Electrons with the highest energy exist in the outermost shell of an atom and are relatively loosely bound to the atom. This outermost shell is known as the **valence** shell, and electrons in this shell are called *valence electrons*.

These valence electrons contribute to chemical reactions and bonding within the structure of a material and determine its electrical properties. **When a valence electron gains sufficient energy from an external source, it can break free from its atom**. This is the basis for conduction in materials.

IONIZATION

When an atom absorbs energy, the valence electrons can easily jump to higher energy shells. If a valence electron acquires a sufficient amount of energy, called *ionization energy,* it can actually escape from the outer shell and the atom's influence.

The departure of a valence electron leaves a previously neutral atom with an excess of positive charge (more protons than electrons). The process of losing a valence electron is known as **ionization**, and the resulting positively charged atom is called a *positive ion.* For example, the chemical symbol for hydrogen is H^+.

Ionization energy

Definition: The ionization energy, E_i, of an atom/ion is the minimum energy which is required to remove an electron of an atom. The unit of ionization energy is kJ/mol.

When a neutral hydrogen atom loses its valence electron and becomes a positive ion, it is designated H^+. The escaped valence electron is called a **free electron**. The reverse process can occur in certain atoms when a free electron collides with the atom and is captured, releasing energy.

The atom that has acquired the extra electron is called a *negative ion*. The ionization process is not restricted to single atoms. In many chemical reactions, a group of atoms that are bonded together can lose or acquire one or more electrons. For some non-metallic materials such as chlorine, a free electron can be captured by the neutral atom, forming a negative ion. In the case of chlorine, the ion is more stable than the neutral atom because it has a filled outer shell. The chlorine ion is designated as Cl^-.

THE QUANTUM MODEL

Although the Bohr model of an atom is widely used because of its simplicity and ease of visualization, it is not a complete model. The quantum model is considered to be more accurate. The quantum model is a statistical model and very difficult to understand or visualize. Like the Bohr model, the quantum model has a nucleus of protons and neutrons surrounded by electrons.

Unlike the Bohr model, the electrons in the quantum model do not exist in precise circular orbits as particles. Three important principles underlie the quantum model: the wave-particle duality principle, the uncertainty principle, and the superposition principle.

Wave-particle duality. Just as light can be thought of as exhibiting both a wave and a particle **(photon),** electrons are thought to exhibit a wave-particle duality. The velocity of an orbiting electron is related to its wavelength, which interferes with neighbouring electron wavelengths by amplifying or canceling each other.

Uncertainly principle. As you know, a wave is characterized by peaks and valleys; therefore, electrons acting as waves cannot be precisely identified in terms of their position. According to a principle ascribed to Heisenberg, it is impossible to determine simultaneously both the position and velocity of an electron with any degree of accuracy or certainty.

Superposition. A principle of quantum theory that describes a challenging concept about the behavior of matter and forces at the subatomic level. Basically, **the principle states that although the state of any object is unknown, it is actually in all possible states simultaneously as long as an observation is not attempted**.

In the quantum model, each shell or energy level consists of up to four subshells called **orbitals**, which are designated s, p, d, and f. Orbital s can hold a maximum of two electrons, orbital p can hold six electrons, orbital d can hold 10 electrons, and orbital f can hold 14 electrons.

Each atom can be described by an electron configuration table that shows the shells or energy levels, the orbitals, and the number of electrons in each orbital. For example, the electron configuration table for the nitrogen atom is given in Table 1–1. The first full-size number is the shell or energy level, the letter is the orbital, and the exponent is the number of electrons in the orbital.

Electron configuration table for nitrogen.

NOTATION	EXPLANATION
$1s^2$	2 electrons in shell 1, orbital s
$2s^2\ 2p^3$	5 electrons in shell 2: 2 in orbital s, 3 in orbital p

Atomic orbitals do not resemble a discrete circular path for the electron as depicted in Bohr's planetary model. In the quantum picture, each shell in the Bohr model is a three dimensional space surrounding the atom that represents the mean (average) energy of the electron cloud. The term **electron cloud** (probability cloud) is used to describe the area around an atom's nucleus where an electron will probably be found.

Using the atomic number from the periodic table in Figure 1–3, describe a silicon (Si) atom using an electron configuration table.

The atomic number of silicon is 14. This means that there are 14 protons in the nucleus. Since there is always the same number of electrons as protons in a neutral atom, there are also 14 electrons.

As you know, there can be up to two electrons in shell 1, eight in shell 2, and eighteen in shell 3. Therefore, in silicon there are two electrons in shell 1, eight electrons in shell 2, and four electrons in shell 3 for a total of 14 electrons. The electron configuration table for silicon is shown in Table 1–2.

NOTATION	EXPLANATION
$1s^2$	2 electrons in shell 1, orbital s
$2s^2$ $2p^6$	8 electrons in shell 2: 2 in orbital s, 6 in orbital p
$3s^2$ $3p^2$	4 electrons in shell 3: 2 in orbital s, 2 in orbital p

ALPHA-PARTICLE SCATTERING AND RUTHERFORD'S NUCLEAR MODEL OF ATOM

At the suggestion of Ernst Rutherford, in 1911, H. Geiger and E. Marsden performed some experiments. In one of their experiments, as shown in Fig. 12.1, they directed a beam of 5.5 MeV α-particles emitted from a $^{214}_{83}Bi$ radioactive source at a thin metal foil made of gold.
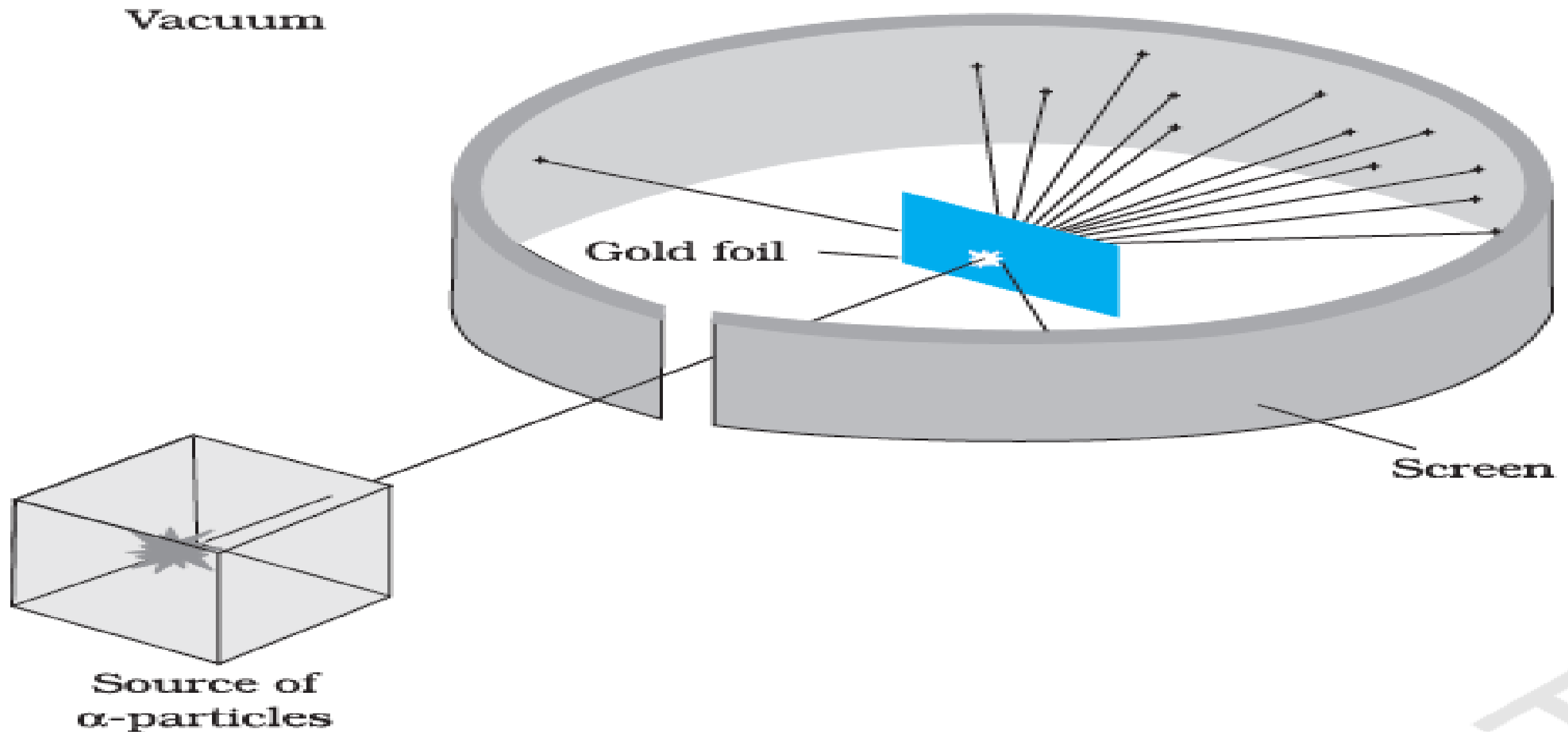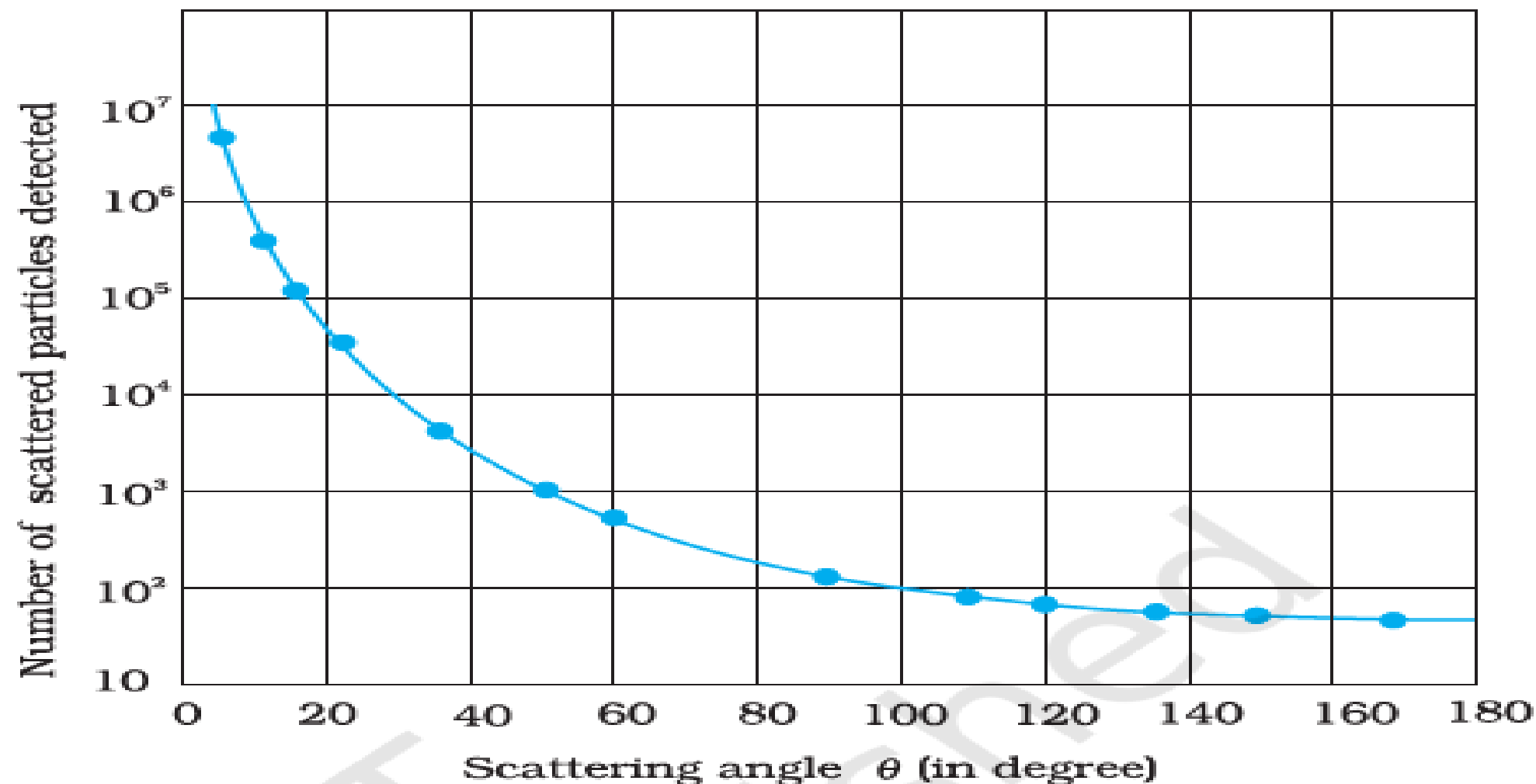
FIGURE 12.1 Geiger–Marsden scattering experiment. The entire apparatus is placed in a vacuum chamber (not shown in this figure).

Figure 12.2 shows a schematic diagram of this experiment. α-particles emitted by a $^{214}_{83}Bi$ radioactive source were collimated into a narrow beam by their passage through lead bricks. The beam was allowed to fall on a thin foil of gold of thickness 2.1×10^{-7} m. The scattered α-particles were observed through a rotatable detector consisting of zinc sulphide screen and a microscope. The scattered α-particles on striking the screen produced brief light flashes.

These flashes may be viewed through a microscope and the distribution of the number of scattered particles may be studied as a function of angle of scattering.

Angle of scattering is defined as the angle by which α-particle gets deviated from its original path around the nucleus.

A typical graph of the total number of α-particles scattered at different angles, in a given interval of time, is shown in Fig. 12.3.

FIGURE 12.3 Experimental data points (shown by dots) on scattering of α-particles by a thin foil at different angles obtained by Geiger and Marsden using the setup shown in Figs. 12.1 and 12.2. Rutherford's nuclear model predicts the solid curve which is seen to be in good agreement with experiment.

The dots in this figure represent the data points and the solid curve is the theoretical prediction based on the assumption that the target atom has a small, dense, positively charged nucleus.

Many of the α-particles pass through the foil. It means that they do not suffer any collisions. Only about 0.14% of the incident α-particles scatter by more than 1°; and about 1 in 8000 deflect by more than 90°.

Rutherford argued that, to deflect the α-particle backwards, it must experience a large repulsive force.

This force could be provided if the greater part of the mass of the atom and its positive charge were concentrated tightly at its centre.

Then the incoming α-particle could get very close to the positive charge without penetrating it, and such a close encounter would result in a large deflection. This agreement supported the hypothesis of the nuclear atom. This is why Rutherford is credited with the *discovery* of the nucleus.

In Rutherford's nuclear model of the atom, the entire positive charge and most of the mass of the atom are concentrated in the nucleus with the electrons some distance away. The electrons would be moving in orbits about the nucleus just as the planets do around the sun. Rutherford's experiments suggested the size of the nucleus to be about 10^{-15} m to 10^{-14} m. From kinetic theory, the size of an atom was known to be 10^{-10}, about 10,000 to 100,000 times larger than the size of the nucleus

The scattering data shown in Fig. 12.3 can be analysed by employing Rutherford's nuclear model of the atom. As the gold foil is very thin, it can be assumed that α-particles will suffer not more than one scattering during their passage through it. Therefore, computation of the trajectory of an α-particle scattered by a single nucleus is enough. α-particles are nuclei of helium atoms and, therefore, carry two units, $2e$, of positive charge and have the mass of the helium atom. The charge of the gold nucleus is Ze, where Z is the atomic number of the atom; for gold $Z = 79$.

Since the nucleus of gold is about 50 times heavier than an α-particle, it is reasonable to assume that it remains stationary throughout the scattering process. Under these assumptions, the trajectory of an α-particle can be computed employing Newton's second law of motion and the Coulomb's law for electrostatic force of repulsion between the α-particle and the positively charged nucleus.

The magnitude of this force is

$$F = \frac{1}{4\pi\varepsilon_o}\frac{(2e)[Ze]}{r^2} \quad [12.1]$$

where r is the distance between the α-particle and the nucleus. The force is directed along the line joining the α-particle and the nucleus. The magnitude and direction of the force on an α-particle continuously changes as it approaches the nucleus and recedes away from it.

ALPHA-PARTICLE TRAJECTORY

The trajectory traced by an α-particle depends on the impact parameter, b of collision. The *impact parameter* is the perpendicular distance of the initial velocity vector of the α-particle from the centre of the nucleus (Fig. 12.4). A given beam of α-particles has a distribution of impact parameters b, so that the beam is scattered in various directions with different probabilities (Fig. 12.4).
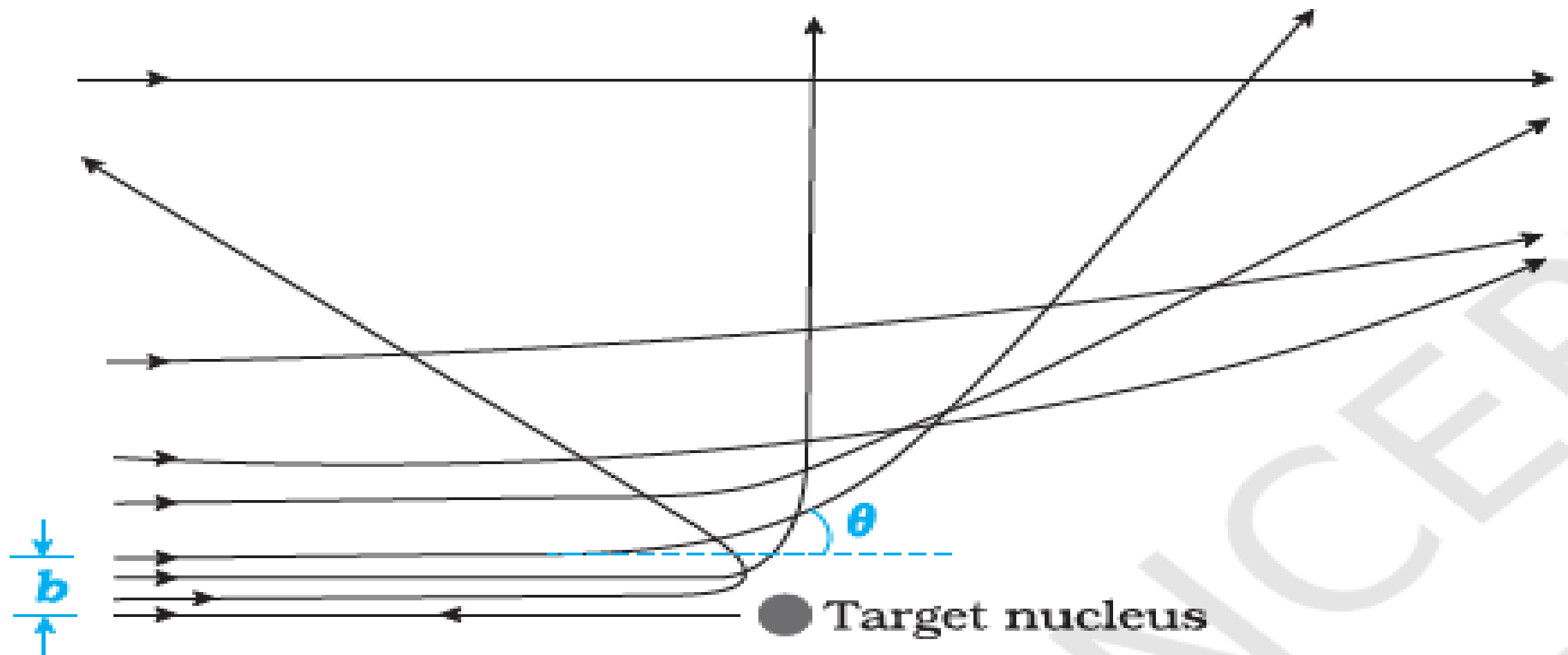
FIGURE 12.4 Trajectory of α-particles in the coulomb field of a target nucleus. The impact parameter, *b* and scattering angle *θ* are also depicted.

(In a beam, all particles have nearly same kinetic energy.) It is seen that an α-particle close to the nucleus (small impact parameter) suffers large scattering. In case of head-on collision, the impact parameter is minimum and the α-particle rebounds back ($\theta \cong \pi$). For a large impact parameter, the α-particle goes nearly undeviated and has a small deflection ($\theta \cong 0$).

The fact that only a small fraction of the number of incident particles rebound back indicates that the number of α-particles undergoing head on collision is small. This, in turn, implies that the mass of the atom is concentrated in a small volume. Rutherford scattering therefore, is a powerful way to determine an upper limit to the size of the nucleus.

ELECTRON ORBITS

The Rutherford nuclear model of the atom which involves classical concepts, pictures the atom as an electrically neutral sphere consisting of a very small, massive and positively charged nucleus at the centre surrounded by the revolving electrons in their respective dynamically stable orbits.

The electrostatic force of attraction, F_e between the revolving electrons and the nucleus provides the requisite centripetal force (F_c) to keep them in their orbits. Thus, for a dynamically stable orbit in a hydrogen atom

$$F_e = F_c$$

$$\frac{mv^2}{r} = \frac{1}{4\pi\varepsilon_o}\frac{e^2}{r^2}$$

Thus the relation between the orbit radius and the electron velocity is

$$r = \frac{e^2}{4\pi\varepsilon_o mv^2} \qquad (eqn.\,12.3)$$

The kinetic energy (K) and electrostatic potential energy (U) of the electron in hydrogen atom are

$$k = \frac{1}{2}mv^2 = \frac{e^2}{8\pi\varepsilon_o r}$$

and

$$u = -\frac{e^2}{4\pi\varepsilon_o r}$$

(The negative sign in U signifies that the electrostatic force is in the $-r$ direction) Thus the total energy of the electron in a hydrogen atom is

$$E = K + U = \frac{e^2}{8\pi\varepsilon_o r} - \frac{e^2}{4\pi\varepsilon_o r}$$

$$= -\frac{e^2}{8\pi\varepsilon_o r} \qquad (12.4)$$

The total energy of the electron is negative. This implies the fact that the electron is bound to the nucleus. If E were positive, an electron will not follow a closed orbit around the nucleus.

Example 12.3 It is found experimentally that 13.6 eV energy is required to separate a hydrogen atom into a proton and an electron. Compute the orbital radius and the velocity of the electron in a hydrogen atom.

Solution Total energy of the electron in hydrogen atom is -13.6 eV $= -13.6 \times 1.6 \times 10^{-19}$ J $= -2.2 \times 10^{-18}$ J. Thus from Eq. (12.4), we have

$$-\frac{e^2}{8\pi\varepsilon_0 r} = -2.2 \times 10^{-18} \text{ J}$$

This gives the orbital radius

$$r = -\frac{e^2}{8\pi\varepsilon_0 E} = -\frac{(9 \times 10^9 \text{ N m}^2/\text{C}^2)(1.6 \times 10^{-19}\text{C})^2}{(2)(-2.2 \times 10^{-18} \text{ J})}$$

$$= 5.3 \times 10^{-11} \text{ m}.$$

The velocity of the revolving electron can be computed from Eq. (12.3) with $m = 9.1 \times 10^{-31}$ kg,

$$v = \frac{e}{\sqrt{4\pi\varepsilon_0 m r}} = 2.2 \times 10^6 \text{ m/s}.$$

ATOMIC SPECTRA

When an atomic gas or vapour is excited at low pressure, usually by passing an electric current through it, the emitted radiation has a spectrum which contains certain specific wavelengths only. A spectrum of this kind is termed as emission consists of bright lines on a dark background. The spectrum emitted by atomic hydrogen is shown in Fig. 12.5.
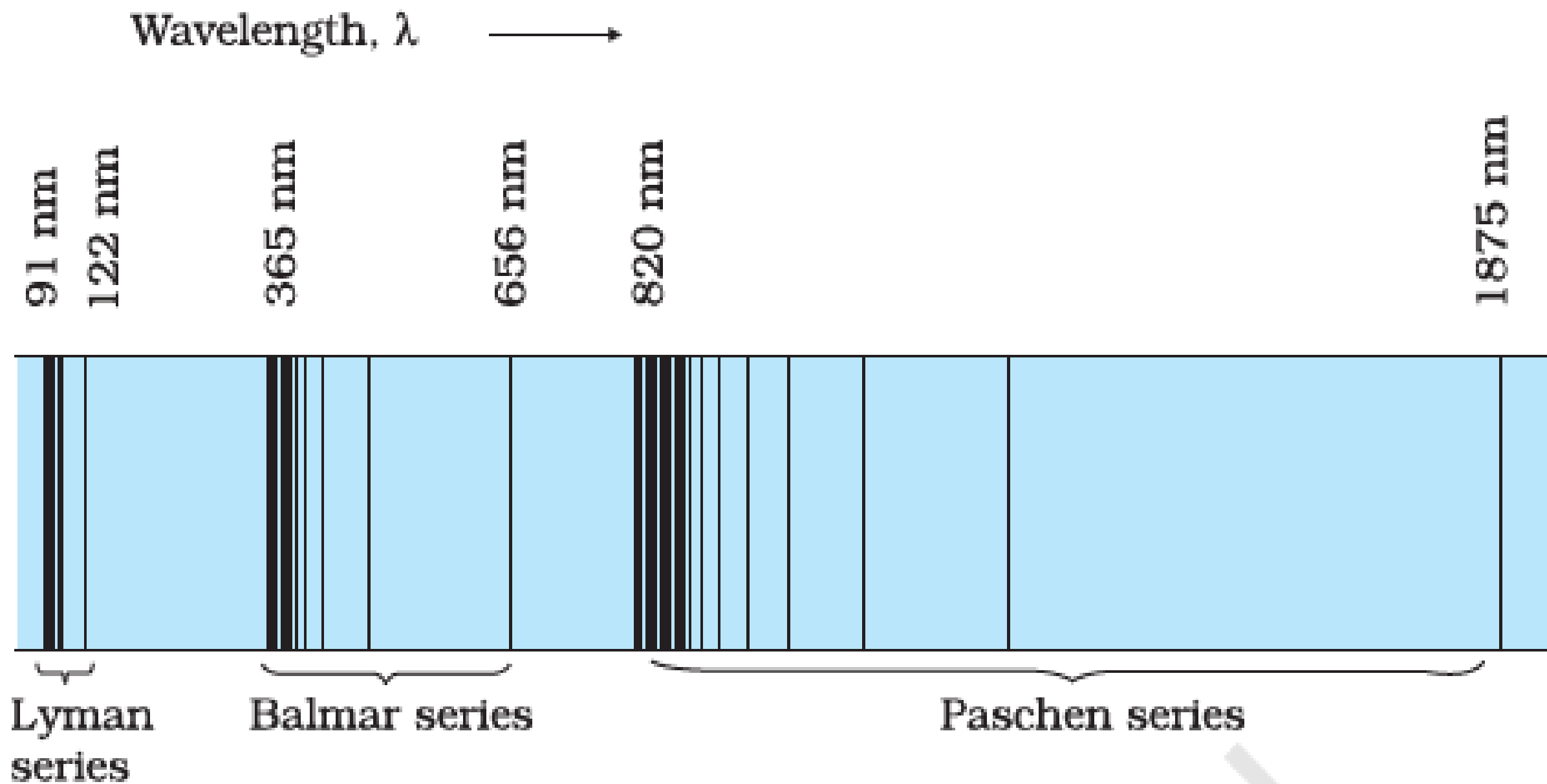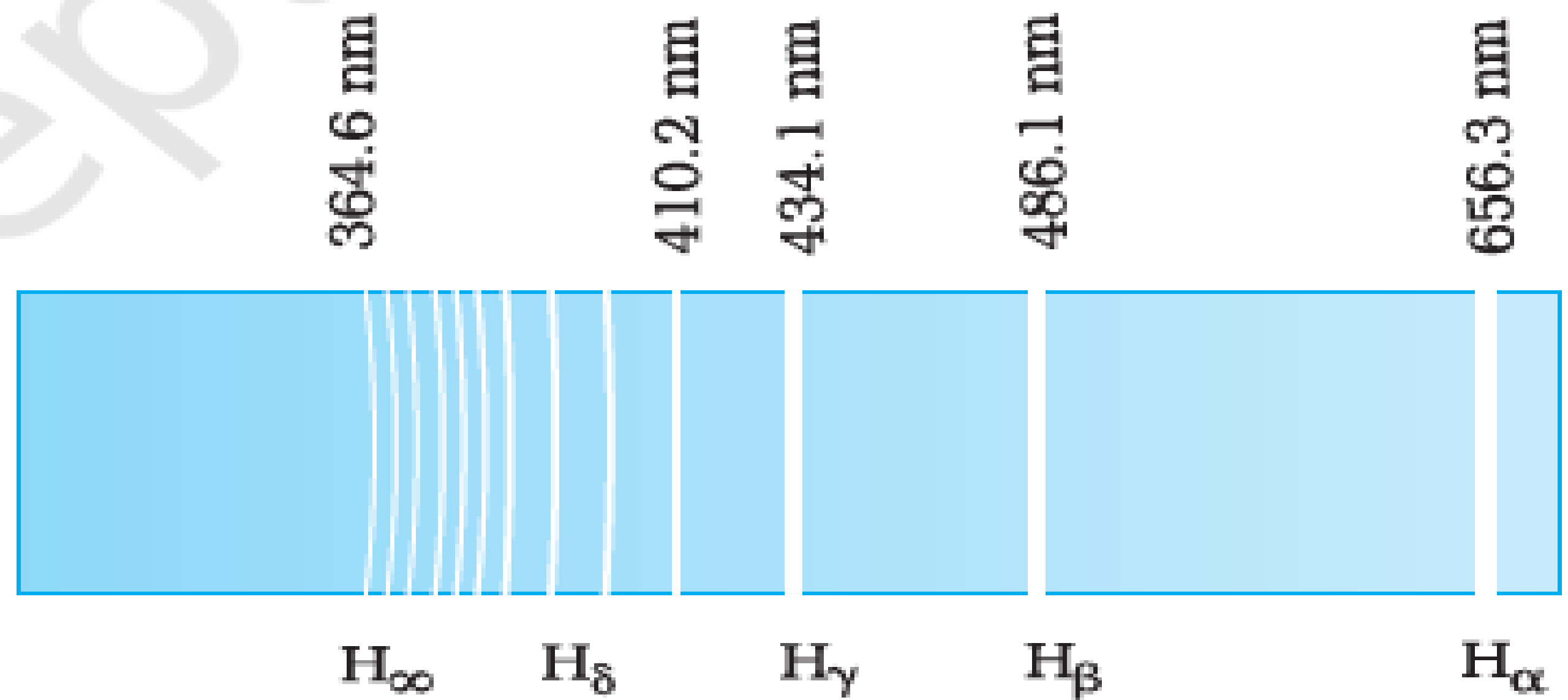
FIGURE 12.5 Emission lines in the spectrum of hydrogen.

When white light passes through a gas and we analyse the transmitted light using a spectrometer we find some dark lines in the spectrum. These dark lines correspond precisely to those wavelengths which were found in the emission line spectrum of the gas. This is called the *absorption spectrum* of the material of the gas.

SPECTRAL SERIES

We might expect that the frequencies of the light emitted by a particular element would exhibit some regular pattern. Hydrogen is the simplest atom and therefore, has the simplest spectrum. In the observed spectrum, however, at first sight, there does not seem to be any resemblance of order or regularity in spectral lines. But the spacing between lines within certain sets of the hydrogen spectrum decreases in a regular way (Fig. 12.5).

Each of these sets is called a *spectral series*. In 1885, the first such series was observed by a Swedish school teacher Johann Jakob Balmer (1825–1898) in the visible region of the hydrogen spectrum. This series is called *Balmer series* (Fig. 12.6).

364.6 nm
410.2 nm
434.1 nm
486.1 nm
656.3 nm

H_∞ H_δ H_γ H_β H_α

FIGURE 12.6 Balmer series in the emission spectrum of hydrogen.

The line with the longest wavelength, 656.3 nm in the red is called H_α; the next line with wavelength 486.1 nm in the bluegreen is called H_β, the third line 434.1 nm in the violet is called H_γ; and so on. As the wavelength decreases, the lines appear closer together and are weaker in intensity.

Balmer found a simple empirical formula for the observed wavelengths

$$\frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \qquad (12.5)$$

where λ is the wavelength, R is a constant called the *Rydberg constant*, and n may have integral values 3, 4, 5, etc. The value of R is 1.097×10^7 m^{-1}. This equation is also called Balmer formula.

Taking $n = 3$ in Eq. (12.5), one obtains the wavelength of the H$_\alpha$ line:

$$\frac{1}{\lambda} = 1.097 \times 10^7 \left(\frac{1}{2^2} - \frac{1}{3^2}\right) \text{m}^{-1}$$

$$= 1.522 \times 10^6 \text{ m}^{-1}$$

i.e., $\lambda = 656.3$ nm

For $n = 4$, one obtains the wavelength of H_β line, etc. For $n = \infty$, one obtains the limit of the series, at $\lambda = 364.6$ nm. This is the shortest wavelength in the Balmer series. Beyond this limit, no further distinct lines appear, instead only a faint continuous spectrum is seen. Other series of spectra for hydrogen were subsequently discovered. These are known, after their discoverers, as Lyman, Paschen, Brackett, and Pfund series. These are represented by the formulae:

Lyman series:

$$\frac{1}{\lambda} = R\left(\frac{1}{1^2} - \frac{1}{n^2}\right) \quad n = 2,3,4... \tag{12.6}$$

Paschen series:

$$\frac{1}{\lambda} = R\left(\frac{1}{3^2} - \frac{1}{n^2}\right) \quad n = 4,5,6... \tag{12.7}$$

Brackett series:

$$\frac{1}{\lambda} = R\left(\frac{1}{4^2} - \frac{1}{n^2}\right) \quad n = 5,6,7... \tag{12.8}$$

Pfund series:

$$\frac{1}{\lambda} = R\left(\frac{1}{5^2} - \frac{1}{n^2}\right) \quad n = 6,7,8... \tag{12.9}$$

The Lyman series is in the ultraviolet, and the Paschen and Brackett series are in the infrared region.

The Balmer formula Eq. (12.5) may be written in terms of frequency of the light, recalling that

$$c = \nu\lambda$$

or $$\frac{1}{\lambda} = \frac{\nu}{c}$$

Thus, Eq. (12.5) becomes

$$\nu = Rc\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \qquad (12.10)$$

BOHR MODEL OF THE HYDROGEN ATOM

The model of the atom proposed by Rutherford assumes that the atom, consisting of a central nucleus and revolving electron is stable much like sun-planet system which the model imitates. However, there are some fundamental differences between the two situations. While the planetary system is held by gravitational force, the nucleus-electron system being charged objects, interact by Coulomb's Law of force.

We know that an object which moves in a circle is being constantly accelerated – the acceleration being centripetal in nature. According to classical electromagnetic theory, an accelerating charged particle emits radiation in the form of electromagnetic waves. The energy of an accelerating electron should therefore, continuously decrease. The electron would spiral inward and eventually fall into the nucleus (Fig. 12.7).
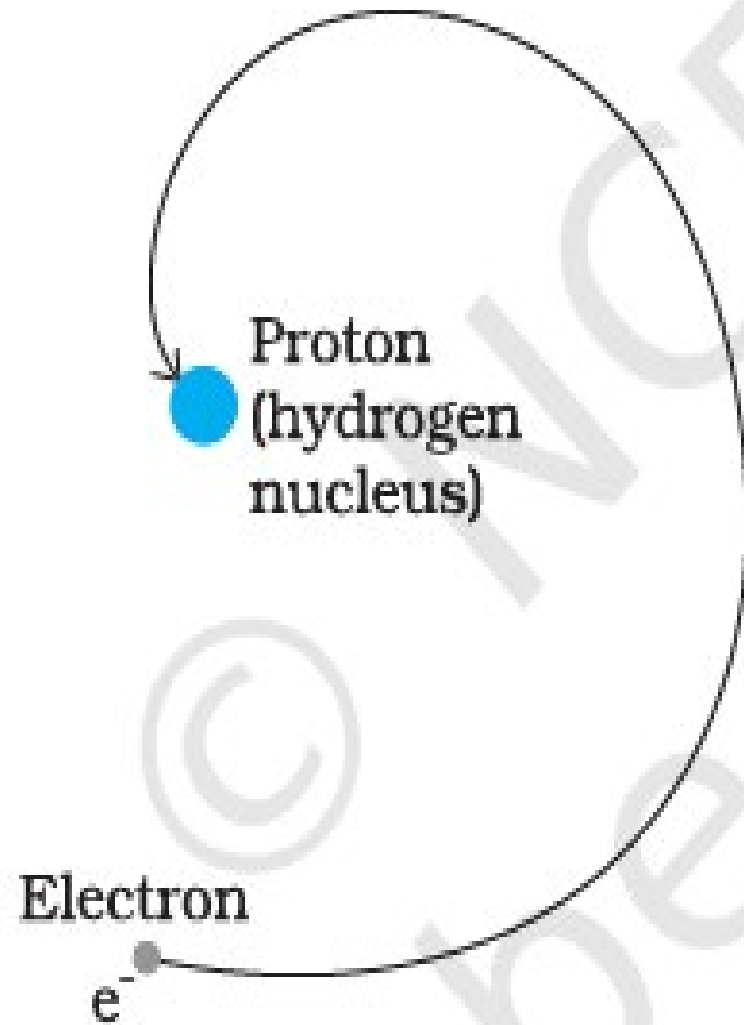
FIGURE 12.7 An accelerated atomic electron must spiral into the nucleus as it loses energy.

Thus, such an atom can not be stable. Further, according to the classical electromagnetic theory, the frequency of the electromagnetic waves emitted by the revolving electrons is equal to the frequency of revolution. As the electrons spiral inwards, their angular velocities and hence their frequencies would change continuously, and so will the frequency of the light emitted. Thus, they would emit a continuous spectrum, in contradiction to the line spectrum actually observed. Clearly Rutherford model tells only a part of the story implying that the classical ideas are not sufficient to explain the atomic structure.

It was Niels Bohr (1885 – 1962) who made certain modifications in this model by adding the ideas of the newly developing quantum hypothesis. Niels Bohr studied in Rutherford's laboratory for several months in 1912 and he was convinced about the validity of Rutherford nuclear model. Faced with the dilemma as discussed above, Bohr, in 1913, concluded that in spite of the success of electromagnetic theory in explaining large-scale phenomena, it could not be applied to the processes at the atomic scale.

It became clear that a fairly radical departure from the established principles of classical mechanics and electromagnetism would be needed to understand the structure of atoms and the relation of atomic structure to atomic spectra. Bohr combined classical and early quantum concepts and gave his theory in the form of three postulates. These are:

(i) Bohr's first postulate was that *an electron in an atom could revolve in certain stable orbits without the emission of radiant energy*, contrary to the predictions of electromagnetic theory. According to this postulate, each atom has certain definite stable states in which it can exist, and each possible state has definite total energy. These are called the stationary states of the atom.

(ii) Bohr's second postulate defines these stable orbits. This postulate states that the *electron* revolves around the nucleus *only in those orbits for which the angular momentum is some integral multiple of $h/2\pi$* where h is the Planck's constant (= 6.6×10^{-34} J s). Thus the angular momentum (L) of the orbiting electron is quantised. That is

$$L = nh/2\pi \qquad (12.11)$$

(iii) Bohr's third postulate incorporated into atomic theory the early quantum concepts that had been developed by Planck and Einstein. It states that *an electron might make a transition from one of its specified non-radiating orbits to another of lower energy. When it does so, a photon is emitted having energy equal to the energy difference between the initial and final states. The frequency of the emitted photon is then given by*

$$hv = E_i - E_f \qquad (12.12)$$

where E_i and E_f are the energies of the initial and final states and $E_i > E_f$.

For a hydrogen atom, Eq. (12.4) gives the expression to determine the energies of different energy states. But then this equation requires the radius r of the electron orbit. To calculate r, Bohr's second postulate about the angular momentum of the electron–the quantisation condition – is used. The angular momentum L is given by

$L = mvr$

Bohr's second postulate of quantisation [Eq. (12.11)] says that the allowed values of angular momentum are integral multiples of $h/2\pi$.

$$L_n = mv_n r_n = \frac{nh}{2\pi} \qquad (12.13)$$

where n is an integer, r_n is the radius of nth possible orbit and v_n is the speed of moving electron in the nth orbit. The allowed orbits are numbered

1, 2, 3 ..., according to the values of n, which is called the *principal quantum number* of the orbit.

From Eq. (12.3), the relation between v_n and r_n is

$$v_n = \frac{e}{\sqrt{4\pi\varepsilon_0 m r_n}}$$

Combining it with Eq. (12.13), we get the following expressions for v_n and r_n,

$$v_n = \frac{1}{n}\frac{e^2}{4\pi\varepsilon_0}\frac{1}{(h/2\pi)} \qquad (12.14)$$

and

$$r_n = \left(\frac{n^2}{m}\right)\left(\frac{h}{2\pi}\right)^2\frac{4\pi\varepsilon_0}{e^2} \qquad (12.15)$$

Eq. (12.14) depicts that the orbital speed in the nth orbit falls by a factor of n. Using Eq. (12.15), the size of the innermost orbit ($n = 1$) can be obtained as

$$r_1 = \frac{h^2\varepsilon_0}{\pi m e^2}$$

This is called the *Bohr radius*, represented by the symbol a_0. Thus,

$$a_0 = \frac{h^2 \varepsilon_0}{\pi m e^2} \qquad (12.16)$$

Substitution of values of h, m, ε_0 and e gives $a_0 = 5.29 \times 10^{-11}$ m. From Eq. (12.15), it can also be seen that the radii of the orbits increase as n^2.

The total energy of the electron in the stationary states of the hydrogen atom can be obtained by substituting the value of orbital radius in Eq. (12.4) as

$$E_n = -\left(\frac{e^2}{8\pi\varepsilon_0}\right)\left(\frac{m}{n^2}\right)\left(\frac{2\pi}{h}\right)^2\left(\frac{e^2}{4\pi\varepsilon_0}\right)$$

or $E_n = -\dfrac{me^4}{8n^2\varepsilon_0^2 h^2} \qquad (12.17)$

Substituting values, Eq. (12.17) yields

$$E_n = -\frac{2.18 \times 10^{-18}}{n^2}\,\text{J} \qquad (12.18)$$

Atomic energies are often expressed in electron volts (eV) rather than joules. Since 1 eV $= 1.6 \times 10^{-19}$ J, Eq. (12.18) can be rewritten as

$$E_n = -\frac{13.6}{n^2}\ \text{eV} \qquad (12.19)$$

The negative sign of the total energy of an electron moving in an orbit means that the electron is bound with the nucleus. Energy will thus be required to remove the electron from the hydrogen atom to a distance infinitely far away from its nucleus (or proton in hydrogen atom).

THE LINE SPECTRA OF THE HYDROGEN ATOM

According to the third postulate of Bohr's model, when an atom makes a transition from the higher energy state with quantum number n_i to the lower energy state with quantum number n_f $(n_f < n_i)$, the difference of energy is carried away by a photon of frequency v_{if} such that

Equation 12.7

$$h\nu_{if} = E_{n_i} - E_{n_f} \qquad (12.20)$$

Using Eq. (12.16), for E_{n_f} and E_{n_i}, we get

$$h\nu_{if} = \frac{me^4}{8e_0^2 h^2}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \qquad (12.21)$$

$$\text{or } \nu_{if} = \frac{me^4}{8e_0^2 h^2}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \qquad (12.22)$$

Equation (12.21) is the Rydberg formula, for the spectrum of the hydrogen atom. In this relation, if we take $n_f = 2$ and $n_i = 3, 4, 5...$, it reduces to a form similar to Eq. (12.10) for the Balmer series. The Rydberg constant R is readily identified to be

$$R = \frac{me^4}{8\varepsilon_0^2 h^3 c} \qquad (12.23)$$

If we insert the values of various constants in Eq. (12.23), we get

$$R = 1.03 \times 10^7 \text{ m}^{-1}$$

This is a value very close to the value $(1.097 \times 10^7 \text{ m}^{-1})$ obtained from the empirical Balmer formula. This agreement between the theoretical and experimental values of the Rydberg constant provided a direct and striking confirmation of the Bohr's model.

Since both n_f and n_i are integers, this immediately shows that in transitions between different atomic levels, light is radiated in various discrete frequencies. For hydrogen spectrum, the Balmer formula corresponds to $n_f = 2$ and $n_i = 3, 4, 5$, etc. The results of the Bohr's model suggested the presence of other series spectra for hydrogen atom–those corresponding to transitions resulting from $n_f = 1$ and $n_i = 2, 3$, etc.; $n_f = 3$ and $n_i = 4, 5$, etc., and so on. Such series were identified in the course of spectroscopic investigations and are known as the Lyman, Balmer, Paschen, Brackett, and Pfund series. The electronic transitions corresponding to these series are shown in Fig. 12.9.
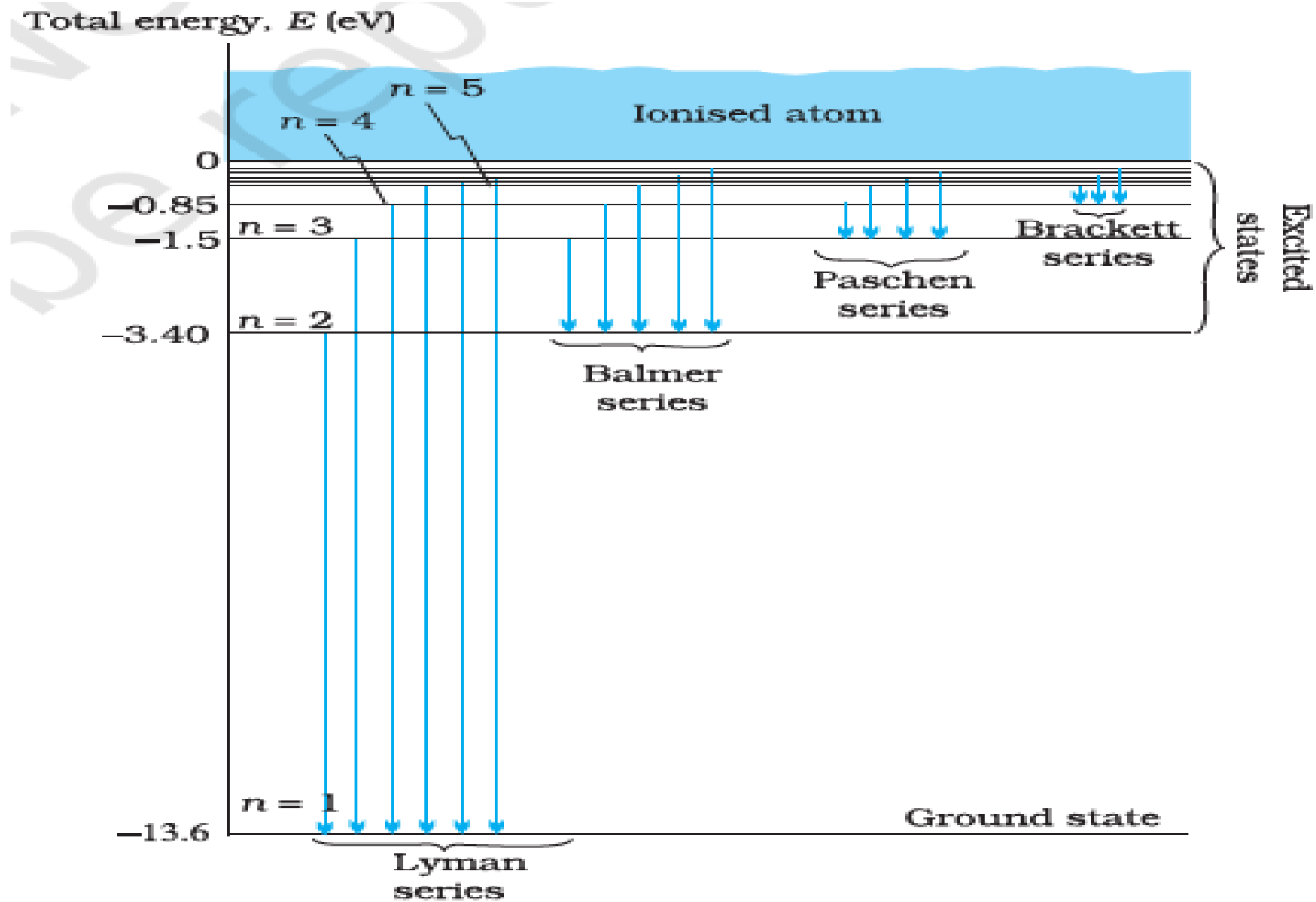
FIGURE 12.9 Line spectra originate in transitions between energy levels.

Example 12.6 Using the Rydberg formula, calculate the wavelengths of the first four spectral lines in the Lyman series of the hydrogen spectrum.

Solution The Rydberg formula is

$$hc/\lambda_{if} = \frac{me^4}{8\varepsilon_0^2 h^2}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

The wavelengths of the first four lines in the Lyman series correspond to transitions from $n_i = 2,3,4,5$ to $n_f = 1$. We know that

$$\frac{me^4}{8\varepsilon_0^2 h^2} = 13.6 \text{ eV} = 21.76 \times 10^{-19} \text{ J}$$

Therefore,

$$\lambda_{i1} = \frac{hc}{21.76 \times 10^{-19}\left(\frac{1}{1} - \frac{1}{n_i^2}\right)} \text{ m}$$

$$= \frac{6.625 \times 10^{-34} \times 3 \times 10^8 \times n_i^2}{21.76 \times 10^{-19} \times (n_i^2 - 1)} \text{ m} = \frac{0.9134 \, n_i^2}{(n_i^2 - 1)} \times 10^{-7} \text{ m}$$

$$= 913.4 \, n_i^2/(n_i^2 - 1) \text{ Å}$$

Substituting $n_i = 2,3,4,5$, we get $\lambda_{21} = 1218$ Å, $\lambda_{31} = 1028$ Å, $\lambda_{41} = 974.3$ Å, and $\lambda_{51} = 951.4$ Å.

THEORY OF ENERGY BANDS IN CONDUCTORS, INSULATORS AND SEMICONDUCTORS

MATERIALS USED IN ELECTRONIC DEVICES

Based on the electrical properties and conductivity all the materials in nature are classified as conductors, semiconductors, and insulators. When atoms combine to form a solid, crystalline material, they arrange themselves in a symmetrical/regular pattern. The atoms within a semiconductor crystal structure are held together by covalent bonds, which are created by the interaction of the valence electrons of the atoms. Silicon is a crystalline material.

Electronic devices such as diodes, transistors, and integrated circuits are made of a semi-conductive material. To understand how these devices work, you should have a basic knowledge of the structure of atoms and the interaction of atomic particles.

INSULATOR

An insulator is a material that offers a very low level (or negligible) of conductivity when voltage is applied. E.g.: Rubber, Plastics, Paper, Mica, glass, quartz. Typical resistivity level of an insulator is of the order of 10^{10} to $10^{12}\Omega$-cm. An insulator is a material that does not conduct electrical current under normal conditions.

Insulators also have many electrons in them, but the electrons cannot move. Some of them are trapped in individual atoms and can't get away from them. Others are nominally free to move about but are locked in place by "gridlock/congestion." There are so many electrons in so much that there is no place for an electron to move, so it stays even in an electric field:
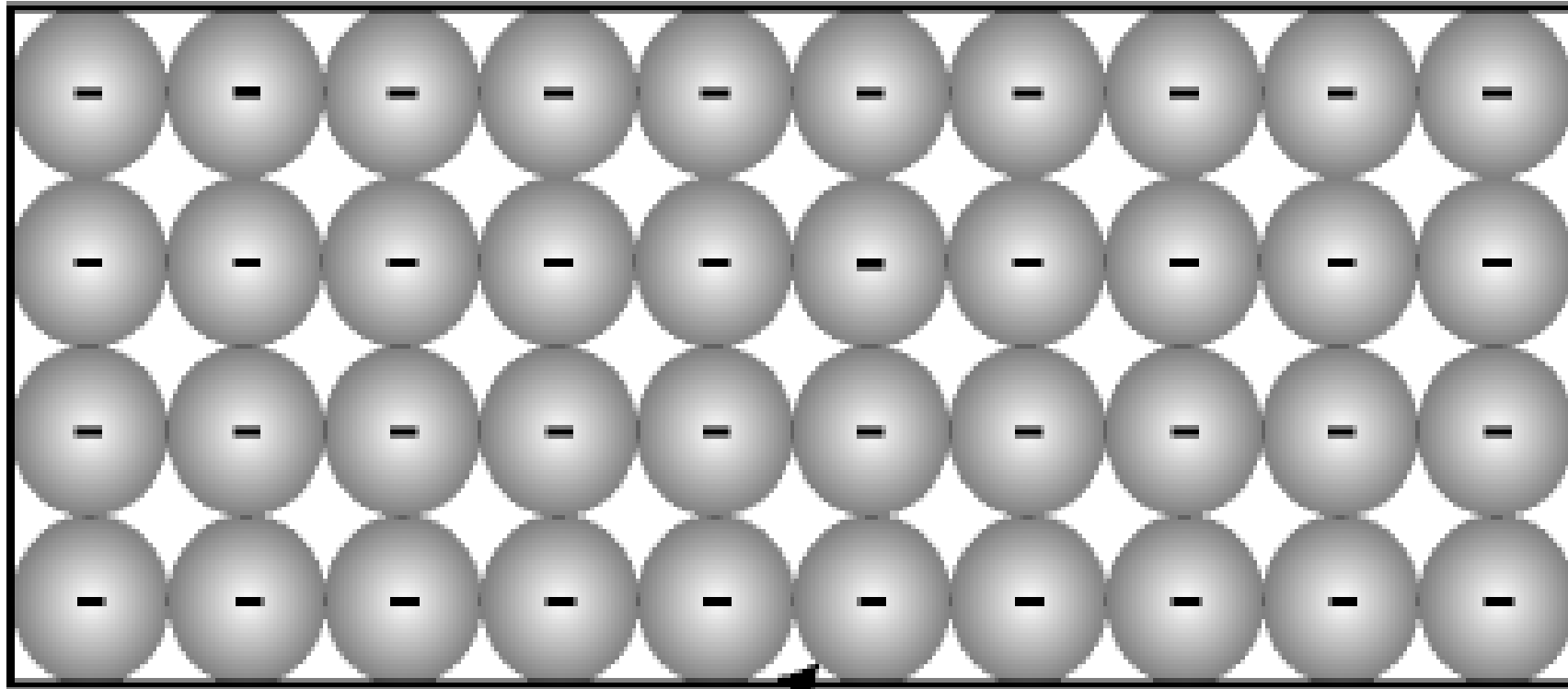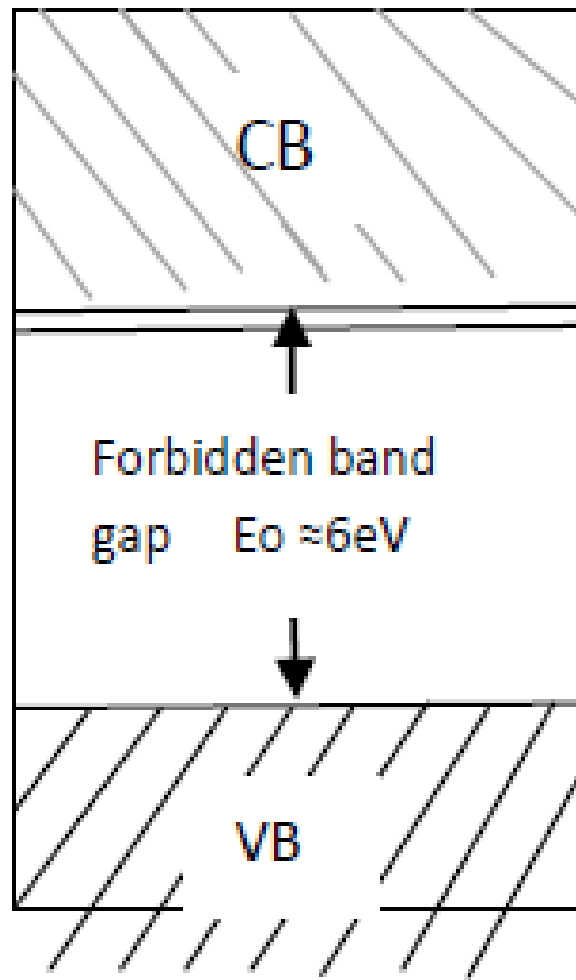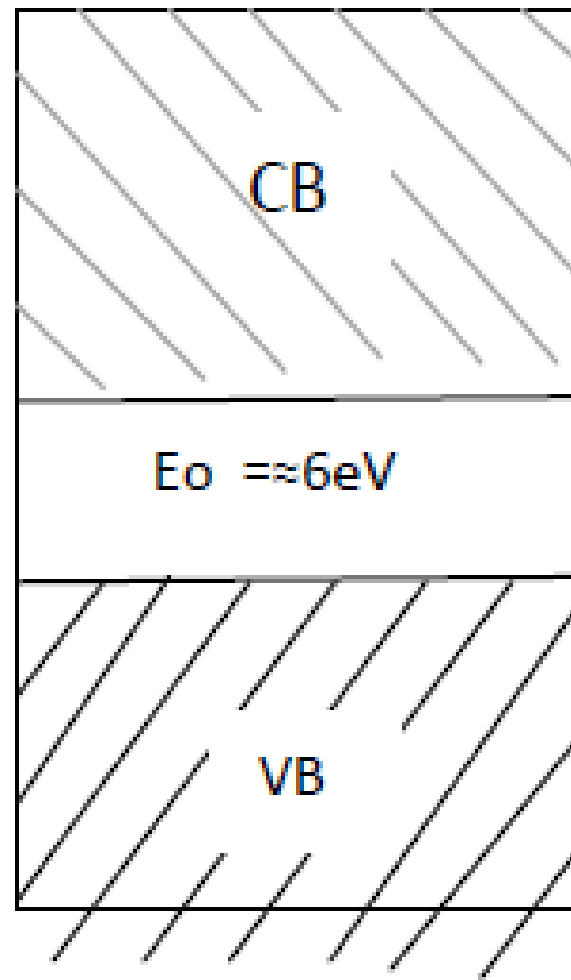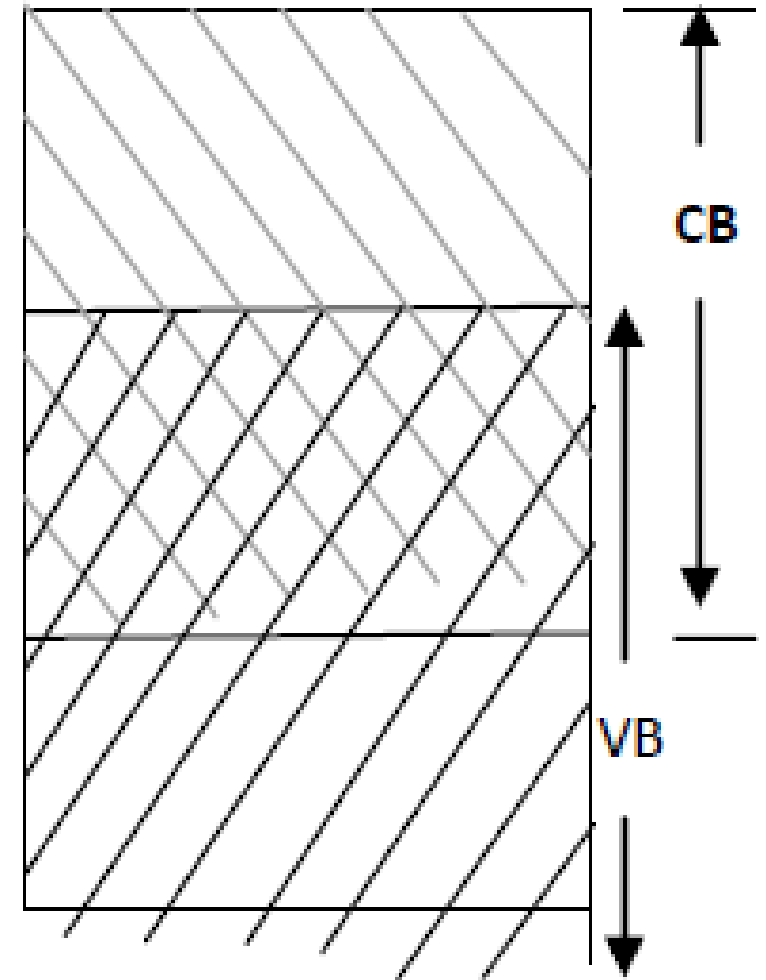
Insulator   Semiconductor   Conductor

FiG:1.1 Energy band diagrams insulator, semiconductor and conductor

Most good insulators are compounds rather than single-element materials and have very high resistivities. Valence electrons are tightly bound to the atoms; therefore, there are very few free electrons in an insulator. The energy band structure of an insulator is shown in the fig.1.1. **Band structure of a material defines the band of energy levels that an electron can occupy**. Valance band is the range of electron energy where the electron remain bounded to the atom and do not contribute to the electric current.

Conduction band (CB) is the range of electron energies higher than valance band where electrons are free to accelerate under the influence of external voltage source resulting in the flow of charge. The band gap between the valance band and conduction band is called a forbidden energy gap. *Forbidden energy band is the energy required by an electron to move from valance band to conduction band i.e. the energy required for a valance electron to become a free electron.*

The following are the characteristics of Insulators.

- The Forbidden energy gap is very large.

- Valance band electrons are bound tightly to atoms.

- The value of forbidden energy gap for an insulator will be of 10eV.

- For some insulators, as the temperature increases, they might show some conduction.

- The resistivity of an insulator will be in the order of 10^7 ohm-meter.

ENERGY BAND DESCRIPTION OF SEMICONDUCTORS

It has already been discussed that a semiconductor is a substance whose resistivity lies between conductors and insulators. The resistivity is of the order of 10^{-4} to 0.5 ohm metre. However, a semiconductor can be defined much more comprehensively on the basis of energy bands as under:
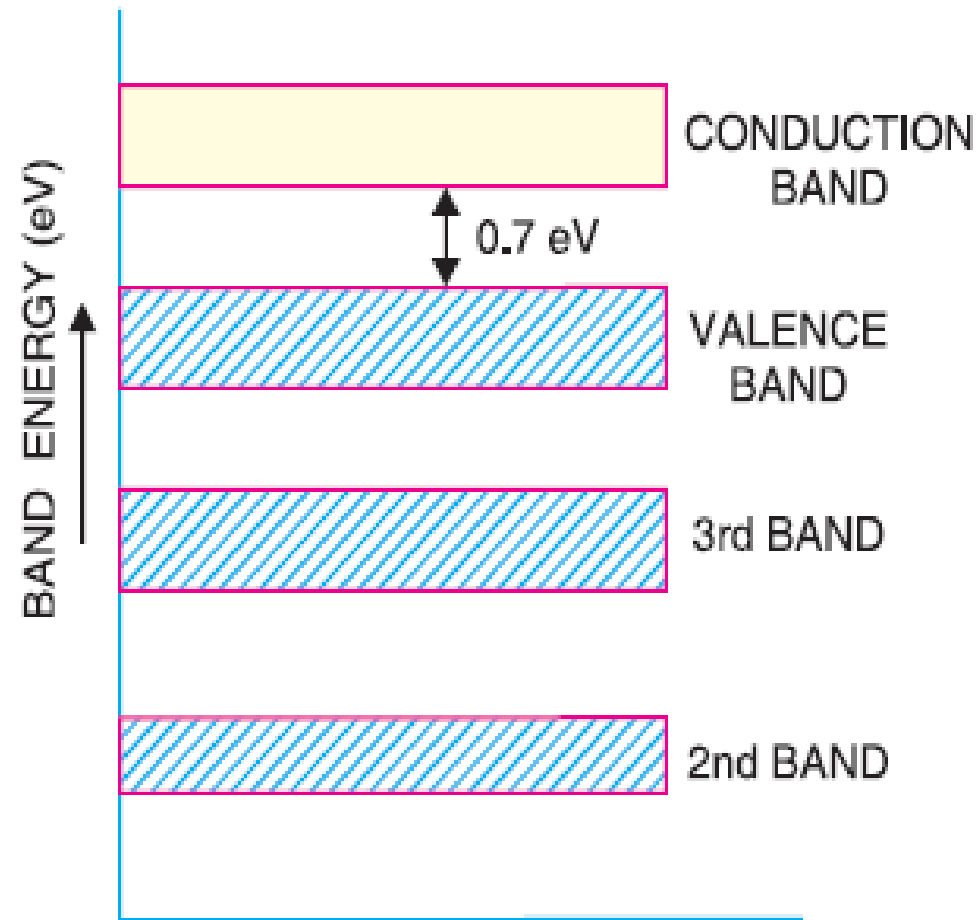
Fig. 5.4
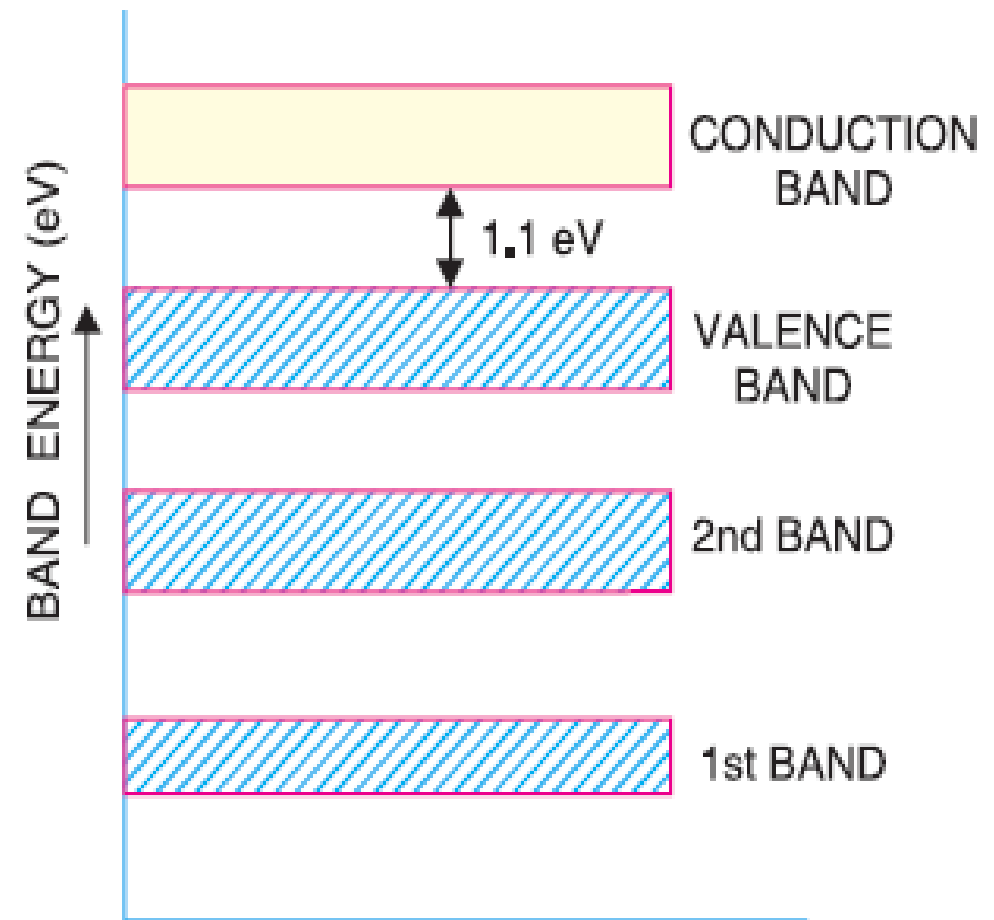
Fig. 5.5

A **semiconductor** *is a substance which has almost filled valence band and nearly empty conduction band with a very small energy gap (≈1 eV) separating the two.*

Figs. 5.4 and 5.5 show the energy band diagrams of germanium and silicon respectively. It may be seen that forbidden energy gap is very small; being 1.1 eV for silicon and 0.7 eV for germanium. Therefore, relatively small energy is needed by their valence electrons to cross over to the conduction band.

At room temperature, some of the valence electrons may acquire sufficient energy to enter into the conduction band and thus become free electrons. However, at this temperature, the number of free electrons available is very small. Therefore, at room temperature, a piece of germanium or silicon is neither a good conductor nor an insulator. For this reason, such substances are called *semiconductors*.

BAND GAP

In solid materials, interactions between atoms "smear" the valence shell into a band of energy levels called the *valence band*. Valence electrons are confined to that band. When an electron acquires enough additional energy, it can leave the valence shell, become a *free electron,* and exist in what is known as the *conduction band.*

The difference in energy between the valence band and the conduction band is called an *energy gap* or **band gap.** This is the amount of energy that a valence electron must have in order to jump from the valence band to the conduction band. Once in the conduction band, the electron is free to move throughout the material and is not tied to any given atom.
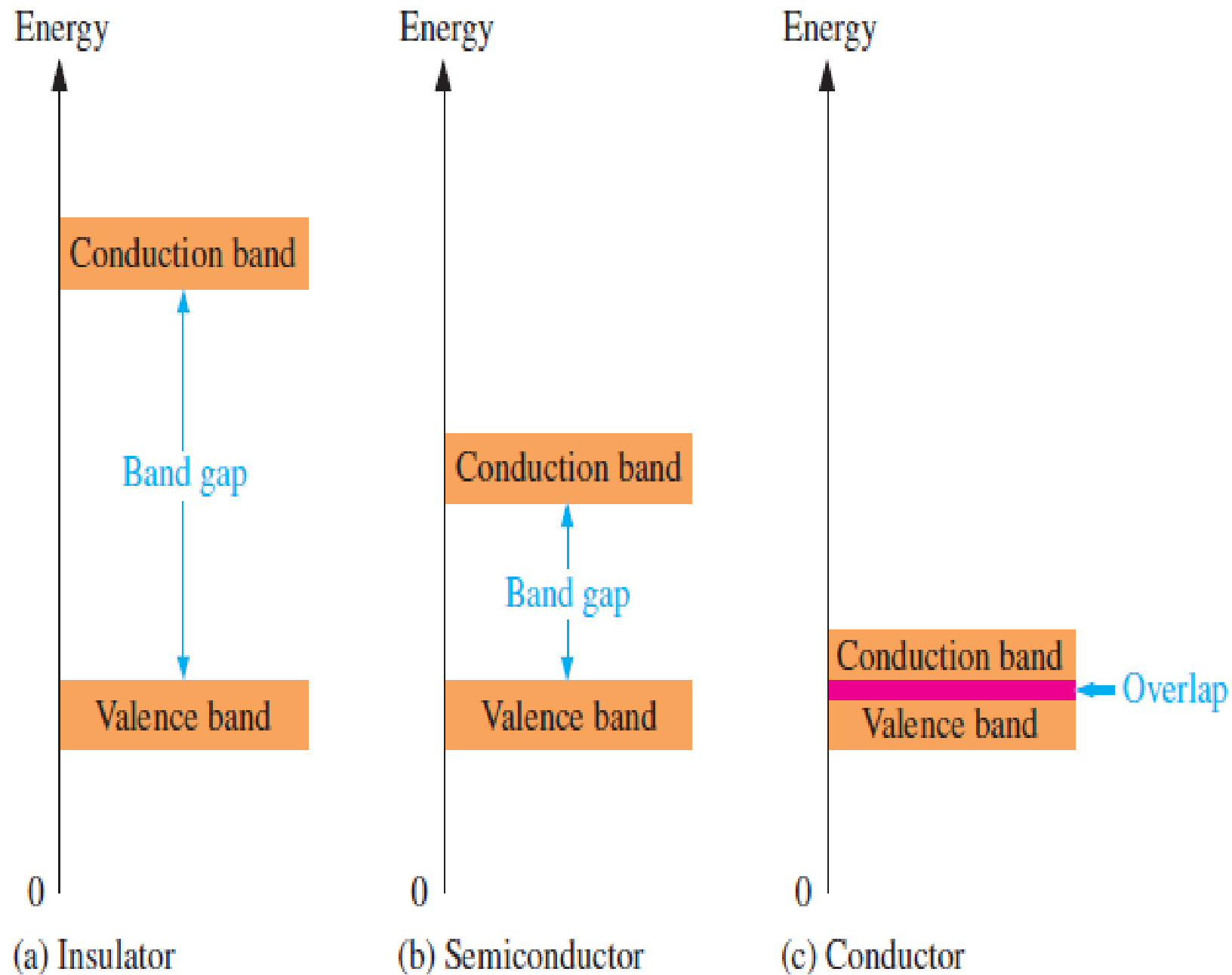
◄ **FIGURE 1–7**

Energy diagrams for the three types of materials.

Figure 1–7 shows energy diagrams for insulators, semiconductors, and conductors. The energy gap or band gap is the difference between two energy levels and electrons are "not allowed" in this energy gap based on quantum theory. Although an electron may not exist in this region, it can "jump" across it under certain conditions. For insulators, the gap can be crossed only when breakdown conditions occur—as when a very high voltage is applied across the material.

The band gap is illustrated in Figure 1–7(a) for insulators. In semiconductors the band gap is smaller, allowing an electron in the valence band to jump into the conduction band if it absorbs a photon. The band gap depends on the semiconductor material. This is illustrated in Figure 1–7(b). In conductors, the conduction band and valence band overlap, so there is no gap, as shown in Figure 1–7(c). This means that electrons in the valence band move freely into the conduction band, so there are always electrons available as free electrons.

VALANCE BAND

The electrons move in the atoms in certain energy levels but the energy of the electrons in the innermost shell is higher than the outermost shell electrons. The electrons that are present in the outermost shell are called as **Valance Electrons**. These valance electrons, containing a series of energy levels, form an energy band which is called as **Valence Band**. The valence band is the *band having the highest occupied energy*.

CONDUCTION BAND

The valence electrons are so loosely attached to the nucleus that even at room temperature, few of the valence electrons leave the band to be free. These are called as **free electrons** as they tend to move towards the neighbouring atoms. These free electrons are the ones which conduct the current in a conductor and hence called as **Conduction Electrons**. The band which contains conduction electrons is called as **Conduction Band**. The conduction band is the *band having the lowest occupied energy*.

FORBIDDEN GAP

The gap between valence band and conduction band is called as **forbidden energy gap**. As the name implies, this band is the forbidden one without energy. Hence no electron stays in this band. The valence electrons, while going to the conduction band, pass through this. The forbidden energy gap if greater, means that the valence band electrons are tightly bound to the nucleus. Now, in order to push the electrons out of the valence band, some external energy is required, which would be equal to the forbidden energy gap. The following figure shows the valance band, conduction band, and the forbidden gap.
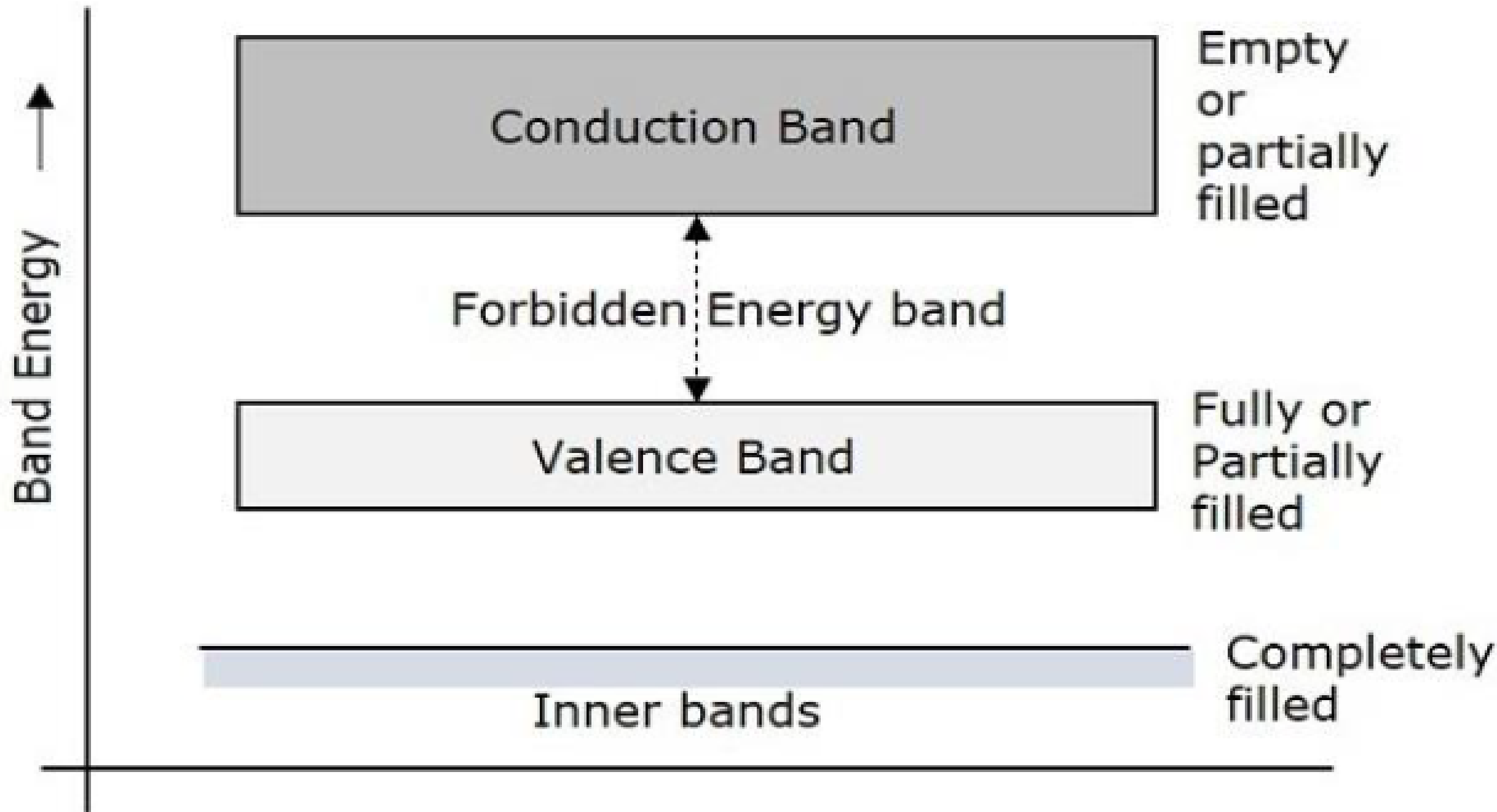
CONDUCTORS

A conductor is a material which supports a generous flow of charge when a voltage is applied across its terminals. i.e. it has a very high conductivity. OR a **conductor** is a material that easily conducts electrical current. Most metals are good conductors The best conductors are single-element materials, such as Copper, Aluminium, Silver, Gold, which are characterized by atoms with only one valence electron very loosely bound to the atom.

These loosely bound valence electrons can become free electrons with the addition of a small amount of energy to free them from the atom. Therefore, in a conductive material the free electrons are available to carry current. The resistivity of a conductor is in the order of 10^{-4} and 10^{-6} Ω-cm. The Valance and conduction bands overlap (Fig. 1-7) and there is no energy gap for the electrons to move from valance band to conduction band. This implies that there are free electrons in CB even at absolute zero temperature (0K). Therefore at room temperature when electric field is applied large current flows through the conductor.

CONDUCTION IN METALS

Metals are filled with electrons. Many of these, typically one or two per atom in the metal, are free to move about throughout the metal. When an electric field is applied, the electrons move in the direction *opposite* the field. Since they are negatively charged, this corresponds to a positive current in the direction *opposite* the motion, that is, in the direction of the electric field:
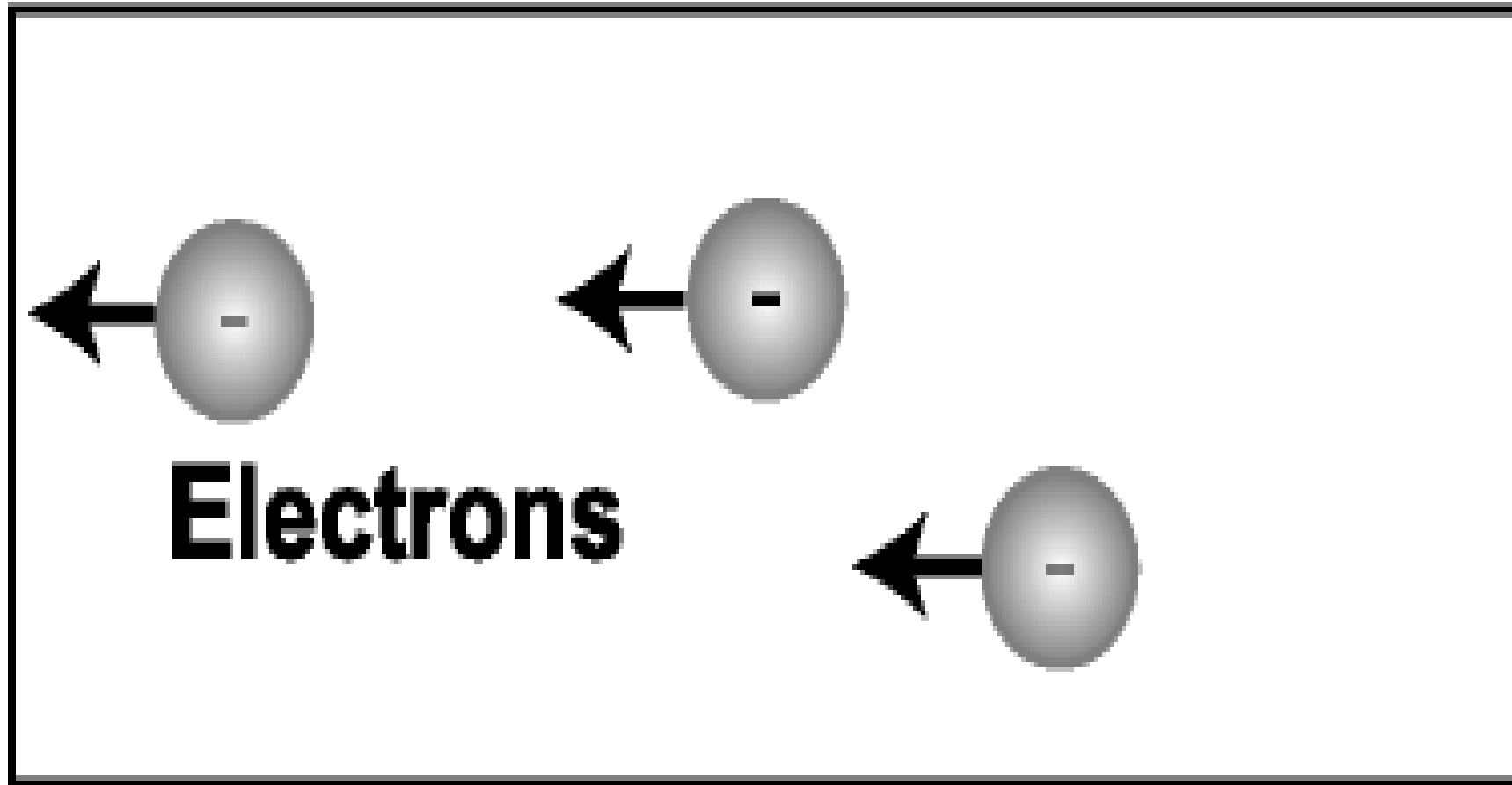
Conductors are such materials in which the forbidden energy gap disappears as the valence band and conduction band become very close that they overlap. Examples: The figure below shows the structure of energy bands in conductors.

Energy bands in Conductors
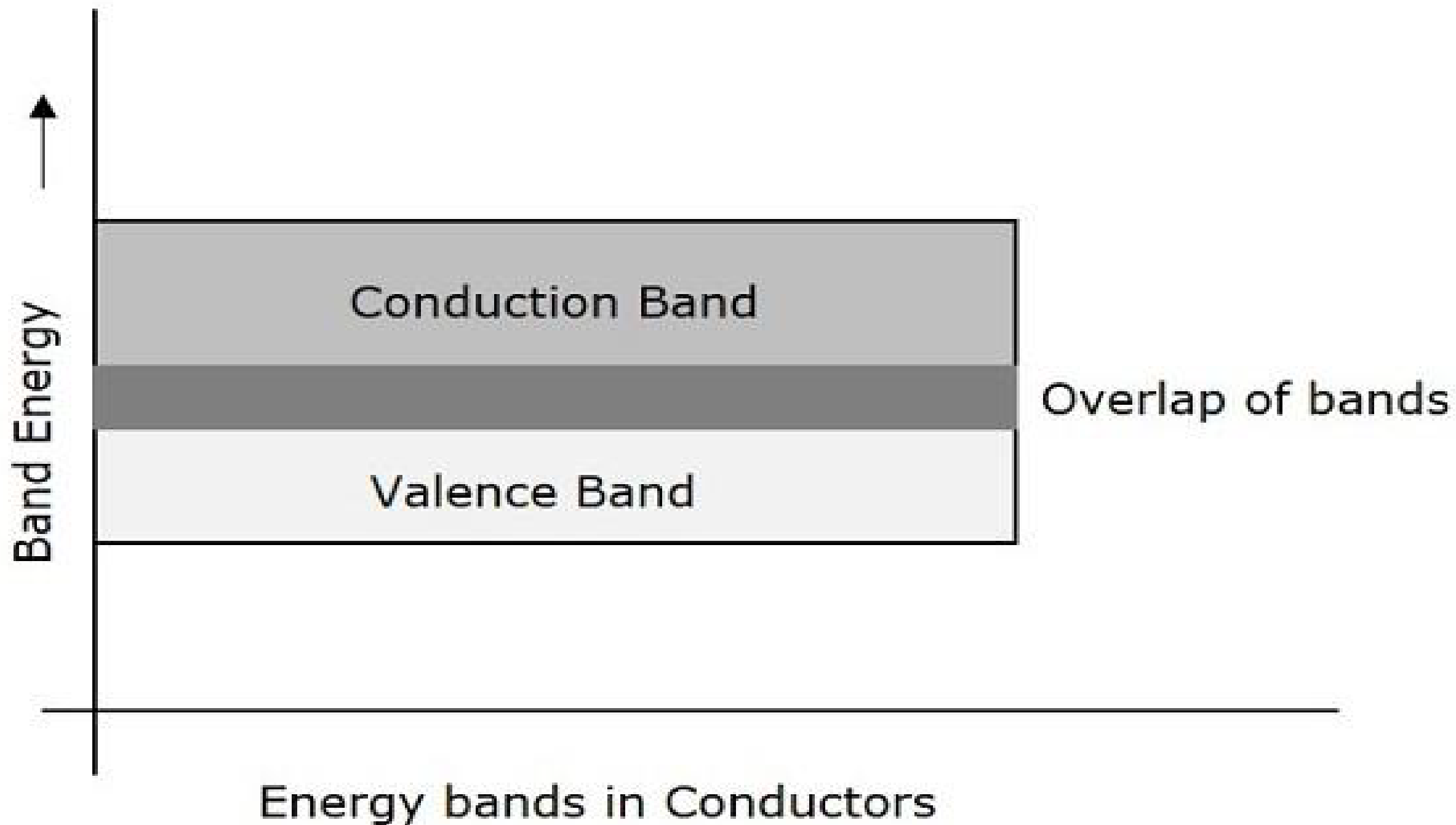
The following are the characteristics of Conductors.

1. There exists no forbidden gap in a conductor.

2. The valance band and the conduction band gets overlapped.

3. The free electrons available for conduction are plenty.

4. A slight increase in voltage, increases the conduction.

5. There is no concept of hole formation, as a continuous flow of electrons contribute the current.

SEMICONDUCTOR

A **semiconductor** is a material which has electrical conductivity to a degree between that of a metal (such as copper) and that of an insulator (such as glass). Semiconductors are the foundation of modern electronics, including transistors, solar cells, light-emitting diodes (LEDs), quantum dots and digital and analog integrated circuits. A semiconductor is a material that has its conductivity somewhere between the insulator and conductor. Two of the most commonly used are Silicon (Si=14 atomic no.) and germanium (Ge=32 atomic no.).

Both have 4 valance electrons. A semiconductor in its pure (intrinsic) state is neither a good conductor nor a good insulator. Single-element semiconductors are antimony (Sb), arsenic (As), astatine (At), boron (B), polonium (Po), tellurium (Te), silicon (Si), and germanium (Ge). Compound semiconductors such as gallium arsenide, indium phosphide, gallium nitride, silicon carbide, and silicon germanium are also commonly used.

The single-element semiconductors are characterized by atoms with four valence electrons. *Silicon is the most commonly used semiconductor. The forbidden band gap is in the order of 1eV*. For e.g., the band gap energy for Si, Ge and GaAs is 1.21, 0.785 and 1.42 eV, respectively at absolute zero temperature (0K). At 0K and at low temperatures, the valance band electrons do not have sufficient energy to move from VB to CB. Thus semiconductors act as insulators at 0K.

As the temperature increases, a large number of valance electrons acquire sufficient energy to leave the VB, cross the forbidden bandgap and reach CB. These are now free electrons as they can move freely under the influence of electric field. At room temperature there are sufficient electrons in the CB and hence the semiconductor is capable of conducting some current at room temperature. Inversely related to the conductivity of a material is its resistance to the flow of charge or current.$[\sigma = \frac{1}{R}]$. Typical resistivity values for various materials' are given as follows.

Insulator	Semiconductor	Conductor
10^{-6} Ω-cm (Cu)	50Ω-cm (Ge)	10^{12} Ω-cm (mica)
	50×10^{3} Ω-cm (Si)	

Typical resistivity values

PROPERTIES OF SEMICONDUCTORS

(*i*) The resistivity of a semiconductor is less than an insulator but more than a conductor.

(*ii*) Semiconductors have *negative temperature co-efficient of resistance i.e.* the resistance of a semiconductor decreases with the increase in temperature and *vice-versa*. For example, germanium is actually an insulator at low temperatures but it becomes a good conductor at high temperatures.

(*iii*) When a suitable metallic impurity (*e.g.* arsenic, gallium etc.) is added to a semiconductor, its current conducting properties change appreciably.

COMPARISON OF SILICON AND GERMANIUM

The atomic structures of silicon and germanium are compared in Figure 1–9. **Silicon** is used in diodes, transistors, integrated circuits, and other semiconductor devices. Notice that both silicon and **germanium** have the characteristic four valence electrons.

Diagrams of the silicon and germanium atoms.

Four valence electrons in the outer (valence) shell

+14

Silicon atom

+32
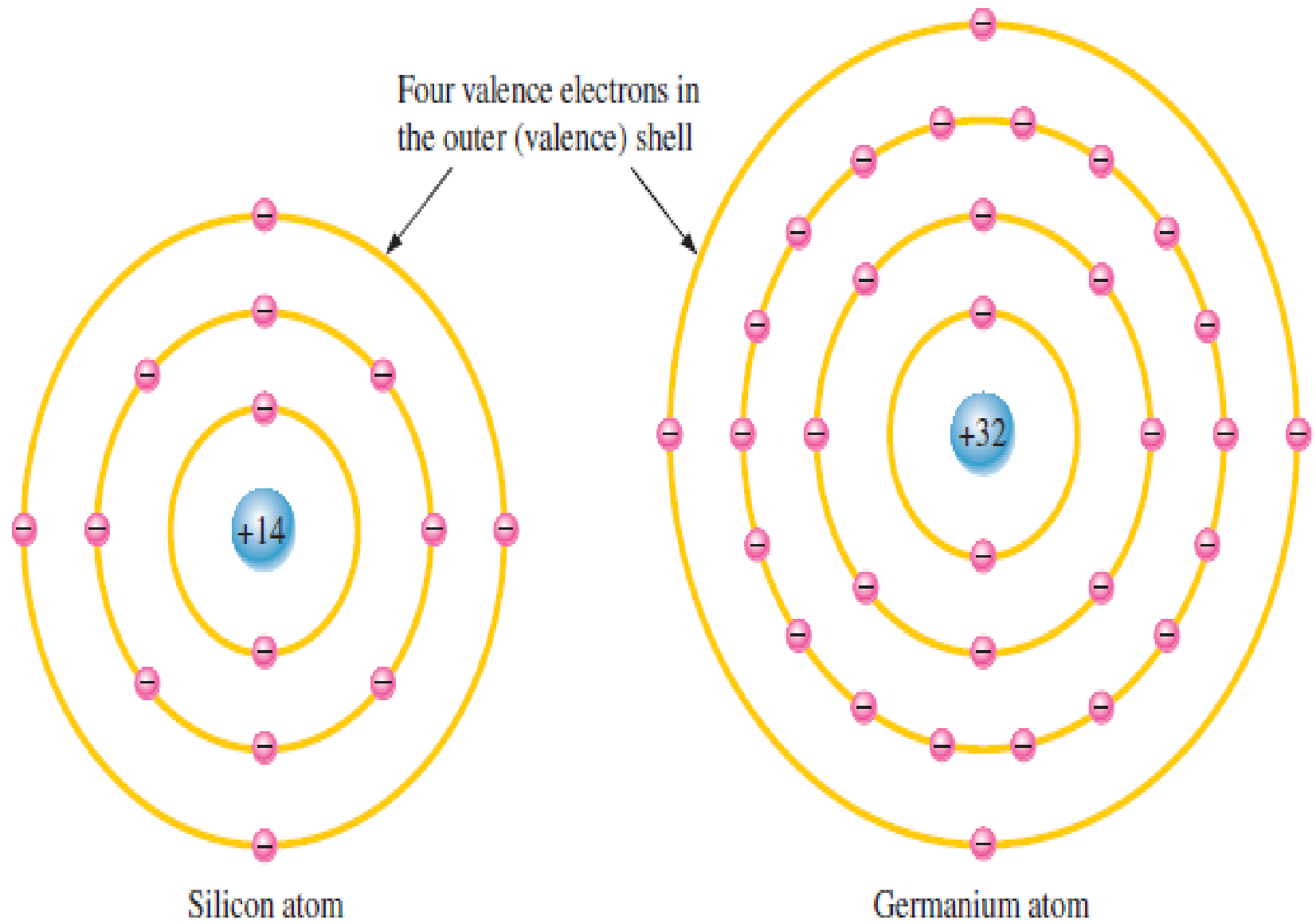
Germanium atom

The valence electrons in germanium are in the fourth shell while those in silicon are in the third shell, closer to the nucleus. <u>This means that the germanium valence electrons are at higher energy levels than those in silicon and, therefore, require a smaller amount of additional energy to escape from the atom.</u> *This property makes germanium more unstable at high temperatures and results in excessive reverse current*. This is why silicon is a more widely used semi conductive material.

Illustration of covalent bonds in silicon.
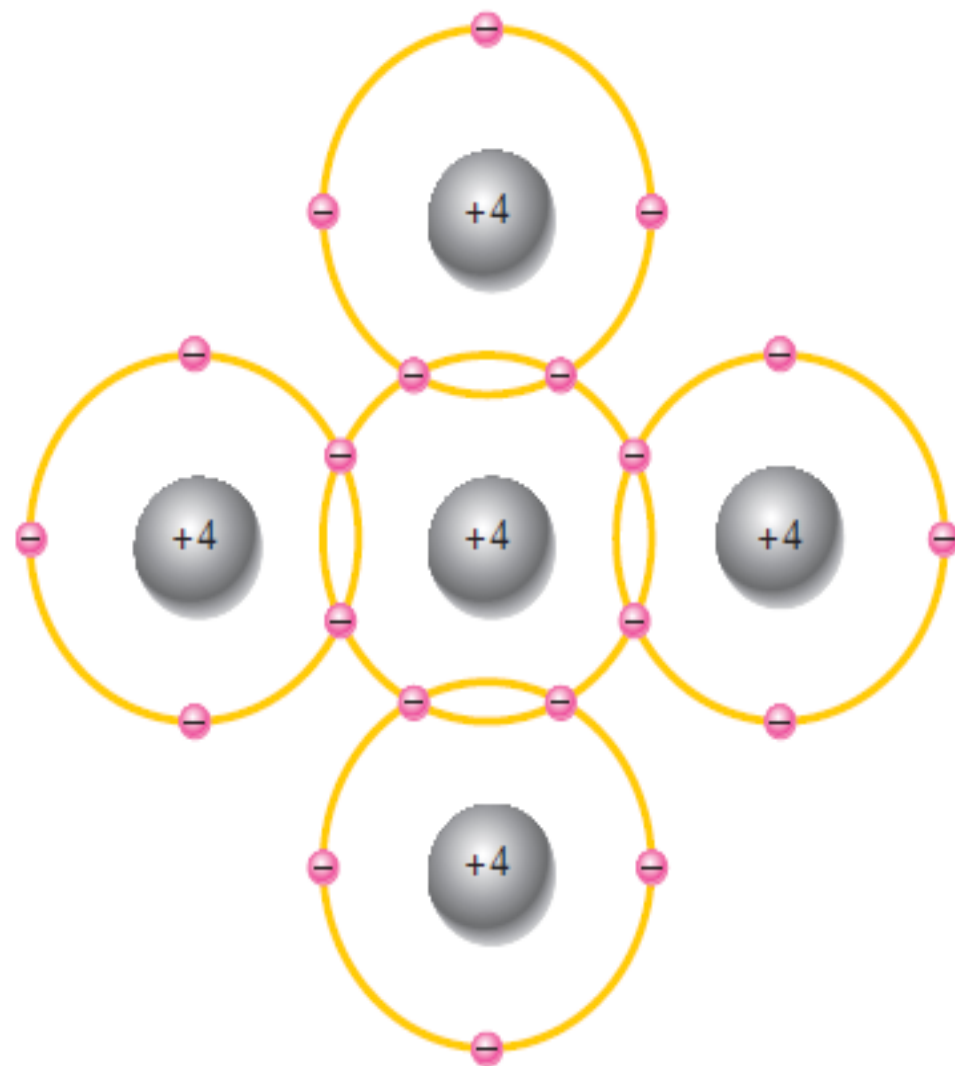
(a) The center silicon atom shares an electron with each of the four surrounding silicon atoms, creating a covalent bond with each. The surrounding atoms are in turn bonded to other atoms, and so on.
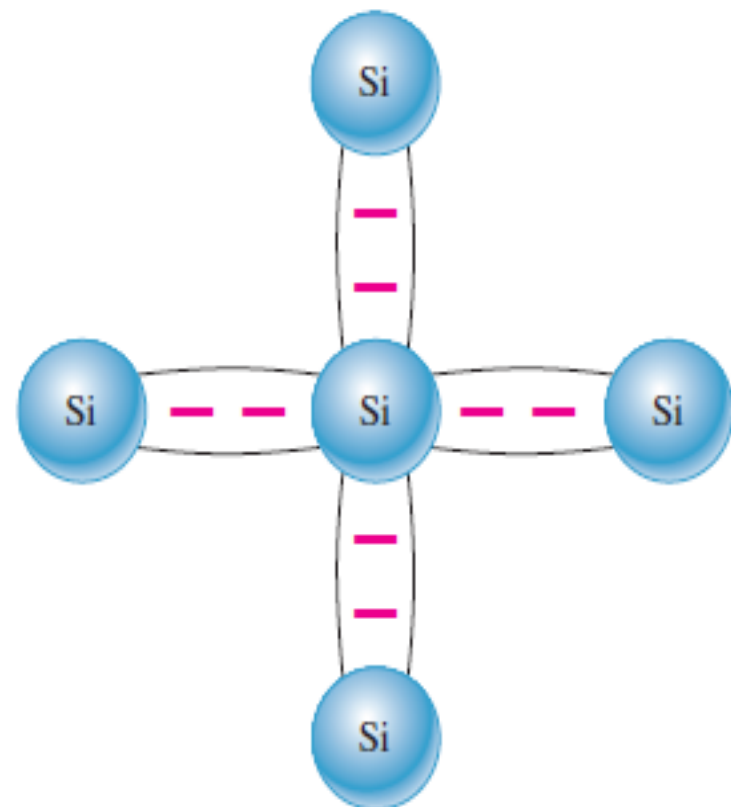
(b) Bonding diagram. The red negative signs represent the shared valence electrons.

COVALENT BONDS

Figure 1–10 shows how each silicon atom positions itself with four adjacent silicon atoms to form a silicon **crystal**, which is a three-dimensional symmetrical arrangement of atoms. A silicon (Si) atom with its four valence electrons shares an electron with each of its four neighbours.

BONDS IN SEMICONDUCTORS

The atoms of every element are held together by the bonding action of valence electrons. This bonding is due to the fact that it is the tendency of each atom to complete its last orbit by acquiring 8 electrons in it.

However, in most of the substances, the last orbit is incomplete *i.e.* the last orbit does not have 8 electrons. This makes the atom active to enter into bargain with other atoms to acquire 8 electrons in the last orbit. To do so, the atom may lose, gain or share valence electrons with other atoms. In semiconductors, bonds are formed by sharing of valence electrons. Such bonds are called *co-valent bonds*. In the formation of a co-valent bond, each atom contributes equal number of valence electrons and the contributed electrons are shared by the atoms engaged in the formation of the bond.
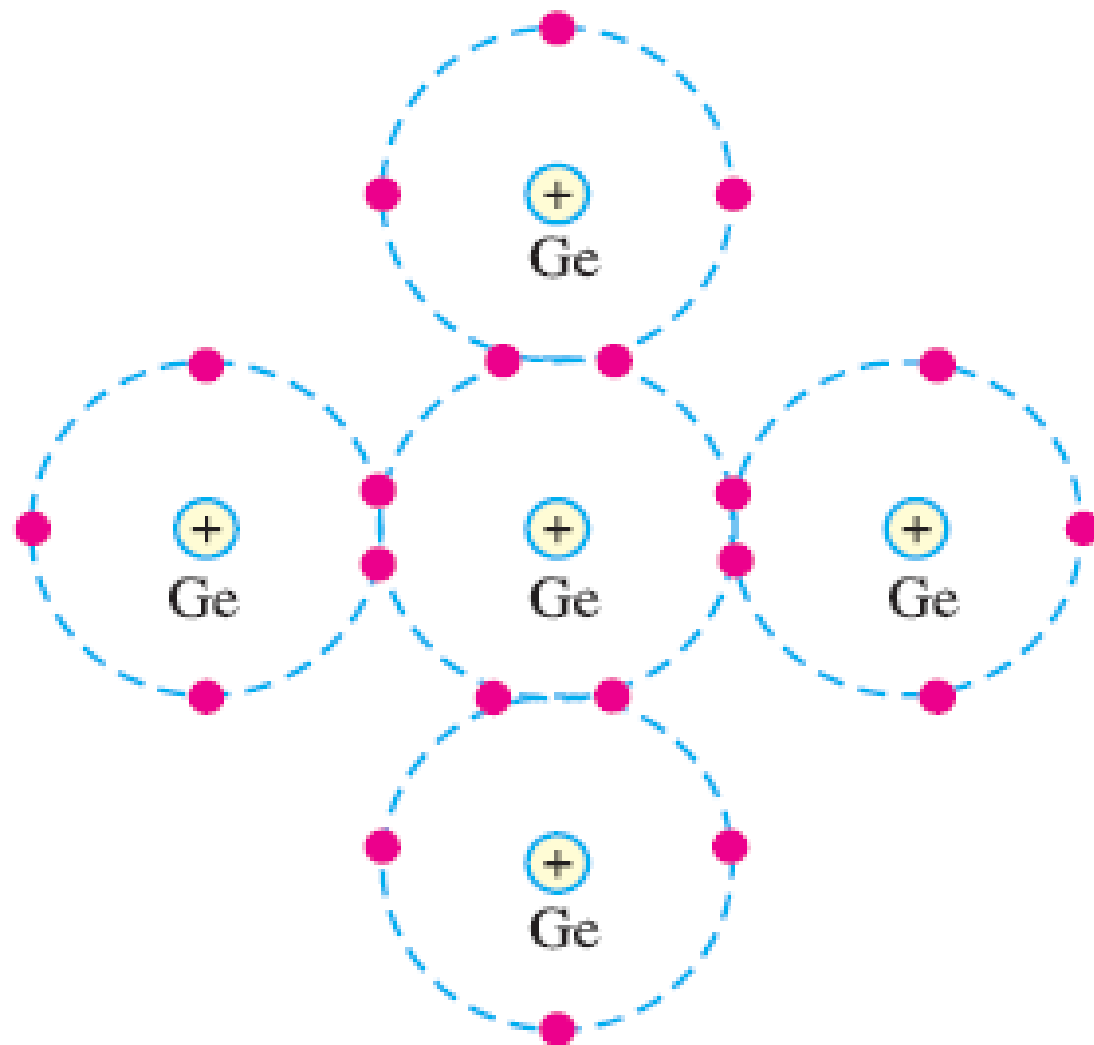
COVALENT BONDS AMONG Ge ATOMS

(i)
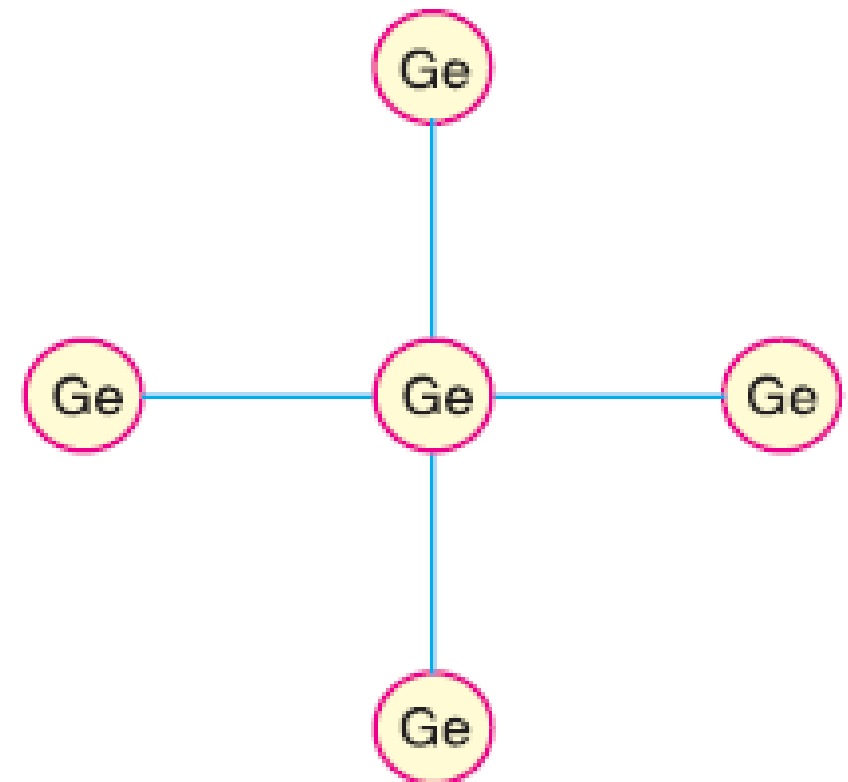
BONDING DIAGRAM

(ii)

Fig. 5.1

Fig. 5.1 shows the co-valent bonds among germanium atoms. A germanium atom has 4 valence electrons. It is the tendency of each germanium atom to have 8 electrons in the last orbit. To do so, each germanium atom positions itself between four other germanium atoms as shown in Fig. 5.1 (*i*). Each neighbouring atom shares one valence electron with the central atom. In this business of sharing, the central atom completes its last orbit by having 8 electrons revolving around the nucleus. In this way, the central atom sets up co-valent bonds. Fig. 5.1 (*ii*) shows the bonding diagram. The following points may be noted regarding the co-valent bonds:

(*i*) Co-valent bonds are formed by sharing of valence electrons.

(*ii*) In the formation of co-valent bond, each valence electron of an atom forms direct bond with the valence electron of an adjacent atom. In other words, valence electrons are associated with particular atoms. For this reason, valence electrons in a semiconductor are not free.

This effectively creates eight shared valence electrons for each atom and produces a state of chemical stability. Also, this sharing of valence electrons produces a strong **covalent bond** that hold the atoms together; each valence electron is attracted equally by the two adjacent atoms which share it. Covalent bonding in an intrinsic silicon crystal is shown in Figure 1–11. An **intrinsic** crystal is one that has no impurities. Covalent bonding for germanium is similar because it also has four valence electrons.
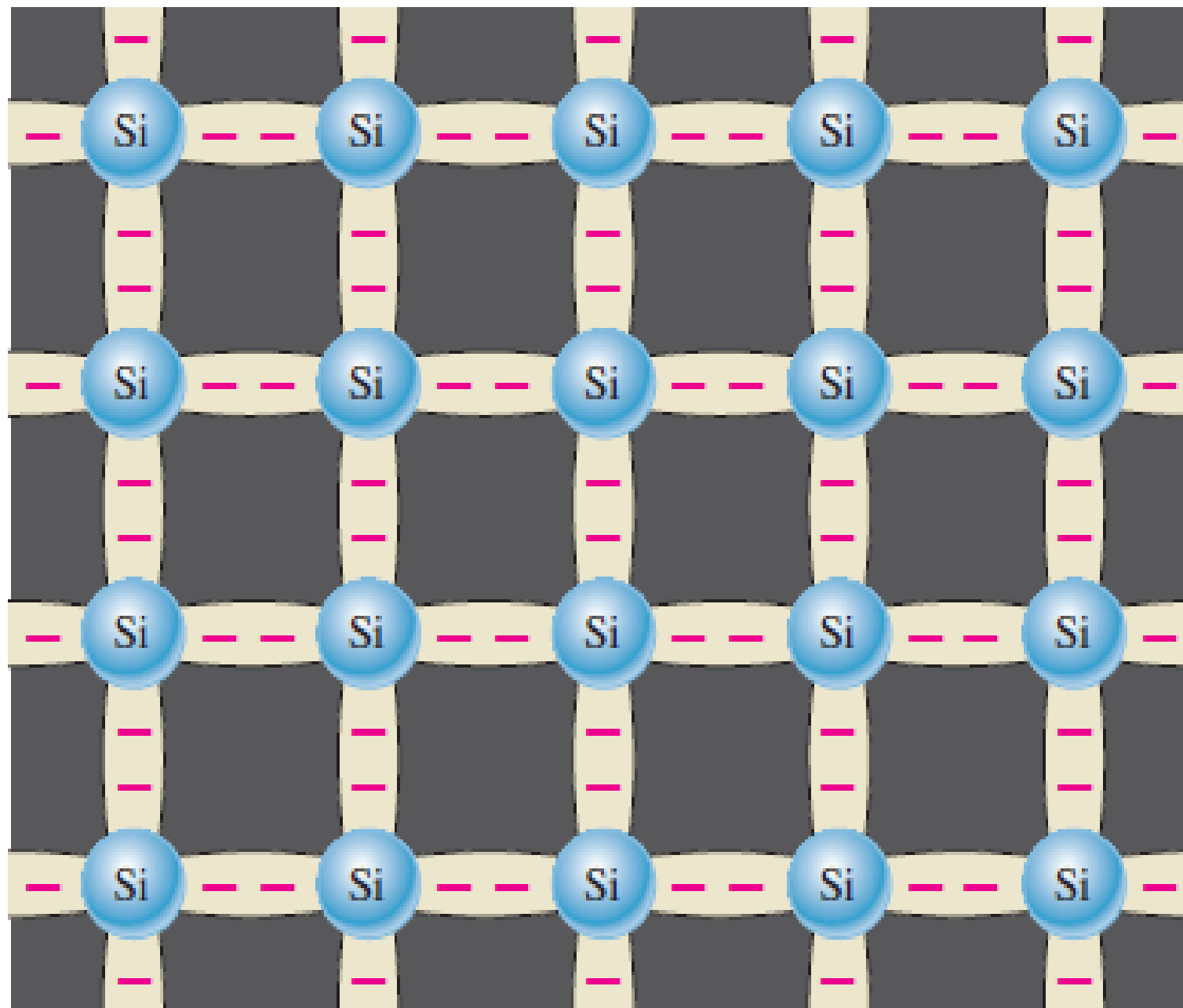
◄ FIGURE 1–11

Covalent bonds in a silicon crystal.

COMMONLY USED SEMICONDUCTORS

There are many semiconductors available, but very few of them have a practical application in electronics. The two most frequently used materials are *germanium* (Ge) and *silicon* (Si). It is because the energy required to break their co-valent bonds (*i.e.* energy required to release an electron from their valence bands) is very small; being about 0.7 eV for germanium and about 1.1 eV for silicon. A germanium atom has 32 electrons. First orbit has 2 electrons, second 8 electrons, third 18 electrons and the fourth orbit has 4 electrons

GERMANIUM

Germanium has become the model substance among the semiconductors; *the main reason being that it can be purified relatively well and crystallised easily*. Germanium is an earth element and was discovered in 1886. It is recovered from the ash of certain coals or from the flue/pipe dust of zinc smelters. Generally, recovered germanium is in the form of germanium dioxide powder which is then reduced to pure germanium.

(i)   (ii)
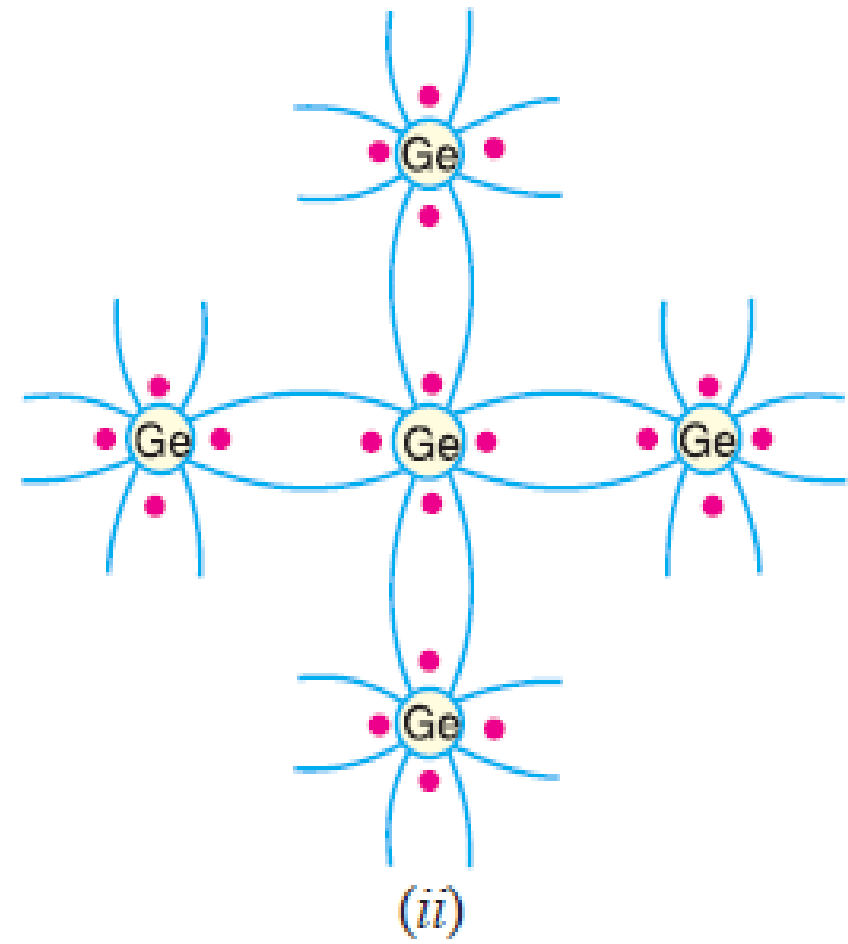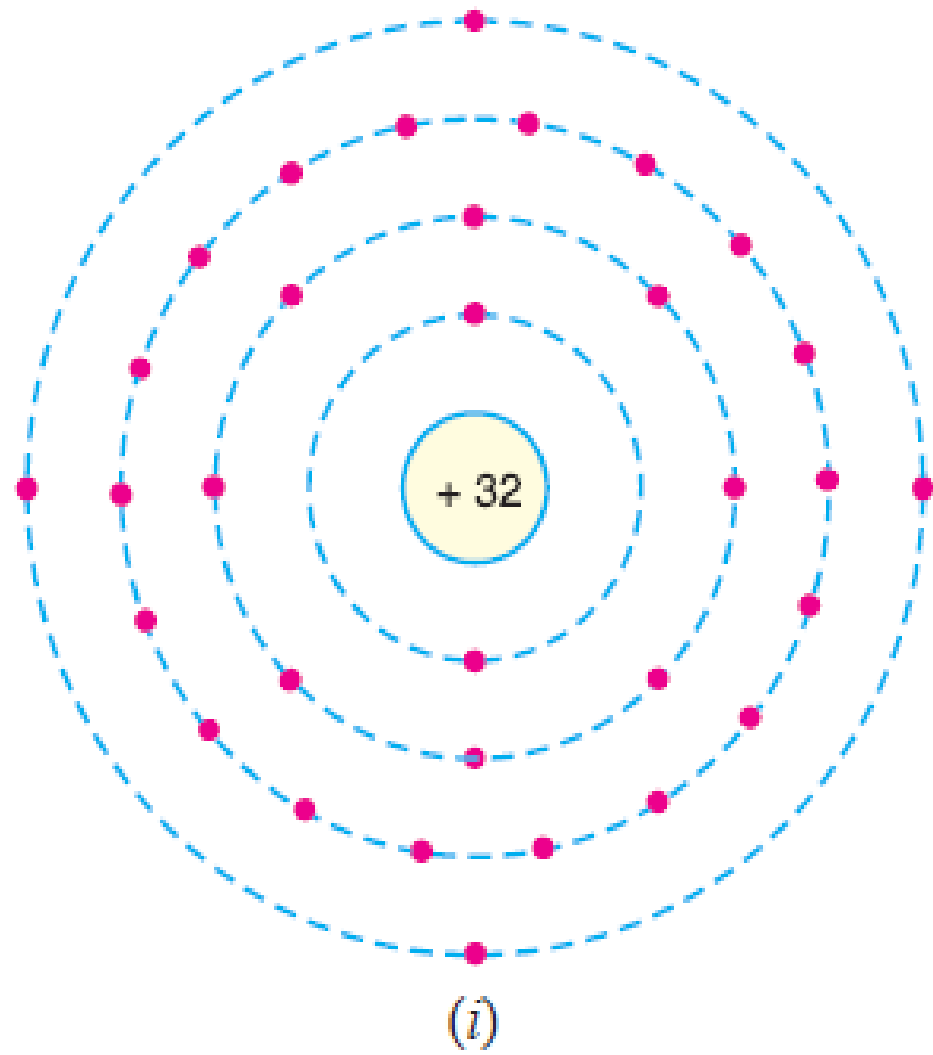
Fig. 5.2

The atomic number of germanium is 32. Therefore, it has 32 protons and 32 electrons. Two electrons are in the first orbit, eight electrons in the second, eighteen electrons in the third and four electrons in the outer or valence orbit [See Fig. 5.2 (*i*)]. It is clear that germanium atom has four valence electrons *i.e.*, it is a tetravalent element. Fig. 5.2 (*ii*) shows how the various germanium atoms are held through co-valent bonds. As the atoms are arranged in an orderly pattern, therefore, germanium has crystalline structure.

SILICON

Silicon is an element in most of the common rocks. Actually, sand is silicon dioxide. The silicon compounds are chemically reduced to silicon which is 100% pure for use as a semiconductor.
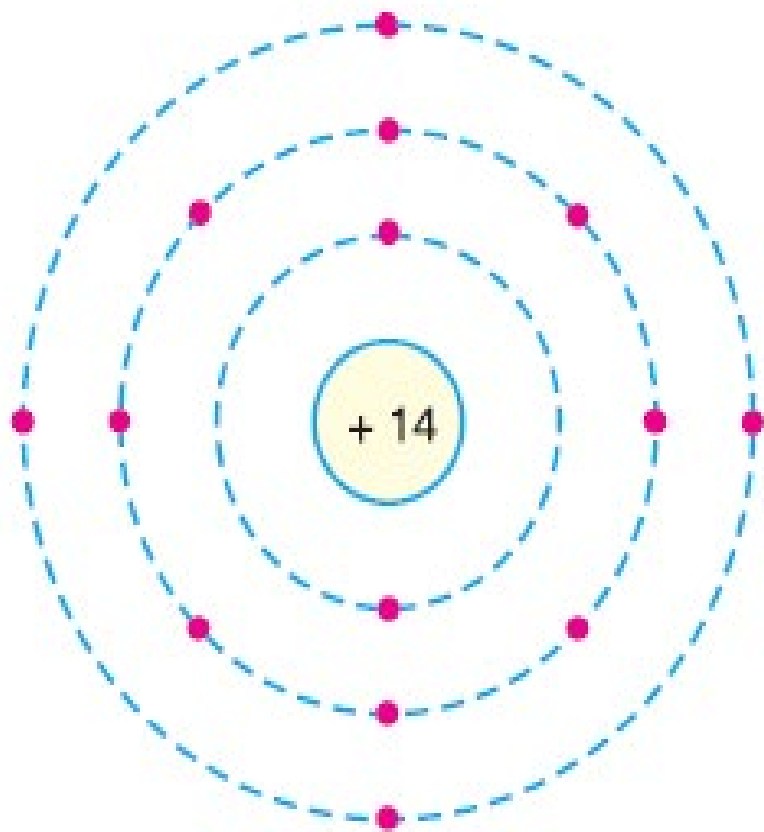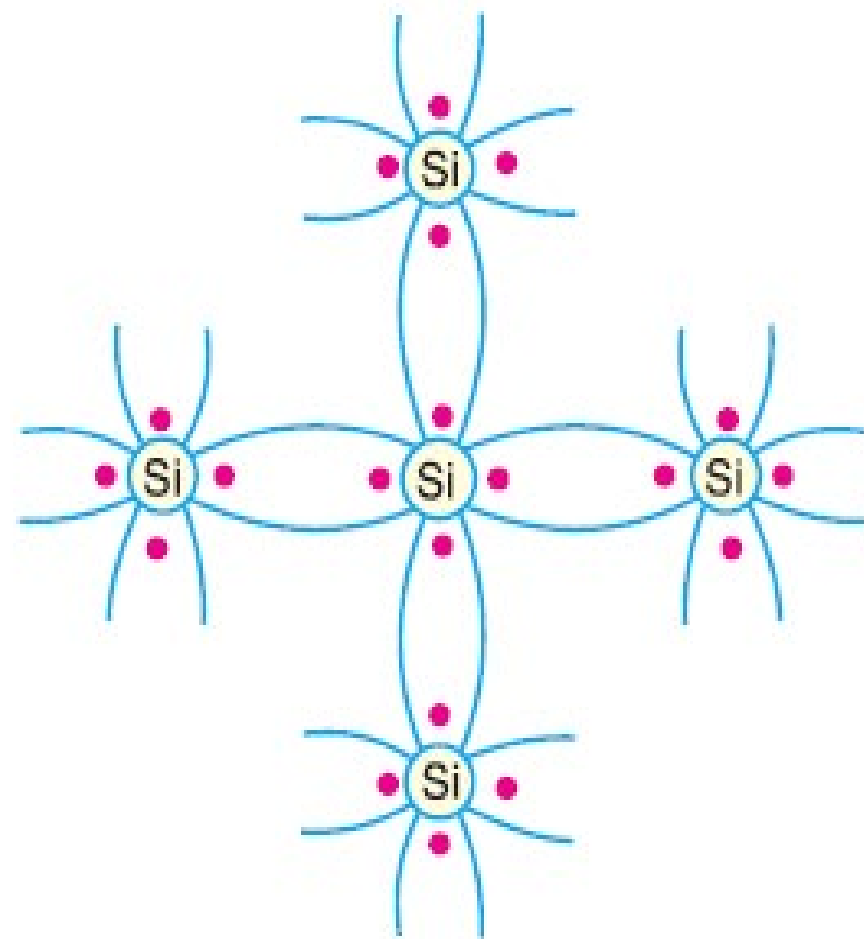
(i)   (ii)

Fig. 5.3

The atomic number of silicon is 14. Therefore, it has 14 protons and 14 electrons. Two electrons are in the first orbit, eight electrons in the second orbit and four electrons in the third orbit [See Fig. 5.3 (*i*)]. It is clear that silicon atom has four valence electrons *i.e.* it is a tetravalent element. Fig. 5.3 (*ii*) shows how various silicon atoms are held through co-valent bonds. Like germanium, silicon atoms are also arranged in an orderly manner. Therefore, silicon has crystalline structure.

SEMICONDUCTOR TYPES

The following illustration shows the classification of semiconductors.
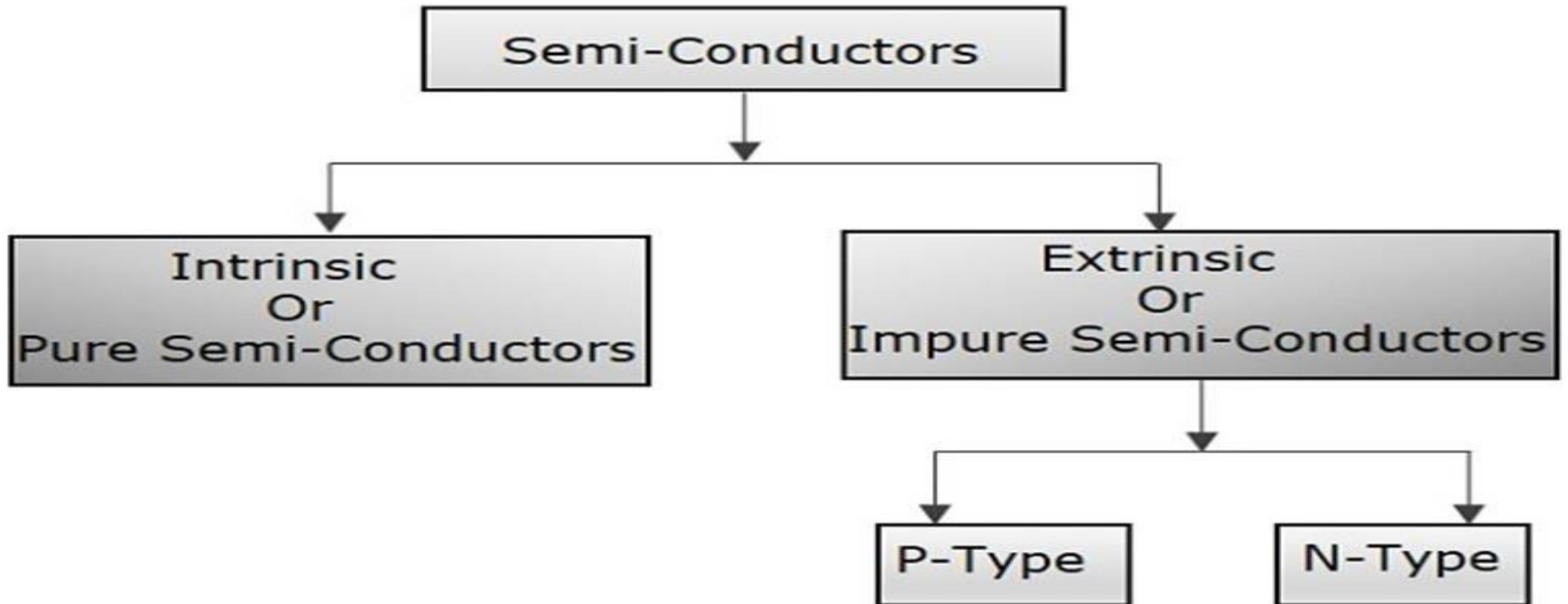
A pure form of semiconductors is called an intrinsic semiconductor. Conduction in intrinsic semiconductor is either due to <u>thermal excitation or crystal defects</u>. Within a semiconductor crystal lattice, **thermal excitation** is a process where lattice vibrations provide enough energy to transfer electrons to a higher energy band such as a more energetic sublevel or energy level. When an excited electron falls back to a state of lower energy, it undergoes electron relaxation. **Crystal defect**, imperfection in the regular geometrical arrangement of the atoms in a **crystalline** solid. These imperfections result from deformation of the solid, rapid cooling from high temperature, or high-energy radiation (X-rays or neutrons) striking the solid.

The single-element semiconductors are characterized by atoms with four valence electrons. Silicon is the most commonly used semiconductor.. Let us consider the structure of Si. A Si atomic no. is 14 and it has 4 valance electrons. These 4 electrons are shared by four neighbouring atoms in the crystal structure by means of covalent bond. Fig.1.2a shows the crystal structure of Si at absolute zero temperature (0K). Hence a pure SC acts has poor conductivity (due to lack of free electrons) at low or absolute zero temperature.

Covalent bond
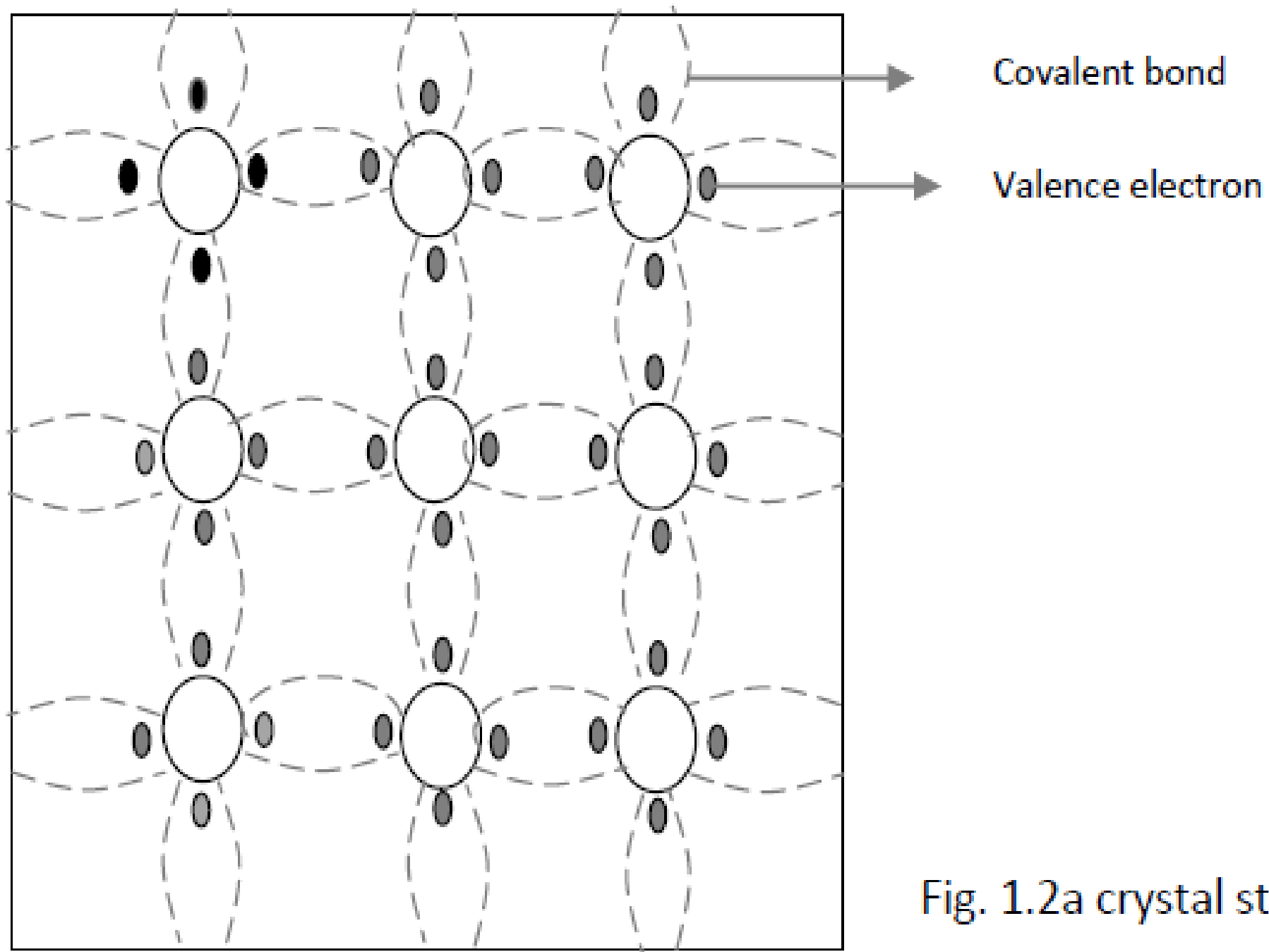
Valence electron

Fig. 1.2a crystal structure of Si at 0K

At room temperature some of the covalent bonds break up to thermal energy as shown in fig.1.2b. The valance electrons that jump into conduction band are called as free electrons that are available for conduction.
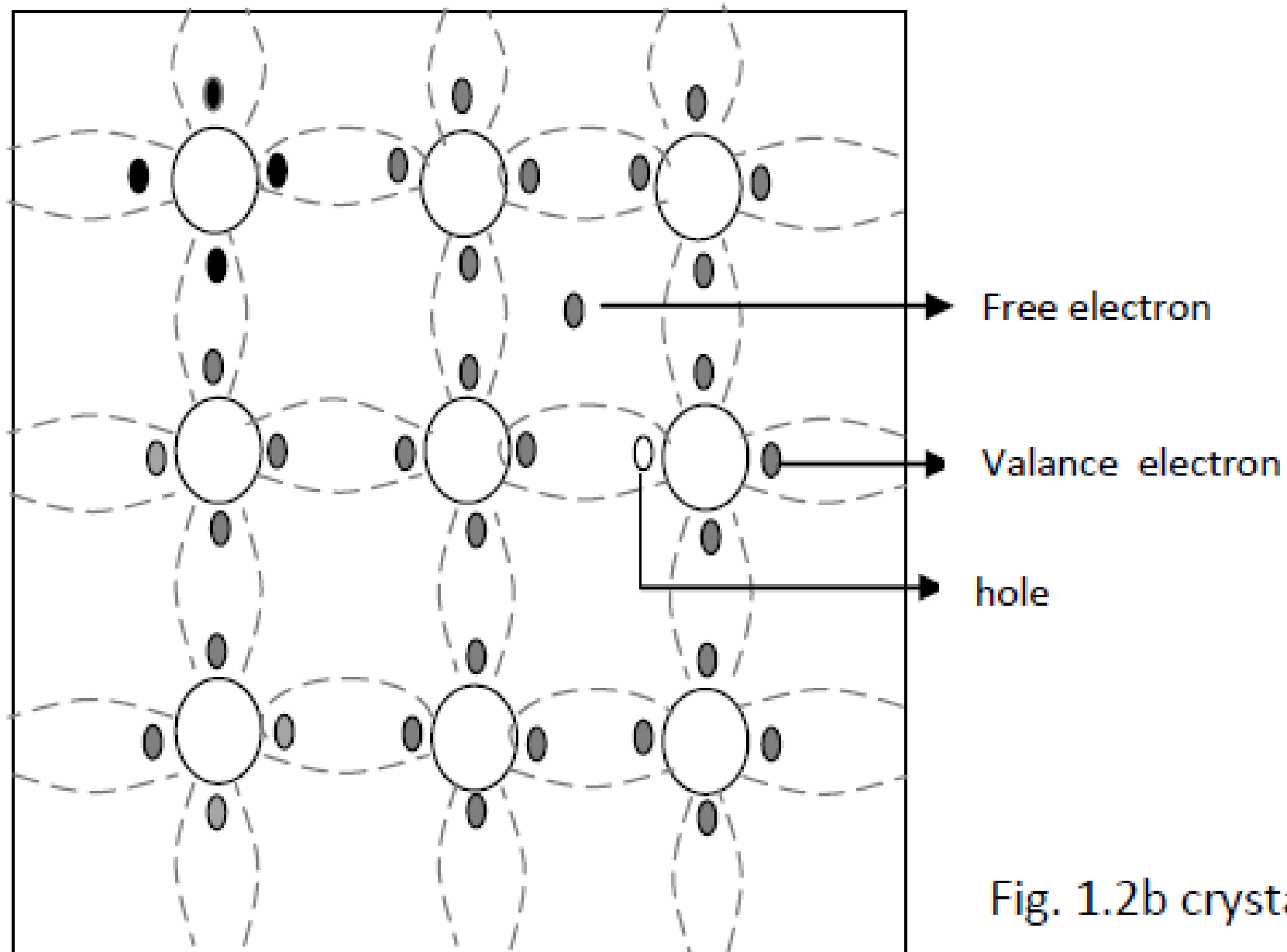
Fig. 1.2b crystal structure of Si at room temperature0K

The absence of electrons in covalent bond is represented by a small circle usually referred to as hole which is of positive charge. Even a hole serves as carrier of electricity in a manner similar to that of free electron. The mechanism by which a hole contributes to conductivity is explained as follows: **When a bond is in complete so that a hole exists, it is relatively easy for a valance electron in the neighbouring atom to leave its covalent bond to fill this hole**.

An electron moving from a bond to fill a hole moves in a direction opposite to that of the electron. This hole, in its new position may now be filled by an electron from another covalent bond and the hole will correspondingly move one more step in the direction opposite to the motion of electron. Here we have a mechanism for conduction of electricity which does not involve free electrons. This phenomenon is illustrated in fig1.3
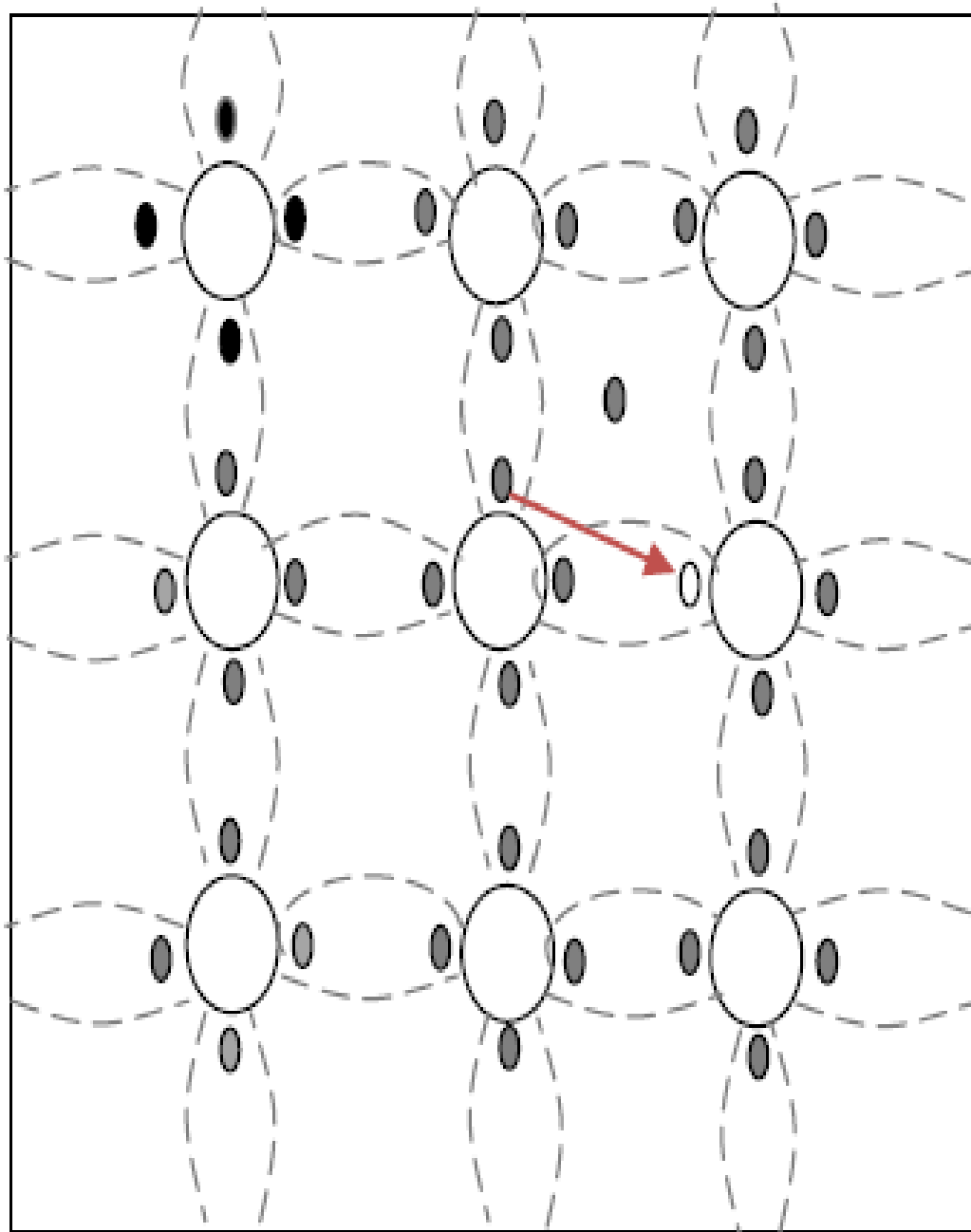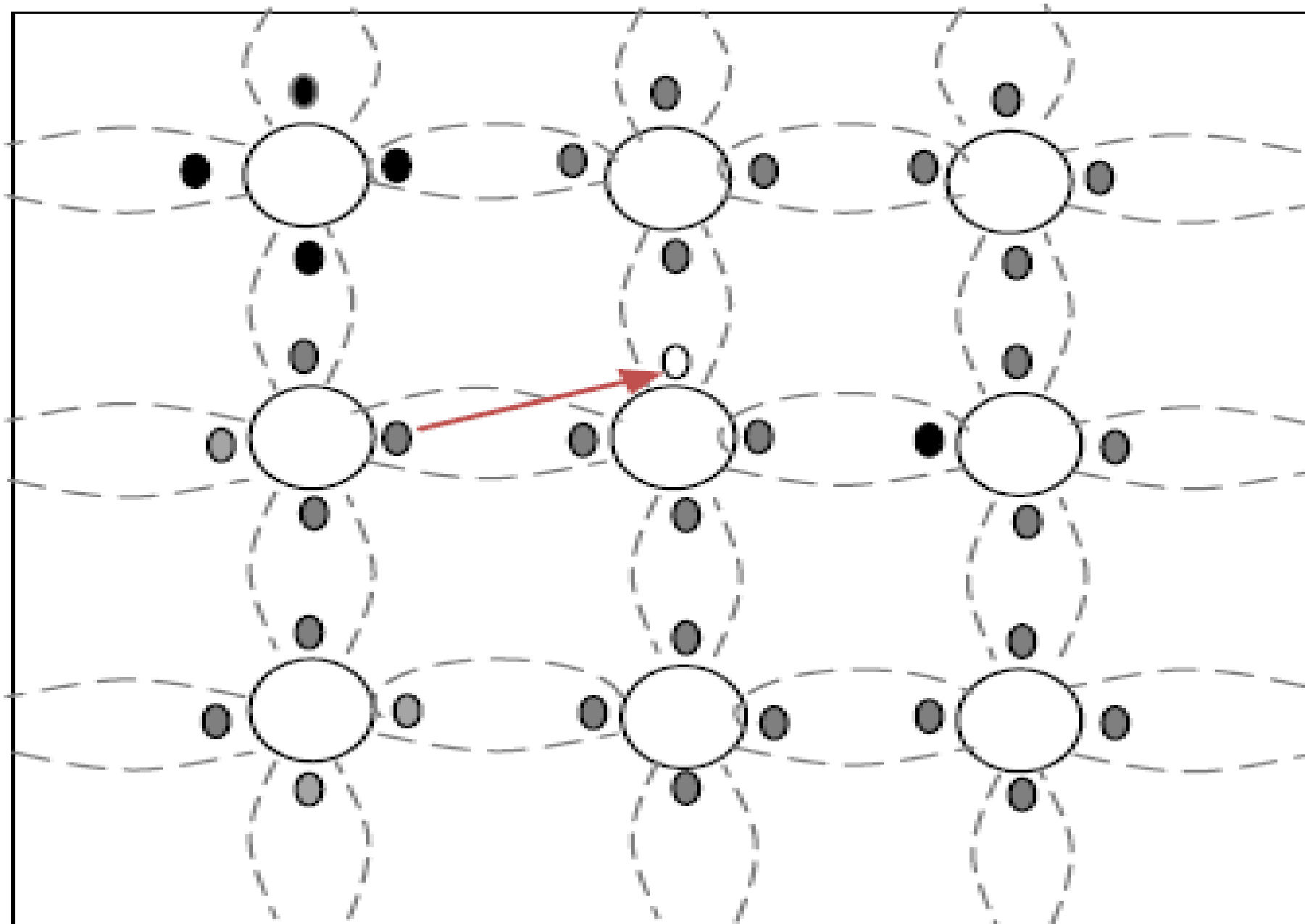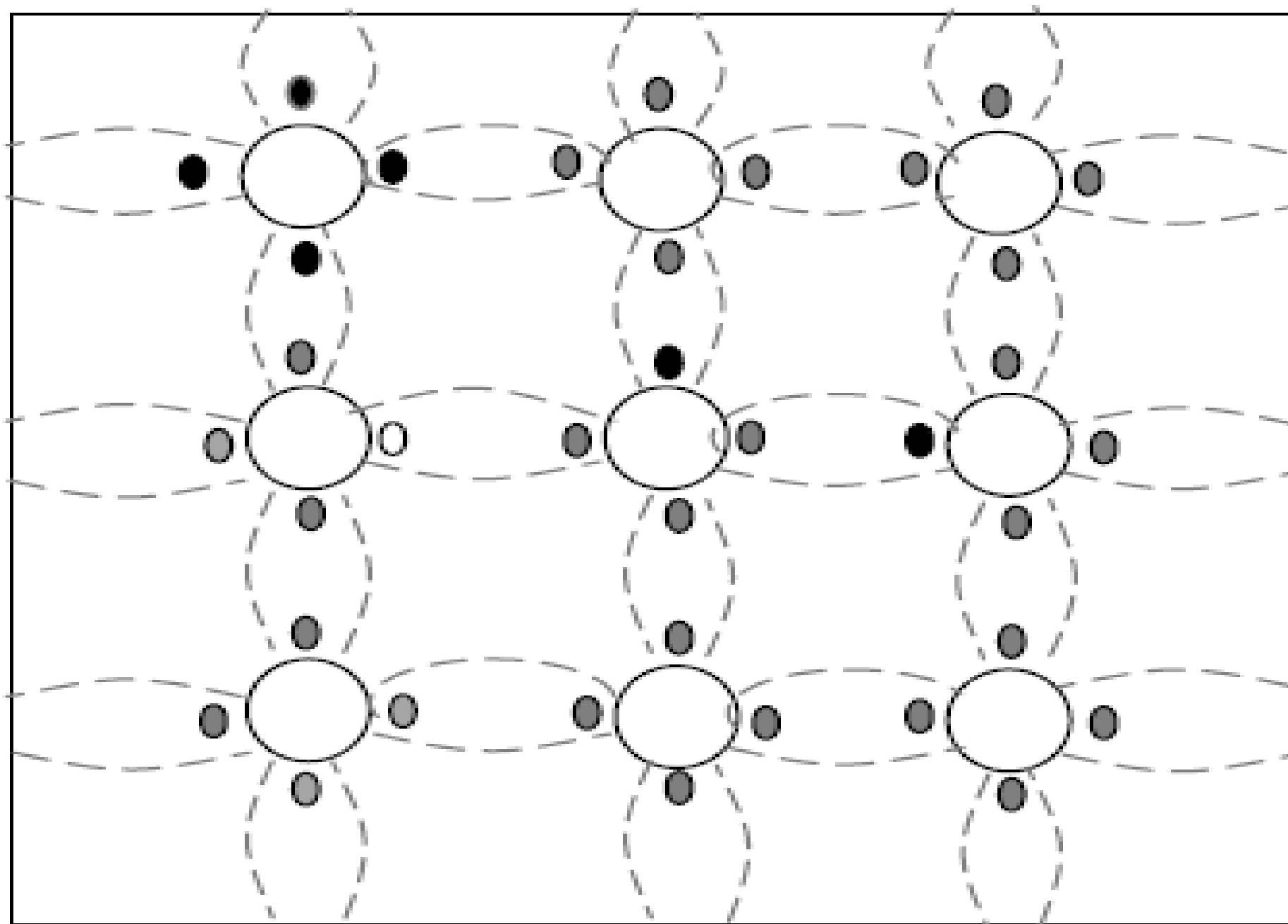
Fig. 1.3a

Fig. 1.3b

Fig. 1.3c

Fig 1.3a show that there is a hole at ion 6.Imagine that an electron from ion 5 moves into the hole at ion 6 so that the configuration of 1.3b results. If we compare both fig1.3a &fig 1.3b, it appears as if the hole has moved towards the left from ion 6 to ion 5. Further if we compare fig 1.3b and fig 1.3c, the hole moves from ion5 to ion 4. This discussion indicates the motion of hole is in a direction opposite to that of motion of electron. Hence we consider holes as physical entities whose movement constitutes flow of current. ***In a pure semiconductor, the number of holes is equal to the number of free electrons***.

CURRENT IN SEMICONDUCTORS

The way a material conducts electrical current is important in understanding how electronic devices operate. You can't really understand the operation of a device such as a diode or transistor without knowing something about current in semiconductors. As you have learned, the electrons in a solid can exist only within prescribed energy bands. Each shell corresponds to a certain energy band and is separated from adjacent shells by band gaps, in which no electrons can exist. Figure 1–12 shows the energy band diagram for the atoms in a pure silicon crystal at its lowest energy level. There are no electrons shown in the conduction band, a condition that occurs *only* at a temperature of absolute 0 Kelvin.

Energy band diagram for an atom in a pure (intrinsic) silicon crystal at it lowest energy state. There are no electrons in the conduction band at a temperature of 0 K.
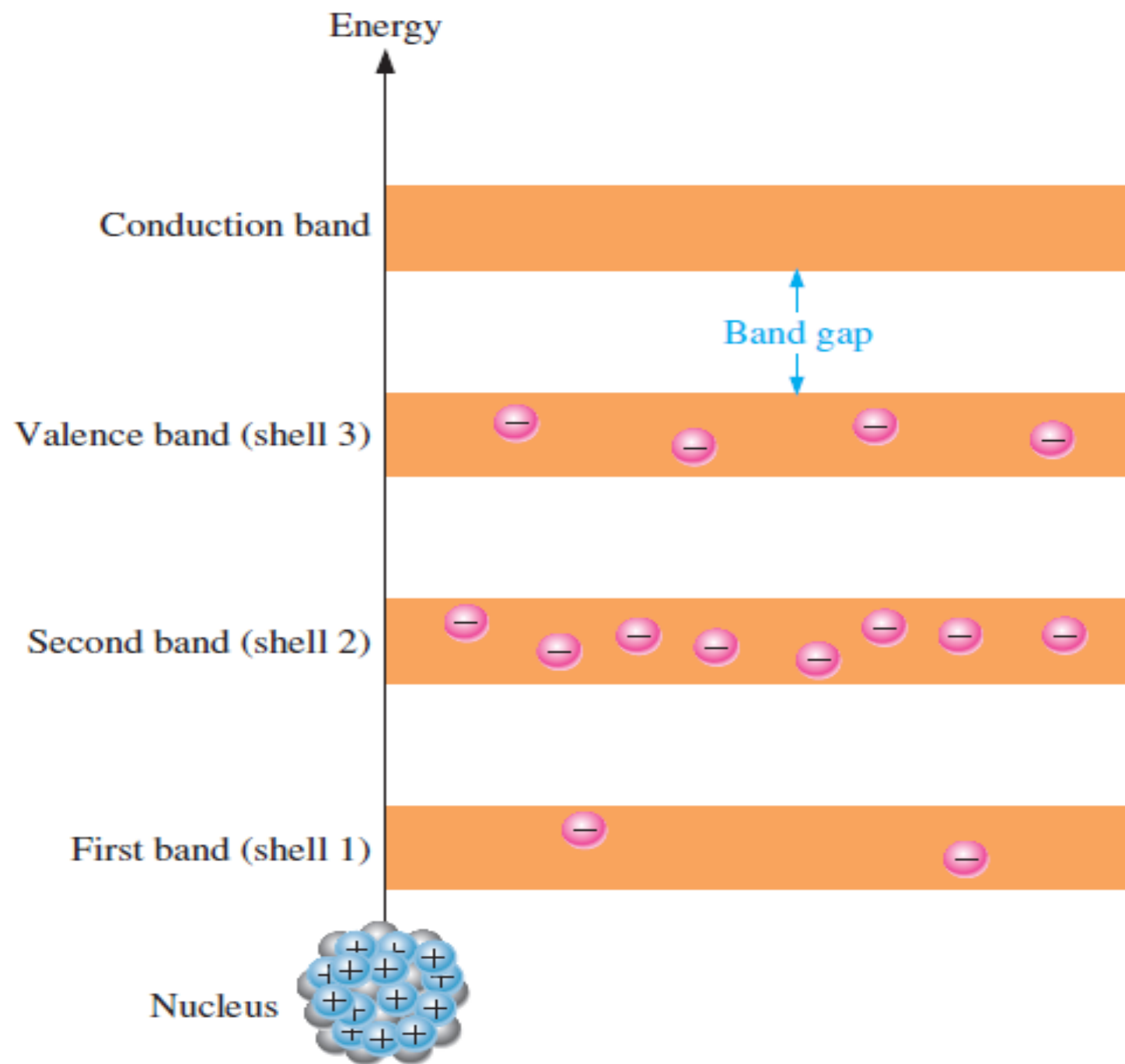
CONDUCTION ELECTRONS AND HOLES

An intrinsic (pure) silicon crystal at room temperature has sufficient heat (thermal) energy for some valence electrons to jump the gap from the valence band into the conduction band, becoming free electrons. Free electrons are also called **conduction electrons.** This is illustrated in the energy diagram of Figure 1–13(a) and in the bonding diagram of Figure 1–13(b).

Creation of electron-hole pairs in a silicon crystal. Electrons in the conduction band are free electrons.
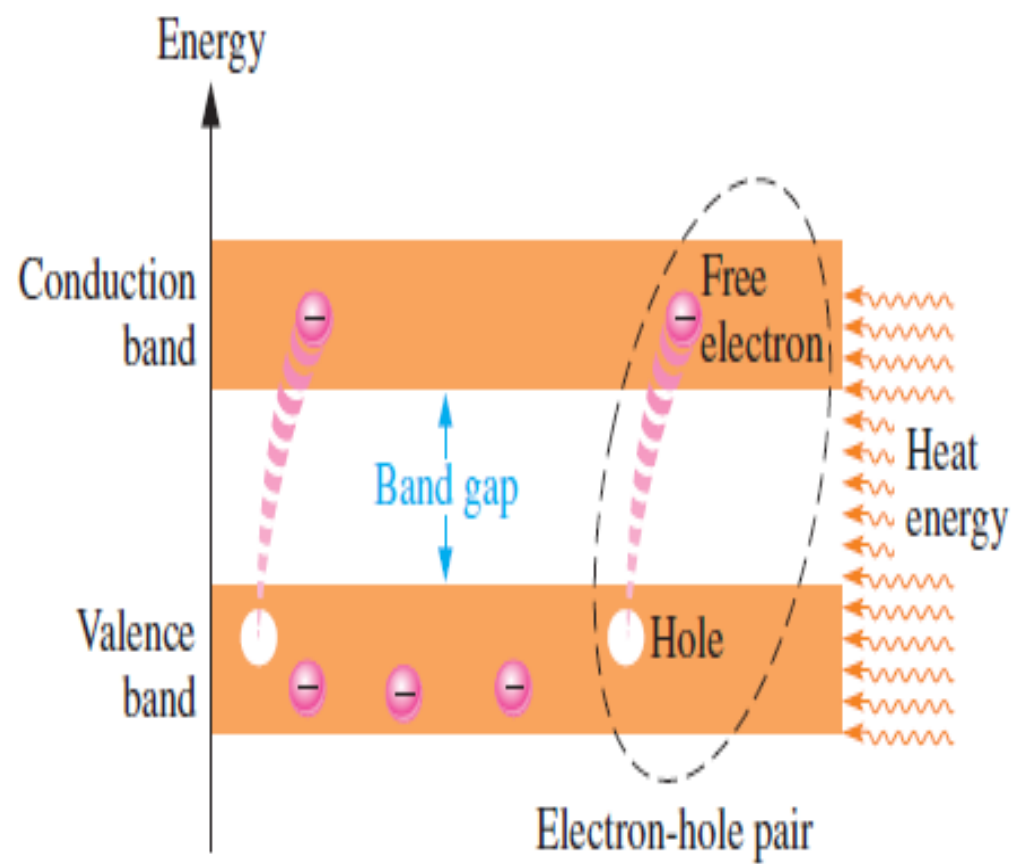
(a) Energy diagram for the outer two bands
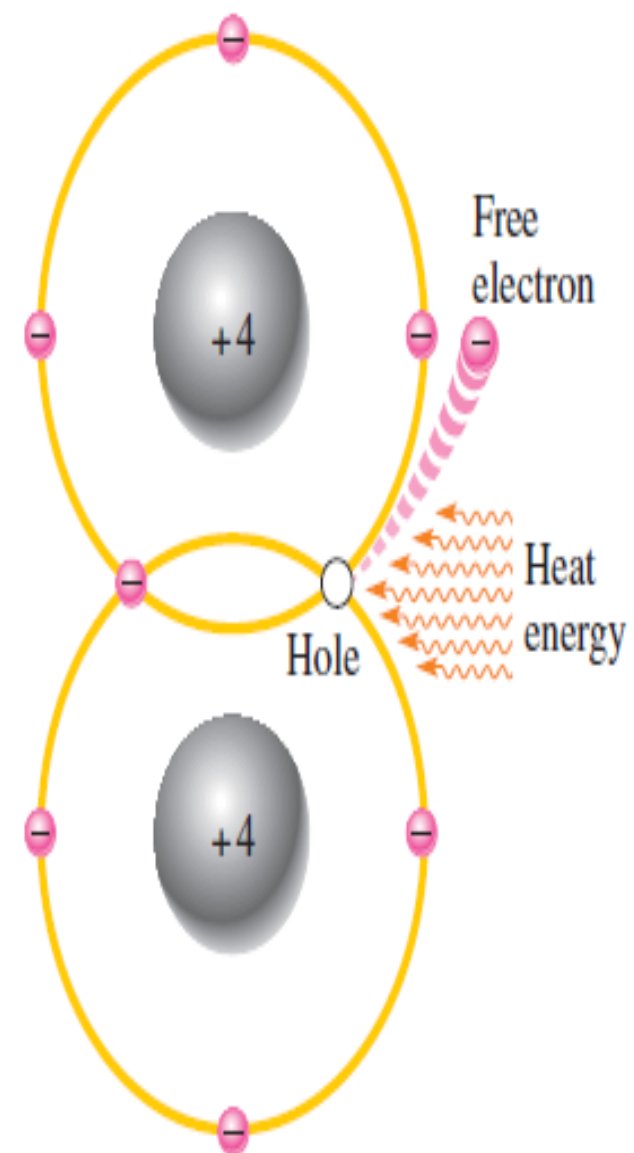
(b) Bonding diagram

When an electron jumps to the conduction band, a vacancy is left in the valence band within the crystal. This vacancy is called a **hole**. For every electron raised to the conduction band by external energy, there is one hole left in the valence band, creating what is called an **electron-hole pair**. <u>Recombination occurs when a conduction-band electron loses energy and falls back into a hole in the valence band.</u> To summarize, a piece of intrinsic silicon at room temperature has, at any instant, a number of conduction-band (free) electrons that are unattached to any atom and are essentially drifting randomly throughout the material. There is also an equal number of holes in the valence band created when these electrons jump into the conduction band. This is illustrated in Figure 1–14.

Generation of an electron-hole pair

Recombination of an electron with a hole

Heat energy
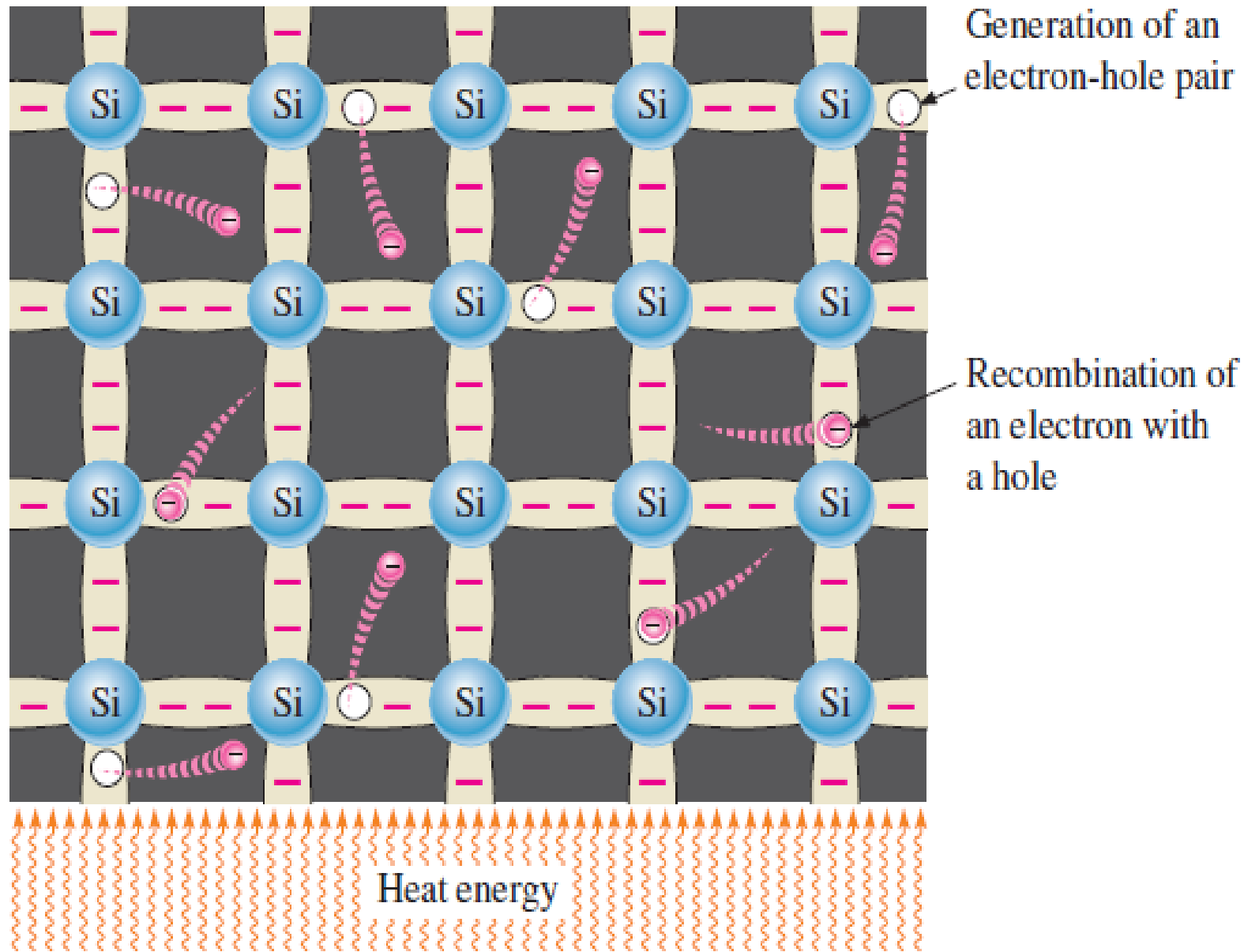
Electron-hole pairs in a silicon crystal. Free electrons are being generated continuously while some recombine with holes.

ELECTRON AND HOLE CURRENT

When a voltage is applied across a piece of intrinsic silicon, as shown in Figure 1–15, the thermally generated free electrons in the conduction band, which are free to move randomly in the crystal structure, are now easily attracted toward the positive end. This movement of free electrons is one type of **current** in a semi-conductive material and is called *electron current.*
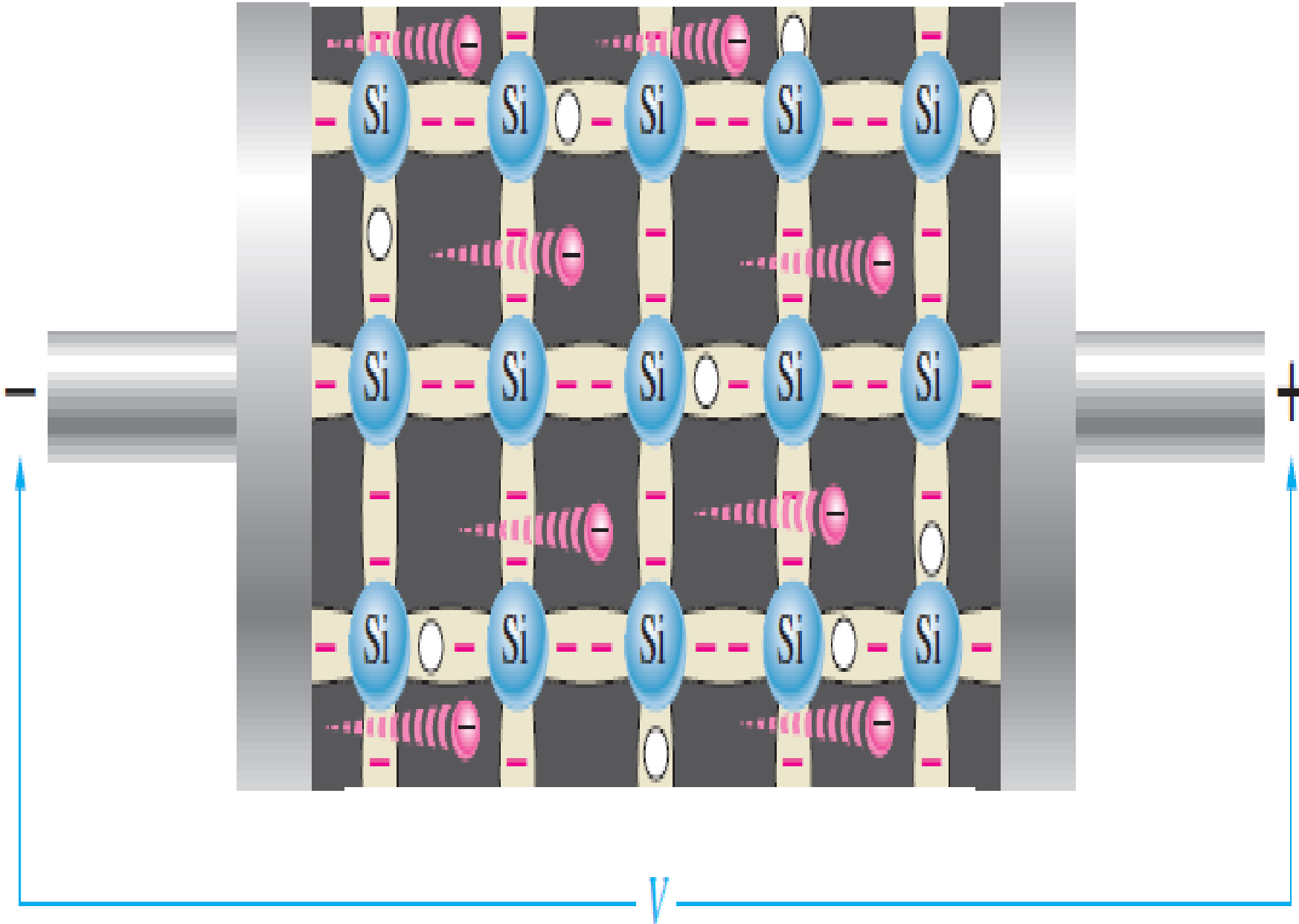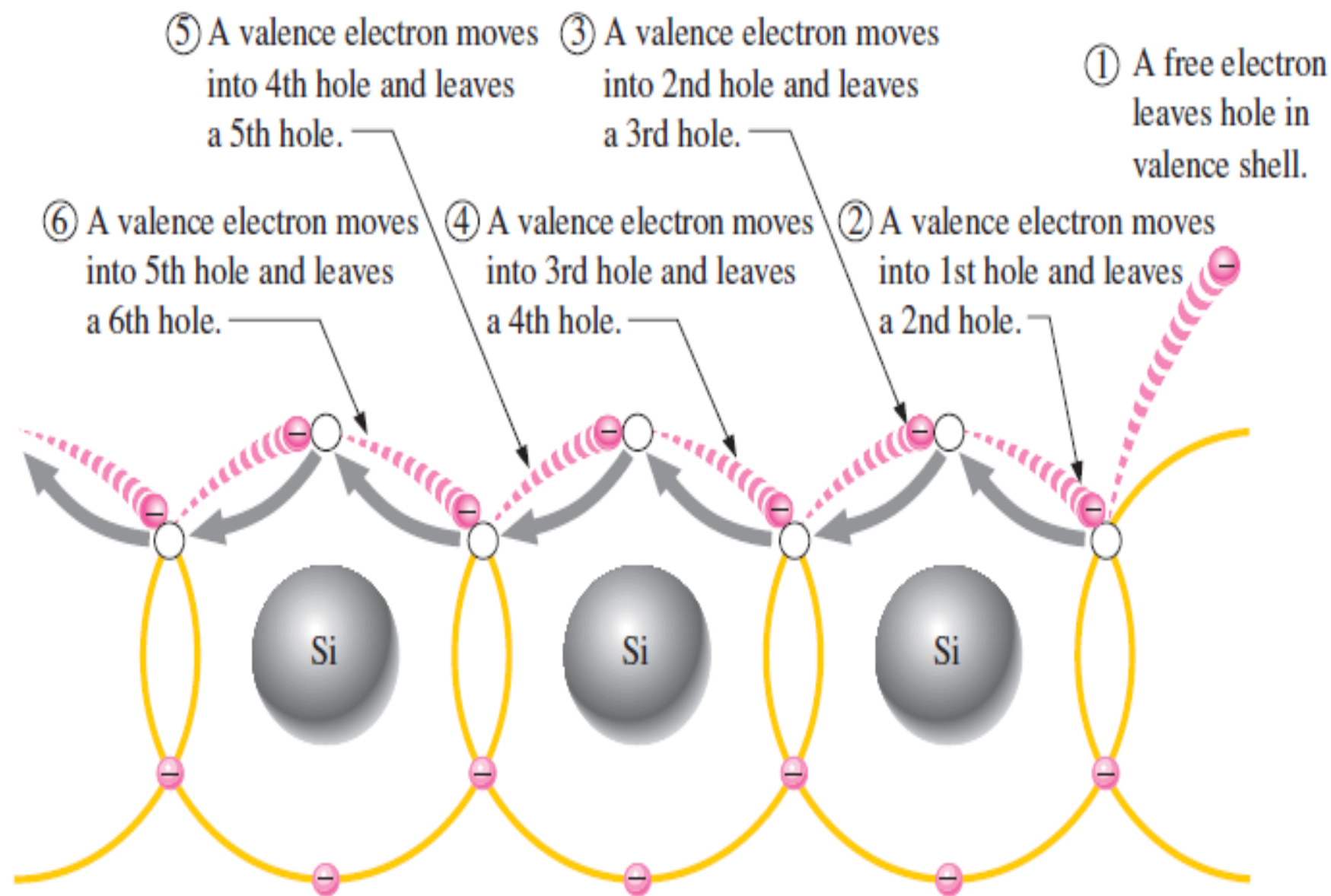
Electron current in intrinsic silicon is produced by the movement of thermally generated free electrons.

Another type of current occurs in the valence band, where the holes created by the free electrons exist. Electrons remaining in the valence band are still attached to their atoms and are not free to move randomly in the crystal structure as are the free electrons. However, a valence electron can move into a nearby hole with little change in its energy level, thus leaving another hole where it came from. Effectively the hole has moved from one place to another in the crystal structure, as illustrated in Figure 1–16. *Although current in the valence band is produced by valence electrons, it is called hole current to distinguish it from electron current in the conduction band.*

FIGURE 1–16

Hole current in intrinsic silicon.

⑤ A valence electron moves into 4th hole and leaves a 5th hole.

③ A valence electron moves into 2nd hole and leaves a 3rd hole.

① A free electron leaves hole in valence shell.

⑥ A valence electron moves into 5th hole and leaves a 6th hole.

④ A valence electron moves into 3rd hole and leaves a 4th hole.

② A valence electron moves into 1st hole and leaves a 2nd hole.

When a valence electron moves left to right to fill a hole while leaving another hole behind, the hole has effectively moved from right to left. Gray arrows indicate effective movement of a hole.

Conduction in semiconductors is considered to be either the movement of free electrons in the conduction band or the movement of holes in the valence band, which is actually the movement of valence electrons to nearby atoms, creating hole current in the opposite direction. It is interesting to contrast the two types of charge movement in a semiconductor with the charge movement in a metallic conductor, such as copper.

Copper atoms form a different type of crystal in which the atoms are not covalently bonded to each other but consist of a "sea" of positive ion cores, which are atoms stripped of their valence electrons. The valence electrons are attracted to the positive ions, keeping the positive ions together and forming the **metallic bond**. The valence electrons do not belong to a given atom, but to the crystal as a whole. Since the valence electrons in copper are free to move, the application of a voltage results in current. There is only one type of current—the movement of free electrons—because there are no "holes" in the metallic crystal structure.

INTRINSIC SEMICONDUCTOR

A semiconductor in an extremely pure form is known as an **intrinsic semiconductor**. An intrinsic semiconductor is one, which is pure enough that impurities do not appreciably affect its electrical behavior. In this case, all carriers are created due to thermally or optically excited electrons from the full valence band into the empty conduction band. Thus equal numbers of electrons and holes are present in an intrinsic semiconductor.

Electrons and holes flow in opposite directions in an electric field, though they contribute to current in the same direction since they are oppositely charged. Hole current and electron current are not necessarily equal in an intrinsic semiconductor, however, because electrons and holes have different effective masses (crystalline analogues to free inertial masses).

The concentration of carriers is strongly dependent on the temperature. At low temperatures, the valence band is completely full making the material an insulator. Increasing the temperature leads to an increase in the number of carriers and a corresponding increase in conductivity. This characteristic shown by intrinsic semiconductor is different from the behavior of most metals, which tend to become less conductive at higher temperatures due to increased phonon scattering.

Both silicon and germanium are tetravalent, i.e. each has four electrons (valence electrons) in their outermost shell. Both elements crystallize with a diamond-like structure, i.e. in such a way that each atom in the crystal is inside a tetrahedron formed by the four atoms which are closest to it. Each atom shares its four valence electrons with its four immediate neighbours, so that each atom is involved in four covalent bonds. ***Intrinsic semiconductor has very limited applications as they conduct very small amounts of current at room temperature***.

The current conduction capability of intrinsic semiconductor can be increased significantly by adding a small amounts impurity to the intrinsic semiconductor. By adding impurities it becomes impure or extrinsic semiconductor. ***This process of adding impurities is called as doping***. The amount of impurity added is 1 part in 10^6 atoms. Since semiconductors are generally poor conductors, their conductivity can be drastically increased by the controlled addition of impurities to the intrinsic (pure) semi-conductive material. This process, called **doping**, increases the number of current carriers (electrons or holes). The two categories of impurities are *n*-type and *p*-type.

In an intrinsic semiconductor, even at room temperature, hole-electron pairs are created. When electric field is applied across an intrinsic semiconductor, the current conduction takes place by two processes, namely ; by *free electrons* and *holes* as shown in Fig. 5.10. The free electrons are produced due to the breaking up of some covalent bonds by thermal energy. At the same time, holes are created in the covalent bonds. Under the influence of electric field, conduction through the semiconductor is by both free electrons and holes. Therefore, the total current inside the semiconductor is the sum of currents due to free electrons and holes.

FREE ELECTRONS

A

B
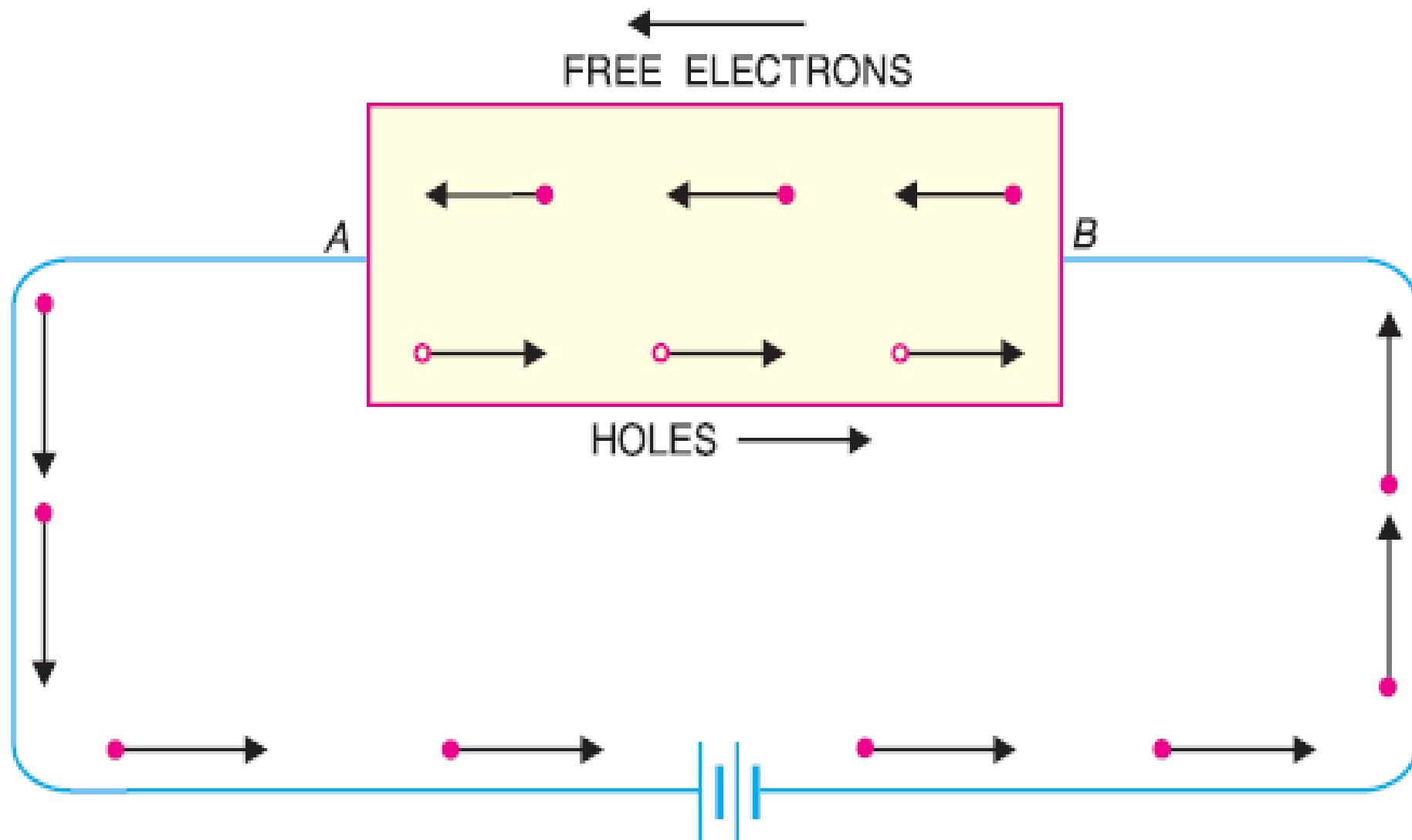
HOLES

Fig. 5.10

It may be noted that current in the external wires is fully electronic *i.e.* by electrons. What about the holes? Referring to Fig. 5.10, holes being positively charged move towards the negative terminal of supply. As the holes reach the negative terminal *B*, electrons enter the semiconductor crystal near the terminal and combine with holes, thus cancelling them. At the same time, the loosely held electrons near the positive terminal *A* are attracted away from their atoms into the positive terminal. This creates new holes near the positive terminal which again drift towards the negative terminal.

Not surprisingly, semiconductors are somewhat like metals and somewhat like insulators. As in atoms, the electrons in solids are in certain energy levels, or, more correctly, certain bands of energy levels, so this is called the band theory of solids. In metals, the uppermost band has a few electrons in it and these can move around. In insulators the uppermost band is completely filled with electrons and they are gridlocked, as in the second figure , above. In semiconductors, the next-to the highest band is *almost* completely filled with electrons and they are *almost* gridlocked/congested. However, the highest band is only a little higher than the next-to-highest, and a few electrons jump into this band and conduct a little electricity.

Conduction band (almost empty)
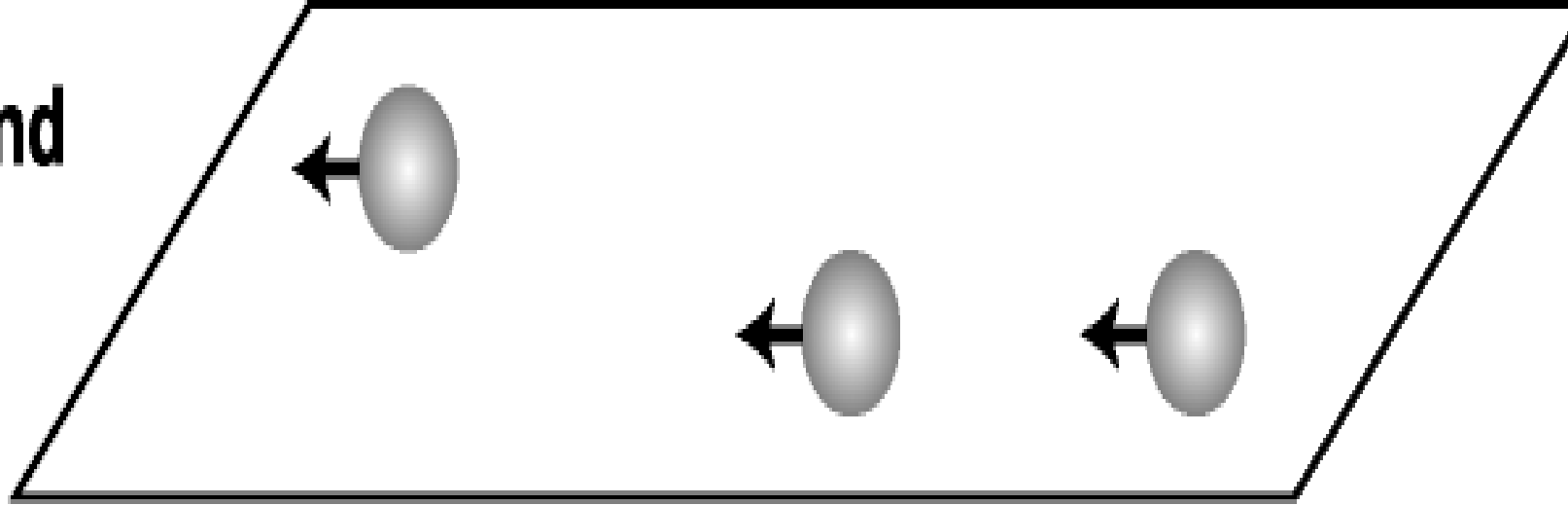
Valence band (almost full)
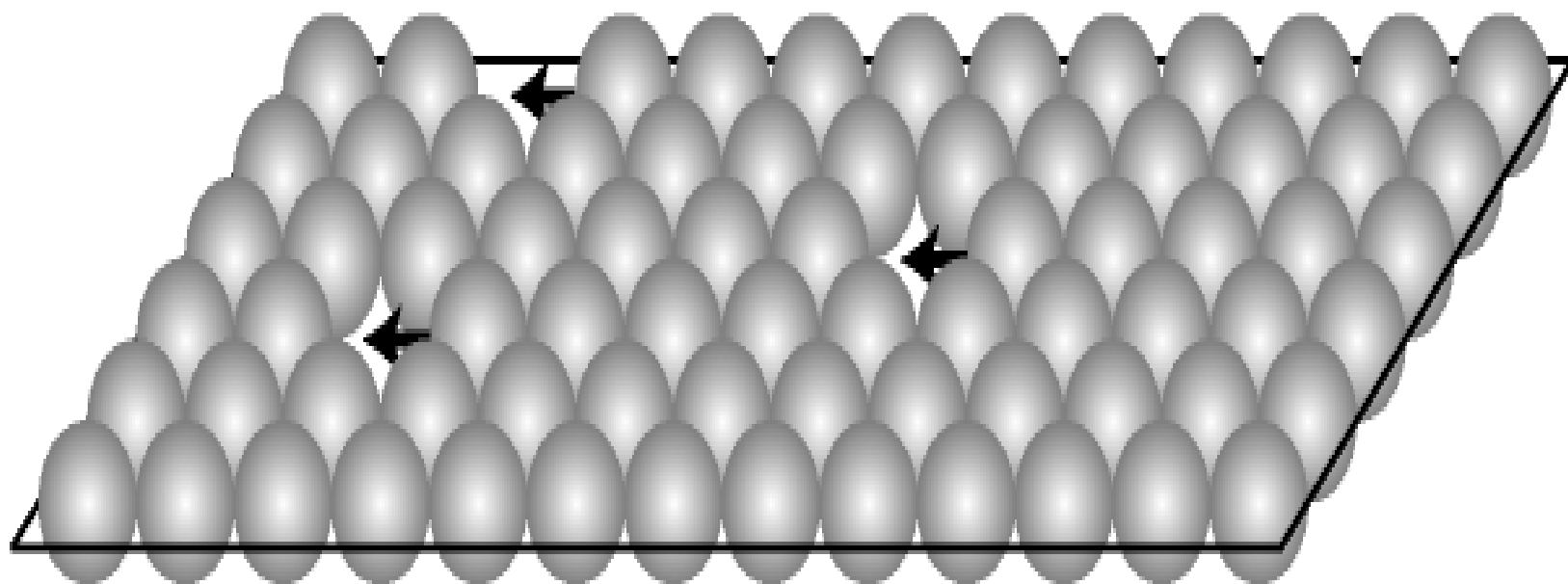
Energy
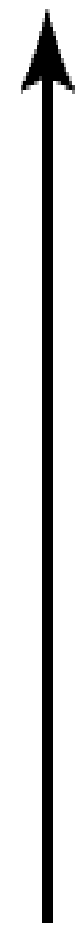
In addition, the holes left in the band below give a little room for the gridlocked electrons to move. Actually, the holes look like positive charges (due to the positive ions left behind when the electrons move away) and these holes appear to move in the direction opposite the electrons. Therefore, the current carried by a semiconductor consists of the motion of the negatively-charged electrons in the conduction band (the uppermost band) and the positively-charged holes in the valence band (the next-to highest band).

In pure semiconductors, electric conduction is due to the thermally-generated electron hole pairs. Hence, if pure semiconductors kept in the dark, thermally-generated charge carriers are the only means of conduction. The number of such charge carriers per unit volume (i.e. intrinsic carrier density is given by

$$n_i = N exp\left(-\frac{E_g}{2KT}\right)$$

Where *N* is constant for a given semiconductor, E_g is the band gap energy in joules, K is Boltzmann's constant and T is the temperature in 0^0K.

51.27. Conductivity of Intrinsic Semiconductors

In their case, current flow is due to the movement of electrons and holes in opposite directions. However, since their charges are of opposite sign, the current due to each is in the same direction. Even though the number of electrons equals the number of holes, hole mobility μ_h is practically half of electron mobility μ_e.

As shown in Fig. 51.31, the total current flow which is due to the sum of electron flow and hole flow, is given by

$$I = I_e + I_h$$

Let
v_e = drift velocity of electrons (m/s)

v_h = drift velocity of holes (m/s)

n_i = density of free electrons in an intrinsic semiconductor (per m^3)
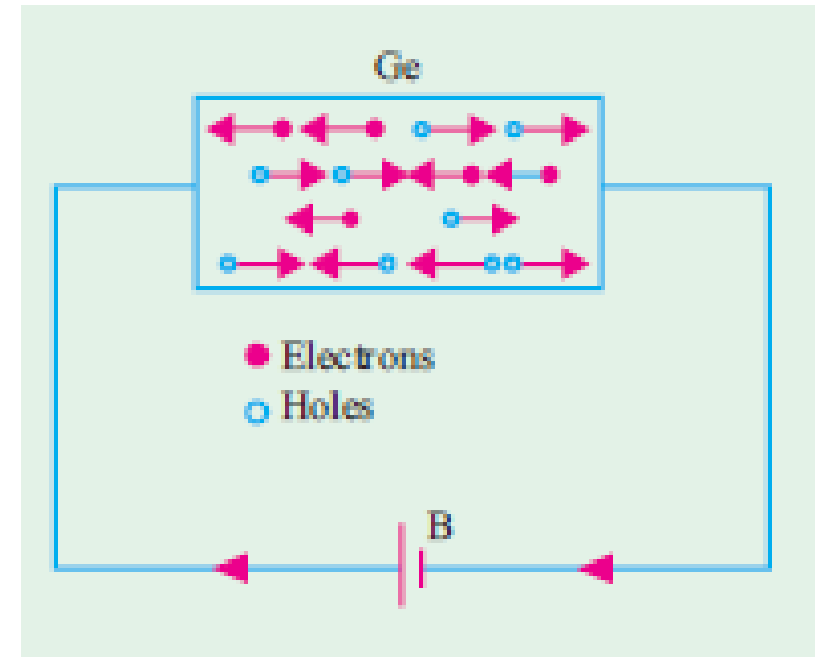
Fig. 51.31
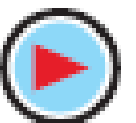
$$p_i = \text{density of holes in an intrinsic semiconductor (per m}^3)$$
$$e = \text{electron charge (coulomb)}$$
$$A = \text{cross-section of the semiconductor (m}^2)$$

Since in an intrinsic semiconductor $n_i = p_i$

$\therefore$
$$I = n_i e (v_e + v_h)A = n_i e (\mu_h + \mu_e) EA$$

where
$$\mu_e = \text{electron mobility} = v_e/E$$
$$\mu_h = \text{hole mobility} = v_h/E$$

Since $E = V/l$ where l is the length of the intrinsic semiconductor,

$\therefore$
$$I = n_i e (\mu_e + \mu_h) AV/l$$

$\therefore$
$$\frac{V}{I} = \frac{l}{A} \cdot \frac{1}{n_i e (\mu_e + \mu_h)} = \rho_i \frac{l}{A}$$

where ρ is the resistivity of the semoconductor. It is given by

$$\rho_i = \frac{1}{n_i e (\mu_e + \mu_h)} \text{ ohm-m}$$

The electrical conductivity which is the reciprocal of resistivity is given by

$$\sigma_i = n_i e (\mu_e + \mu_h) \text{ S/m}$$

Now, current density, $J = I/A$ $\therefore$ $J = n_i e (\mu_e + \mu_h)E = \sigma_i E$ $\therefore$ $\sigma_i = J/E$

Obviously, conductivity of semiconductors depends on two factors (*i*) nunber of current carriers present per unit volume and (*ii*) the mobility of the current carriers. It is found that with increase in temperature, n_i as well p_i increase and correspondingly, the conductivity of intrinsic semiconductors increases *i.e.* resistivity decreases.

Find the intrinsic carrier concentration in Silicon at 300^0K for which $N = 3 \times 10^{25} m^{-3}$, $E_g = 1.1 ev, \mu_e = 0.14 m^2/V$ -s and $\mu_h = 0.05 m^2/V$ -s.

Also, find the conductivity of silicon

$$n_i = N exp \left(-\frac{E_g}{2kT} \right)$$

Now,

$$N = 3 \times 10^{25} m^{-3}, E_g = 1.1 \times 1.6 \times 10^{-19} = 1.76 \times 10^{-19} J$$
$$k = 1.38 \times 10^{-23} J/K, T = 300^o K$$

Therefore,

$$n_i = \frac{3 \times 10^{25}(-1.76 \times 10^{-19})}{3 \times (1.38 \times 10^{-23} \times 300)} = 2 \times 10^{16} m^{-3}$$

$$\sigma = n_i(\mu_e + \mu_h) = 2 \times 10^{16} \times 1.6 \times 10^{-19}(0.14 + 0.05)$$
$$= 0.61 \times 10^{-3} S/m$$

51.21. Hole Formation in Semiconductors

The formation of a hole which is a positive charge carrier is explained below :

As shown in Fig. 51.21, suppose the covalent bond is broken at *A* and the electron has moved through the crystal lattice leaving behind a hole in the covalent bond. An electron at *B* may jump into the vacant hole at *A* and later, an electron at *C* may jump into the hole at *B* and so on. In this way, by a succession of electron movements, a hole will appear at *G* and a negative charge would have moved from *G* to *A*. It would, however, be more convenient to regard positive charge to have moved from *A* to *G* and this conception gives rise to a hole as a positive charge carrier as if it were an electron with a positive charge. It should be clearly understood that these holes are due to the movement of electrons in the valence band and that each electron movement corresponds to a collision. The drift velocity of holes is, obviously, much less than the drift velocity of electrons.

Alternatively, an intrinsic semiconductor may be defined as one *in which the number of conduction electrons is equal to the number of holes.*
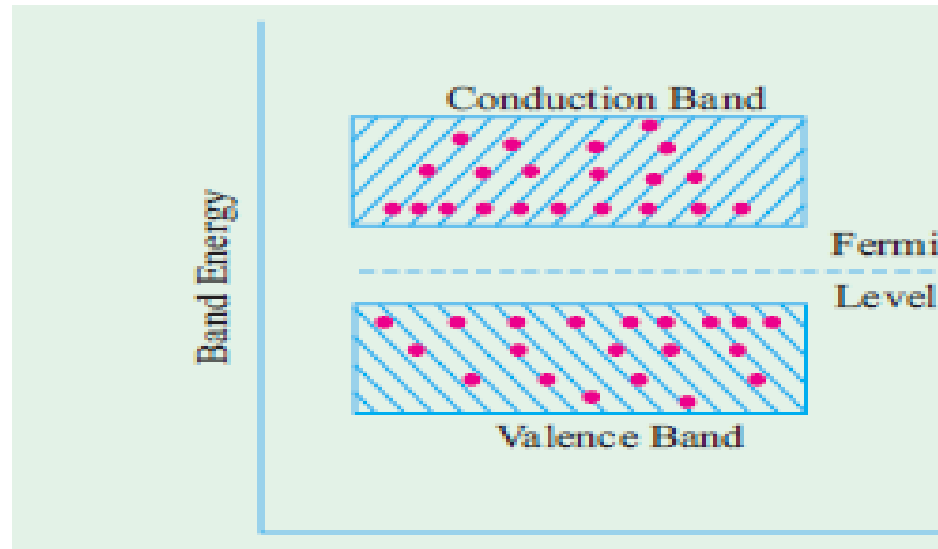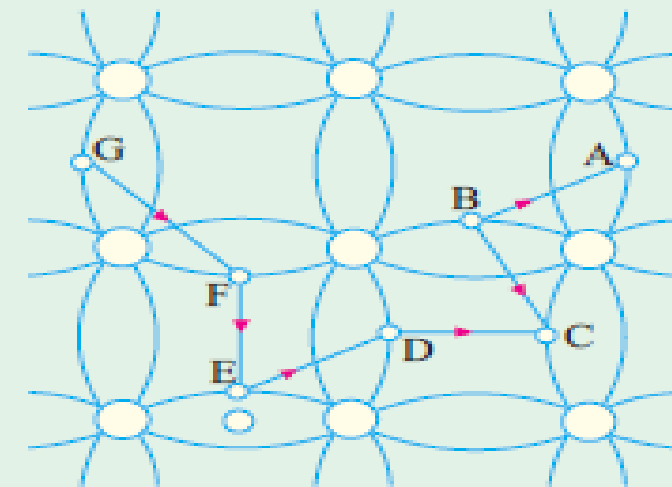

Fig. 51.20

Fig. 51.21

Schematic energy band diagram of an intrinsic semiconductor at room temperature is shown in Fig. 51.20. Only two bands *i.e.* valence and conduction bands have been shown since lower filled bands are not of any consequence. Here, Fermi level* lies exactly in the middle of the forbidden energy gap.

51.23. Extrinsic Semiconductors

Those intrinsic semiconductors to which some suitable impurity or doping agent or dopant has been added in extremely small amounts (about 1 part in 10^8) are called *extrinsic* or impurity semiconductors.

The usual doping agents are :

1. *pentavalent* atoms having five valence electrons (arsenic, antimony, phosphorus) or
2. *trivalent* atoms having three valence electrons (gallium, indium, aluminium, boron).

Pentavalent doping atom is known as *donor* atom because it donates or contributes one electron to the conduction band of pure germanium. The trivalent atom, on the other hand, is called *acceptor* atom because it accepts one electron from the germanium atom.

Depending on the type of doping material used, extrinsic semiconductors can be sub-divided into two classes :

(*i*)   N-type semiconductors and   (*ii*) P-type semiconductors.

(*a*) **N-type Extrinsic Semiconductor.** This type of semiconductor is obtained when a pentavalent material like antimonty (Sb) is added to pure germanium crystal. As shown in Fig. 51.24 (*a*), each antimony atom forms covalent bonds with the surrounding four germanium atoms with the help of four of its five electrons. The fifth electron is superfluous and is loosely bound to the antimony atom. Hence, it can be easily excited from the valence band to the conduction band by the application of electric field or increase in thermal energy. Thus, practically every antimony atom introduced into the germanium lattice, contributes one conduction electron into the germanium lattice *without creating a positive hole*. Antimony is called *donor* impurity and makes the pure germanium an N-type (N for negative) extrinsic semi-conductor. As an aid to memory, the student should associate the N in doNor with N in the N-type material and N in Negative charge carrier.
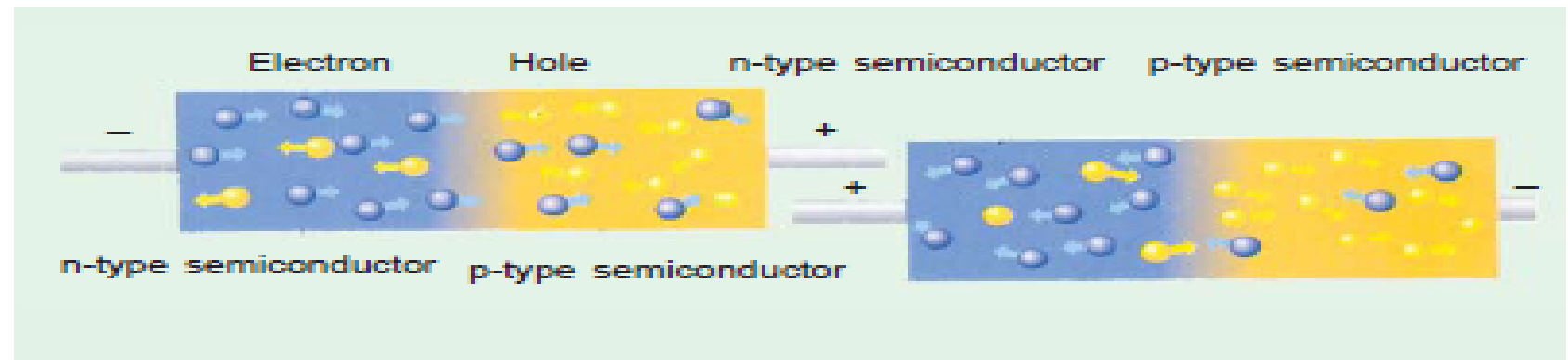
It may be noted that by giving away its one valence electron, the donor atom becomes a positively-charged *ion*. But it cannot take part in conduction because it is firmly fixed or tied into the crystal lattice. It will be seen that apart from electrons and holes *intrinsically available in germanium*, the addition of antimony greatly increases the number of conduction electrons. Hence, concentration of electrons in the conduction band is increased and exceeds the concentration of holes in the valence band. Because of this, Fermi level shifts upwards towards the bottom of the conduction band as shown in Fig. 51.24 (b),* because the number of charge carriers has become more in conduction band than in valence band.

In terms of energy levels, the fifth antimony electron has an energy level (called donor level) just below the conduction band. Usually, the donor level is 0.01 eV below conduction band for germanium and 0.054 eV for silicon.

It is seen from the above description that in *N*-type semiconductors, electrons *are the majority carriers* while holes constitute the minority carriers. Hence, *N*-type semiconductor conducts principally by electrons in the nearly empty conduction band and the process is called '*excess*' conduction.

Another point worth noting is that even though *N*-type semiconductor has excess of electrons,

--- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

* Since the number of electrons as compared to the number of holes increases with temperature, the position of Fermi level also changes considereably with temperature.

still it is electrically neutral. It is so because by the addition of donor impurity, number of electrons available for conduction purposes becomes more than the number of holes available intrinsically. But the total charge of the semiconductor does not change because the donor impurity brings in as much negative charge (by way of electrons) as positive charge (by way of protons in its nucleus).
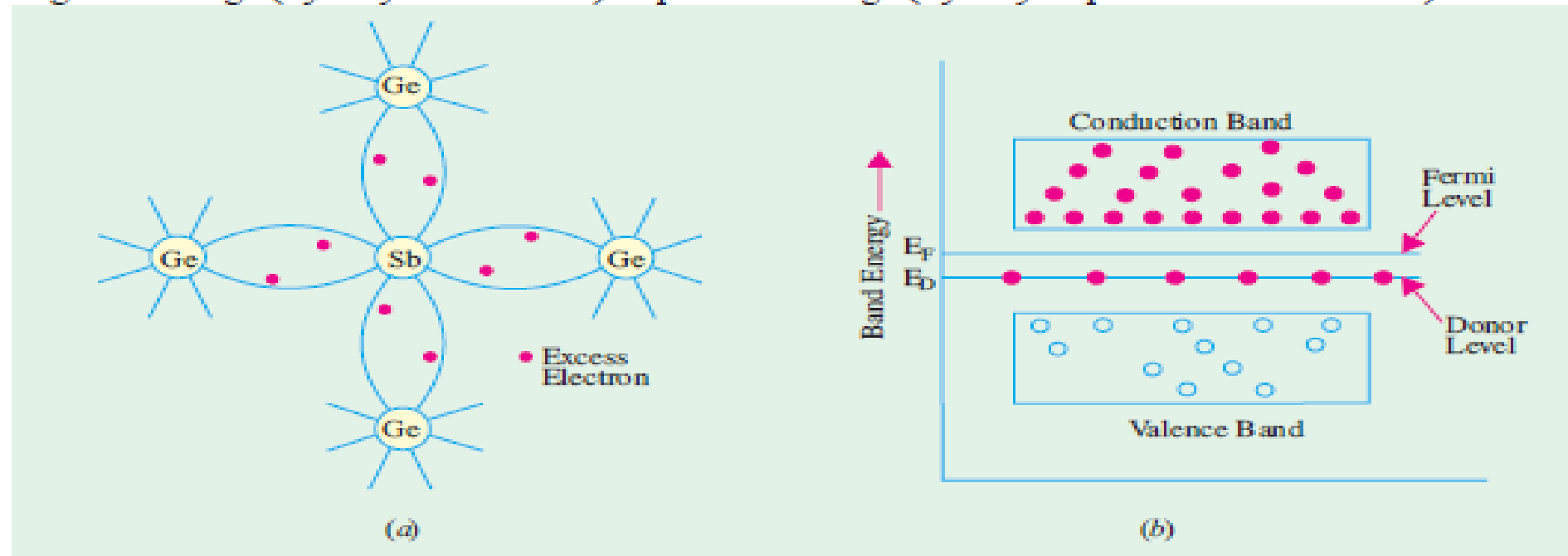

Fig. 51.24

(b) *P-type Extrinsic Semiconductor.* This type of semiconductor is obtained when traces of a trivalent like boron (B) are added to a pure germanium crystal.

In this case, the three valence electrons of boron atom form covalent bonds with four surrounding germanium atoms but one bond is left incomplete and gives rise to a hole as shown in Fig. 51.25 (a).

Thus, boron which is called an *acceptor* impurity causes as many positive holes in a germanium crystal as there are boron atoms thereby producing a P-type (P for positive) extrinsic semiconductor. As an aid to memory, the student should associate the **P** in acce**P**tor with **P** in P-type material and **P** with **P**ositive charge carrier.

In this type of semiconductor, conduction is by the movement of holes in the valence band. Accordingly, holes form the majority carriers *whereas electrons constitute minority carriers*. The process of conduction is called 'deficit' conduction.
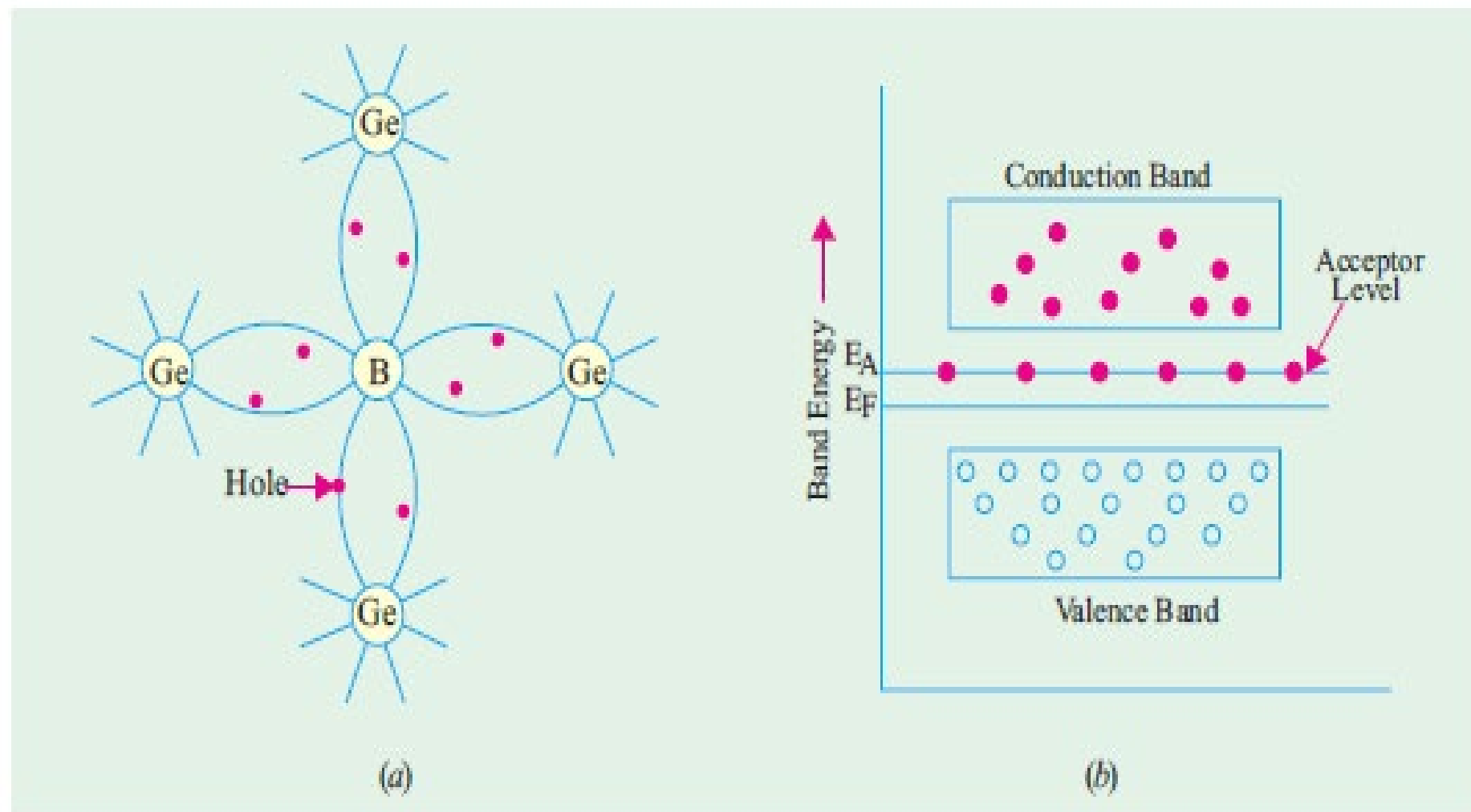
Fig. 51.25

Since concentration of holes in the valence band is more than the concentration of electrons in the conduction band, Fermi level shifts nearer to the valence band [Fig. 51.25 (b)]. The acceptor level lies immediately above the Fermi level. Conduction is by means of hole movement at the top of valence band, acceptor level readily accepting electrons from the valence band.

Again, it may be noted that even though P-type semiconductor has excess of holes for conduction purposes, on the whole it is electrically neutral for the same reasons as given above.

51.24. Majority and Minority Carriers

In a piece of *pure* germanium or silicon, no free charge carriers are available at 0°K. However, as its temperature is raised to room temperature, some of the covalent bonds are broken by heat energy and as a result, electron-hole pairs are produced. These are called thermally-generated charge carriers. They are also known as intrinsically-available charge carriers. Ordinarily, their number is quite small.

Fig. 51.26

An intrinsic of pure germanium can be converted into a P-type semiconductor by the addition of an acceptor impurity which adds a large number of holes to it. Hence, a P-type material contains following charge carriers :

(a)   large number of positive holes—most of them being the added impurity holes with only a very small number of thermally generated ones;

(b)   a very small number of thermally-generated electrons (the companions of the thermally generated holes mentioned above).

Obviously, in a P-type material, the number of holes (both added and thermally-generated) is much more than that of electrons. Hence, in such a material, holes constitute *majority* carriers and electrons form *minority* carriers as shown in Fig. 51.26 (a).

Similarly, in an N-type material, the number of electrons (both added and thermally-generated) is much larger than the number of thermally-generated holes. Hence, in such a material, electrons are majority carriers whereas holes are minority carriers as shown in Fig. 51.26 (b).
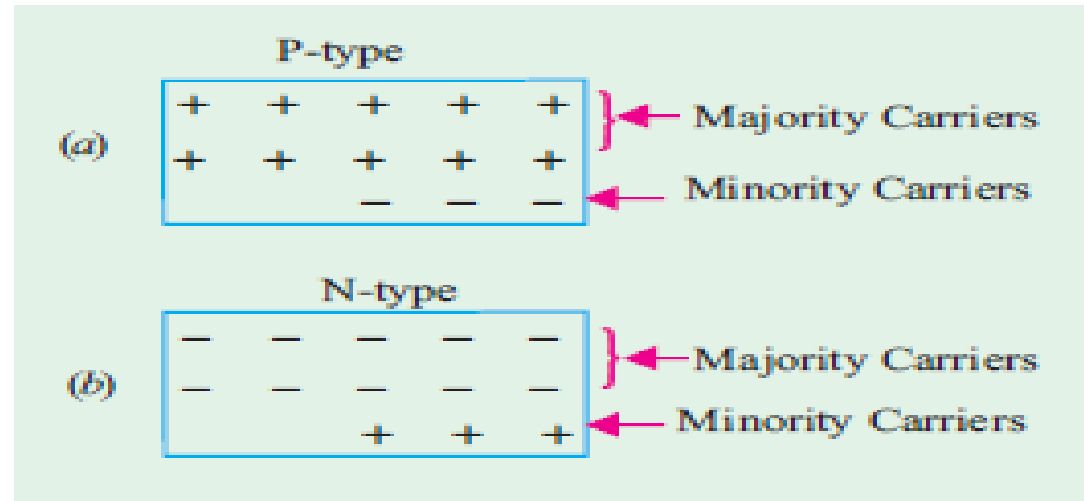
EXTRINSIC SEMICONDUCTOR

An extrinsic semiconductor is one that has been doped with impurities to modify the number and type of free charge carriers. An extrinsic semiconductor is a semiconductor that has been *doped,* that is, into which a doping agent has been introduced, giving it different electrical properties than the intrinsic (pure) semiconductor. Doping involves adding dopant atoms to an intrinsic semiconductor, which changes the electron and hole carrier concentrations of the semiconductor at thermal equilibrium. Dominant carrier concentrations in an extrinsic semiconductor classify it as either an n-type or p-type semiconductor.

The intrinsic semiconductor has little current conduction capability at room temperature. To be useful in electronic devices, the pure semiconductor must be altered so as to significantly increase its conducting properties. This is achieved by adding a small amount of suitable impurity to a semiconductor. It is then called *impurity* or *extrinsic semiconductor*. The process of adding impurities to a semiconductor is known as *doping*. The amount and type of such impurities have to be closely controlled during the preparation of extrinsic semiconductor. Generally, for 10^8 atoms of semiconductor, one impurity atom is added.

The purpose of adding impurity is to increase either the number of free electrons or holes in the semiconductor crystal. As we shall see, if a pentavalent impurity (having 5 valence electrons) is added to the semiconductor, a large number of free electrons are produced in the semiconductor. On the other hand, addition of trivalent impurity (having 3 valence electrons) creates a large number of holes in the semiconductor crystal. Depending upon the type of impurity added, extrinsic semiconductors are classified into: **(i)** *n*-type semiconductor **(ii)** *p*-type semiconductor

The electrical properties of extrinsic semiconductors make them essential components of many electronic devices. A pure or intrinsic conductor has thermally generated holes and electrons. However these are relatively few in number. An enormous increase in the number of charge carriers can by achieved by introducing impurities into the semiconductor in a controlled manner. The result is the formation of an extrinsic semiconductor. This process is referred to as doping. There are basically two types of impurities: donor impurities and acceptor impurities. Donor impurities are made up of atoms (arsenic for example) which have five valence electrons. Acceptor impurities are made up of atoms (gallium for example) which have three valence electrons.

The two types of extrinsic semiconductor

N-TYPE AND *P*-TYPE SEMICONDUCTORS

Semi conductive materials do not conduct current well and are of limited value in their intrinsic state. This is because of the limited number of free electrons in the conduction band and holes in the valence band. Intrinsic silicon (or germanium) must be modified by increasing the number of free electrons or holes to increase its conductivity and make it useful in electronic devices. This is done by adding impurities to the intrinsic material. Two types of extrinsic (impure) semi conductive materials, *n*-type and *p*-type, are the key building blocks for most types of electronic devices.

Since semiconductors are generally poor conductors, their conductivity can be drastically increased by the controlled addition of impurities to the intrinsic (pure) semi conductive material. This process, called **doping**, increases the number of current carriers (electrons or holes). The two categories of impurities are *n*-type and *p*-type.

N-TYPE SEMICONDUCTORS

When a small amount of pentavalent impurity is added to a pure semiconductor, it is known as **n-type semiconductor.** To increase the number of conduction-band electrons in intrinsic silicon, **pentavalent** impurity atoms are added. These are atoms with five valence electrons such as arsenic (As), phosphorus (P), bismuth (Bi), and antimony (Sb).

Extrinsic semiconductors with a larger electron concentration than hole concentration are known as n-type semiconductors. The phrase 'n-type' comes from the negative charge of the electron.

In n-type semiconductors, electrons are the majority carriers and holes are the minority carriers. N-type semiconductors are created by doping an intrinsic semiconductor with donor impurities. In an n-type semiconductor, the Fermi energy level is greater than that of the intrinsic semiconductor and lies closer to the conduction band than the valence band. Arsenic has 5 valence electrons, however, only 4 of them form part of covalent bonds.

The 5th electron is then free to take part in conduction. The electrons are said to be the majority carriers and the holes are said to be the minority carriers. If the added impurity is a pentavalent atom then the resultant semiconductor is called N-type semiconductor. Examples of pentavalent impurities are Phosphorus, Arsenic, Bismuth, Antimony etc. A pentavalent impurity has five valance electrons. Fig 1.4a shows the crystal structure of N-type semiconductor material where four out of five valance electrons of the impurity atom(antimony) forms covalent bond with the four intrinsic semiconductor atoms. The fifth electron is loosely bound to the impurity atom. This loosely bound electron can be easily
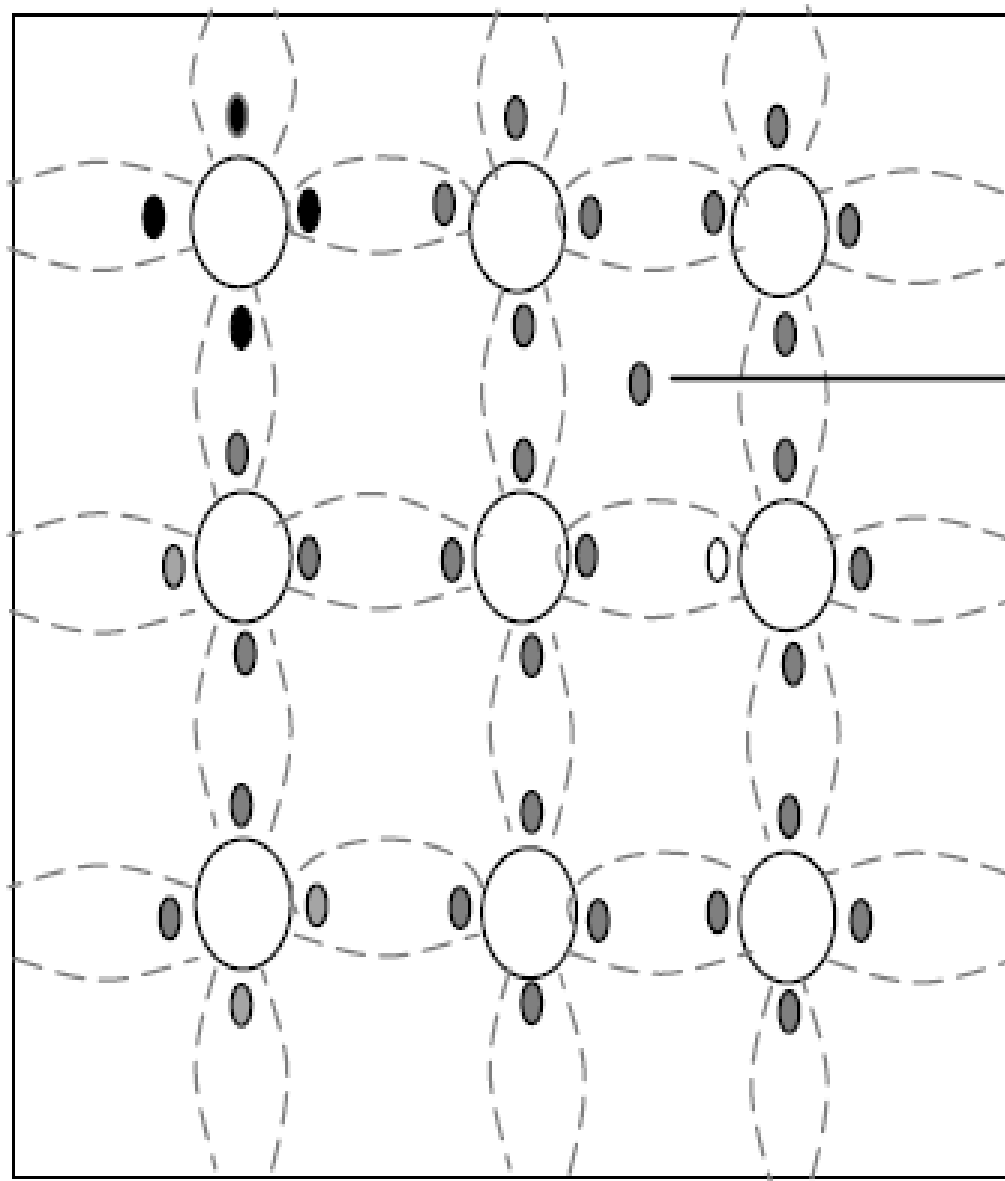
Fifth valance electron of SB

E_c

CB

E_d

Donor energy level

E_v

VB
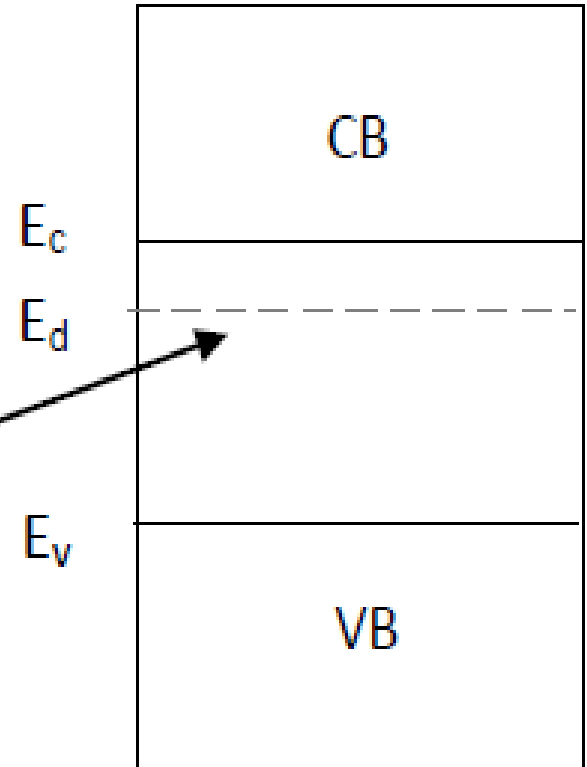
Fig. 1.4a crystal structure of N type SC

Fig. 1.4bEnergy band diagram of N type

Excited from the valance band to the conduction band by the application of electric field or increasing the thermal energy. The energy required to detach the fifth electron form the impurity atom is very small of the order of 0.01ev for Ge and 0.05 eV for Si. The effect of doping creates a discrete energy level called donor energy level in the forbidden band gap with energy level Ed slightly less than the conduction band (fig 1.4b).

The difference between the energy levels of the conducting band and the donor energy level is the energy required to free the fifth valance electron (0.01 eV for Ge and 0.05 eV for Si). At room temperature almost all the fifth electrons from the donor impurity atom are raised to conduction band and hence the number of electrons in the conduction band increases significantly. Thus every antimony atom contributes to one conduction electron without creating a hole. In the N-type sc the no. of electrons increases and the no. of holes decreases compared to those available in an intrinsic sc. The reason for decrease in the no. of holes is that the larger no. of electrons present increases the recombination of electrons with holes. Thus current in N type sc is dominated by electrons which are referred to as majority carriers. Holes are the minority carriers in N type sc

Free (conduction) electron from Sb atom
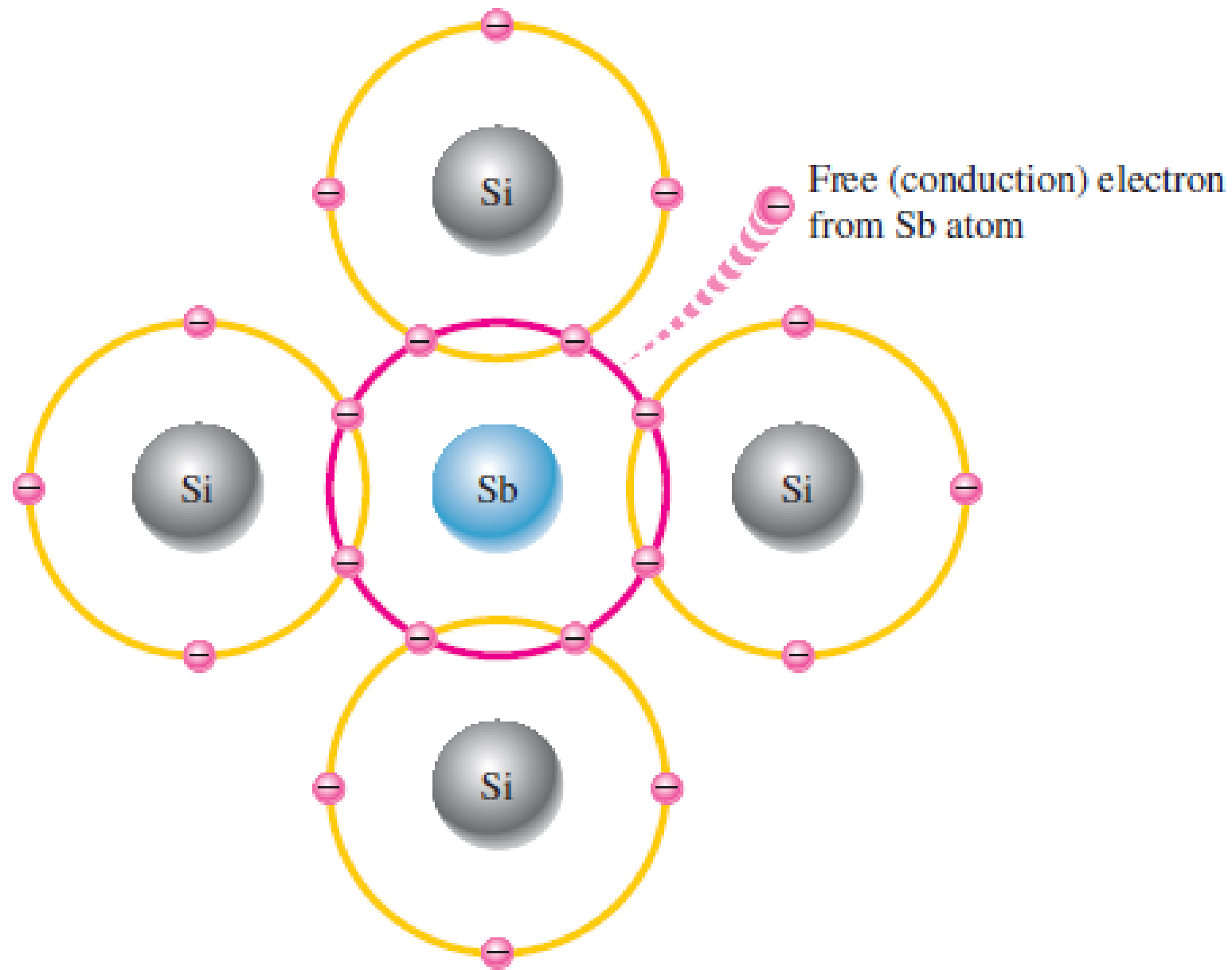
Pentavalent impurity atom in a silicon crystal structure. An antimony (Sb) impurity atom is shown in the center. The extra electron from the Sb atom becomes a free electron.

As illustrated in Figure 1–17, each pentavalent atom (antimony, in this case) forms covalent bonds with four adjacent silicon atoms. Four of the antimony atom's valence electrons are used to form the covalent bonds with silicon atoms, leaving one extra electron. This extra electron becomes a conduction electron because it is not involved in bonding. Because the pentavalent atom gives up an electron, it is often called a *donor atom.* The number of conduction electrons can be carefully controlled by the number of impurity atoms added to the silicon. A conduction electron created by this doping process does not leave a hole in the valence band because it is in excess of the number required to fill the valence band.

51.28. Conductivity of Extrinsic Semiconductors

The general expression for current density (derived above) in the case of an extrinsic semiconductor when an electric field is employed is

$$J = (ne\mu_e + pe\mu_h)E \qquad \qquad ...(i)$$

(a) If it is an N-type semiconductor, then the above expression becomes

$$J_n = e(n_n\mu_e + p_n\mu_h)E$$

where n_n and p_n represent the electron and hole densities in the N-type semiconductor *after doping*.

(b) If it is a P-type semiconductor, then

$$J_p = e(n_p\mu_e + p_p\mu_h)E$$

where n_p and p_p represent similar quantities in a P-type semiconductor after doping.

The conductivity is given by $\sigma = J/E$

$$\therefore \qquad \qquad \sigma = ne\mu_e + pe\mu_h$$

or $$\qquad \qquad \sigma_n = e(n_n\mu_e + p_n\mu_h) \qquad \qquad \text{— for } N\text{-type}$$

and $$\qquad \qquad \sigma_p = e(n_p\mu_e + p_p\mu_h) \qquad \qquad \text{— for } P\text{-type}$$

(i) In N-type semiconductors, electrons form the majority carriers although holes are also available as minority carriers.

The current density in such a semiconductor is given by Eq. (i) above. However, since electron density in such extrinsic semiconductors is much more than hole density *i.e.* $n_n \gg p_n$, the above expressions are simplified to

$$J_n = n_n e\mu_e E \qquad \text{and} \qquad \sigma_n = n_n e \mu_e \qquad \qquad ...(ii)$$

(ii) In P-type semiconductors, conduction is by means of holes in the valence band which form majority carriers in this case although electrons are available as minority carriers.

Since in such extrinsic semiconductors, $n_p \ll p_p$, the above expressions become

$$J_p = p_p e \mu_h E \qquad \text{and} \qquad \sigma_p = p_p e \mu_h \qquad \qquad ...(iii)$$

The addition of pentavalent impurity provides a large number of free electrons in the semiconductor crystal. Typical examples of pentavalent impurities are *arsenic* and *antimony*. Such impurities which produce *n*-type semiconductor are known as *donor impurities* because they donate or provide free electrons to the semiconductor crystal.

To explain the formation of *n*-type semiconductor, consider a pure germanium crystal. We know that germanium atom has four valence electrons. When a small amount of pentavalent impurity like arsenic is added to germanium crystal, a large number of free electrons become available in the crystal. The reason is simple. Arsenic is pentavalent *i.e.* its atom has five valence electrons. An arsenic atom fits in the germanium crystal in such a way that its four valence electrons form covalent bonds with four germanium atoms. The *fifth* valence electron of arsenic atom finds no place in co-valent bonds and is thus free as shown in Fig. 5.11. Therefore, for each arsenic atom added, one free electron will be available in the germanium crystal. Though each arsenic atom provides one free electron, yet an extremely small amount of arsenic impurity provides enough atoms to supply millions of free electrons.

FREE ELECTRON

PENTAVALENT IMPURITY ATOM
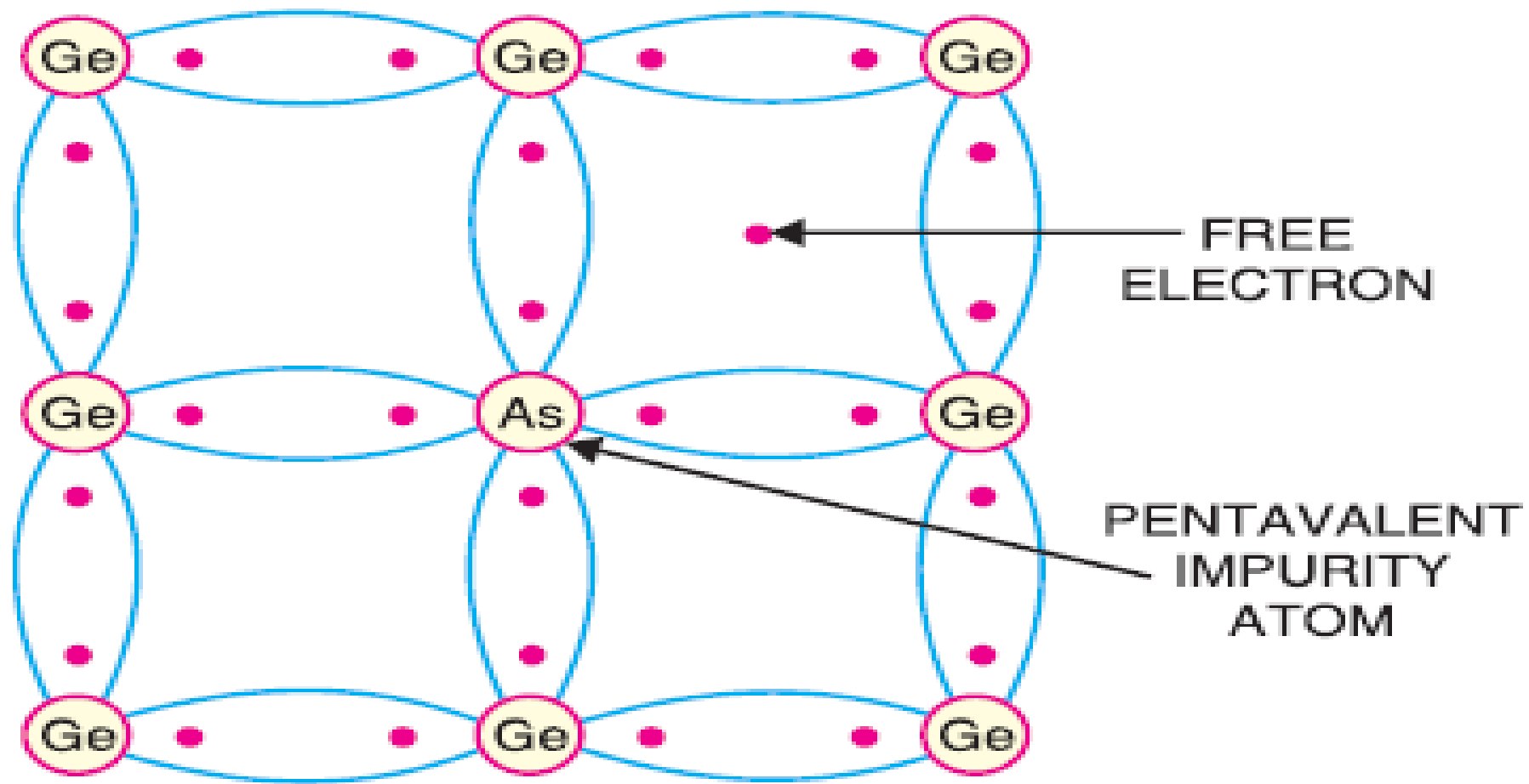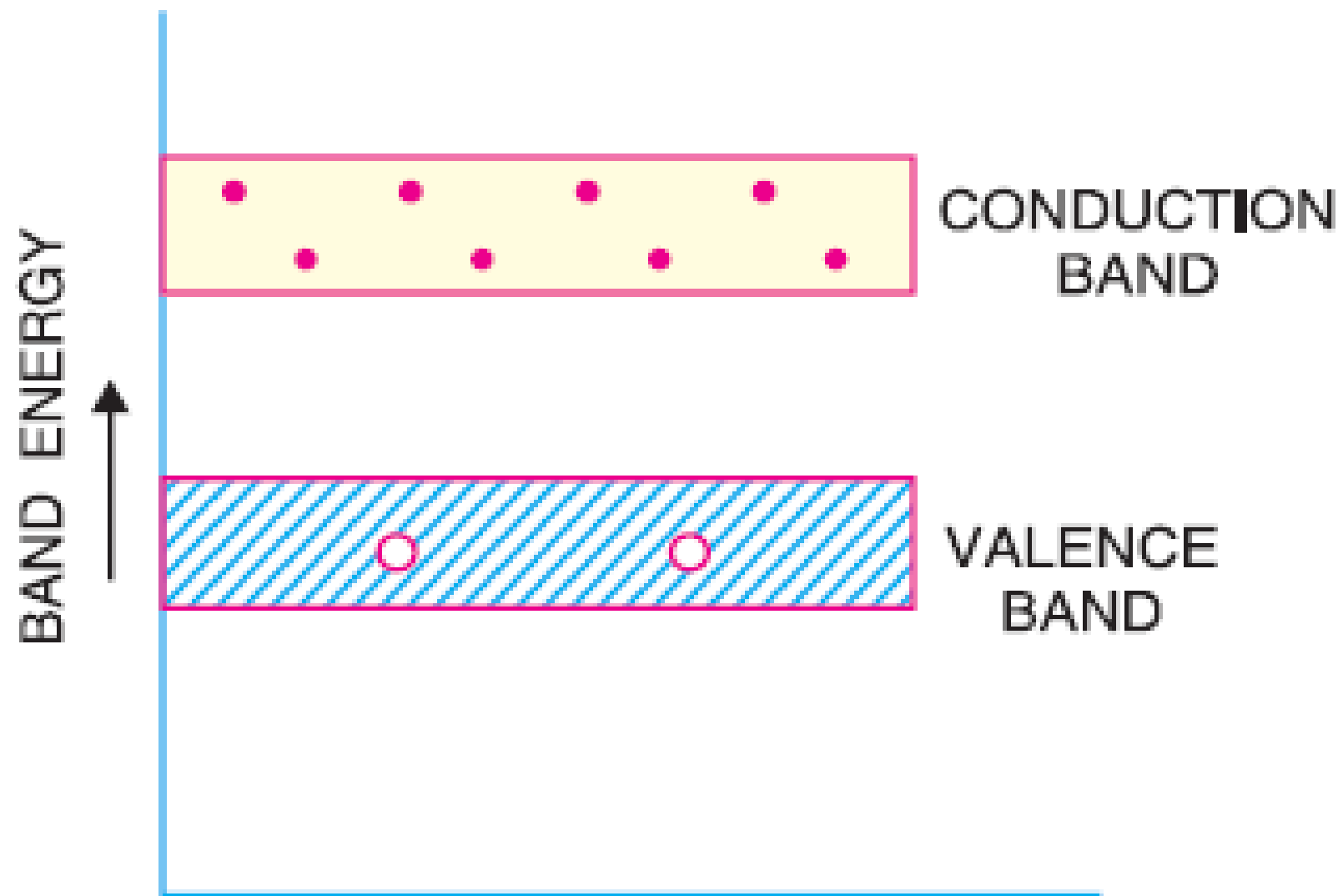
Fig. 5.11

Fig. 5.12

Fig. 5.12 shows the energy band description of *n*-type semi-conductor. The addition of pentavalent impurity has produced a number of conduction band electrons *i.e.*, free electrons. The four valence electrons of pentavalent atom form covalent bonds with four neighbouring germanium atoms. The fifth left over valence electron of the pentavalent atom cannot be accommodated in the valence band and travels to the conduction band. The following points may be noted carefully :

(*i*) Many new free electrons are produced by the addition of pentavalent impurity.

(*ii*) Thermal energy of room temperature still generates a few hole-electron pairs. However, the number of free electrons provided by the pentavalent impurity far exceeds the number of holes. It is due to this predominance of electrons over holes that it is called *n*-type semiconductor (*n* stands for negative).

N-TYPE CONDUCTIVITY

The current conduction in an *n*-type semiconductor is *predominantly* by free electrons *i.e.* negative charges and is called *n-type* or *electron type conductivity*. To understand *n*-type conductivity, refer to Fig. 5.13. When p.d. is applied across the *n*-type semiconductor, the free electrons (donated by impurity) in the crystal will be directed towards the positive terminal, constituting electric current. As the current flow through the crystal is by free electrons which are carriers of negative charge, therefore, this type of conductivity is called negative or *n*-type conductivity. It may be noted that conduction is just as in ordinary metals like copper.

POSITIVELY CHARGED IMPURITY ATOM

FREE ELECTRON

n-type
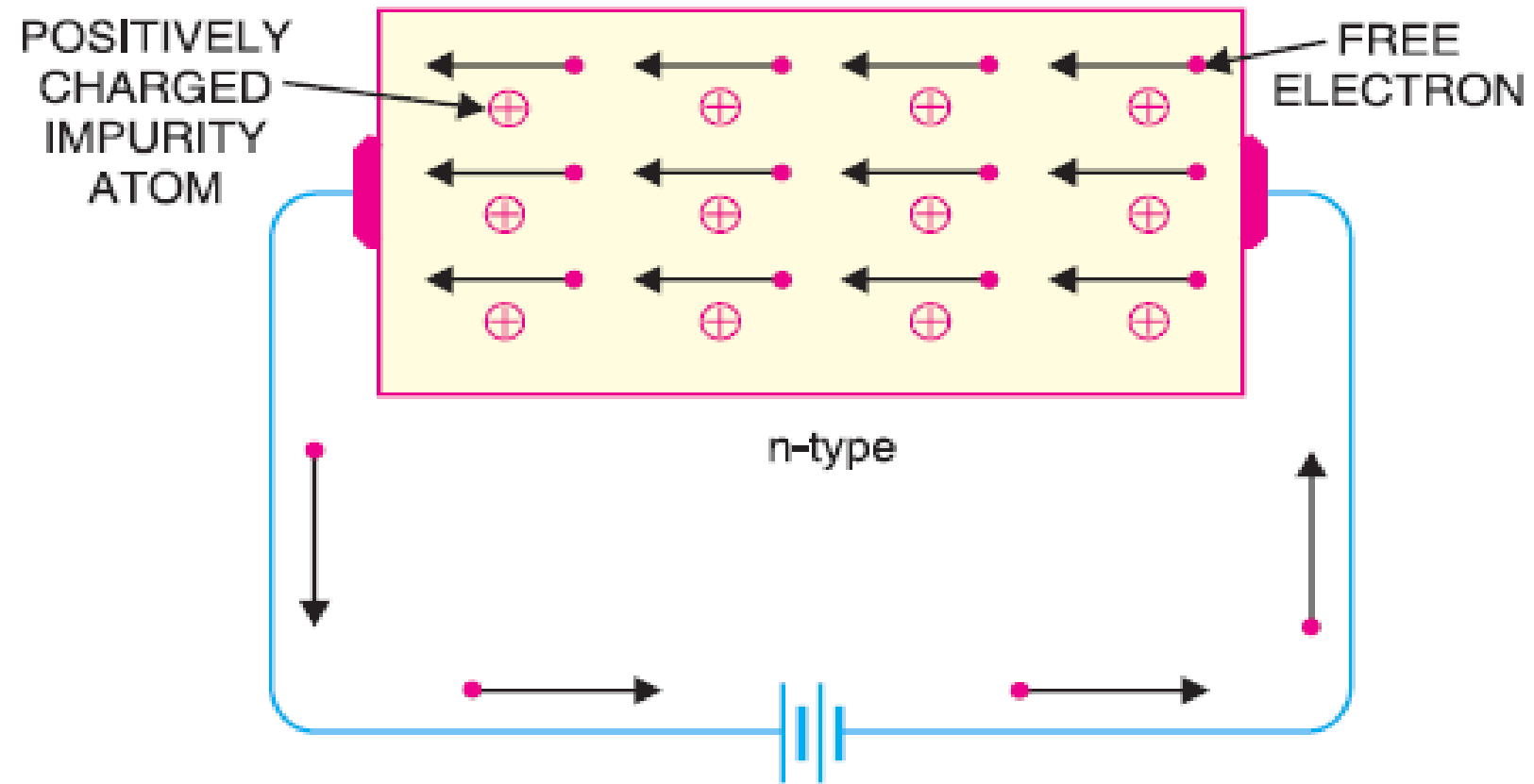
Fig. 5.13

Example 51.4. *What length of a round copper wire of diameter 1 mm will have a resistance of 1 k Ω if copper conductivity is 60 MS/m. A cylindrical piece of silicon having a diameter of 1 mm is doped with 10^{20} m^{-3} atoms of phosphorous which are fully ionized. What length of this silicon would be required to give a resistance of 1 k Ω if electronic mobility in silicon is 0.1 m^2 / V-s ?*

(Electronic Devices & Circuits, Pune Univ. 1994)

Solution. $R = 1$ k $\Omega = 1000\ \Omega$, $\sigma = 60 \times 10^6$ S/m, $A = \pi d^2/4 = \pi \times (1 \times 10^{-3})^2/4 m^2$

$R = l/\sigma A$

$l = \sigma A R = 60 \times 10^6 \times (\pi \times 10^{-6}/4) \times 1000 = 47{,}100$ m $= \mathbf{47.1\ km}$

For Silicon Wire

$\sigma = n_i e \mu_e = 10^{20} \times 1.6 \times 10^{-19} \times 0.1 = 1.6$ S/m

$l = \sigma A R = 1.6 \times (\pi \times 10^{-6}/4) \times 1000$

$= 1.26 \times 10^{-3} = \mathbf{1.26\ mm}$

Example 51.5. *Calculate the intrinsic conductivity of silicon at room temperature if $n = 1.41 \times 10^{16}$ m^{-3}, $\mu_e = 0.145$ m^2/V-s, $\mu_h = 0.05$ m^2/V-s and $e = 1.6 \times 10^{-19}$ C. What are the individual contributions made by electrons and holes ?* *(Electronic Engg., Nagpur Univ. 1991)*

Solution. As seen from Art. 1.27, the conductivity of an intrinsic semiconductor is given by

$\sigma_i = n_i e\, \mu_e + n_i e\, \mu_h$

$= 1.41 \times 10^{16} \times 1.6 \times 10^{-19} \times 0.145 + 1.41 \times 10^{16} \times 1.6 \times 10^{-19} \times 0.05$

$= 0.325 \times 10^{-3} + 0.112 \times 10^{-3}$ S/m $= \mathbf{0.437 \times 10^{-3}}$ **S/m**

Contribution by electrons $= \mathbf{0.325 \times 10^{-3}}$ **S/m**

Contribution by holes $= \mathbf{0.112 \times 10^{-3}}$ **S/m**

Example 51.6. *Calculate the donor concentration in N-type germanium having resistivity of 100 Ω-m. Derive the formula you use. Take $e = 1.6 \times 10^{-19}$ C, $\mu_e = 0.36$ m^2 $V^{-1} s^{-1}$.*

(Electronics ; Nagpur Univ. 1990)

Solution. As seen from Art. 1.29, $\qquad \rho_n = 1/N_d e\, \mu_e$

$\therefore \quad 100 = 1/N_d \times 1.6 \times 10^{-19} \times 0.36 ; \qquad N_d = \mathbf{1.74 \times 10^{17}}$ **atoms/m^3**

Example 51.7. *An N-type silicon has a resistivity of 1500 Ω-m at a certain temperature. Compute the electron-hole concentration given that $n_i = 1.5 \times 10^{16}$ m^{-3}, $\mu_e = 0.14$ m^2/V-s, $\mu_h = 0.05$ m^2/V-s and $e = 1.6 \times 10^{-19}$ C.*

Solution. Being N-type silicon, it is assumed that $n \gg p$

$\therefore \qquad \sigma = e\,(n\mu_e + p\mu_h) = n_e\, \mu_e$

$\therefore \qquad \rho = 1/n_e \mu_e$ or $15 = 1/n \times 1.6 \times 10^{-19} \times 0.14$ or $n = 3.1 \times 10^{20}$ m^{-3}

Now, $\qquad np = n_i^2$ or $p = n_i^2/n = (1.5 \times 10^{16})^2/(3.1 \times 10^{20}) = 2 \times 10^{12}$ m^{-3}

Example 51.8. *A specimen of pure germanium at 300° K has a density of charge carriers of $2.5 \times 10^{19}/m^3$. It is doped with donor impurity atoms at the rate of one impurity atom for every 10^6 atoms of germanium. All impurity atoms may be supposed to be ionized. The density of germanium atom is 4.2×10^{28} atoms/m^3.*

Find the resistivity of the doped germanium if electron mobility is 0.36 m^2/V-s.

Solution. Density of added impurity atoms is
$$N_d = 4.2 \times 10^{28}/10^6 = 4.2 \times 10^{22} \text{ atoms/m}^3$$

As seen, it is very large as compared to the intrinsic charge carrier density of $2.5 \times 10^{19}/m^3$ which will, therefore, be neglected. Now, as seen from Art. 1.29.

$$\sigma_n = N_d e \mu_e = 4.2 \times 10^{22} \times 1.6 \times 10^{-19} \times 0.36 = 2.42 \times 10^3 \text{ S/m}$$
$$\rho_n = 1/\sigma_n = 1/2.42 \times 10^3 = 0.413 \times 10^{-3} \text{ } \Omega\text{-m.}$$

Example 51.9. *Compute the relative concentration of silicon atoms and electron-hole pairs at 300°K. Also, calculate intrinsic resistivity of silicon. Given Avogadro's number = 6.02×10^{23} atoms/ g-atom, density = 2.33×10^6 g/m^3, atomic Wt. = 28.1, intrinsic carrier density = 1.5×10^{16} m^{-3}, $\mu_e = 0.14$ m^2/V-s, $\mu_h = 0.05$ m^2/V-s.*

Solution. n_A = Avogadro's No. × density/atomic Wt.
$$= 6.02 \times 10^{23} \times 2.33 \times 10^{16}/28.1 \cong 5 \times 10^{28} \text{ atoms/m}^3$$

Since intrinsic concentration n_i *i.e.*, electron-hole pairs/m^3 is 1.5×10^{16}

$$\therefore \quad \frac{n_A}{n_i} = \frac{5 \times 10^{28}}{1.5 \times 10^{16}} \cong 3.3 \times 10^{12}$$

It means that there are 3.3×10^{12} Si atoms for each electron-hole pair.

Since in a pure semiconductor, $n = p = n_i$, intrinsic conductivity is given by

$$\sigma_i = n_i e (\mu_e + \mu_h) = 1.5 \times 10^{16} \times 1.6 \times 10^{-19} (0.14 + 0.05) = 0.456 \times 10^{-3} \text{ S/m}$$

Hence, intrinsic resistivity is given by

$$\rho_i = 1/\sigma_i = 1/0.456 \times 10^{-3} = 2193 \text{ } \Omega\text{-m}$$

Example 51.10. *Silicon is doped with acceptor atoms to a density of 10^{22} m^{-3}. If it is assumed that all acceptor centres are ionized, calculate the conductivity of the extrinsic silicon. Given that intrinsic density is 1.4×10^{16} m^{-3}, $\mu_e = 0.145$ m^{-3} and $\mu_h = 0.05$ m^{-3}.*

(Electronic Devices & Circuits, Pune Univ. 1992)

Solution. The minority carrier density can be found from the equation given in Art. 1.30, *i.e* $np = n_i^2$.

Now, $p = 10^{22}$ $\therefore$ $n \times 10^{22} = (1.4 \times 10^{16})^2$; $n = 1.96 \times 10^{10}/m^3$

Now, $\sigma = ne \mu_e + pe \mu_h$
$$= 1.96 \times 10^{10} \times 0.145 \times 1.6 \times 10^{-19} + 10^{22} \times 0.05 \times 1.6 \times 10^{-19} = 80 \text{ S/m}$$

51.35. Recombination

Apart from drift and diffusion, a third phenomenon which occurs in semiconductors is called *recombination* that results from the collision of an electron with a hole.

The process is essentially the return of a free conduction electron to the valence band and is accompanied by the emission of energy. Obviously, the recombination rate is directly proportional to the carrier concentration for the simple reason that larger the number of carriers, the more likely is the occurrence of electron-hole recombination. This phenomenon is important in describing minority carrier flow.

As is well-known, in a semiconductor, thermal generation of electron-hole pairs also takes place continuously. Hence, there is *net recombination* rate given by the difference between the recombination and generation rates.

51.40. Junction or Barrier Voltage

Even though depletion layer is cleared of charge carriers, it has oppositely-charged fixed rows of ions on its two sides. Because of this charge separation, an electric potential difference V_B is established across the junction even when the *junction is externally isolated* (Fig. 51.34). It is known as *junction or barrier potential*. It stops further flow of carriers across the junction unless supplied by energy from an external source. At room temperature of 300°K, V_B is about 0.3 V for Ge and 0.7 V for Si.

The value of barrier voltage is given by $V_B = V_T \log_e N_a N_d / n_i^2$ where N_a, N_d, n_i^2 and V_T have the meanings explained in Art. 51.29 and 34. The value of V_T at room temperature of 300°K is given by

$$V_T = V_{300} = \frac{kT}{e} = \frac{1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}} = 26 \text{ mV}$$

$$\therefore \qquad V_B = 26 \log_e (N_a N_d / n_i^2) \text{ mV}$$

Barrier voltage depends on doping density, electronic charge and temperature. For a given junction, the first two factors are constant, thus making V_B dependent on temperature. With increase in temperature, more minority charge carriers are produced, leading to their increased drift across the junction. As a result, equilibrium occurs at a lower barrier potential. It is found that both for Ge and Si, V_B decreases by about 2 mV / °C.

$$\therefore \qquad \Delta V_B = -0.002 \, \Delta t \qquad \text{where } \Delta t \text{ is the rise in temperature in °C.}$$

The strong field set up by V_B causes drift of carriers through depletion layer. As seen from Fig. 51.35 (b), under the influence of this field, *holes drift from N-to P-region and electrons from P- to -N region*. This drift current must be equal and opposite to the diffusion current [Fig. 51.35 (a)] because under condition of equilibrium and with no external supply, net current through the crystal is zero.

$V_B = 0.3V$ for Ge
= 0.7V for Si

Potential
Barrier

Fig. 51.34

Difusion
Currents

(a)

Drift
Current

(b)
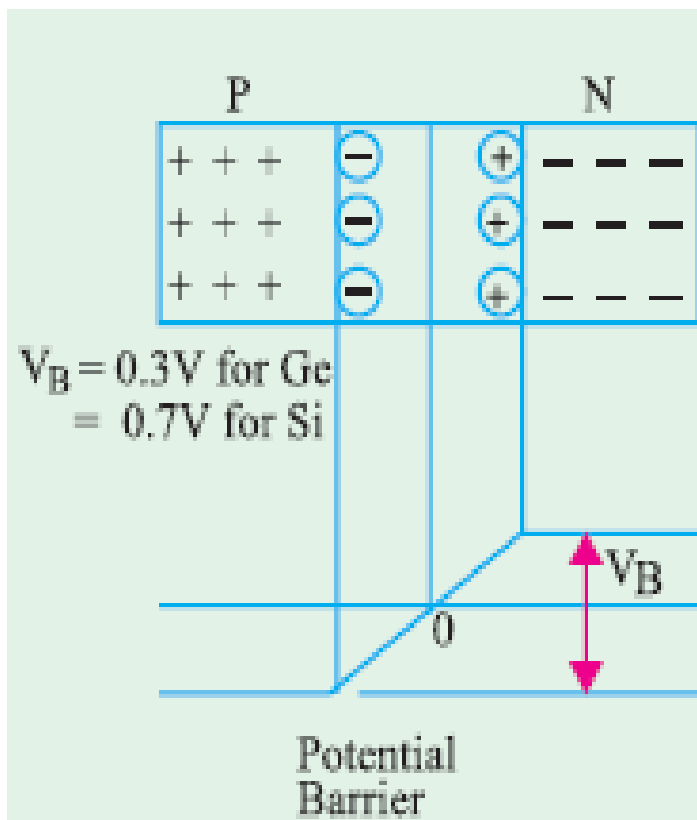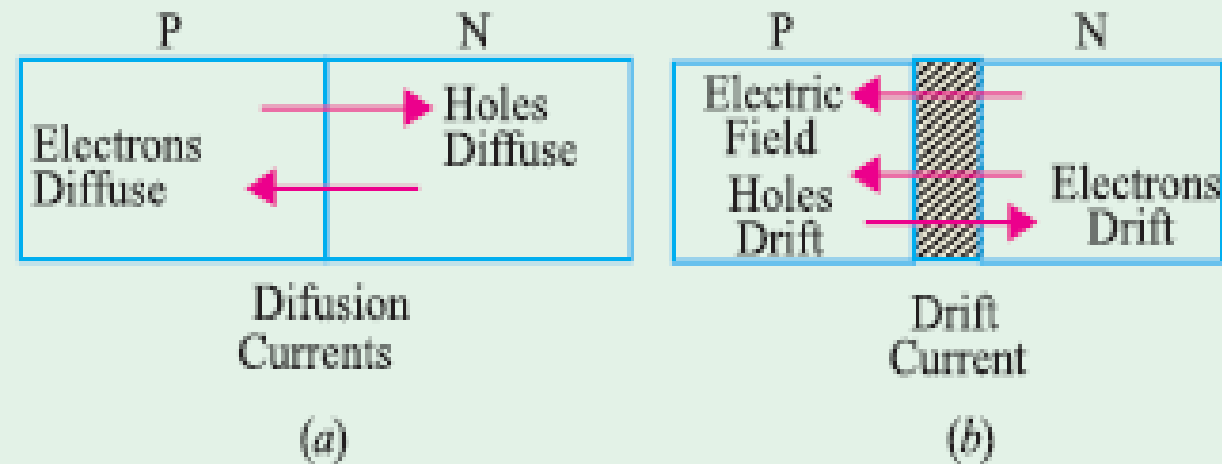
Fig. 51.35

To summarize the main points of Art. 51.39 and 51.40, we have :

1. As soon as P-N junction is formed, free electrons and holes start diffusing across the junction and recombining.

2. Their recombination leads to the appearance of a depletion layer across the junction which contains no *mobile* carriers but only immobile ions.

3. These immobile ions set up a barrier potential and hence an electric field which sets up drift current that is equal and opposite to the diffusion current when final equilibrium is reached.

Example 51.13. *Calculate the barrier potential at room temperature for P-N junction in silicon which is doped to a carrier density of 10^{21} m^{-3} on the P-side and 10^{22} m^{-3} on the N-side. The intrinsic carrier density for silicon is 1.4×10^{16} m^{-3}.* (Electronics-I, Bangalore Univ. 1992)

Solution. Using the relation given in Art 1.40, we have

$$V_B = 26 \log_e (N_a N_d / n_i^2) \, mV$$
$$= 26 \log_e 10^{21} \times 10^{22} / (1.4 \times 10^{16})^2 = 641 \, mV = 0.641 \, V$$

Example 51.14. *Calculate the change in barrier potential of a P-N junction at 300°K if doping on the N-side is increased 1000 times while keeping doping on P-side unchanged.*

Solution. As seen from Art. 51.40, at 300°K

$$V_B = 26 \log_e (N_a N_d / n^2) \, mV$$
$$\therefore \quad V_{B1} = 26 \ln (N_{a1} N_{d1} / n_i^2); \quad V_{B2} = 26 \ln (N_{a2} N_{d2} / n_i^2)$$
$$\therefore \quad V_{B2} - V_{B1} = 26 \log_e (N_{a2} N_{d2}) / (N_{a1} N_{d1})$$
$$= 26 \log_e (N_{d2} / N_{d1}) = 26 \ln 1000 = 179 \, mV$$

51.26. Electron Conductivity of a Metal

According to free electron model of an atom, the valence electrons are not attached to individual atoms but are free to move about in all directions among the atoms. These electrons are called *conduction* electrons and are said to form 'free electron cloud' or free electron 'gas' or the Fermi gas. For example in copper there is one such free electron per atom, the other 28 electrons remaining bound to the copper nuclei to form positive ion cores.

When no external field is applied to the metal, the free electrons move randomly in all directions as shown in Fig. 51.29 (*a*). However, when an external electric field is applied to the metal, the free electron motion becomes directed a shown in Fig. 51.29 (*b*). This directed flow of electrons results in a net charge displacement in a definite direction. This type of motion is known as *drift* and the phenomenon is referred to as process of conduction by drift charge. The drift velocity (*v*) of the electrons is dependent upon the electron mobility (μ_e) and the applied electric field E. The actual relation is $v = \mu_e E$
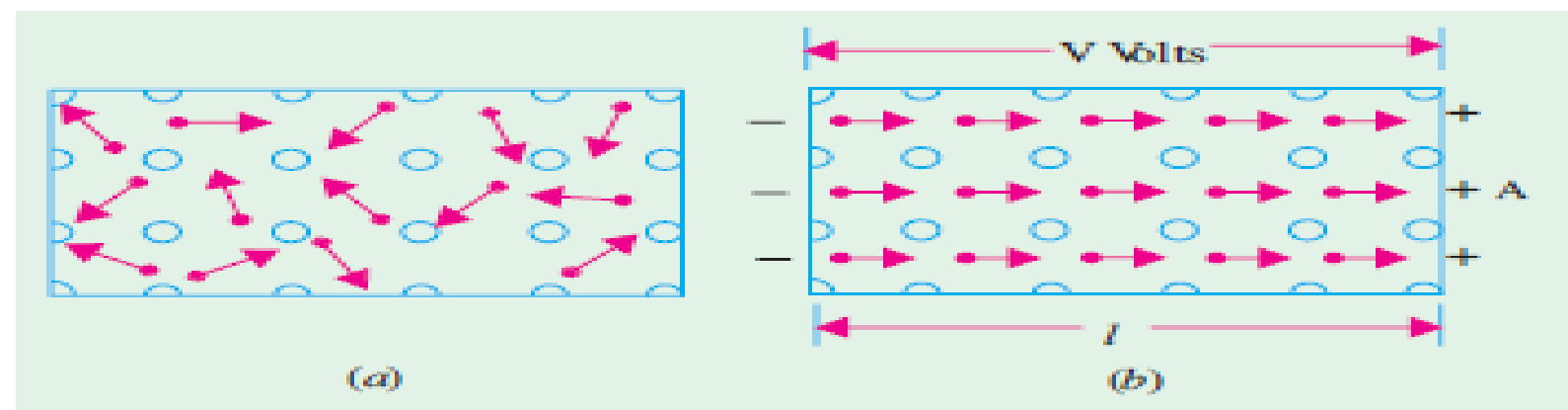
Fig. 51.29

Let,

e = electron charge (coloumb)
v = electron drift velocity (m/s)
A = conductor cross-section (m²)
n = number of free electrons per unit volume of the conductor i.e. electron density (per m³)
l = length of the conductor (m)
E = V/l — applied electric field (V/m)

Now, electric current flowing in any conductor is given by the amount of charge which flows in one second across any plane of the conductor (Fig. 51.30). The *total* number of electrons which cross the plane P of cross-section A in one second $= n \times (v \times A)$.

Charge carried by them per second is = $envA$. Hence, $I = venA$.

Substituting the value of v, we get

$$I = neA\mu_e E = nAe\mu_e V/l$$

$$\therefore \quad R = \frac{V}{I} = \frac{l}{A}\left(\frac{1}{ne\mu_e}\right) = \rho\frac{l}{A}$$

$\therefore$ resistivity $\rho = 1/ne\,\mu_e$ ohm-m and conductivity $\sigma = ne\mu_e$ Siemens/m
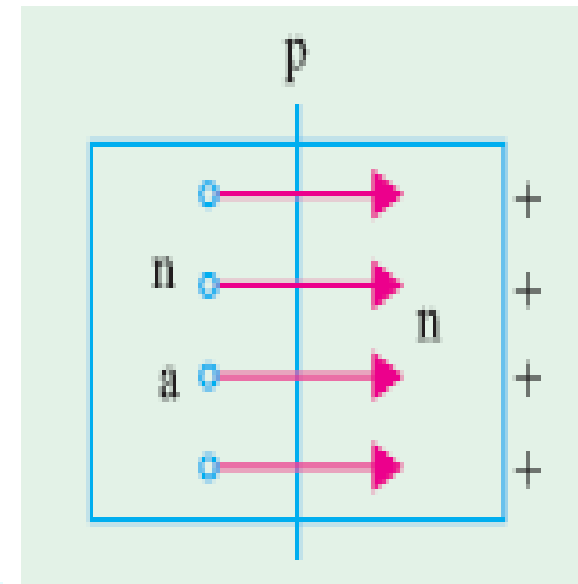
Fig. 51.30

Incidentally, it may be noted that conductivity of a semiconductor differs from that of a metal in one important respect i.e. in a semiconductor, charge carriers are both holes as well as electrons whereas in metals, electrons are the only charge carriers (Art. 51.27).

FORWARD BIASED P-N JUNCTION

Holes   Electrons

P   N

V

(a)

P   N

V

0   V+V$_s$

(b)

P   N

V   I$_0$

(c)
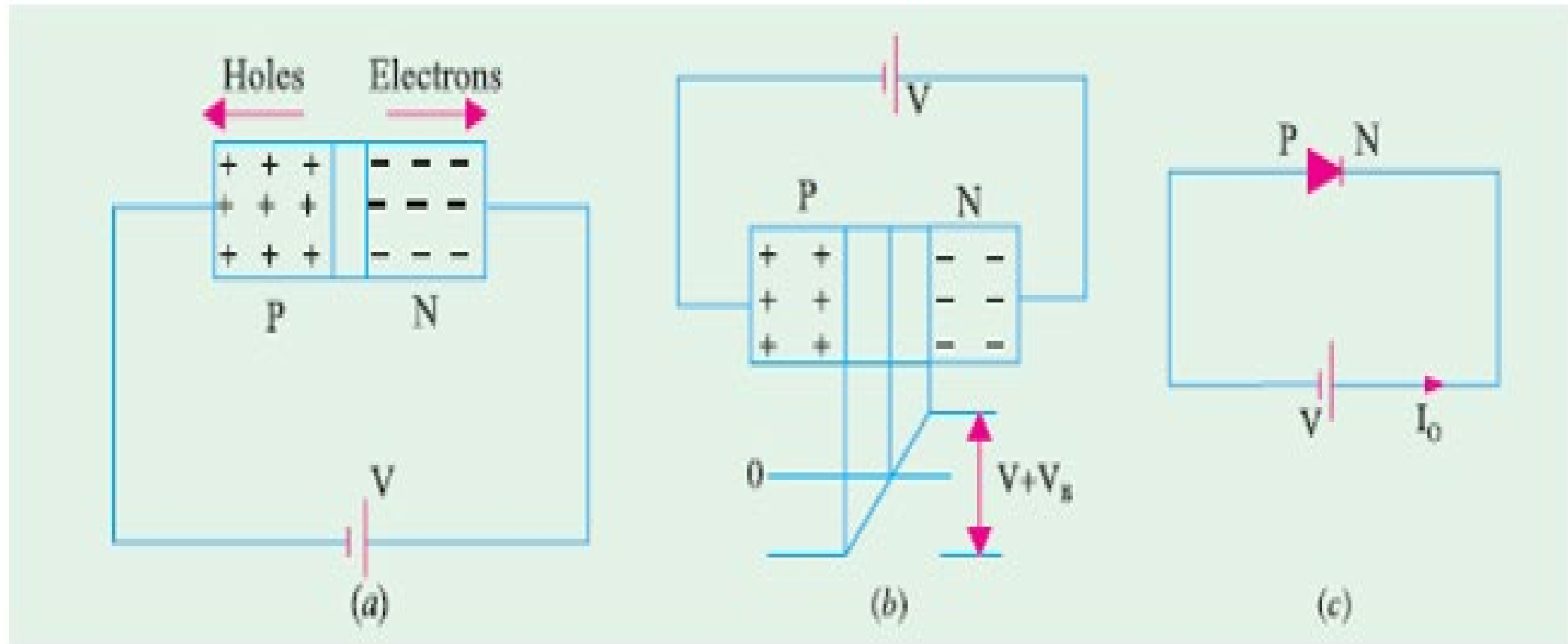
Fig. 51.37

Suppose, a positive battery terminal is connected to $P-$ region of a semiconductor and negative battery terminal to the $N-$ region as shown in Fig. 51.37 (a). In that case the junction is said to be biased in the *forward direction* because it permits easy flow of current across the junction. This current flow may be explained in the following two ways:

As soon as battery connection is made, holes are *repelled* by the positive battery terminal and electrons by the negative battery terminal with the result that both the electrons and the holes are driven *towards* the junction where they recombine. This *en masse* movement of electrons to the left and that of holes to the right of the junction constitutes a large current flow through the semiconductor. Obviously, the junction offers *low resistance* in the forward direction.

As free electrons move to the left, new *free* electrons are injected by the negative battery terminal into the N − region of the semiconductor. Thus, a flow of electrons is set up in the wire connected to the negative battery terminal. As holes are driven towards the junction, more holes are created in the P − region.

These newly – created holes are driven towards the junction to keep up a continuous supply. But the electrons so produced are attracted to the left by the positive battery terminal from where they go to the negative terminal and finally to the N – region of the crystal. Incidentally, it may be noted that though there is movement of both electrons and holes *inside* the crystal, *only free electrons move in the external circuit i.e.* Note. There is also presence of an extremely small amount of current I_o due to minority carriers on either side of the junction. But it is negligible as compared to forward current which is due to majority carriers. This current I_o depends on temperature but is almost independent of applied voltage.
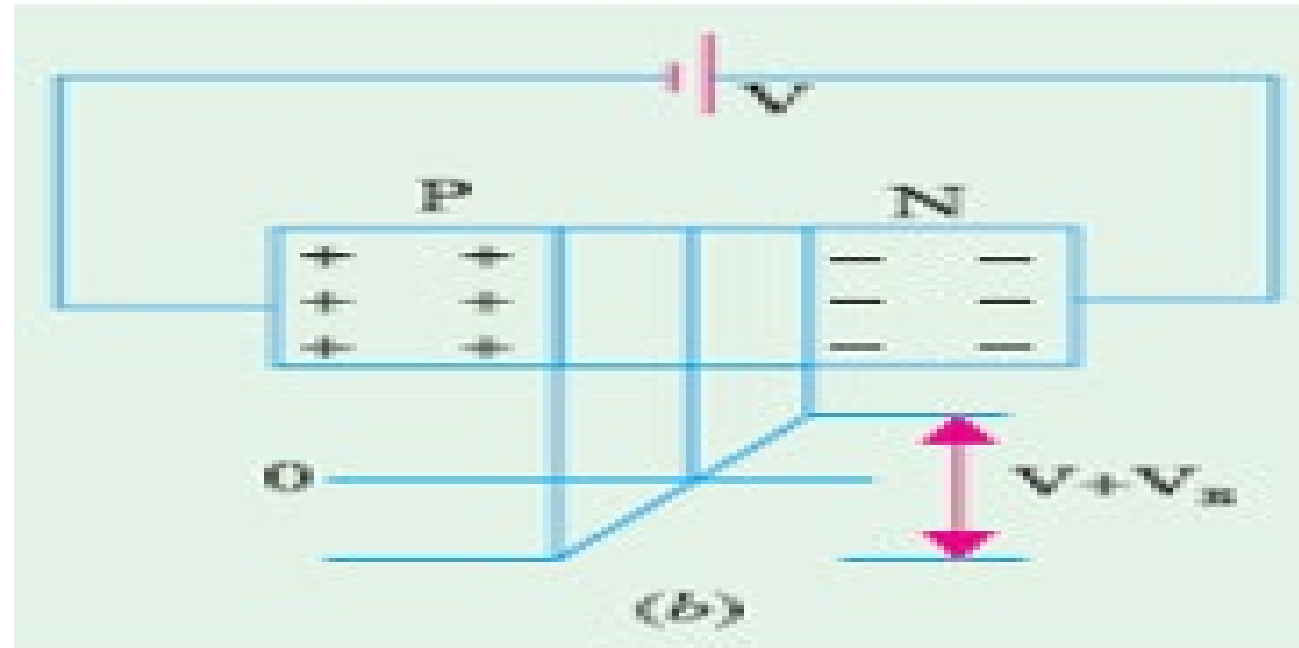
Fig. 51.37

Another way to explain current flow in forward direction is to say that forward bias of V volts lowers the barrier potential to $(V-V_B)$ which now allows more current to flow across the junction [Fig. 51.37 (b)].

$$W < W$$
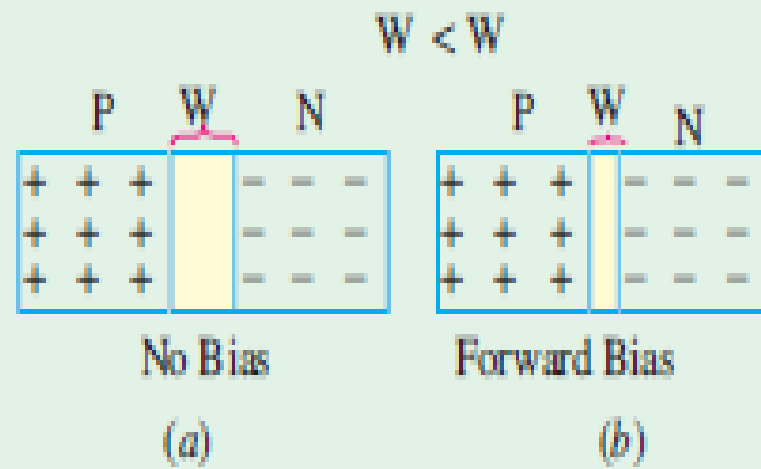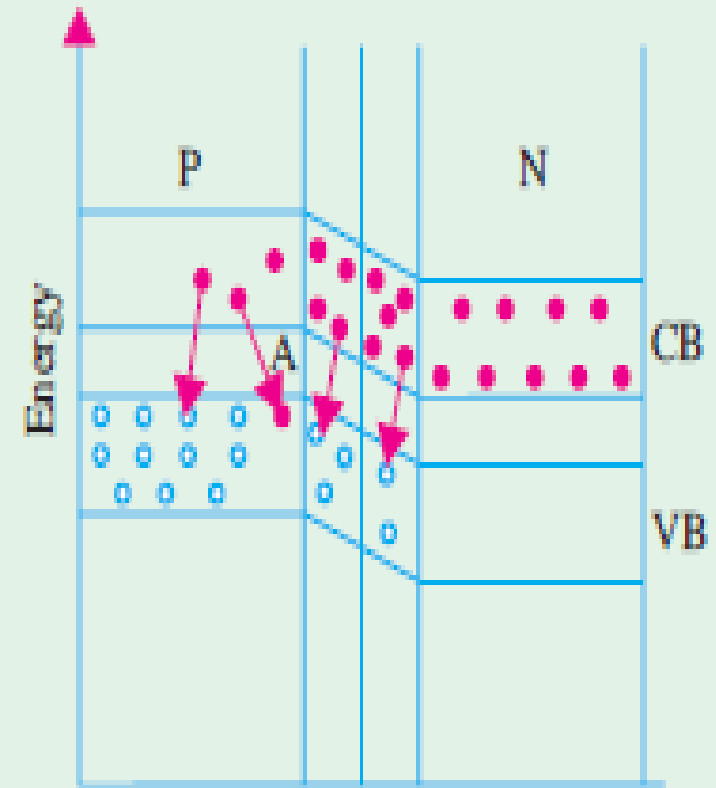
Fig. 51.38

Fig. 51.39

Incidentally, it may be noted that forward bias reduces the thickness of the depletion layer as shown in Fig. 51.38. Energy band diagram for forward bias is shown in Fig. 51.39.

REVERSE BIASED P-N JUNCTION
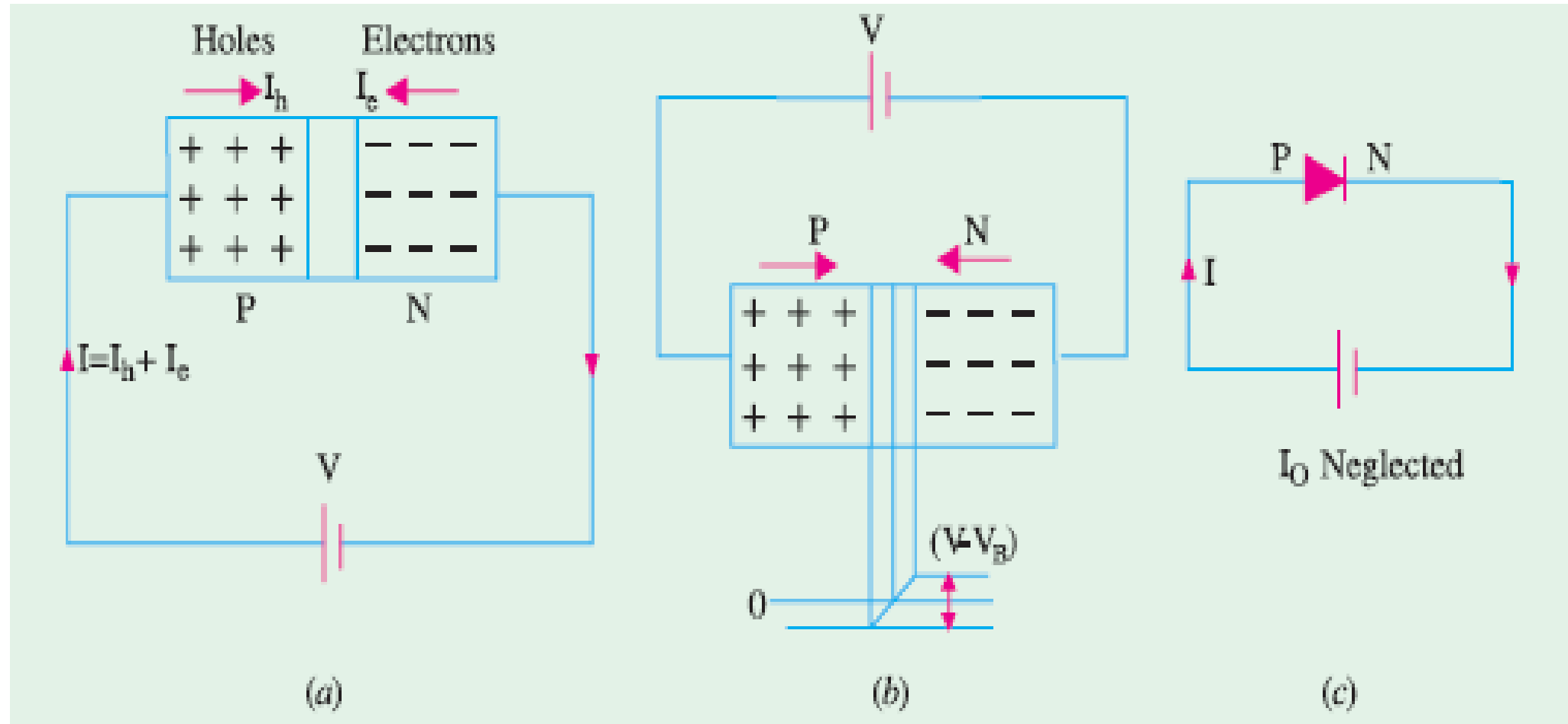
Fig. 51.41

When a battery connections to the semiconductor are made as shown in Fig. 51.41 (a), the junction is said to *reverse-biased*. In this case, holes are attracted by the negative battery terminal and electrons by the positive terminal so that both holes and electrons move *away* from the junction and *away* from each other. Since there is no electron-hole combination, no current flows and the junction offers high resistance. Another way of looking at the process is that in this case, the applied voltage increases the barrier potential to $(V+V_b)$, thereby blocking the flow of majority carriers.
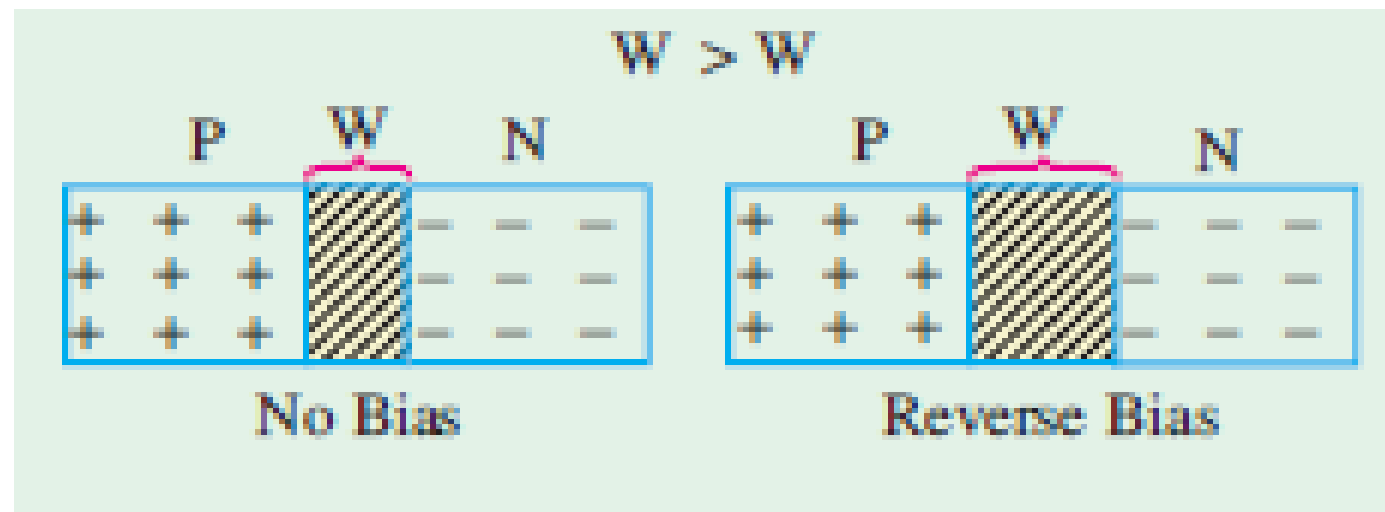
Fig. 51.42

Incidentally, it may be noted that under reverse bias condition, width of depletion layer is increased because of increased barrier potential as shown in Fig. 51.42.

Although, in this case, there is practically no current due to *majority* carriers, yet there is a small amount of current (a few μA only) due to the flow of *minority* carriers across the junction. Due to thermal energy, there are always generated some holes in the $N-$ type region and some electrons in the $P-$ type region of the semiconductor. The battery drives these minority carriers across the junction thereby producing a small current called *reverse current or reverse saturation current* I_o *or* I_s. Since minority carriers are thermally-generated, I_o is extremely *temperature dependent*.

For the same reason, forward current is also temperature dependent but to a much less degree because minority current forms a very small percentage of the majority current. *I_s does not increase with increase in reverse bias.* I_S is found to increase approximately 7 percent per 0^oC rise in temperature both for Ge and Si. Since, $(1.07)^{10} = 2$, it means that reverse current doubles for every 10^oC rise in temperature. It is worth noting that reverse saturation current is also referred to as leakage current of the *P – N* junction diode.

51.43. Forward V/I Characteristic

A typical V/I characteristic for a forward-biased P-N junction is shown in Fig. 51.40. It is seen that forward current rises exponentially with the applied forward voltage. However, at ordinary room temperature, a p.d. of about 0.3 V is required before a reasonable amount of forward current starts flowing in a germanium junction. This voltage is known as *threshold voltage* (V_{th}) or *cut-in voltage* or *knee voltage* V_K. It is practically the same as barrier voltage V_B. Its value for silicon junction is about 0.7 volt. For $V < V_{th}$, current flow is negligible. But as applied voltage increases beyond the threshold value, the forward current increases sharply. If forward voltage is increased beyond a certain safe value, it will produce an extremely large current which may destroy the junction due to overheating.

Fig. 51.40

Ge devices can stand junction temperatures around 100°C whereas *Si* units can function upto 175°C.

Obviously, the *forward-biased junction has a low resistance*. For point B in Fig. 51.40, the forward resistance for Si is

$$R_F = 0.8V/20 \text{ mA} = 40 \ \Omega$$

Similarly, for point A on the *Ge* curve, $R_F = 0.36 \text{ V}/20 \text{ mA} = 18 \ \Omega$

In practice, this static forward resistance is not used. Instead, the **dynamic** resistance or **incremental** resistance or **ac resistance** of the juncion is used. It is given by the reciprocal of the slope of the forward characteristic.

$$r_{ac} = \frac{1}{\Delta I_F / \Delta V_F} = \frac{\Delta V_F}{\Delta I_F}$$

Here, $\Delta V_F = 0.19 \text{ V}$ and $\Delta I_F = 37.6 \text{ mA}$
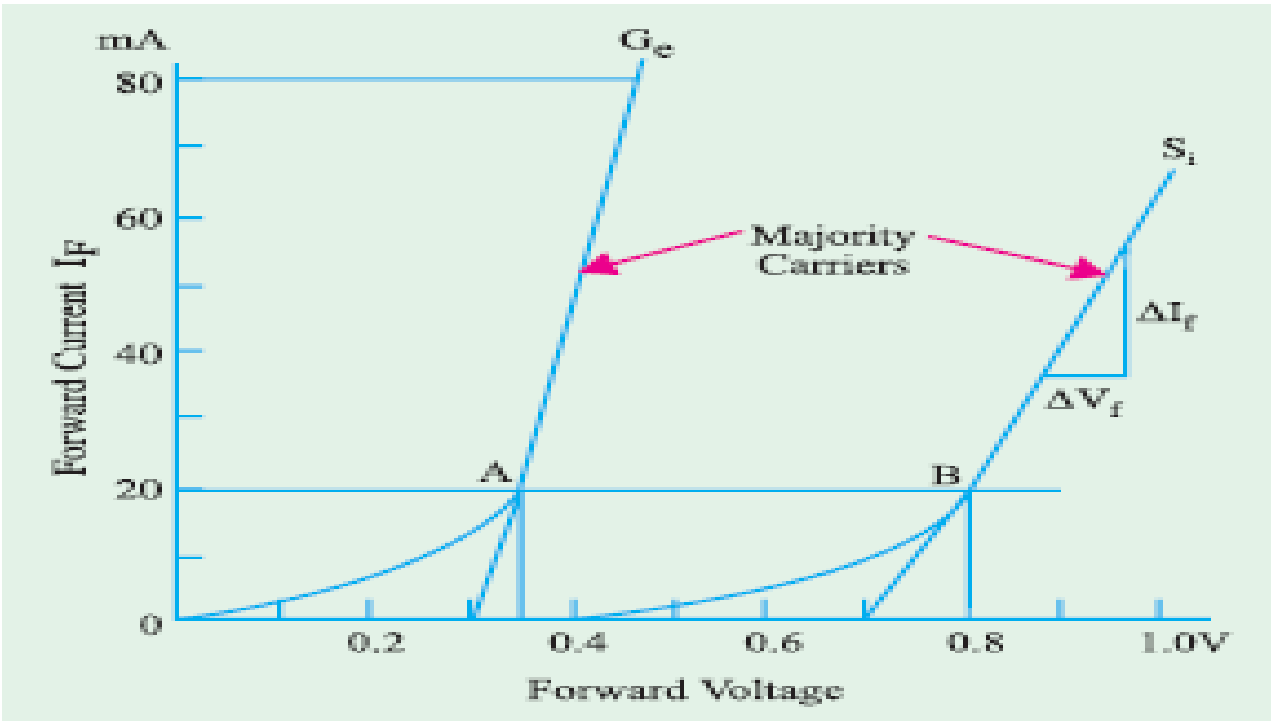
51.45. Reverse V/I Characteristic

As said earlier, the reverse saturation current is also referred to as *leakage current* of the P-N junction. Fig. 51.43 shows V/I characteristics of a reverse-biased P-N junction. It is seen that as reverse voltage is increased from zero, the reverse current quickly rises to its maximum or saturation value. Keeping temperature constant as the reverse voltage is increased, I_o is found to increase only slightly. This slight increase is due to the impurities on the surface of the semiconductor which behaves as a resistor and hence obeys Ohm's law. This gives rise to a very small current called *surface leakage current*. Unlike the main leakage (or saturation) current, this surface leakage current is independent of temperature but depends on the magnitude of the reverse voltage.
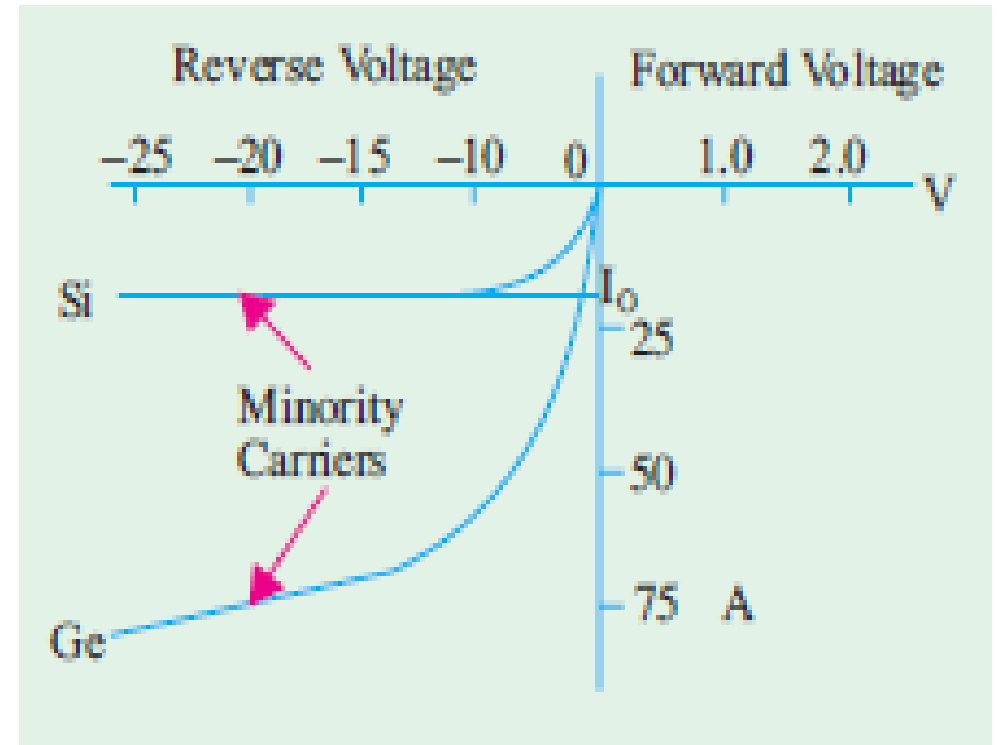
Fig. 51.43

HALL EFFECT

If a specimen (whether of a metal or a semiconductor) carrying a current I is placed in a *transverse* magnetic field of flux density B, an electric field is developed along a direction perpendicular to both B and I. This phenomenon is known as *Hall Effect* and is used for the following purposes:

- To determine whether a semiconductor is of N – type or P – type;

- To find carrier concentration;

- To measure the conductivity of the material;

- To find carrier mobility;

- To detect and measure magnetic fields one million times smaller than that of earth with the help of Hall – effect magnetometers.
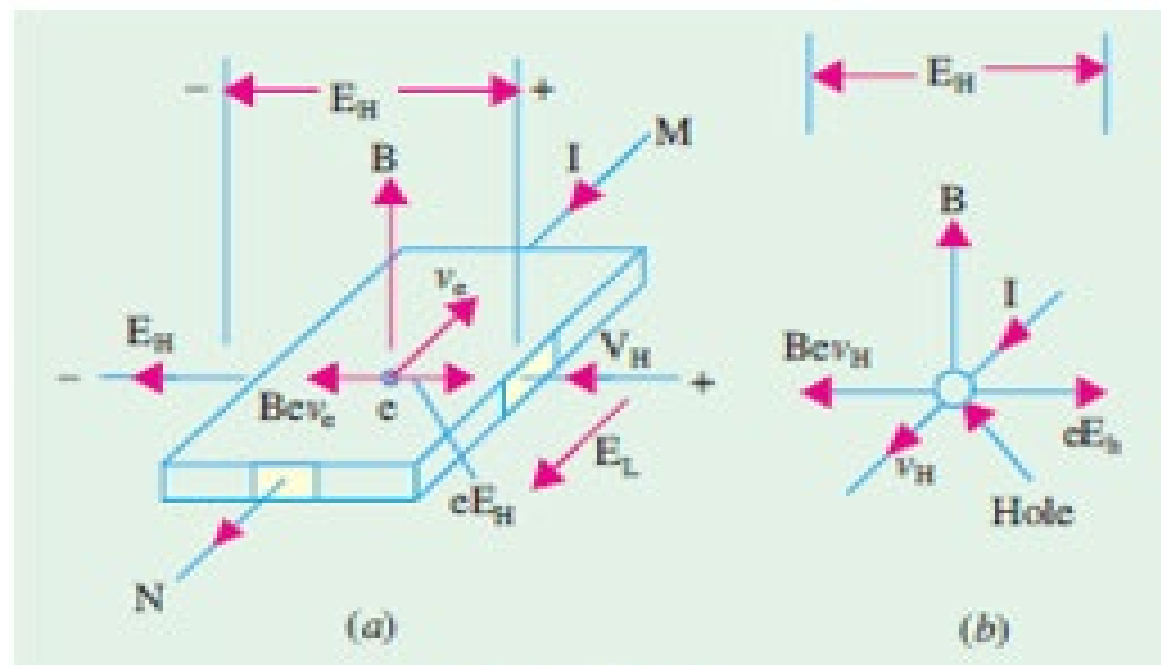
Fig. 51.44

As shown in Fig. 51.44 (a), a current I is flowing through the semiconductor in the direction MN under the influence of an external applied field E_L. Obviously, electrons comprising this current move along NM with a velocity of v_e. Consider one such electron shown in the figure. The direction of the force exerted on it by the magnetic field B can be found by using Fleming's left-hand rule and is as shown. The magnitude of force is Bev_e. Under the influence of this force, electrons will tend to crowd towards the left side of semiconductor.

This collection of electrons to one side gives rise to electric potential difference V_H (called Hall voltage) and hence to an electric field E_H.

This field begins to prevent additional electrons from arriving there. Force exerted by this field on the electrons is eE_H. Equilibrium is established when the two oppositely – directed magnetic and electric forces acting on the electron become equal in magnitude. While using the rule, the middle finger must point in the direction of conventional current *i.e.* along *MN* and not along the direction of electronic current *i.e.* along *NM*.

$$BeV_e = eE_H$$
$$\text{or } B_e = E_H \qquad\qquad (i)$$
$$\text{Also } V_e = \mu_e E_L$$
$$therefore, B\mu_e E_L = E_H \text{ or } \mu_e B = \frac{E_H}{E_L}$$

Hence, knowing E_L and B and measuring E_H, we can find electron mobility.

Now, current density

$$J = \frac{I}{A} = \frac{nev_e A}{A} = nev_e$$

Substituting the value of v_e from (i) above, we get

$$J = \frac{neE_H}{B} \text{ or } n = \frac{JB}{eE_H}$$

This equation may be used to find electron density.

Now, Hall coefficient,

$$R_H = \frac{1}{charge\ density} = \frac{1}{ne}\ or\ R_H = \frac{E_H}{JB}$$

From our previous slide,

$$\mu_e = \frac{1}{ne\rho} = \frac{R_H}{\rho}$$

Where ρ is the resistivity of the semiconductor.

It may be noted that the above treatment is equally applicable to $p-$ type semiconductors where current flow is made up of hole movement. If holes were to move in the same direction as that of the electrons in Fig. 51.44 (a), then polarity of Hall voltage E_H would be reversed as shown in Fig. 51.44 (b). In fact, it was the observation of the polarity of Hall voltage associated with $p-$ type material which led to the concept of a positive hole.

The Hall – coefficient of a specimen of doped semiconductor is $3.66 \times 10^{-4} m^3 C^{-1}$ and the resistivity of the specimen is $8.93 \times 10^{-3} \, \Omega - m$. Determine the carrier mobility in $m^2 V^{-1} S^{-1}$.

Using the equation

$$\mu_e = \frac{R_H}{\rho}$$

$$\mu_e = \frac{3.66 \times 10^{-4}}{8.93 \times 10^{-3}} = 0.041 m^2 V^{-1} S^{-1}$$

Conduction electron

Conduction band

Donor impurity
(missing an electron)

Energy

Valence band
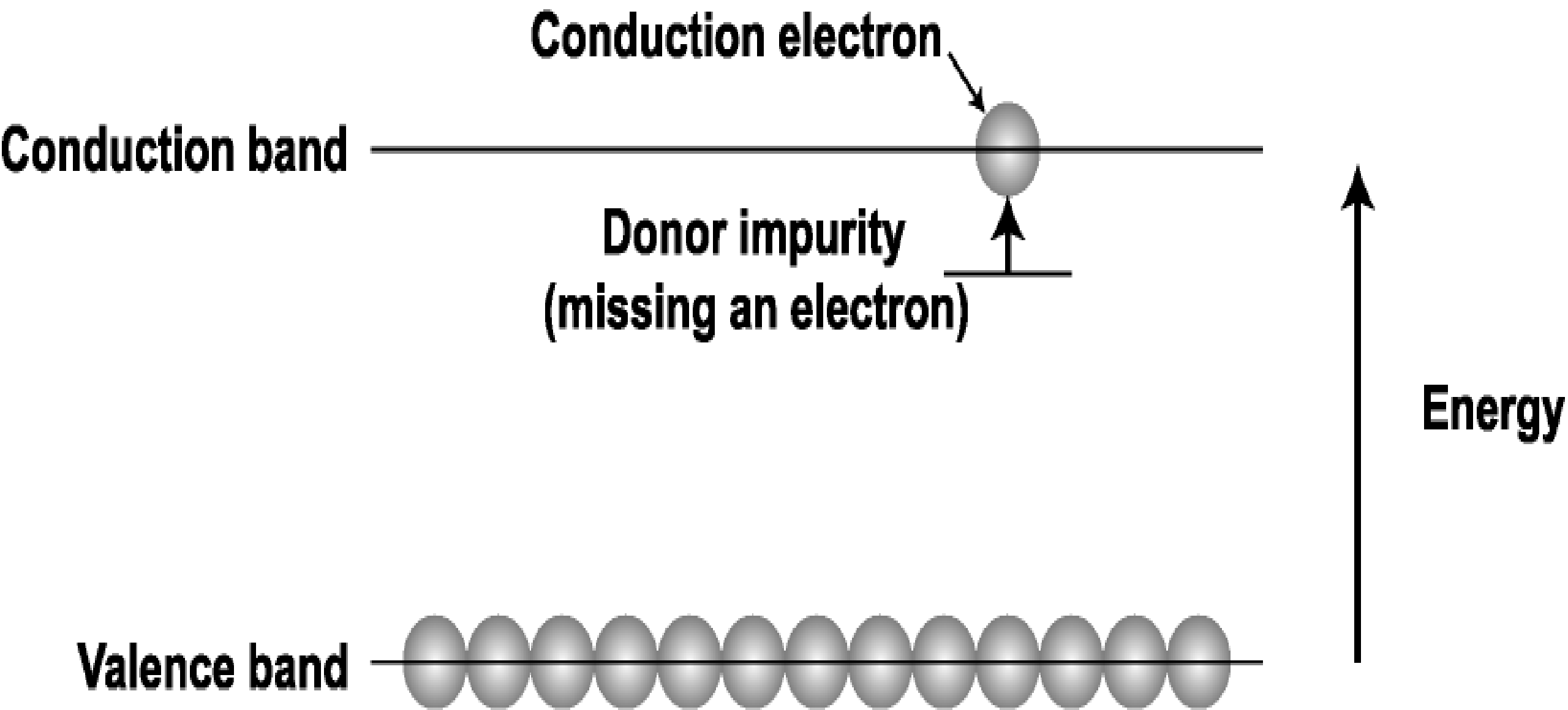
DOPED SEMICONDUCTORS

The electrical conduction in semiconductors can be improved by doping the semiconductor with a small amount of some impurity. In a so-called n-type semiconductor, atoms are added which contribute a few extra electrons to the conduction band.
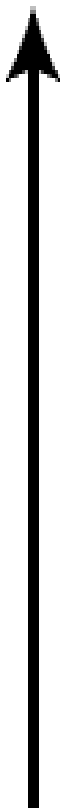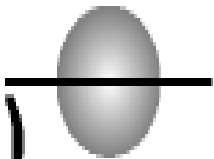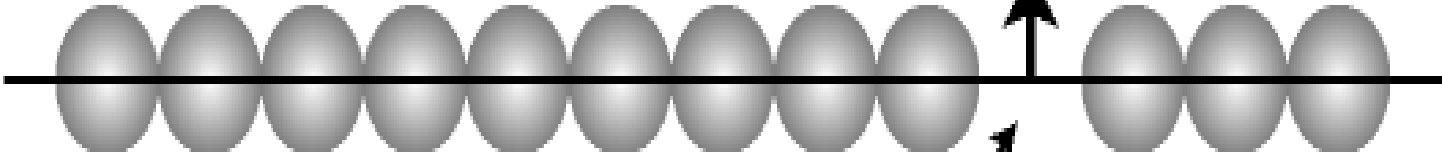
In a so-called p-type semiconductor, atoms are added that "suck up" a few electrons from the valence band, leaving behind holes in the valence band.

Thus, the dopant (the impurity added to the semiconductor) in a p-type semiconductor adds positive charge carriers (holes in the valence band), while the dopant in an n-type semiconductor adds negative charge carriers (electrons in the conduction band).

MAJORITY AND MINORITY CARRIERS

Since most of the current carriers are electrons, silicon (or germanium) doped with pentavalent atoms is an *n*-type semiconductor (the *n* stands for the negative charge on an electron). The electrons are called the **majority carriers** in *n*-type material. Although the majority of current carriers in *n*-type material are electrons, there are also a few holes that are created when electron-hole pairs are thermally generated. These holes are *not* produced by the addition of the pentavalent impurity atoms. Holes in an *n*-type material are called **minority carriers**.

P-TYPE SEMICONDUCTORS

When a small amount of trivalent impurity is added to a pure semiconductor, it is called **p-type semiconductor. To** increase the number of holes in intrinsic silicon, **trivalent** impurity atoms are added. These are atoms with three valence electrons such as boron (B), indium (In), and gallium (Ga).

As opposed to n-type semiconductors, p-type semiconductors have a larger hole concentration than electron concentration.

The phrase 'p-type' refers to the positive charge of the hole. In p-type semiconductors, holes are the majority carriers and electrons are the minority carriers. P-type semiconductors are created by doping an intrinsic semiconductor with acceptor impurities.

P-type semiconductors have Fermi energy levels below the intrinsic Fermi energy level. The Fermi energy level lies closer to the valence band than the conduction band in a p-type semiconductor. Gallium has 3 valence electrons, however, there are 4 covalent bonds to fill. The 4th bond therefore remains vacant producing a hole. The holes are said to be the majority carriers and the electrons are said to be the minority carriers.
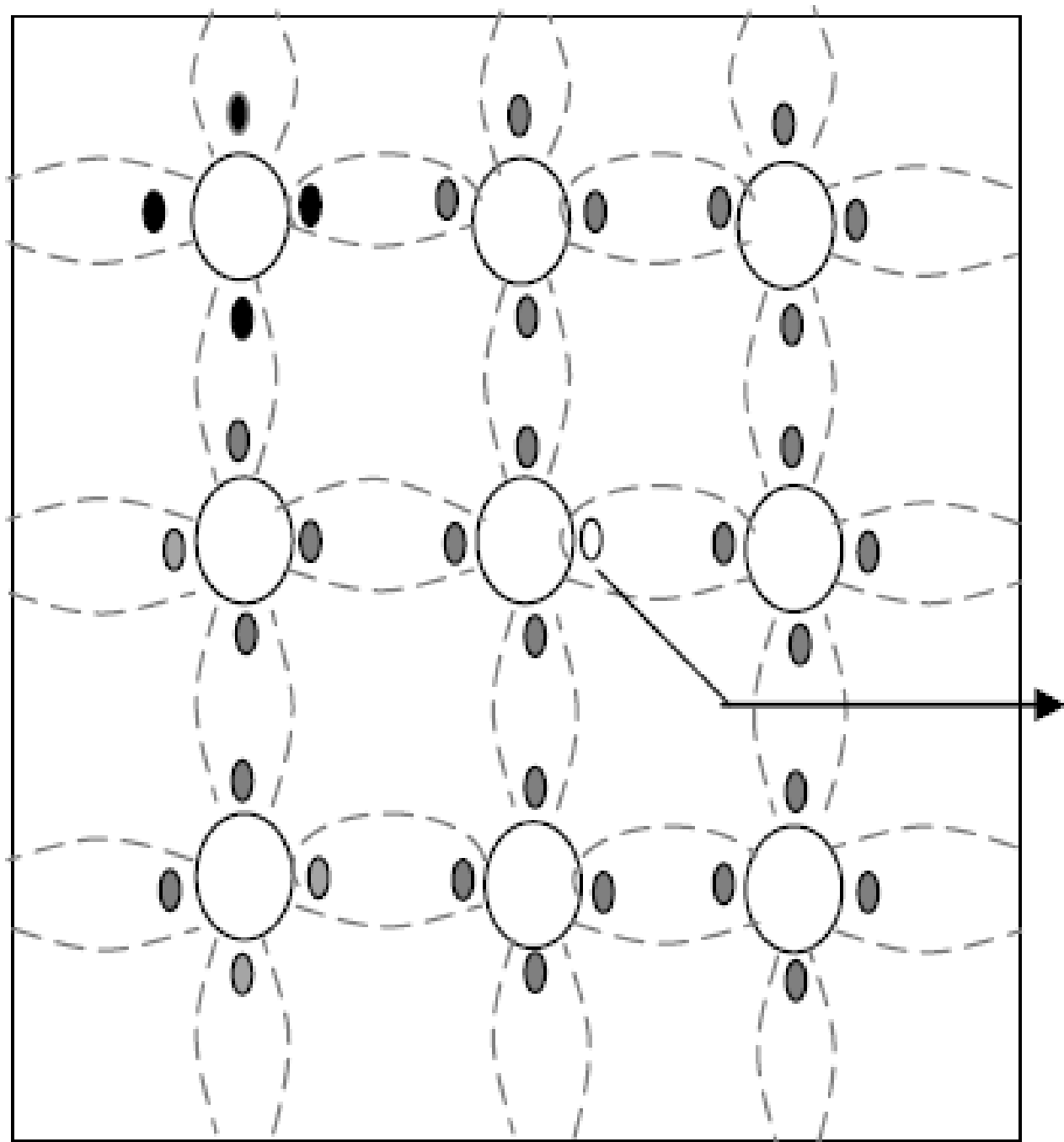
Fig. 1.5a crystal structure of P type sc

If the added impurity is a trivalent atom then the resultant semiconductor is called P-type semiconductor. Examples of trivalent impurities are Boron, Gallium , indium etc. The crystal structure of p type sc is shown in the fig1.5a. The three valance electrons of the impurity (boron) forms three covalent bonds with the neighbouring atoms and a vacancy exists in the fourth bond giving rise to the holes. The hole is ready to accept an electron from the neighbouring atoms. Each trivalent atom contributes to one hole generation and thus introduces a large no. of holes in the valance band. At the same time the no. electrons are decreased compared to those available in intrinsic sc because of increased recombination due to creation of additional holes.
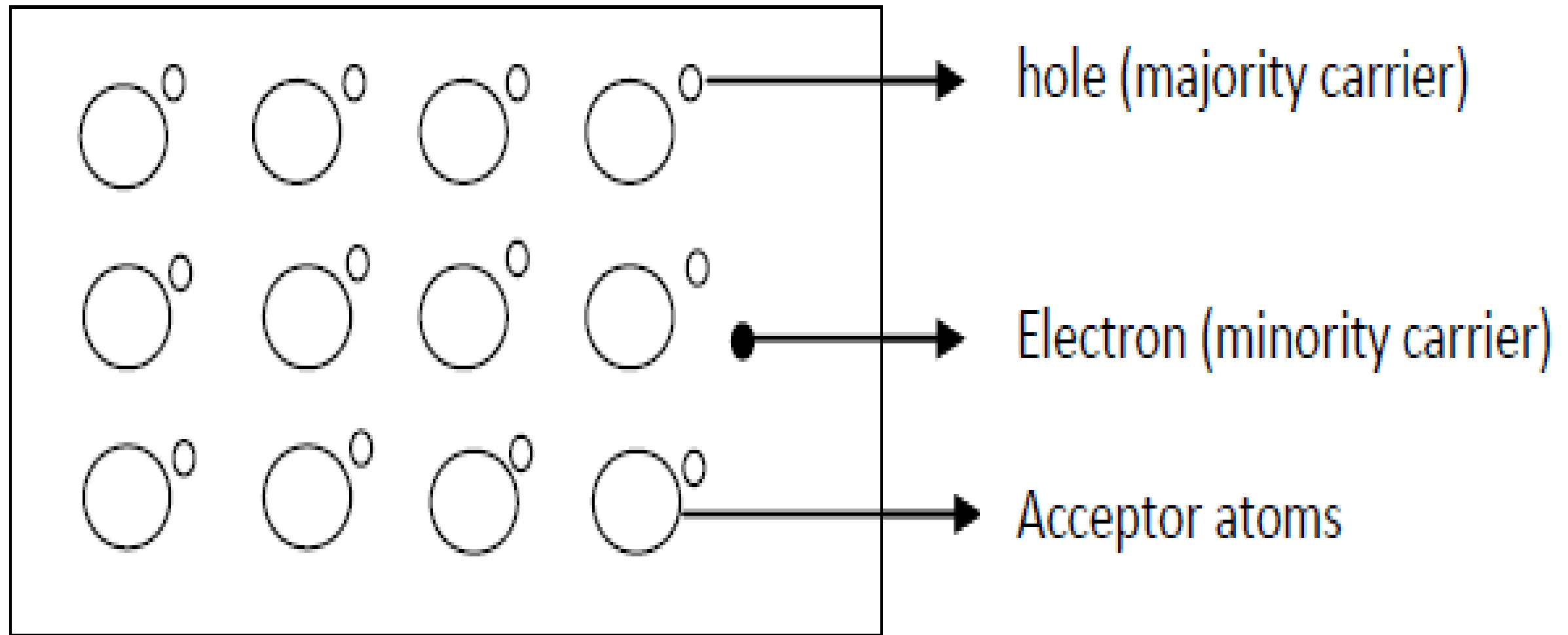
Fig. 1.5b crystal structure of P type sc

Thus in P type sc, holes are majority carriers and electrons are minority carriers. Since each trivalent impurity atoms are capable accepting an electron, these are called as acceptor atoms. The following fig 1.5b shows the pictorial representation of P type sc

The conductivity of N type sc is greater than that of P type sc as the mobility of electron is greater than that of hole. For the same level of doping in N type sc and P type sc, the conductivity of an N type sc is around twice that of a P type sc

► FIGURE 1–18

Trivalent impurity atom in a silicon crystal structure. A boron (B) impurity atom is shown in the center.

Hole from B atom

Si

Si

B

Si

Si
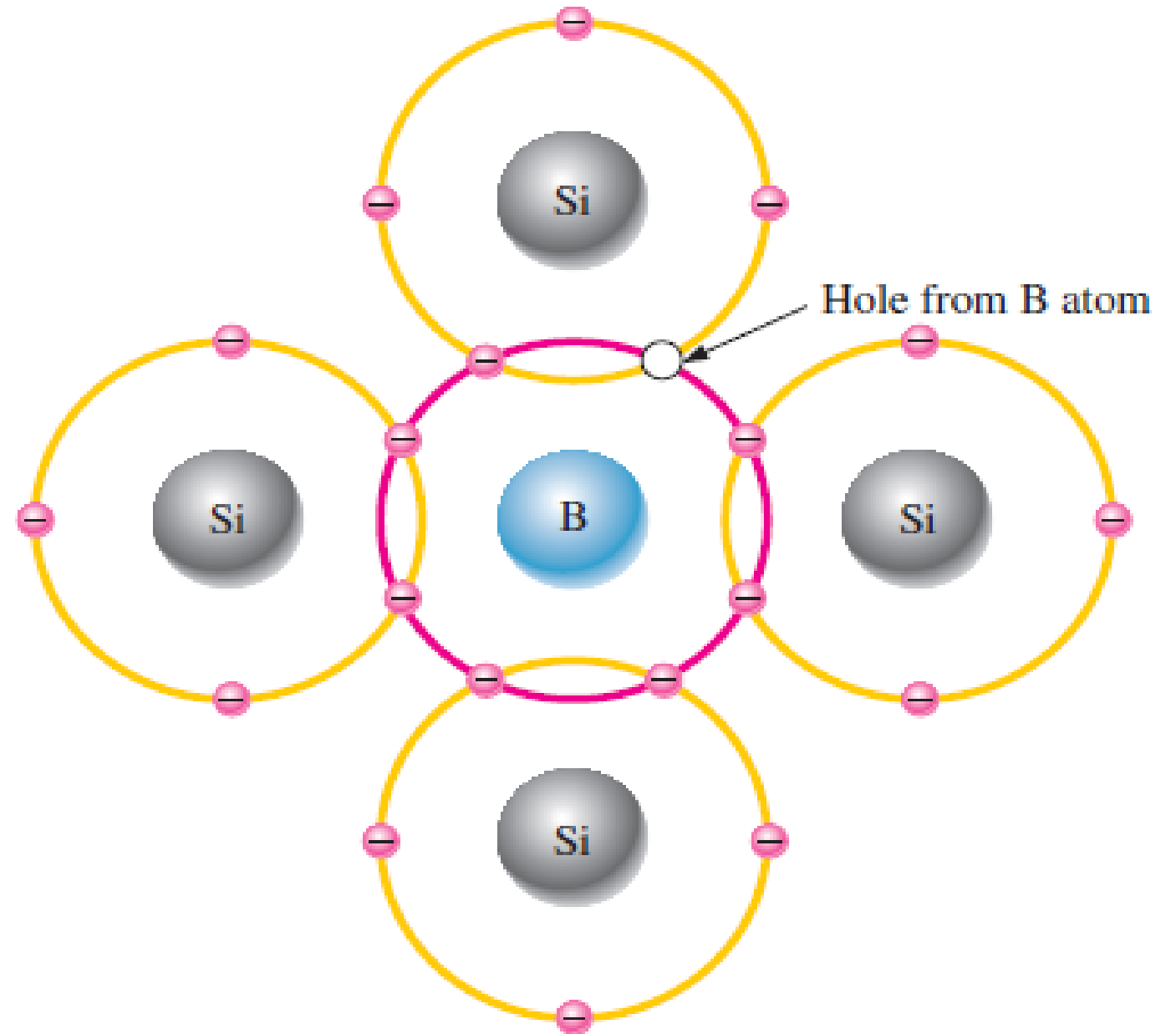
As illustrated in Figure 1–18, each trivalent atom (boron, in this case) forms covalent bonds with four adjacent silicon atoms. All three of the boron atom's valence electrons are used in the covalent bonds; and, since four electrons are required, a hole results when each trivalent atom is added. Because the trivalent atom can take an electron, it is often referred to as an *acceptor atom.* The number of holes can be carefully controlled by the number of trivalent impurity atoms added to the silicon. A hole created by this doping process is *not* accompanied by a conduction (free) electron.

TRIVALENT IMPURITY ATOM
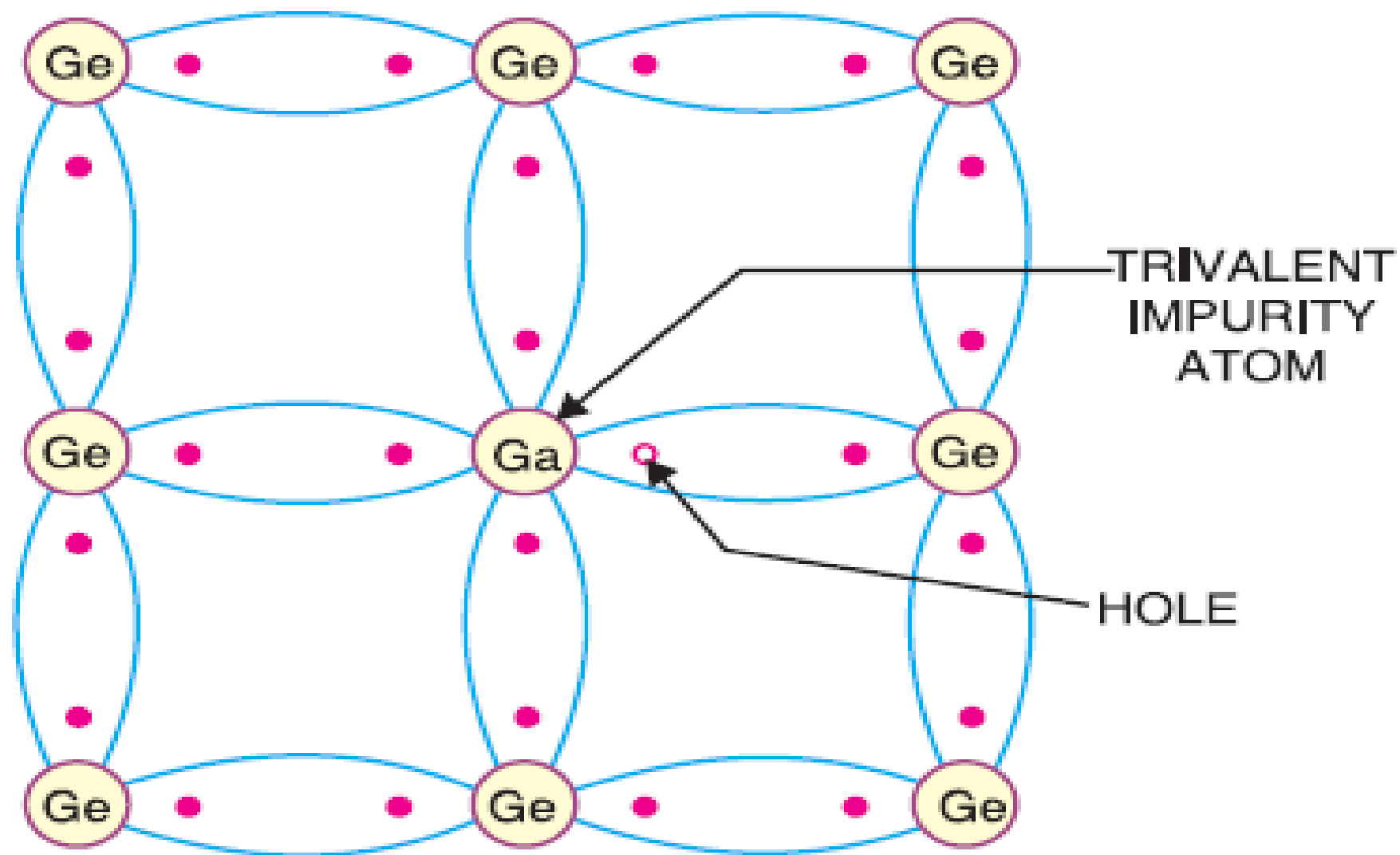
HOLE

Fig. 5.14

The addition of trivalent impurity provides a large number of holes in the semiconductor. Typical examples of trivalent impurities are *gallium* and *indium*. Such impurities which produce *p*-type semiconductor are known as *acceptor impurities* because the holes created can accept the electrons. To explain the formation of *p*-type semiconductor, consider a pure germanium crystal. When a small amount of trivalent impurity like gallium is added to germanium crystal, there exists a large number of holes in the crystal. The reason is simple. Gallium is trivalent *i.e.* its atom has three valence electrons. Each atom of gallium fits into the germanium crystal but now only three co-valent bonds can be formed. It is because three valence electrons of gallium atom can form only three single co-valent bonds with three germanium atoms as shown in Fig. 5.14.

In the fourth co-valent bond, only germanium atom contributes one valence electron while gallium has no valence electron to contribute as all its three valence electrons are already engaged in the co-valent bonds with neighbouring germanium atoms. In other words, fourth bond is incomplete; being short of one electron. This missing electron is called a *hole*. Therefore, for each gallium atom added, one hole is created. A small amount of gallium provides millions of holes.
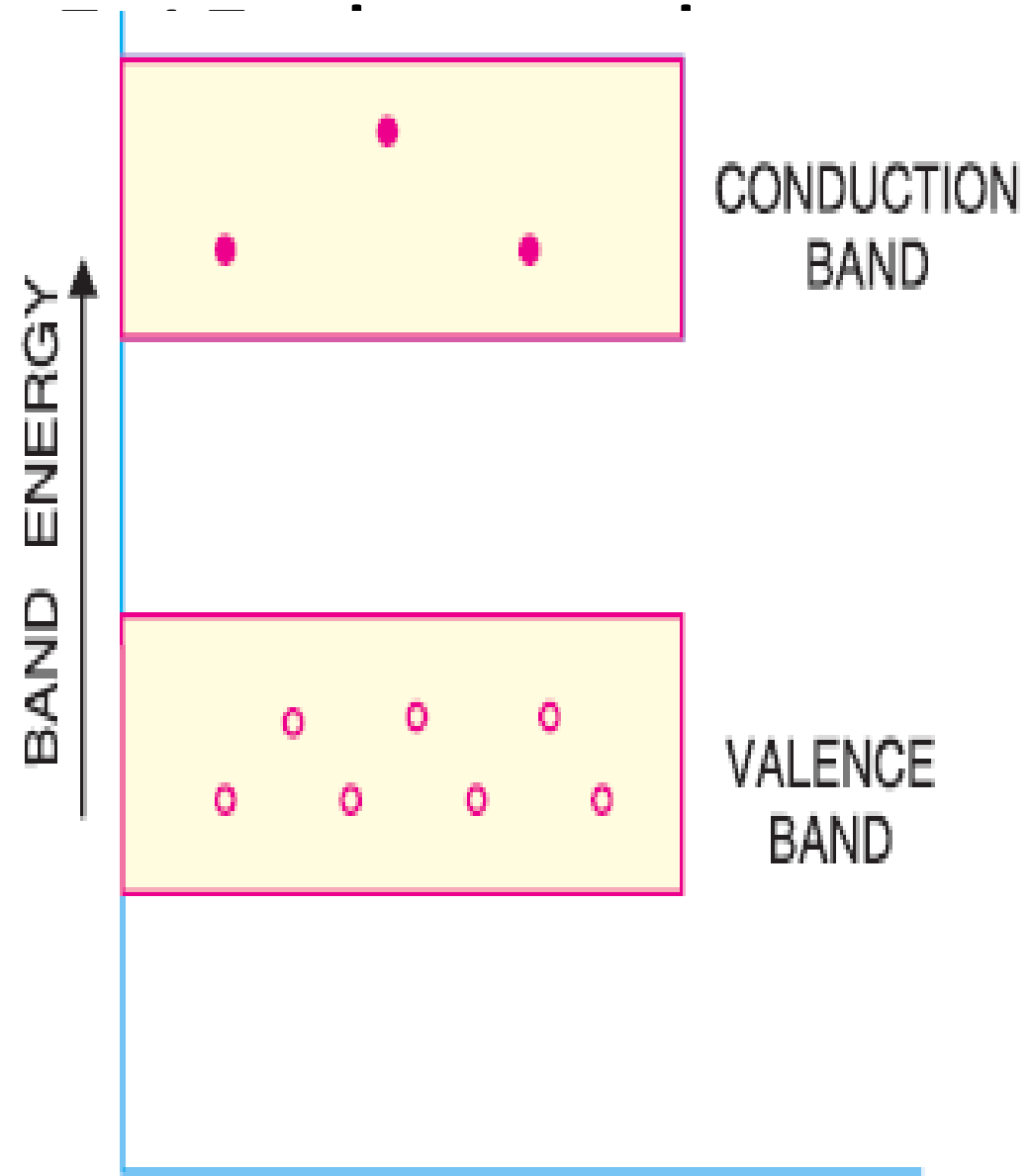
Fig. 5.15
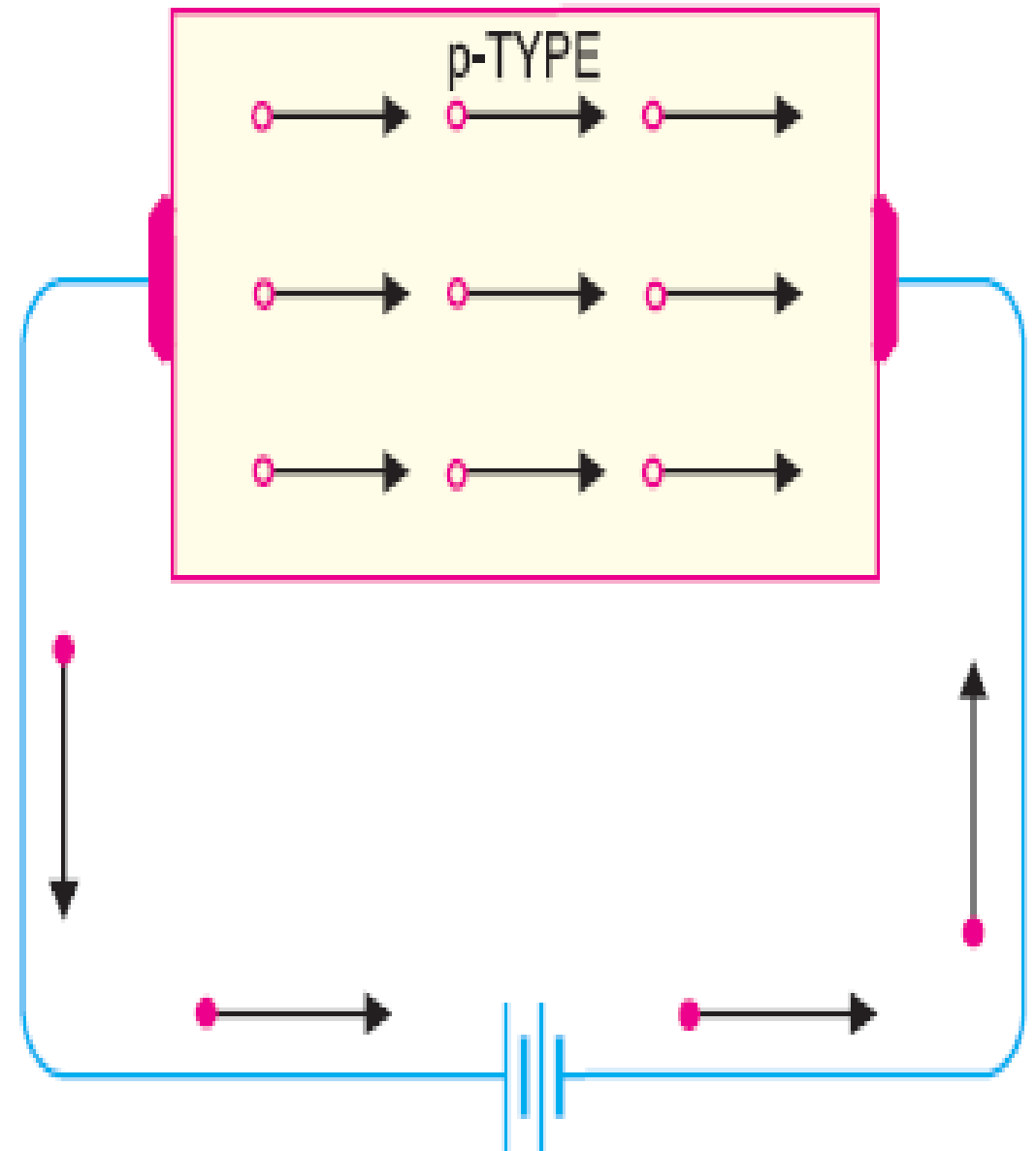
Fig. 5.16

P–TYPE CONDUCTIVITY

The current conduction in *p*-type semiconductor is predominantly by holes *i.e.* positive charges and is called *p-type* or *hole-type conductivity*. To understand *p*-type conductivity, refer to Fig. 5.16. When *p.d.* is applied to the *p*-type semiconductor, the holes (donated by the impurity) are shifted from one co-valent bond to another. As the holes are positively charged, therefore, they are directed towards the negative terminal, constituting what is known as hole current. It may be noted that in *p*-type conductivity, the valence electrons move from one co-valent bond to another unlike the *n*-type where current conduction is by free electrons.

CHARGE ON N-TYPE AND P-TYPE SEMICONDUCTORS

As discussed before, in *n*-type semiconductor, current conduction is due to excess of electrons whereas in a *p*-type semiconductor, conduction is by holes. The reader may think that *n*-type material has a net negative charge and *p*-type a net positive charge. But this conclusion is wrong. It is true that *n*-type semiconductor has excess of electrons but these extra electrons were supplied by the atoms of donor impurity and each atom of donor impurity is electrically neutral.

When the impurity atom is added, the term "excess electrons" refers to an excess with regard to the number of electrons needed to fill the co-valent bonds in the semiconductor crystal. The extra electrons are free electrons and increase the conductivity of the semiconductor. The situation with regard to *p*-type semiconductor is also similar. *It follows, therefore, that n-type as well as p-type semiconductor is electrically neutral.*

MAJORITY AND MINORITY CARRIERS

Since most of the current carriers are holes, silicon (or germanium) doped with trivalent atoms is called a *p*-type semiconductor. The holes are the majority carriers in *p*-type material. Although the majority of current carriers in *p*-type material are holes, there are also a few conduction-band electrons that are created when electron-hole pairs are thermally generated. These conduction-band electrons are *not* produced by the addition of the trivalent impurity atoms. Conduction-band electrons in *p*-type material are the minority carriers.

CONDUCTIVITY OF SEMICONDUCTOR

In a pure sc, the no. of holes is equal to the no. of electrons. Thermal agitation continue to produce new electron- hole pairs and the electron hole pairs disappear because of recombination. With each electron hole pair created, two charge carrying particles are formed. One is negative which is a free electron with mobility μ_n. The other is a positive i.e., hole with mobility μ_p. The electrons and hole move in opposite direction in a an electric field E, but since they are of opposite sign, the current due to each is in the same direction. Hence the total current density J within the intrinsic sc is given by

$$J = J_n + J_p$$

$$= q\, n\, \mu_n E + q\, p\, \mu_p E$$

$$= (n\, \mu_n + p\, \mu_p) qE$$

$$= \sigma E$$

Where n=no. of electrons / unit volume i.e., concentration of free electrons

P= no. of holes / unit volume i.e., concentration of holes

E=applied electric field strength, V/m

q= charge of electron or hole I n Coulombs

Hence, σ is the conductivity of sc which is equal to $(n\,\mu_n + p\,\mu_p)q$. he resistivity of sc is reciprocal of conductivity.

$P = 1/\sigma$

It is evident from the above equation that current density with in a sc is directly proportional to applied electric field E.

For pure sc, $n=p= n_i$ where n_i = intrinsic concentration. The value of n_i is given by

$n_i^2 = AT^3 \exp(-E_{GO}/KT)$

therefore, $\quad J = n_i\,(\mu_n + \mu_p)\,q\,E$

Hence conductivity in intrinsic sc is $\quad \sigma i = n_i\,(\mu_n + \mu_p)\,q$

Intrinsic conductivity increases at the rate of 5% per $^\circ$C for Ge and 7% per $^\circ$C for Si.

Conductivity in extrinsic sc (N Type and P Type):

The conductivity of intrinsic sc is given by $\sigma i = n_i (\mu_n + \mu_p) q = (n \mu_n + p \mu_p)q$

For N type , n>>p

Therefore $\sigma = q \, n \, \mu_n$

For P type ,p>>n

Therefore $\sigma = q \, p \, \mu_p$

1.0.4 CHARGE DENSITIES IN P TYPE AND N TYPE SEMICONDUCTOR:

Mass Action Law:

Under thermal equilibrium for any semiconductor, the product of the no. of holes and the concentration of electrons is constant and is independent of amount of donor and acceptor impurity doping.

$$n.p = n_i^2$$

where n= eleetron concentration

p = hole concentration

n_i^2 = intrinsic concentration

Hence in N type sc , as the no. of electrons increase the no. of holes decreases. Similarly in P type as the no. of holes increases the no. of electrons decreases. Thus the product is constant and is equal to n_i^2 in case of intrinsic as well as extrinsic sc.

The law of mass action has given the relationship between free electrons concentration and hole concentration. These concentrations are further related by the law of electrical neutrality as explained below.

LAW OF ELECTRICAL NEUTRALITY

Sc materials are electrically neutral. According to the law of electrical neutrality, in an electrically neutral material, the magnitude of positive charge concentration is equal to that of negative charge concentration. Let us consider a sc that has N_D donor atoms per cubic centimeter and N_A acceptor atoms per cubic centimeter i.e., the concentration of donor and acceptor atoms are N_D and N_A respectively. Therefore N_D positively charged ions per cubic centimeter are contributed by donor atoms and N_A negatively charged ions per cubic centimeter are contributed by the acceptor atoms. Let n, p is concentration of free electrons and holes respectively. Then according to the law of neutrality

$$N_D + p = N_A + n \qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\text{eq } 1.1$$

For N type sc, $N_A = 0$ and $n \gg p$. Therefore $N_D \approx n$ $\dots\dots\dots\dots\dots\dots\dots\dots\dots$eq 1.2

Hence for N type sc the free electron concentration is approximately equal to the concentration of donor atoms. In later applications since some confusion may arise as to which type of sc is under consideration a the given moment, the subscript n or p is added for Ntype or P type respectively. Hence eq1.2 becomes $N_D \approx n_n$

Therefore current density in N type sc is $J = N_D \, \mu_n \, q \, E$

And conductivity $\sigma = N_D \, \mu_n \, q$

For P type sc, $N_D = 0$ and $p \gg n$. Therefore $N_A \approx p$

$$\text{Or } N_A \approx p_p$$

Hence for P type sc the hole concentration is approximately equal to the concentration of acceptor atoms.

Therefore current density in *N* type *SC* is $J = N_A \mu_q q E$

And conductivity $\sigma = N_A \mu_p q$

Mass action law for N type, $n_n p_n = n_i^2$

$$p_n = \frac{n_i^2}{N_D}$$

Since $(n_n \approx N_D)$

Mass action law for P type, $n_p p_p = n_i^2$

$$n_n = \frac{n_i^2}{N_A}$$

Since $(p_p \approx N_A)$

PURE Ge

CONDUCTION BAND

VALENCE BAND

3rd BAND

2nd BAND

(i)

(ii)
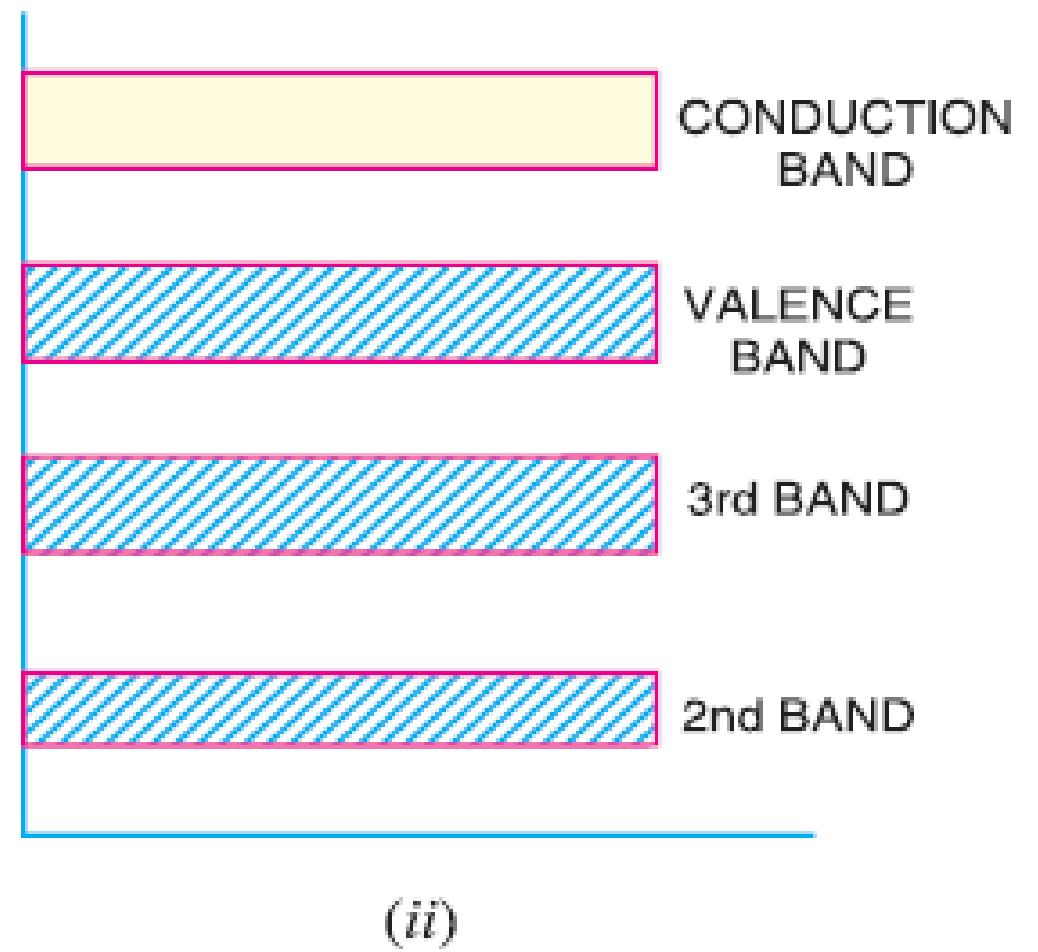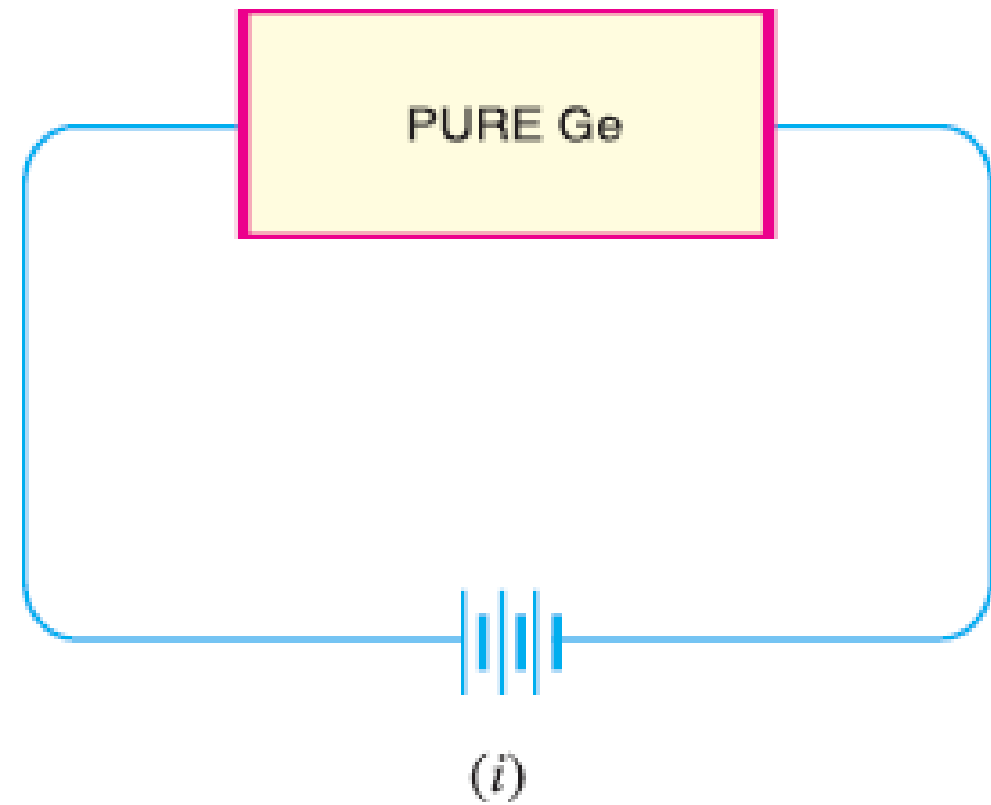
Fig. 5.6

EFFECTS OF TEMPERATURE ON CONDUCTIVITY OF SEMICONDUCTOR

The electrical conductivity of a semiconductor changes appreciably with temperature variations. This is a very important point to keep in mind.

(*i*) **At absolute zero.** At absolute zero temperature, all the electrons are tightly held by the semiconductor atoms. The inner orbit electrons are bound whereas the valence electrons are engaged in co-valent bonding. At this temperature, the co-valent bonds are very strong and there are no free electrons. Therefore, the semiconductor crystal behaves as a perfect insulator [See Fig. 5.6 (*i*)].

In terms of energy band description, the valence band is filled and there is a large energy gap between valence band and conduction band. Therefore, no valence electron can reach the conduction band to become free electron. It is due to the non-availability of free electrons that a semiconductor behaves as an insulator.

PURE Ge

CONDUCTION BAND

VALENCE BAND

3rd BAND

2nd BAND
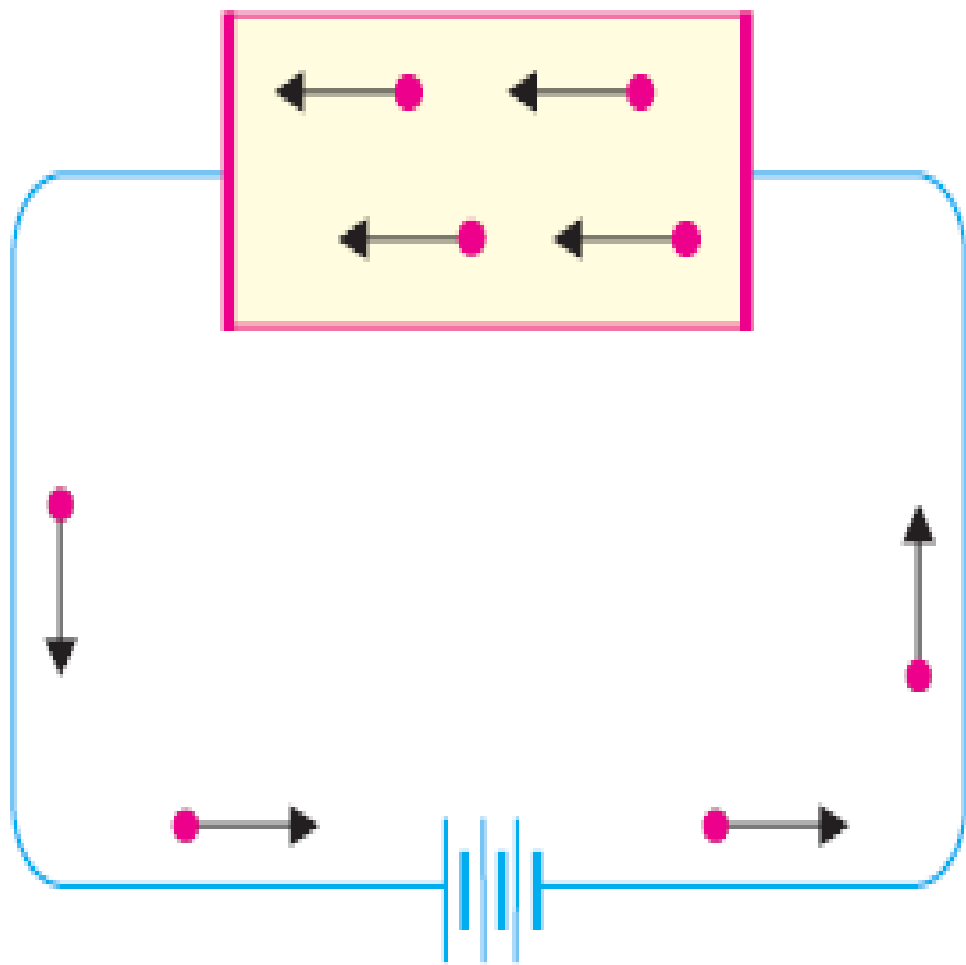
(i)
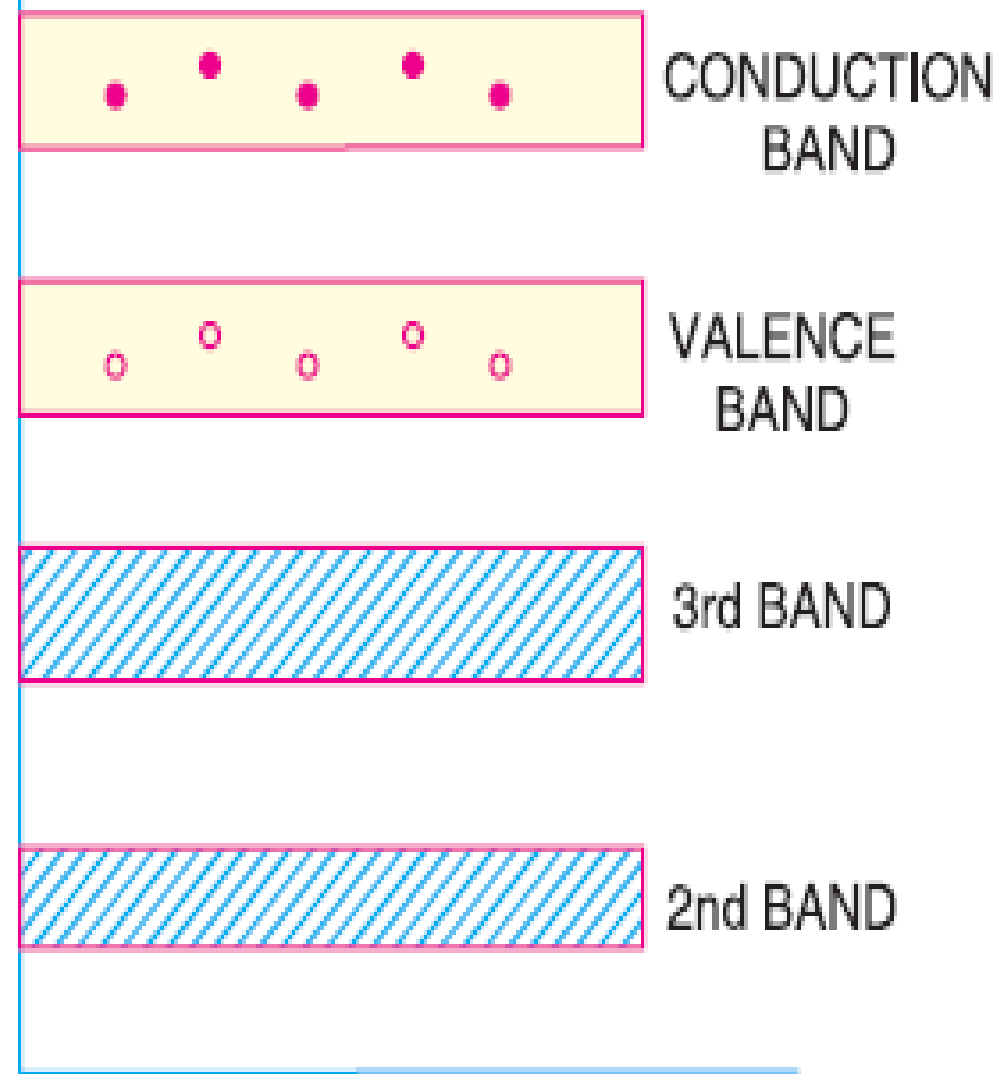
(ii)

Fig. 5.7

(**ii**) **Above absolute zero.** When the temperature is raised, some of the covalent bonds in the semiconductor break due to the thermal energy supplied. The breaking of bonds sets those electrons *free* which are engaged in the formation of these bonds. The result is that a few free electrons exist in the semiconductor. These free electrons can constitute a tiny electric current if potential difference is applied across the semiconductor crystal [See Fig. 5.7 (*i*)]. *This shows that the resistance of a semiconductor decreases with the rise in temperature i.e.* it has negative temperature coefficient of resistance. It may be added that at room temperature, current through a semiconductor is too small to be of any practical value.

PURE Ge

CONDUCTION
BAND

VALENCE
BAND

3rd BAND

2nd BAND
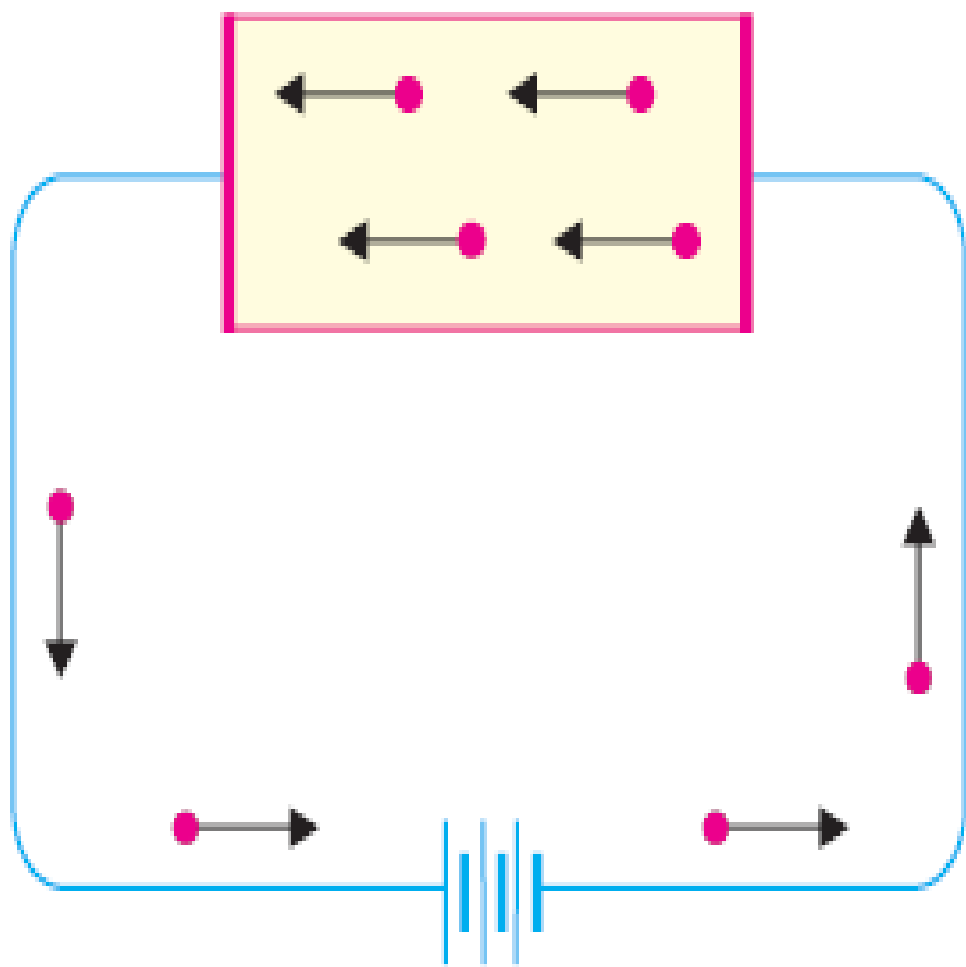
(i)
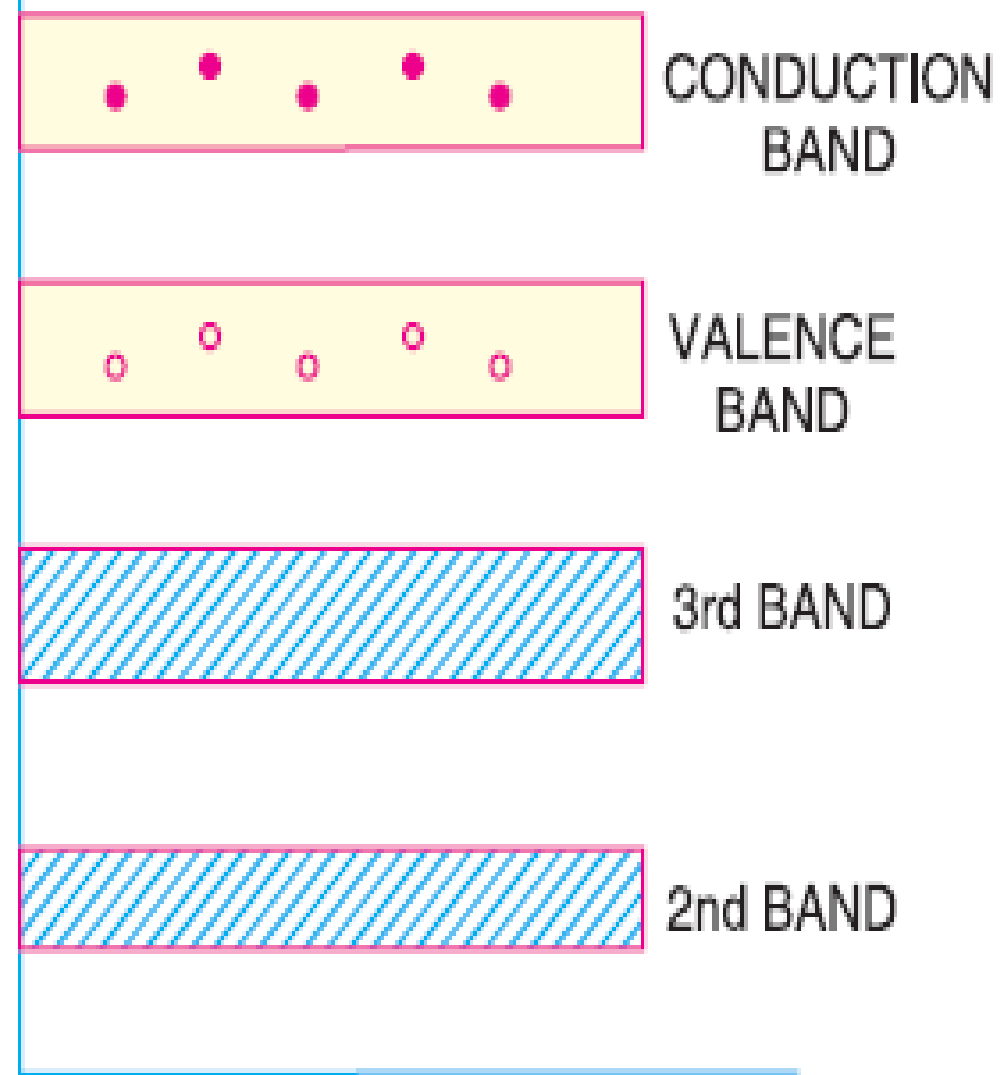
(ii)

Fig. 5.7

Fig. 5.7 (*ii*) shows the energy band diagram. As the temperature is raised, some of the valence electrons acquire sufficient energy to enter into the conduction band and thus become free electrons. Under the influence of electric field, these free electrons will constitute electric current. It may be noted that each time a valence electron enters into the conduction band, a *hole* is created in the valence band. As we shall see in the next article, holes also contribute to current. In fact, hole current is the most significant concept in semiconductors.

3.7. Drift Current

The flow of current in the semiconductor constituted by the drift electrons available in the conduction band and holes available in the valence band, which are formed due to external (heat) energy supplied to them, is known as drift current.

3.8. Diffusion Current

When the two pieces are joined together and suitably treated, they form a pn junction. The moment they form a pn junction, some of the conduction electrons from n- type material diffuse over to the p-type material and undergo electrons holes recombination with the holes available in the valence band. Simultaneously holes from p-type material diffuse over to the n-type material and undergo hole-electron combination with the electron available in the conduction band. This process is called diffusion.

Thus the current which obtained while having diffusion is called diffusion current.

(a) Total Current

It is possible that a potential gradient and a concentration gradient may exist within semiconductor. In such a case the total current is the sum of drift current due to potential gradient and the diffusion current due to charge carrier concentration gradient.

(b) Mobility of Charges

The mobility of charge carriers (elctrons and holes) varies as T^{-m} over a temperature range of 100 and 400 k. for silicon m = 2.5 for electrons amd 2.7 for holes. For germanium m = 1.66 for electrons and 2.33 for holes.

The carriers currents are also due to concentration gradients in the doped material which leads to diffusion of carriers from high concentration region to low concentration region.

HOLE CURRENT

At room temperature, some of the co-valent bonds in pure semiconductor break, setting up free electrons. Under the influence of electric field, these free electrons constitute electric current. At the same time, another current – the hole current – also flows in the semiconductor. When a covalent bond is broken due to thermal energy, the removal of one electron leaves a vacancy *i.e.* a missing electron in the covalent bond. This missing electron is called a hole which acts as a positive charge. For one electron set free, one hole is created. Therefore, thermal energy creates *hole-electron pairs*; there being as many holes as the free electrons. The current conduction by holes can be explained as follows:

The hole shows a missing electron. Suppose the valence electron at *L* (See Fig. 5.8) has become free electron due to thermal energy. This creates a hole in the co-valent bond at *L*. The hole is a strong centre of attraction for the electron. A valence electron (say at *M*) from nearby co-valent bond comes to fill in the hole at *L*. This results in the creation of hole at *M*. Another valence electron (say at *N*) in turn may leave its bond to fill the hole at *M*, thus creating a hole at *N*. Thus the hole having a positive charge has moved from *L* to *N* *i.e.* towards the negative terminal of supply. This constitutes *hole current*.
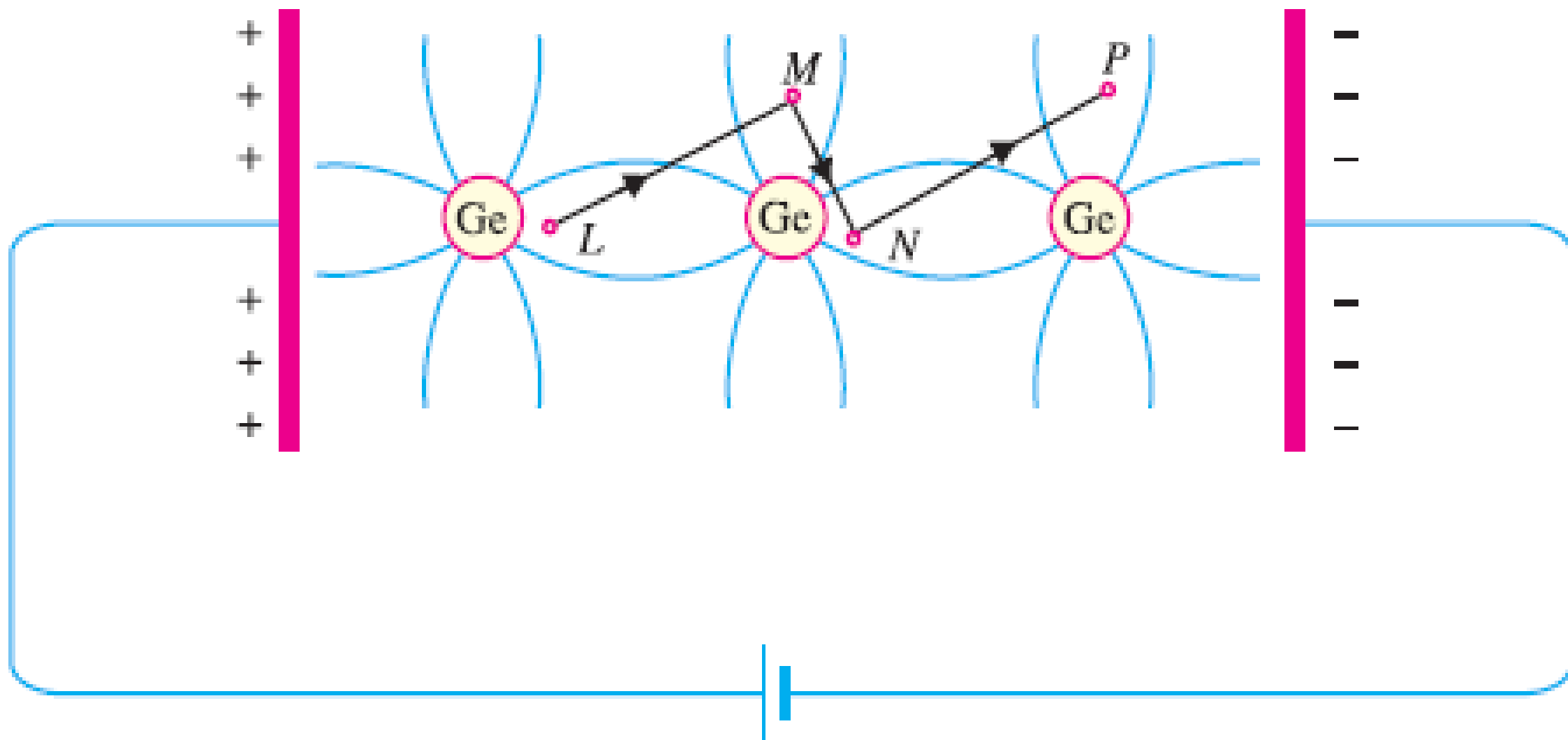
Fig. 5.8

It may be noted that hole current is due to the movement of valence electrons from one covalent bond to another bond. The reader may wonder why to call it a hole current when the conduction is again by electrons (of course *valence electrons*!). The answer is that the basic reason for current flow is the presence of holes in the co-valent bonds. Therefore, it is more appropriate to consider the current as the movement of holes.

3.10. PN Junction Diode (Semiconductor Diode)

A pn junction is known as a semiconductor diode.
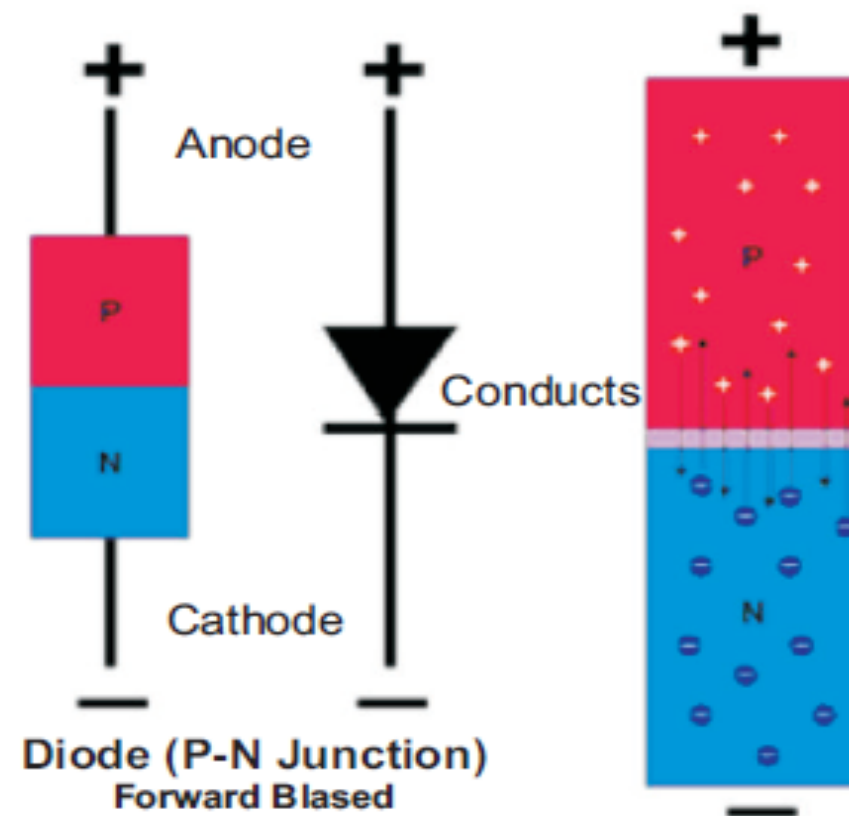
Figure-6

It is also known as crystal diode since it is grown out of a crystal (like germanium of silicon). A semiconductor diode has two terminals. Its symbol is shown in fig.-6. It conducts only when it is forward biased i.e when terminal connected with overhead is at higher potential than the terminal connected to the bar. However when it is reversed biased practically it does not conduct any current through it.

3.11. Depletion Layer

A region around the junction from which the charge carriers (free electrons and holes) are depleted is called depletion layer.
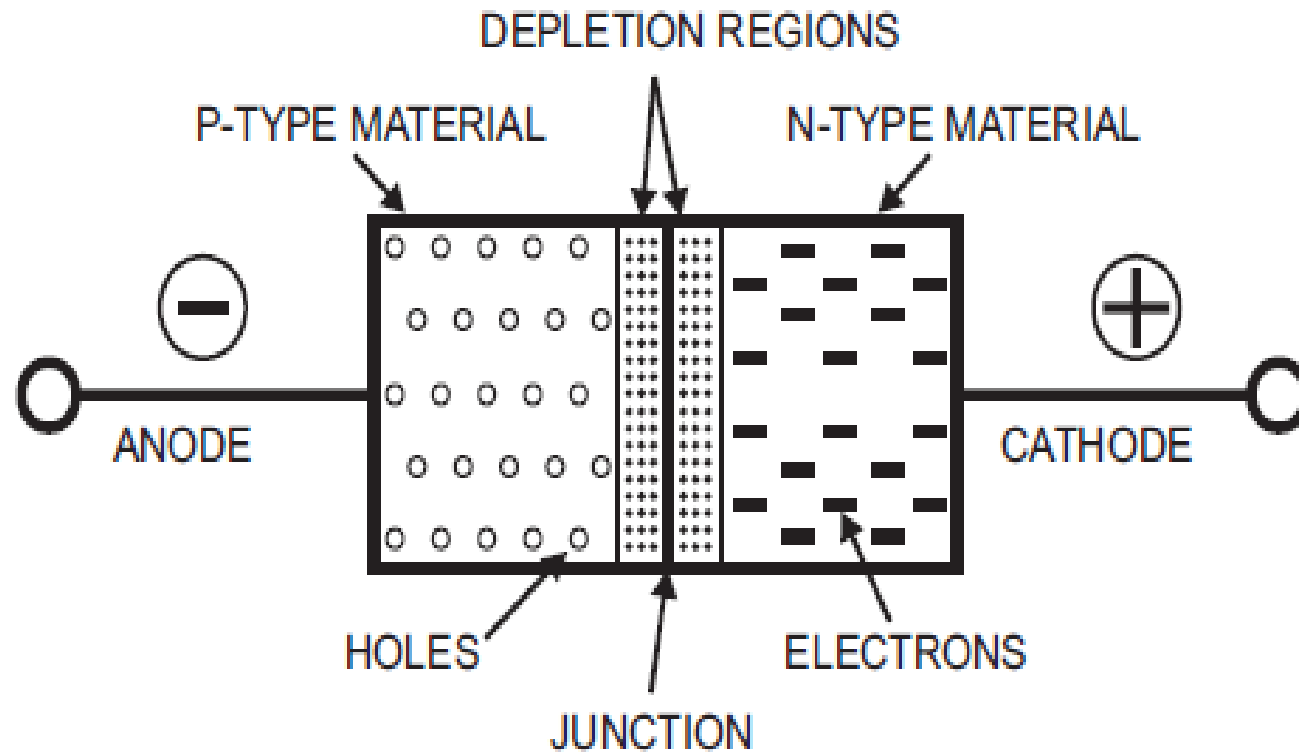
Figure-7

3.12. Potential Barrier

A potential difference built up across the pn junction which restricts further movement of charge carriers across the junction is known as potential barriers.

3.13. Forward & Reverse Biasing

When a pn junction is connected across an electric supply (potential difference) the junction is said to be under biasing. The type of biasing can be
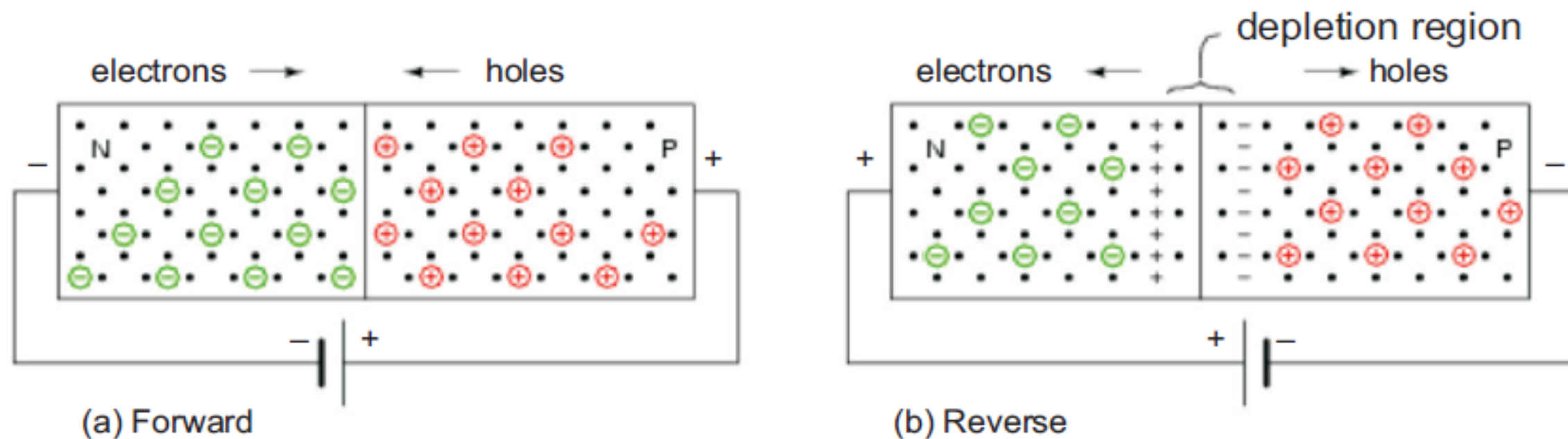
Figure-8

(a) Forward Biasing

When the positive terminal of a d.c. source or battery is connected to p-type and negative terminal is connected to n-type semiconductor of a pn junction, the junction is said to be in forward biasing the following points are worth noting, when a junction is forward biased:

(i) The junction potential barrier is reduced and at some forward voltage (0.3 v for germanium and 0.7 v for silicon). It is eliminated altogether.

(ii) The junction offers low resistance to the flow of current through it.

(iii) The magnitude of flow of current through the circuit depends upon the applied forward voltage.

(b) Reverse Biasing

When the positive terminal of a d.c. source or battery is connected to n-type and negative terminal is connected to p-type semiconductor of a pn junction, the junction is said to be in reverse biasing.

The following points are worth noting, when a junction is forward biased:

(i) The junction potential barrier is strenghened.

(ii) The junction offers high resistance to the flow of current through it.

(iii) The magnitude of flow of current through the circuit depends upon the applied reverse voltage.

3.14. V-I Characteristic of Semiconductor

The volt ampere (v-i) characteristics of a pn junction is just a curve between voltage across the junction and the circuit current. To draw the curve the circuit is arranged. In the circuit it is important to note that a resistor R is connected in series with the pn junction which limits the forward diode current from exceeding the permitted value. The characteristics are studied under three heads viz. zero external voltage, forward biasing and reverse biasing.
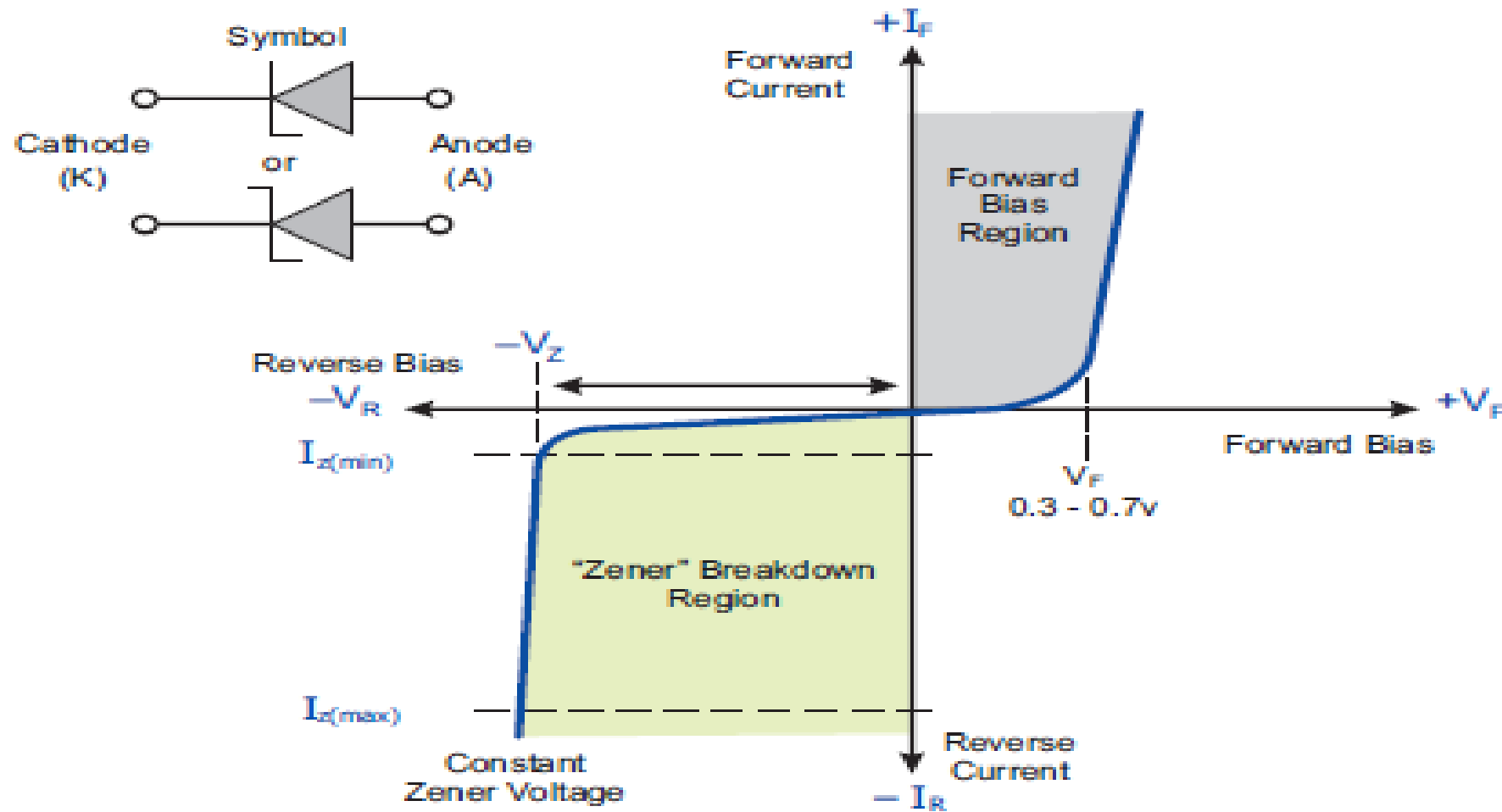
Figure - 9

(i) **Zero External Voltage:** When no external voltage is applied i.e circuit is open at key k. no current flow through the circuit. It is indicated by points 0 on the graph.

(ii) **Forward Biasing:** When key k is closed and double pole double throw switch is thrown to position 1. the pn junction is forward biasd as p-type semiconductor is connected to the positive terminal and n-type to the negative terminal of the supply. Now when supply voltage is increased by changing the variable resistor R the circuit current increases very slowly and the curve is non linear. The slow rise in current in this region is because the external applied voltage is used to overcome the potential barrier (0.3V for Ge and 0.7V for Si) of the pn junction.

However once the potential barrier is eliminated and external voltage is increased further the pn junction behaves like an ordinary conductor and the circuit current rises very sharply. At this instant the circuit current is limited by the series resistance R and a small value of the junction forward resistance R. The curve is almost linear. If the current rises more than the rated value of the diode the diode may change.

(iii) **Reverse Biasing:** When the double pole double throw switch is thrown to position 2, the pn junction is reverse biased as p- type semiconductor is connected to the negative terminal and n-type to the positive terminal of the supply. Under this condition the potential barrier at the junction is increased. Therefore the junction resistance becomes very high and practically no current flows through the circuit. However in actual practice a very small current flows in the circuit as shown, this current is called reverse current and is due to minority carriers available at room temperature in the two types of semiconductor. The reverse bias appears as a forward biased for these undesirable minority carriers and thus they constitue a minor current in reverse direction. The reverse current increases slightly with the increase in reverse bias supply voltage.

The reverse voltage at which pn junction breaks in known as breakdown voltage.

3.15. Resistance Level

An actual diode offers a very small resistance when forward biased and is called a forward resistance whereas it offers a very high resistance when reverse biased and is called a reverse resistance.
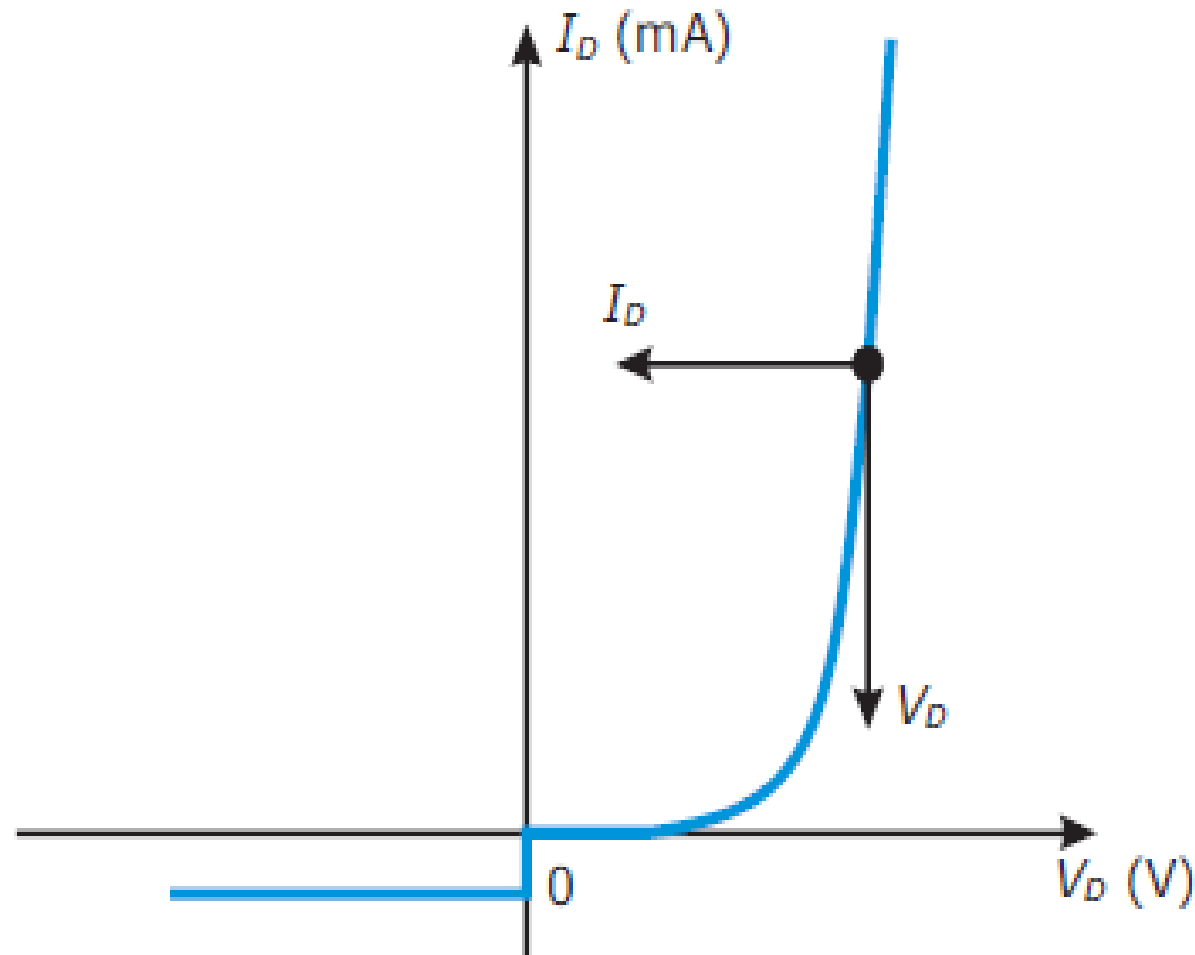
Figure-10 : Determining the DC Resistance of a Diode at a Particular Operating Point

3.16. Breakdown in Junction Diode

The breakdown of the pn junction can be of two types, these are

(i) Avalanche Breakdown

For thicker junctions the breakdown mechanism is by the process of avalanche breakdown. In this mechanism when the electric field existing in the depletion layer is sufficiently high, the velocity of the carriers (minority carriers) crossing the depletion layer increases. These carriers collide with the crystal atoms. Some collisions are so violent that electrons are knocked off the crystal atoms, thus creating electron hole pairs as the pair of electron hole is created in the midst of the high field, they quickly separate and attain high velocities to cause further pair generation through more collisions. This is cumulative process and as we approach the breakdown voltage, the field becomes so large that the chain of collisions can give rise to an almost infinite current with very slight additional increase in voltage. The process is known as avalanche breakdown. Once this breakdown occurs, the junction cannot regain its original position. Thus the diode is said to be burnt off.

(i) Zener Breakdown

This breakdown takes place in a very thin junction i.e. when both sides of the junction are very heavily doped and consequently the depletion layer is narrow. In the zener breakdown mechanism, the electric field becomes as high as 10^7 v/m in the depletion layer with only a small applied reverse bias voltage.

In his process it becomes possible for some electrons to jump across the barrier from the valence band in p- type material to some of the unfilled conduction band in n-material. This process is known as zener breakdown. In this process the junction is not damaged.

PN JUNCTION BREAKDOWN CHARACTERISTICS
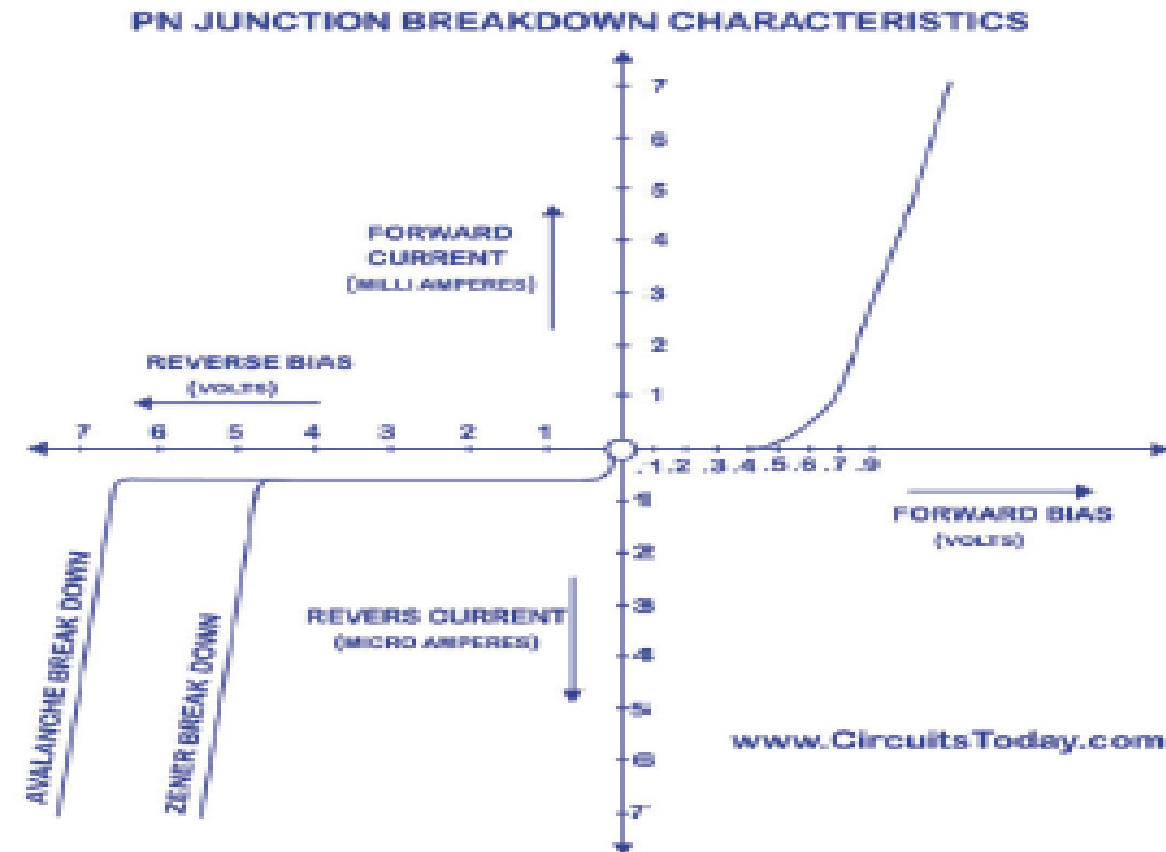
Figure-11

3.17. Zener Diode

A specially designed silicon diode which is optimised to operate in the breakdown region is known as as zener diode.

Appearance

Schematic Symbol
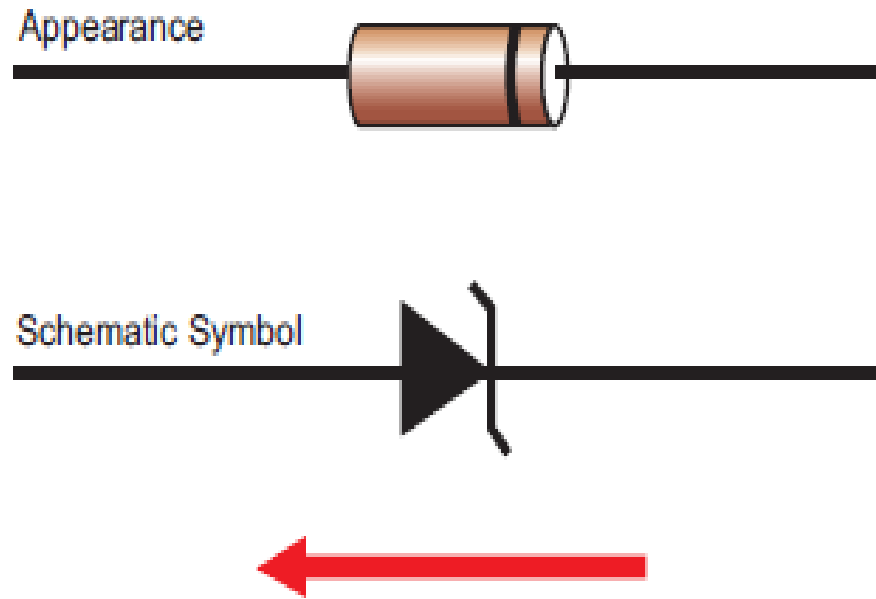
Backwards Current Flow Too, but only Past the "Zener" Breakdown Voltage

Figure-12

Characteristics of zener diode

1) Its characteristics are similar to an ordinary diode with the exception that it has a sharp (or distinct) breakdown voltage called zener voltage v_z.

2) It can be operated in any of the three region i.e. forward, leakage or breakdown. But usually it is operated in the breakdown region.

3) The voltage is almost constant (v_z) over the operating range.

4) During operation it will not burn as long as the external circuit limits the current flowing through it below the burn out value i.e. I_{zm} (the maximum rated zener current)

Applications

i.   Meter protection

ii. Voltage regulator

iii. Wave shaping circuit.

3.19. LED (Light Emitting Diode)

ANODE   CATHODE

(a)

(b)

(c)
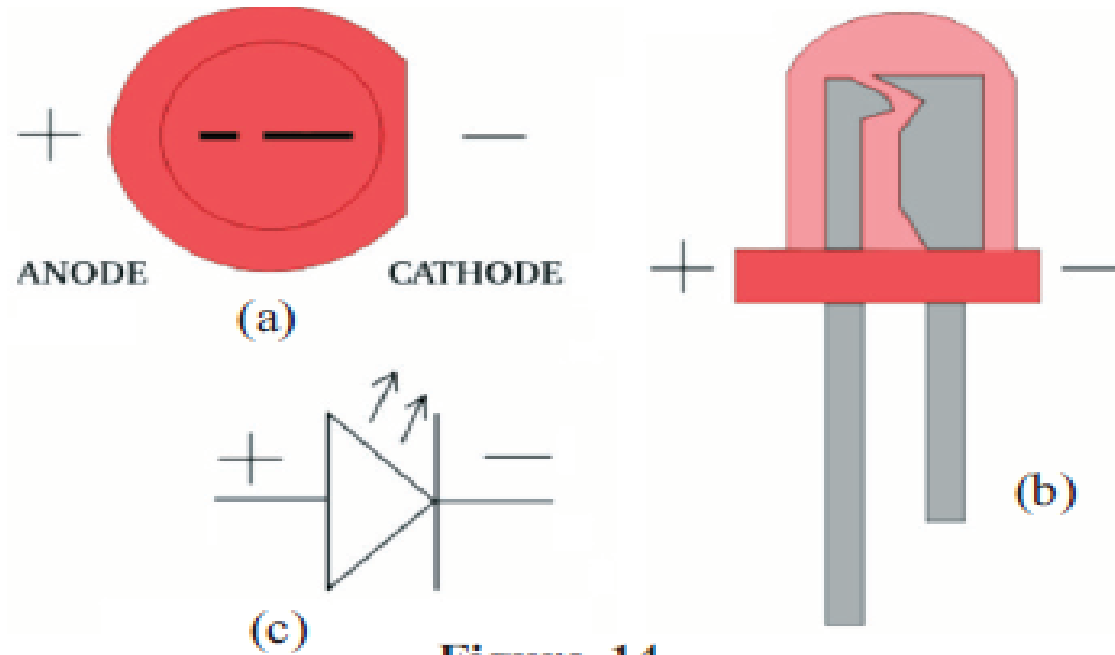
Figure-14

When a diode is forward biased the potential barrier is lowered. The conduction band free electrons from n- region cross the barrier and enter the p-region, as these electrons enter the p- region they fall into the holes lying in the valence band. Hence they fall from a higher energy level to a lower energy level in the process they radiate energy.

The LED are different. These are made of gallium arsenide phosphide (GaAsP) and gallium phosphide (GaP). In LED the energy is radiated in the form of light and hence they glow.

A manufacturer can produce LED that radiate red, green, yellow, blue, orange light.

Application

Instrument display, panel indicators, digital watches, calculator etc.

PN JUNCTION

*When a p-type semiconductor is suitably joined to n-type semiconductor, the contact surface is called **pn junction.*** Most semiconductor devices contain one or more *pn* junctions. The *pn* junction is of great importance because it is in effect, the *control element* for semiconductor devices. A thorough knowledge of the formation and properties of *pn* junction can enable the reader to understand the semiconductor devices.

PROPERTIES OF PN JUNCTION

At the instant of *pn*-junction formation, the free electrons near the junction in the *n* region begin to diffuse across the junction into the *p* region where they combine with holes near the junction. The result is that *n* region loses free electrons as they diffuse into the junction. This creates a layer of positive charges (pentavalent ions) near the junction. As the electrons move across the junction, the *p* region loses holes as the electrons and holes combine.

The result is that there is a layer of negative charges (trivalent ions) near the junction. These two layers of positive and negative charges form the *depletion region* (or *depletion layer*). The term depletion is due to the fact that near the junction, the region is depleted (*i.e.* emptied) of *charge carries* (free electrons and holes) due to diffusion across the junction. It may be noted that depletion layer is formed very quickly and is very thin compared to the n region and the p region. For clarity, the width of the depletion layer is shown exaggerated.

DEPLETION LAYER

p n

(i)

ELECTRIC FIELD

$\leftarrow V_0 \rightarrow$

p n

(ii)

Fig. 5.19

p $- V_0 +$ n

DEPLETION LAYER

V_0
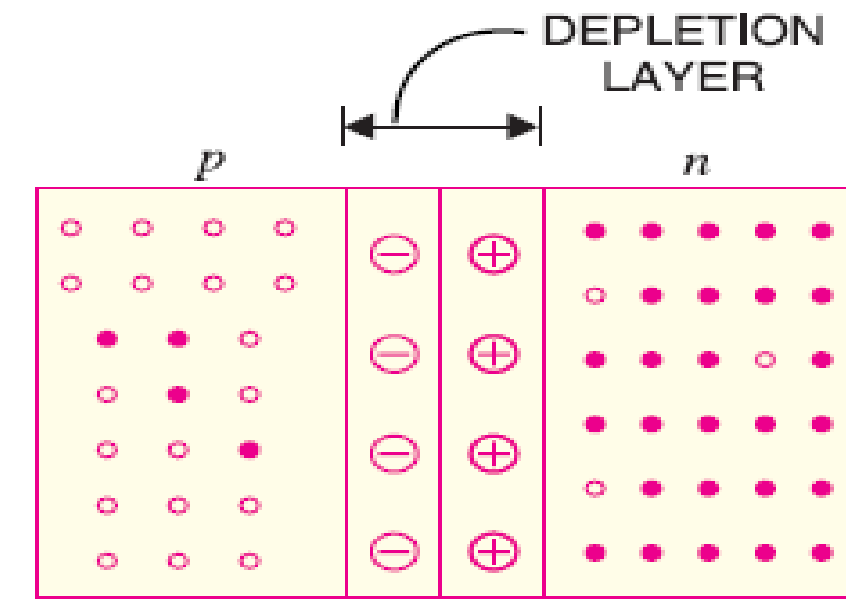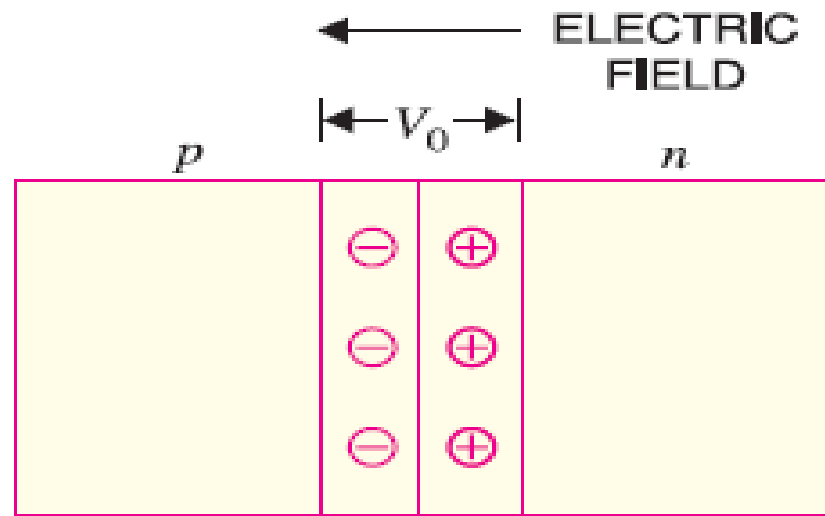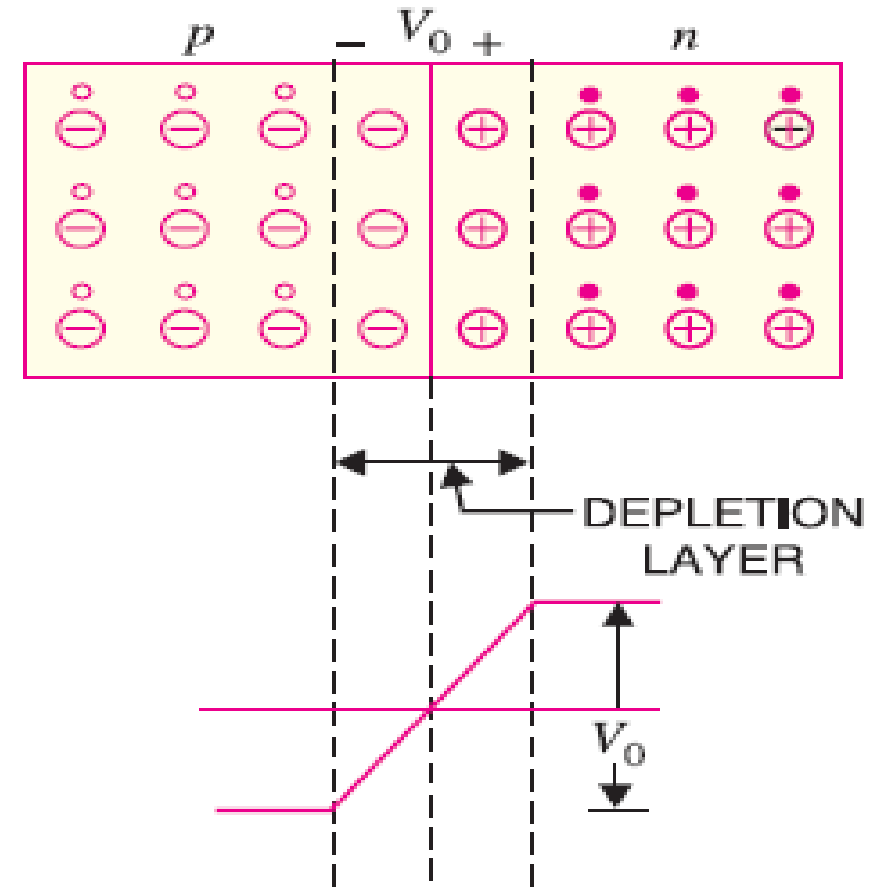
Fig. 5.20

Once *pn* junction is formed and depletion layer created, the diffusion of free electrons stops. In other words, the depletion region acts as a barrier to the further movement of free electrons across the junction. The positive and negative charges set up an electric field. This is shown by a black arrow in Fig. 5.19 (*i*). The electric field is a barrier to the free electrons in the *n*-region. There exists a potential difference across the depletion layer and is called **barrier potential** (V_o). The barrier potential of a *pn* junction depends upon several factors including the type of semiconductor material, the amount of doping and temperature. The typical barrier potential is approximately: For silicon, $V_o = 0.7$ V; For germanium, $V_o = 0.3$ V Fig. 5.20 shows the potential (V_o) distribution curve.

APPLYING D.C. VOLTAGE ACROSS PN JUNCTION OR BIASING A PN JUNCTION

In electronics, the term bias refers to the use of d.c voltage to establish certain operating conditions for an electronic device. In relation to a *pn* junction, there are following two bias conditions:

1. Forward biasing 2. Reverse biasing
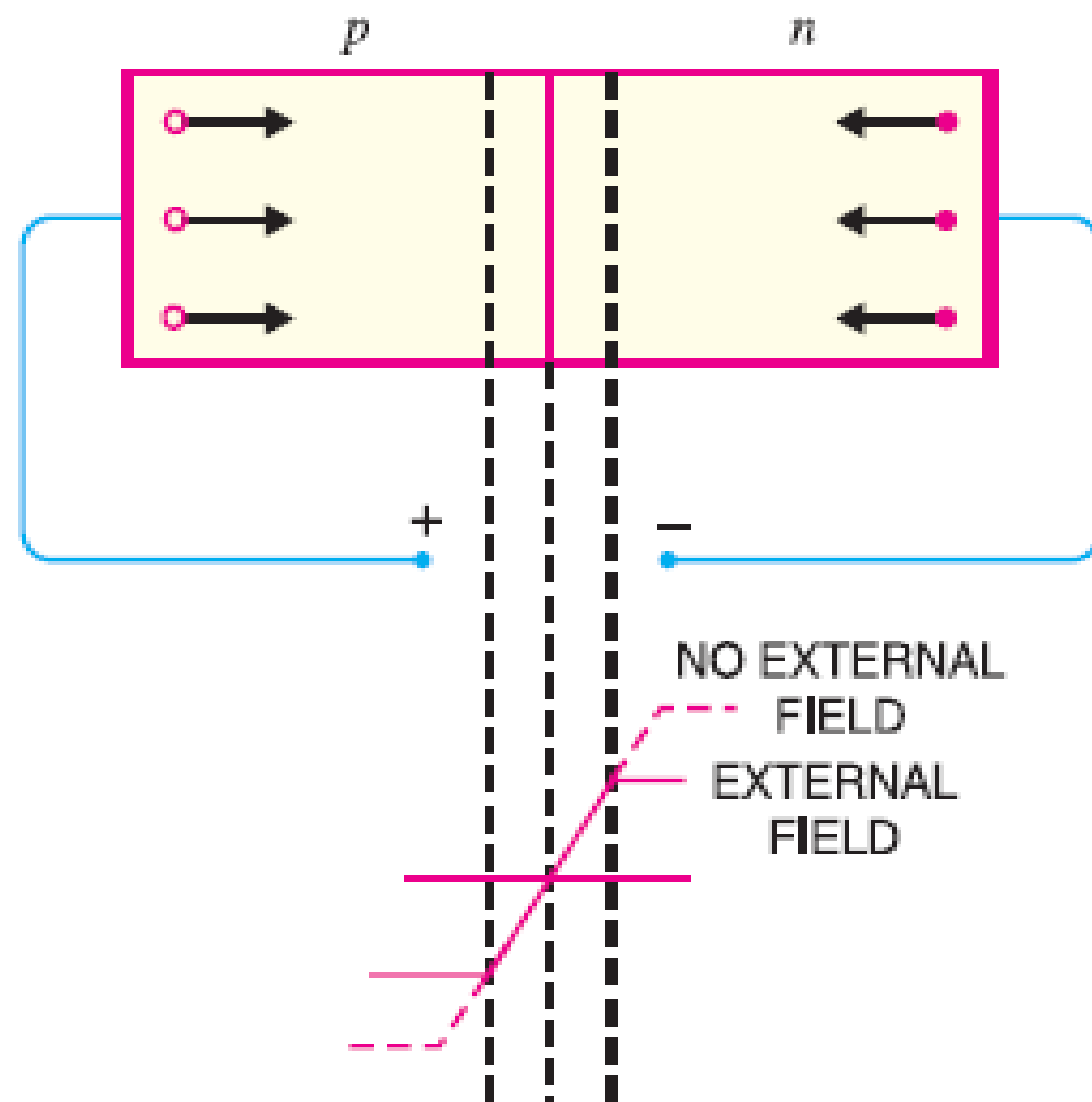
Fig. 5.21

FORWARD BIASING

When external d.c voltage applied to the junction is in such a direction that it cancels the potential barrier, thus permitting current flow, it is called **forward biasing.** To apply forward bias, connect positive terminal of the battery to p-type and negative terminal to n-type as shown in Fig. 5.21. The applied forward potential establishes an electric field which acts against the field due to potential barrier. Therefore, the resultant field is weakened and the barrier height is reduced at the junction as shown in Fig. 5.21. As potential barrier voltage is very small (0.1 to 0.3 V), therefore, a small forward voltage is sufficient to completely eliminate the barrier.

Once the potential barrier is eliminated by the forward voltage, junction resistance becomes almost zero and a low resistance path is established for the entire circuit. Therefore, current flows in the circuit. This is called *forward current*.

With forward bias to *pn* junction, the following points are worth noting:

(***i***) The potential barrier is reduced and at some forward voltage (0.1 to 0.3 V), it is eliminated altogether.

(***ii***) The junction offers low resistance (called *forward resistance, R_f*) to current flow.

(***iii***) Current flows in the circuit due to the establishment of low resistance path. The magnitude of current depends upon the applied forward voltage.

REVERSE BIASING

When the external d.c voltage applied to the junction is in such a direction that potential barrier is increased, it is called **reverse biasing.** To apply reverse bias, connect negative terminal of the battery to *p*-type and positive terminal to *n*-type. It is clear that applied reverse voltage establishes an electric field which acts in the same direction as the field due to potential barrier. Therefore, the resultant field at the junction is strengthened and the barrier height is increased. The increased potential barrier prevents the flow of charge carriers across the junction. Thus, a high resistance path is established for the entire circuit and hence the current does not flow.

With reverse bias to *pn* junction, the following points are worth noting :

(*i*) The potential barrier is increased. (*ii*) The junction offers very high resistance (called *reverse resistance, R_r*) to current flow.

(*iii*) No current flows in the circuit due to the establishment of high resistance path.

Conclusion. From the above discussion, it follows that with reverse bias to the junction, a high resistance path is established and hence no current flow occurs. On the other hand, with forward bias to the junction, a low resistance path is set up and hence current flows in the circuit.
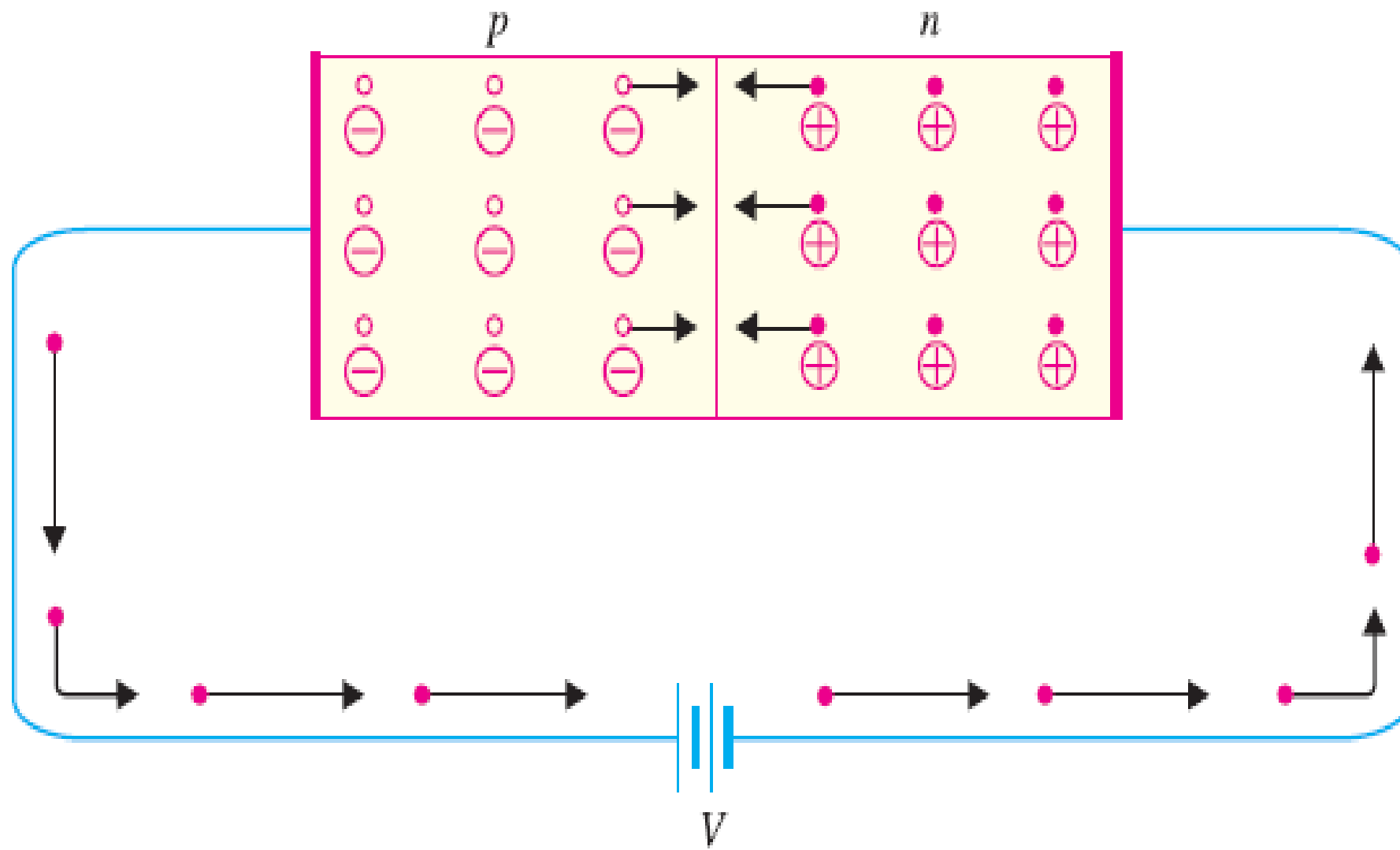
Fig. 5.23

CURRENT FLOW IN A FORWARD BIASED PN JUNCTION

We shall now see how current flows across *pn* junction when it is forward biased. Fig. 5.23 shows a forward biased *pn* junction. Under the influence of forward voltage, the free electrons in *n*-type move towards the junction, leaving behind positively charged atoms. However, more electrons arrive from the negative battery terminal and enter the *n*-region to take up their places.

As the free electrons reach the junction, they become valence electrons. As valence electrons, they move through the holes in the *p*-region. The valence electrons move towards left in the *p*-region which is equivalent to the holes moving to right. When the valence electrons reach the left end of the crystal, they flow into the positive terminal of the battery.

The mechanism of current flow in a forward biased *pn* junction can be summed up as under:

(*i*) The free electrons from the negative terminal continue to pour into the *n*-region while the free electrons in the *n*-region move towards the junction.

(*ii*) The electrons travel through the *n*-region as free-electrons *i.e.* current in *n*-region is by free electrons. Note that negative terminal of battery is connected to *n*-type. It repels the free electrons in *n*-type towards the junction. A hole is in the co-valent bond. When a free electron combines with a hole, it becomes a valence electron.

(*iii*) When these electrons reach the junction, they combine with holes and become valence electrons.

(*iv*) The electrons travel through *p*-region as valence electrons *i.e.* current in the *p*-region is by holes.

(*v*) When these valence electrons reach the left end of crystal, they flow into the positive terminal of the battery.

From the above discussion, it is concluded that in *n*-type region, current is carried by free electrons whereas in *p*-type region, it is carried by holes. However, in the external connecting wires, the current is carried by free electrons.
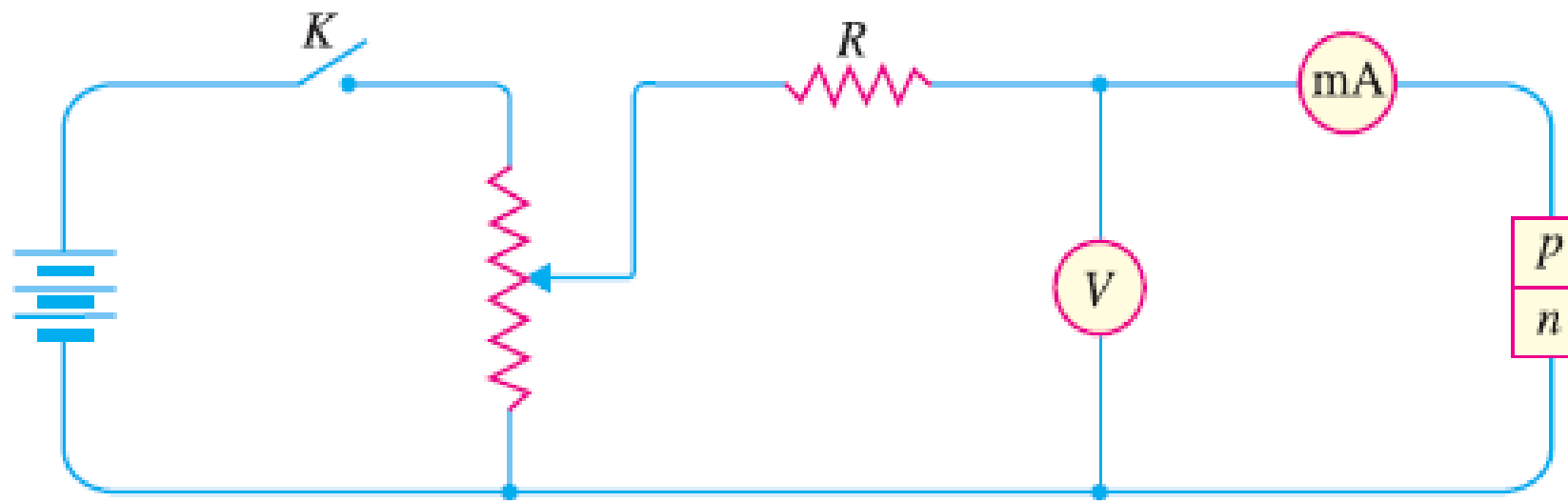
Fig. 5.24

VOLT-AMPERE CHARACTERISTICS OF PN JUNCTION

Volt-ampere or *V-I* characteristic of a *pn* junction (also called a *crystal or semiconductor diode*) is the curve between voltage across the junction and the circuit current. Usually, voltage is taken along *x*-axis and current along *y*-axis. Fig. 5.24 shows the circuit arrangement for determining the *V-I* characteristics of a *pn* junction. The characteristics can be studied under three heads, namely; *zero external voltage, forward bias* and *reverse bias*.
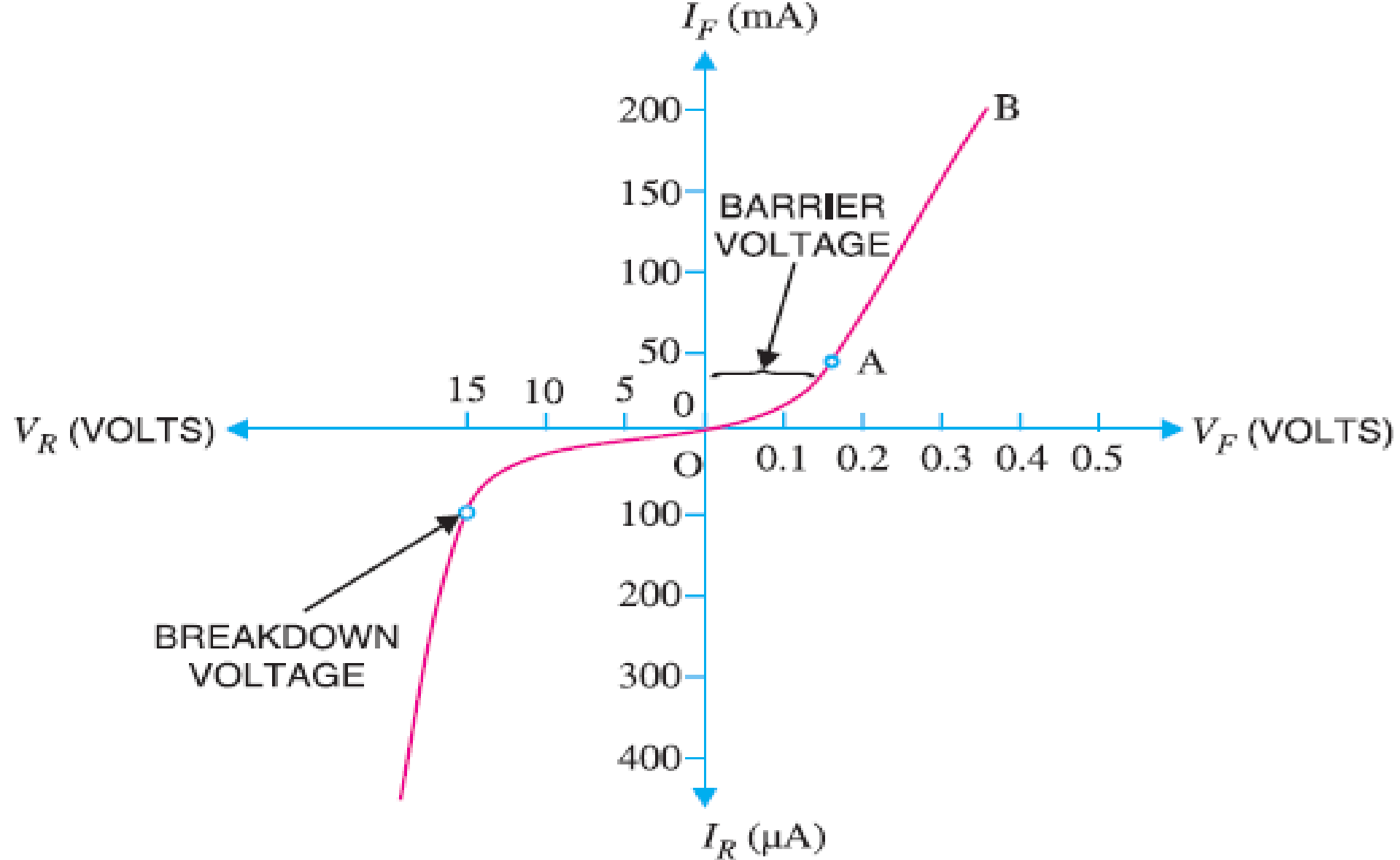
Fig. 5.25

---

* R is the current limiting resistance. It prevents the forward current from exceeding the permitted value.

ZERO EXTERNAL VOLTAGE

When the external voltage is zero, *i.e.* circuit is open at K, the potential barrier at the junction does not permit current flow. Therefore, the circuit current is zero as indicated by point O in Fig. 5.25.
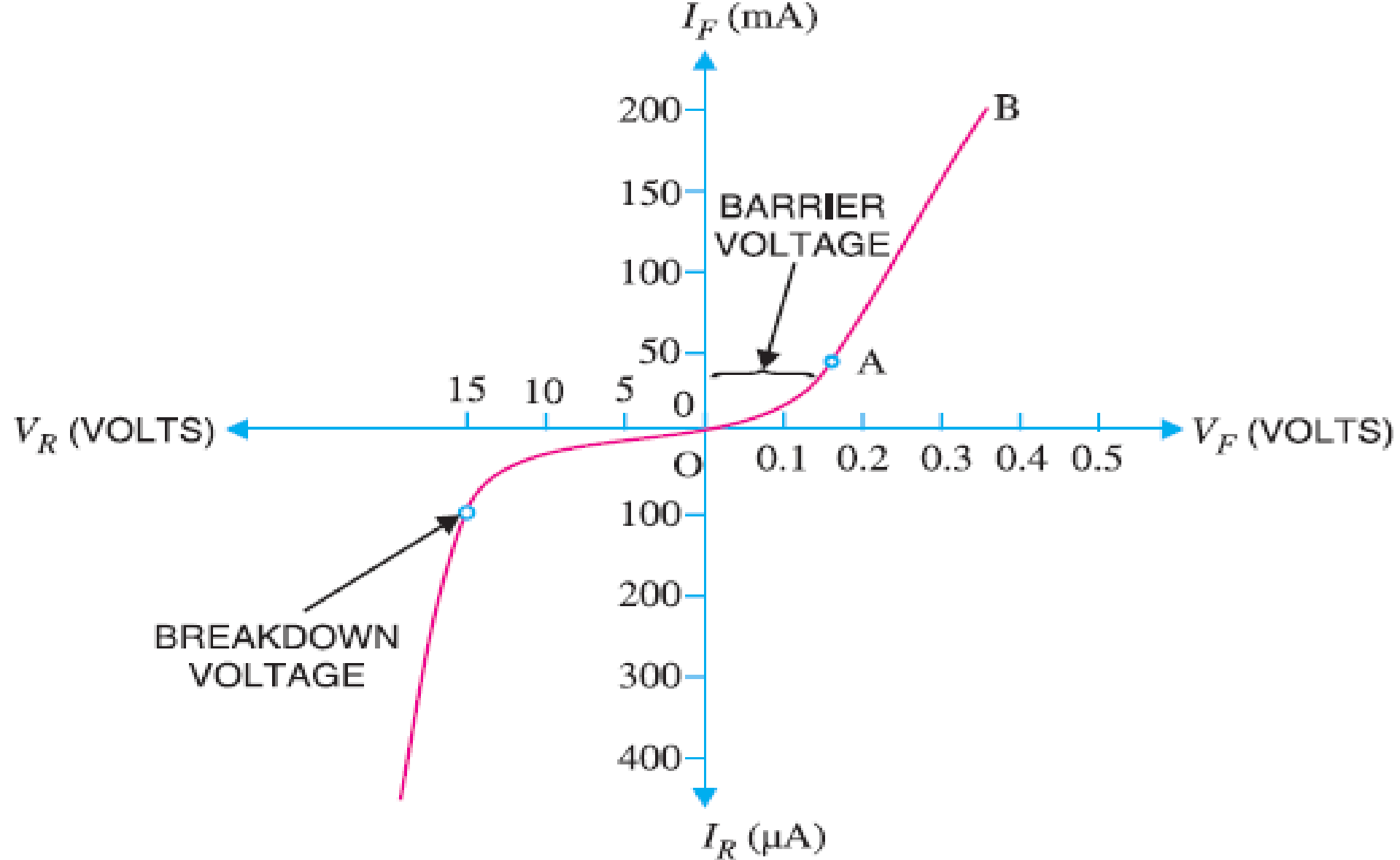
Fig. 5.25

* R is the current limiting resistance. It prevents the forward current from exceeding the permitted value.

FORWARD BIAS

With forward bias to the *pn* junction *i.e. p*-type connected to positive terminal and *n*-type connected to negative terminal, the potential barrier is reduced. At some forward voltage (0.7 V for Si and 0.3 V for Ge), the potential barrier is altogether eliminated and current starts flowing in the circuit. From now onwards, the current increases with the increase in forward voltage. Thus, a rising curve *OB* is obtained with forward bias as shown in Fig. 5.25.

From the forward characteristic, it is seen that at first (*region OA*),the current increases very slowly and the curve is non-linear. It is because the external applied voltage is used up in overcoming the potential barrier. However, once the external voltage exceeds the potential barrier voltage, the *pn* junction behaves like an ordinary conductor. Therefore, the current rises very sharply with increase in external voltage (*region AB on the curve*). The curve is almost linear.

REVERSE BIAS

With reverse bias to the *pn* junction *i.e. p*-type connected to negative terminal and *n*-type connected to positive terminal, potential barrier at the junction is increased. Therefore, the junction resistance becomes very high and practically no current flows through the circuit. However, in practice, a very small current (of the order of μ*A*) flows in the circuit with reverse bias as shown in the reverse characteristic.
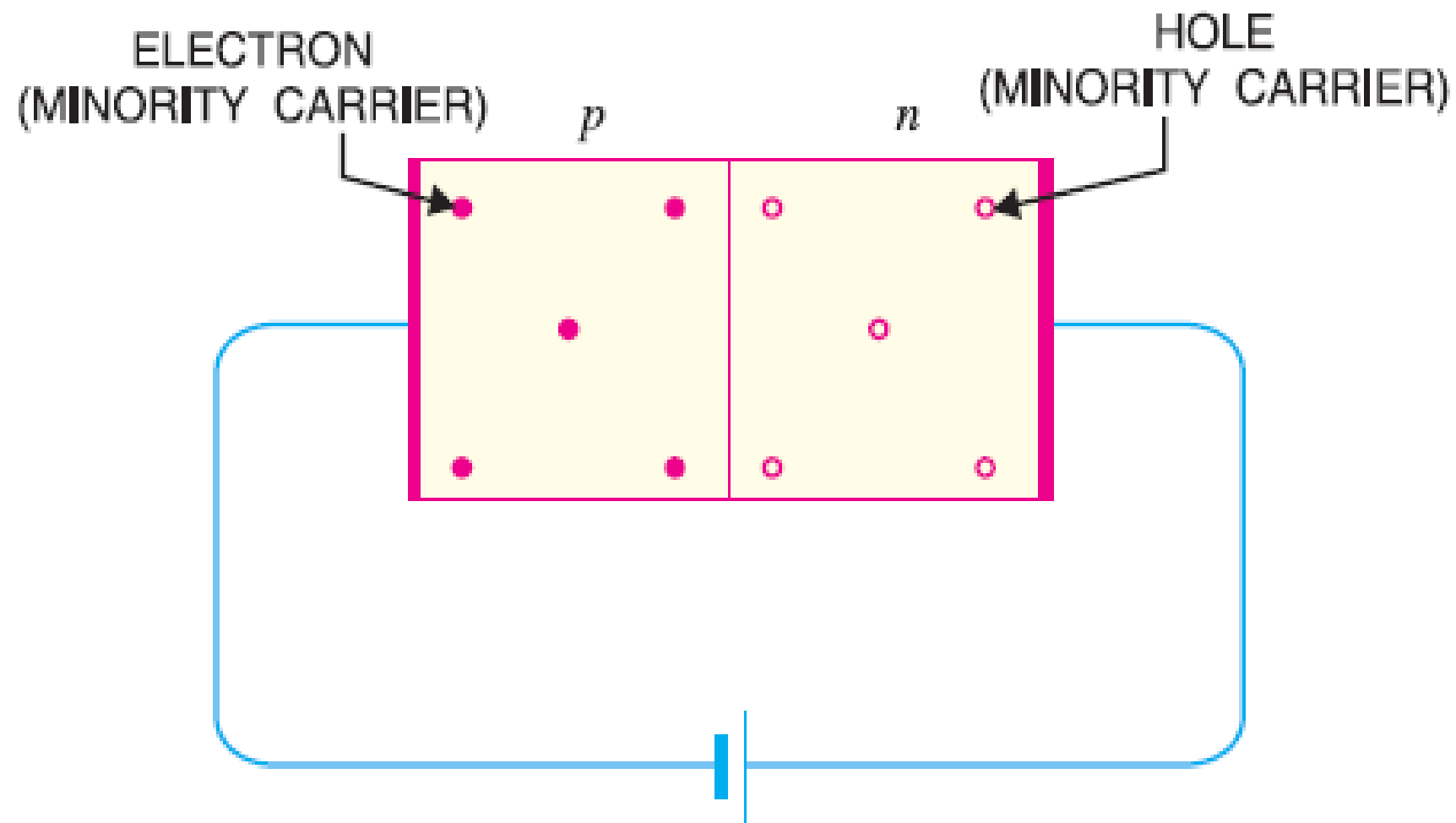
Fig. 5.26

This is called *reverse saturation current* (I_S) and is due to the minority carriers. It may be recalled that there are a few free electrons in *p*-type material and a few holes in *n*-type material. These undesirable free electrons in *p*-type and holes in *n*-type are called *minority carriers*. As shown in Fig. 5.26, to these minority carriers, the applied reverse bias appears as forward bias. Therefore, a small current flows in the reverse direction.

If reverse voltage is increased continuously, the kinetic energy of electrons (minority carriers) may become high enough to knock out electrons from the semiconductor atoms. At this stage *breakdown* of the junction occurs, characterised by a sudden rise of reverse current and a sudden fall of the resistance of barrier region. This may destroy the junction permanently. **Note.** The forward current through a *pn* junction is due to the *majority carriers* produced by the impurity. However, reverse current is due to the *minority carriers* produced due to breaking of some co-valent bonds at room temperature.

FREE ELECTRON

VALENCE ELECTRON
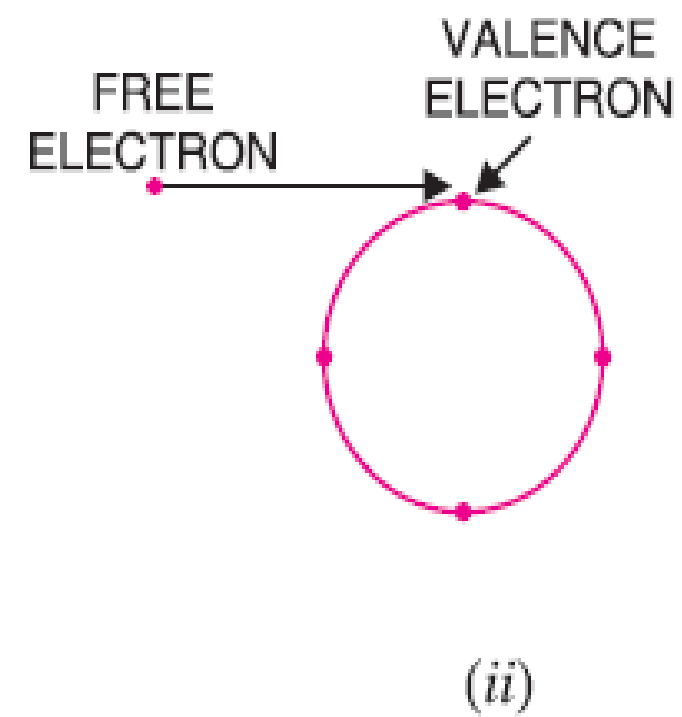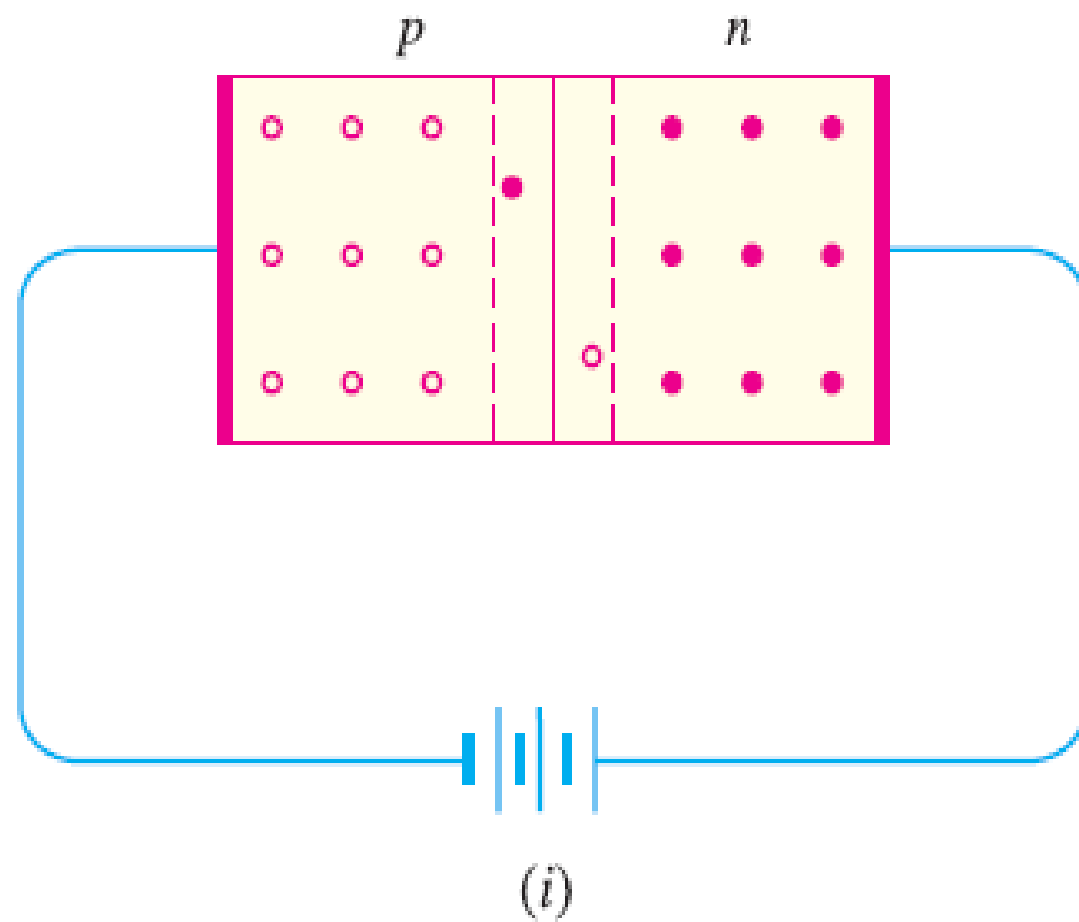
Fig. 5.27

IMPORTANT TERMS

Two important terms often used with *pn* junction (*i.e.* crystal diode) are *breakdown voltage* and *knee voltage*. We shall now explain these two terms in detail. **(i) Breakdown voltage.** *It is the minimum reverse voltage at which pn junction breaks down with sudden rise in reverse current*. Under normal reverse voltage, a very little reverse current flows through a *pn* junction. However, if the reverse voltage attains a high value, the junction may break down with sudden rise in reverse current. For understanding this point, refer to Fig. 5.27.

Even at room temperature, some hole-electron pairs (minority carriers) are produced in the depletion layer as shown in Fig. 5.27 (*i*). With reverse bias, the electrons move towards the positive terminal of supply. At large reverse voltage, these electrons acquire high enough velocities to dislodge valence electrons from semiconductor atoms as shown in Fig. 5.27 (*ii*). The newly liberated electrons in turn free other valence electrons. In this way, we get an *avalanche* of free electrons. Therefore, the *pn* junction conducts a very large reverse current. Once the breakdown voltage is reached, the high reverse current may damage the junction. Therefore, care should be taken that reverse voltage across a *pn* junction is always less than the breakdown voltage.
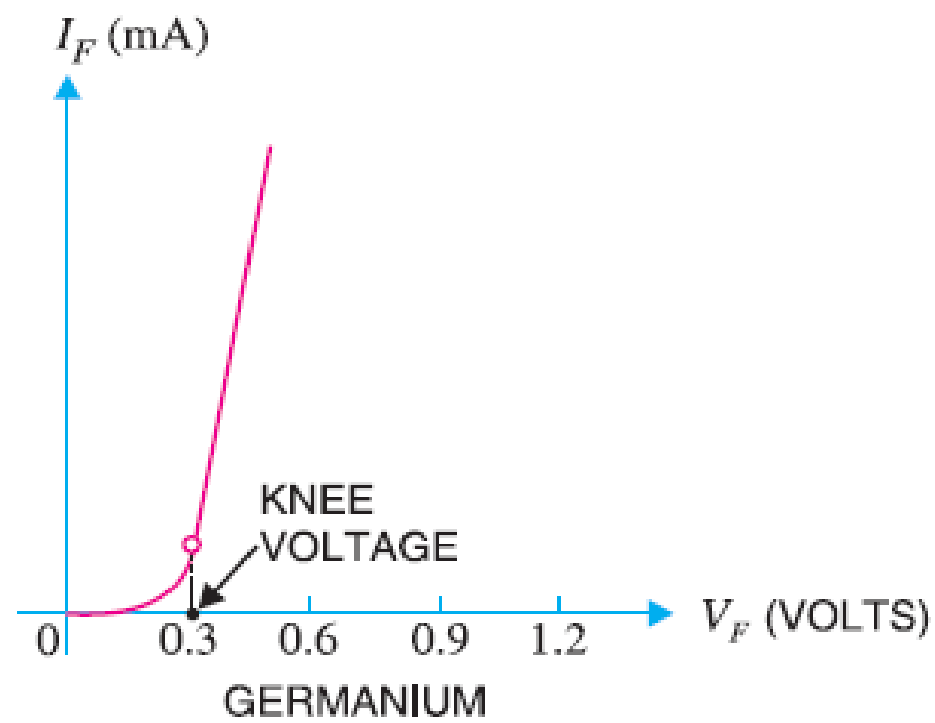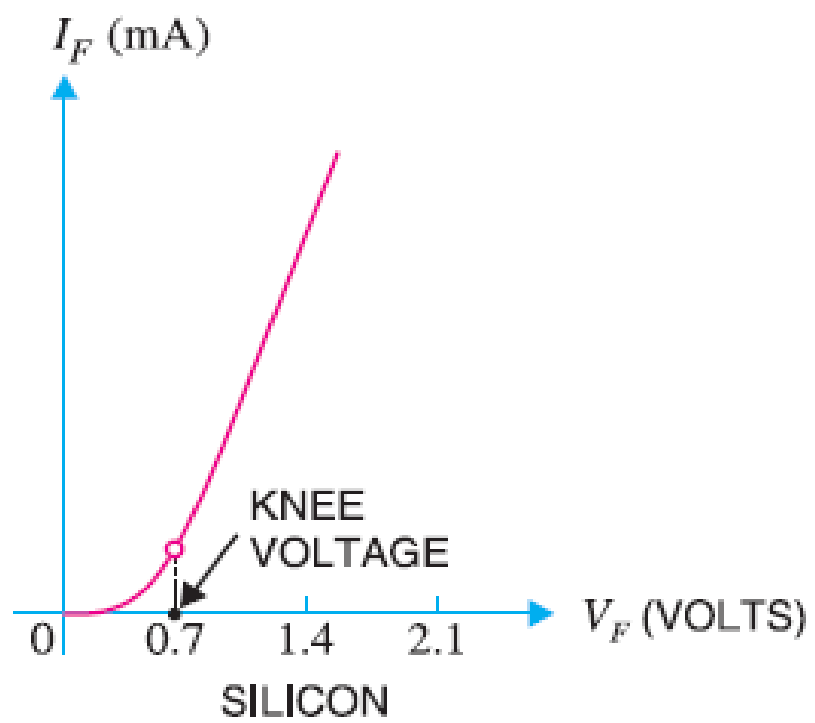
Fig. 5.28

(*ii*) **Knee voltage.** *It is the forward voltage at which the current through the junction starts to increase rapidly.* When a diode is forward biased, it conducts current very slowly until we overcome the potential barrier. For silicon *pn* junction, potential barrier is 0.7 V whereas it is 0.3 V for germanium junction. It is clear from Fig. 5.28 that knee voltage for silicon diode is 0.7 V and 0.3 V for germanium diode. Once the applied forward voltage exceeds the knee voltage, the current starts increasing rapidly. It may be added here that in order to get useful current through a *pn* junction, the applied voltage must be more than the knee voltage. **Note.** The potential barrier voltage is also known as *turn-on voltage*. This is obtained by taking the straight line portion of the forward characteristic and extending it back to the horizontal axis.

LIMITATIONS IN THE OPERATING CONDITIONS OF PN JUNCTION

Every *pn* junction has limiting values of <u>*maximum forward current, peak inverse voltage,* and *maximum power rating*</u>. The *pn* junction will give satisfactory performance if it is operated within these limiting values. However, if these values are exceeded, the *pn* junction may be destroyed due to excessive heat.

MAXIMUM FORWARD CURRENT It is the highest instantaneous forward current that a _pn_ junction can conduct without damage to the junction. Manufacturer's data sheet usually specifies this rating. If the forward current in a _pn_ junction is more than this rating, the junction will be destroyed due to overheating.

PEAK INVERSE VOLTAGE (PIV) It is the maximum reverse voltage that can be applied to the *pn* junction without damage to the junction. If the reverse voltage across the junction exceeds its PIV, the junction may be destroyed due to excessive heat. The peak inverse voltage is of particular importance in rectifier service. A *pn* junction *i.e.* a crystal diode is used as a rectifier to change alternating current into direct current. In such applications, care should be taken that reverse voltage across the diode during negative half-cycle of a.c. does not exceed the PIV of diode.

MAXIMUM POWER RATING It is the maximum power that can be dissipated at the junction without damaging it. The power dissipated at the junction is equal to the product of junction current and the voltage across the junction. This is a very important consideration and is invariably specified by the manufacturer in the data sheet.

A *p*-type material consists of silicon atoms and trivalent impurity atoms such as boron. The boron atom adds a hole when it bonds with the silicon atoms. However, since the number of protons and the number of electrons are equal throughout the material, there is no net charge in the material and so it is neutral.

An *n*-type silicon material consists of silicon atoms and pentavalent impurity atoms such as antimony. An impurity atom releases an electron when it bonds with four silicon atoms. Since there is still an equal number of protons and electrons (including the free electrons) throughout the material, there is no net charge in the material and so it is neutral.

(a) The basic silicon structure at the instant of junction formation showing only the majority and minority carriers. Free electrons in the *n* region near the *pn* junction begin to diffuse across the junction and fall into holes near the junction in the *p* region.

(b) For every electron that diffuses across the junction and combines with a hole, a positive charge is left in the *n* region and a negative charge is created in the *p* region, forming a barrier potential. This action continues until the voltage of the barrier repels further diffusion. The blue arrows between the positive and negative charges in the depletion region represent the electric field.

▲ FIGURE 1–19

Formation of the depletion region. The width of the depletion region is exaggerated for illustration purposes.

If a piece of intrinsic silicon is doped so that part is *n*-type and the other part is *p*-type, a ***pn* junction** forms at the boundary between the two regions and a diode is created, as indicated in Figure 1–19(a). The *p* region has many holes (majority carriers) from the impurity atoms and only a few thermally generated free electrons (minority carriers). The *n* region has many free electrons (majority carriers) from the impurity atoms and only a few thermally generated holes (minority carriers).

(a) The basic silicon structure at the instant of junction formation showing only the majority and minority carriers. Free electrons in the *n* region near the *pn* junction begin to diffuse across the junction and fall into holes near the junction in the *p* region.

(b) For every electron that diffuses across the junction and combines with a hole, a positive charge is left in the *n* region and a negative charge is created in the *p* region, forming a barrier potential. This action continues until the voltage of the barrier repels further diffusion. The blue arrows between the positive and negative charges in the depletion region represent the electric field.

▲ FIGURE 1–19

Formation of the depletion region. The width of the depletion region is exaggerated for illustration purposes.

FORMATION OF THE DEPLETION REGION

The free electrons in the n region are randomly drifting in all directions. At the instant of the pn junction formation, the free electrons near the junction in the n region begin to diffuse across the junction into the p region where they combine with holes near the junction, as shown in Figure 1–19(b).

Before the *pn* junction is formed, recall that there are as many electrons as protons in the *n*-type material, making the material neutral in terms of net charge. The same is true for the *p*-type material. When the *pn* junction is formed, the *n* region loses free electrons as they diffuse across the junction. This creates a layer of positive charges (pentavalent ions) near the junction. As the electrons move across the junction, the *p* region loses holes as the electrons and holes combine. This creates a layer of negative charges (trivalent ions) near the junction.

These two layers of positive and negative charges form the **depletion region,** as shown in Figure 1–19(b). The term *depletion* refers to the fact that the region near the *pn* junction is depleted of charge carriers (electrons and holes) due to diffusion across the junction. Keep in mind that the depletion region is formed very quickly and is very thin compared to the *n* region and *p* region. After the initial surge of free electrons across the *pn* junction, the depletion region has expanded to a point where equilibrium is established and there is no further diffusion of electrons across the junction.

This occurs as follows: As electrons continue to diffuse across the junction, more and more positive and negative charges are created near the junction as the depletion region is formed. A point is reached where the total negative charge in the depletion region repels any further diffusion of electrons (negatively charged particles) into the p region (like charges repel) and the diffusion stops. In other words, the depletion region acts as a barrier to the further movement of electrons across the junction.

BARRIER POTENTIAL

Any time there is a positive charge and a negative charge near each other, there is a force acting on the charges as described by Coulomb's law. In the depletion region there are many positive charges and many negative charges on opposite sides of the *pn* junction. The forces between the opposite charges form an *electric field,* as illustrated in Figure 1–19(b) by the blue arrows between the positive charges and the negative charges. This electric field is a barrier to the free electrons in the *n* region, and energy must be expended to move an electron through the electric field. That is, external energy must be applied to get the electrons to move across the barrier of the electric field in the depletion region.

The potential difference of the electric field across the depletion region is the amount of voltage required to move electrons through the electric field. This potential difference is called the **barrier potential** and is expressed in volts. Stated another way, a certain amount of voltage equal to the barrier potential and with the proper polarity must be applied across a *pn* junction before electrons will begin to flow across the junction. The barrier potential of a *pn* junction depends on several factors, including the type of semi conductive material, the amount of doping, and the temperature. The typical barrier potential is approximately 0.7 V for silicon and 0.3 V for germanium.
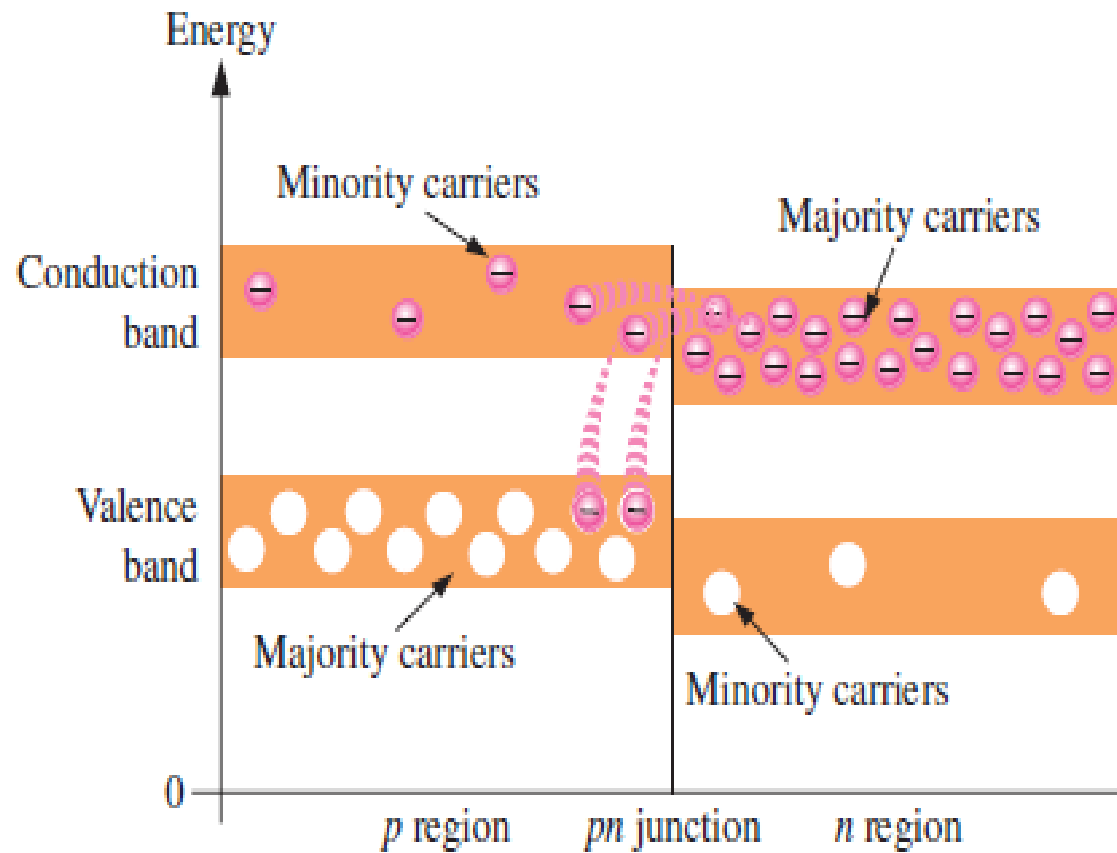
(a) At the instant of junction formation
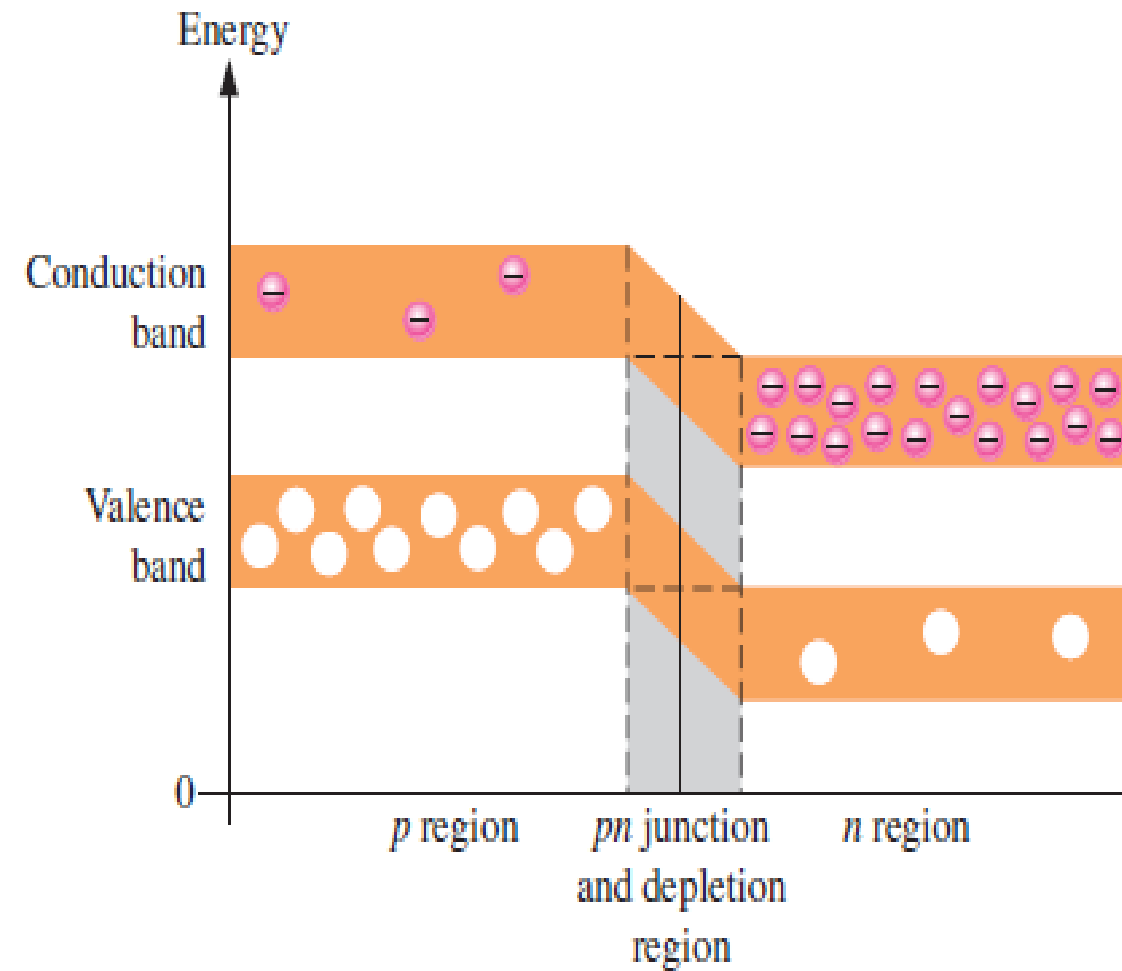
(b) At equilibrium

▲ FIGURE 1–20

Energy diagrams illustrating the formation of the *pn* junction and depletion region.

ENERGY DIAGRAMS OF THE *PN* JUNCTION AND DEPLETION REGION

The valence and conduction bands in an *n*-type material are at slightly lower energy levels than the valence and conduction bands in a *p*-type material. Recall that *p*-type material has trivalent impurities and *n*-type material has pentavalent impurities. The trivalent impurities exert lower forces on the outer-shell electrons than the pentavalent impurities. The lower forces in *p*-type materials mean that the electron orbits are slightly larger and hence have greater energy than the electron orbits in the *n*-type materials. An energy diagram for a *pn* junction at the instant of formation is shown in Figure 1–20(a).
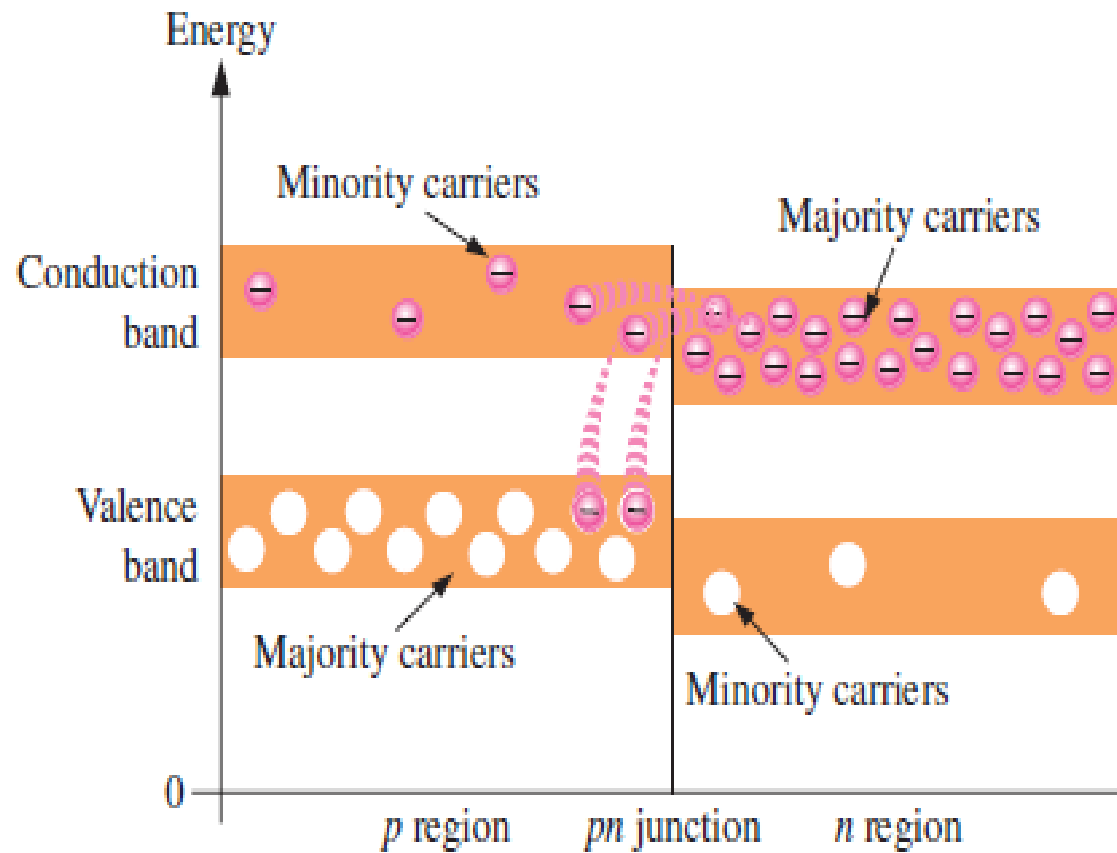
(a) At the instant of junction formation
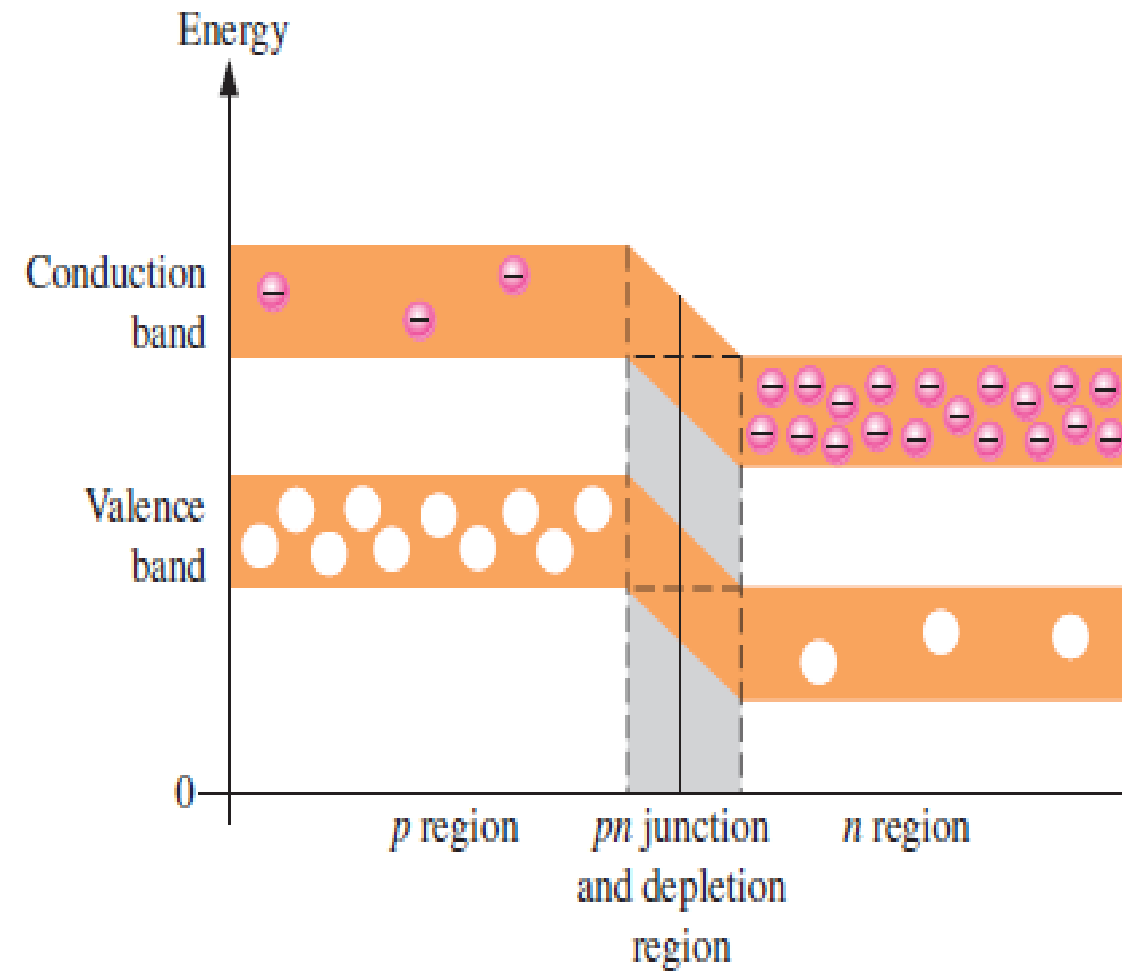
(b) At equilibrium

▲ FIGURE 1–20

Energy diagrams illustrating the formation of the *pn* junction and depletion region.

As you can see, the valence and conduction bands in the *n* region are at lower energy levels than those in the *p* region, but there is a significant amount of overlapping. The free electrons in the *n* region that occupy the upper part of the conduction band in terms of their energy can easily diffuse across the junction (they do not have to gain additional energy) and temporarily become free electrons in the lower part of the *p*-region conduction band.

After crossing the junction, the electrons quickly lose energy and fall into the holes in the *p*-region valence band as indicated in Figure 1–20(a). As the diffusion continues, the depletion region begins to form and the energy level of the *n*-region conduction band decreases.
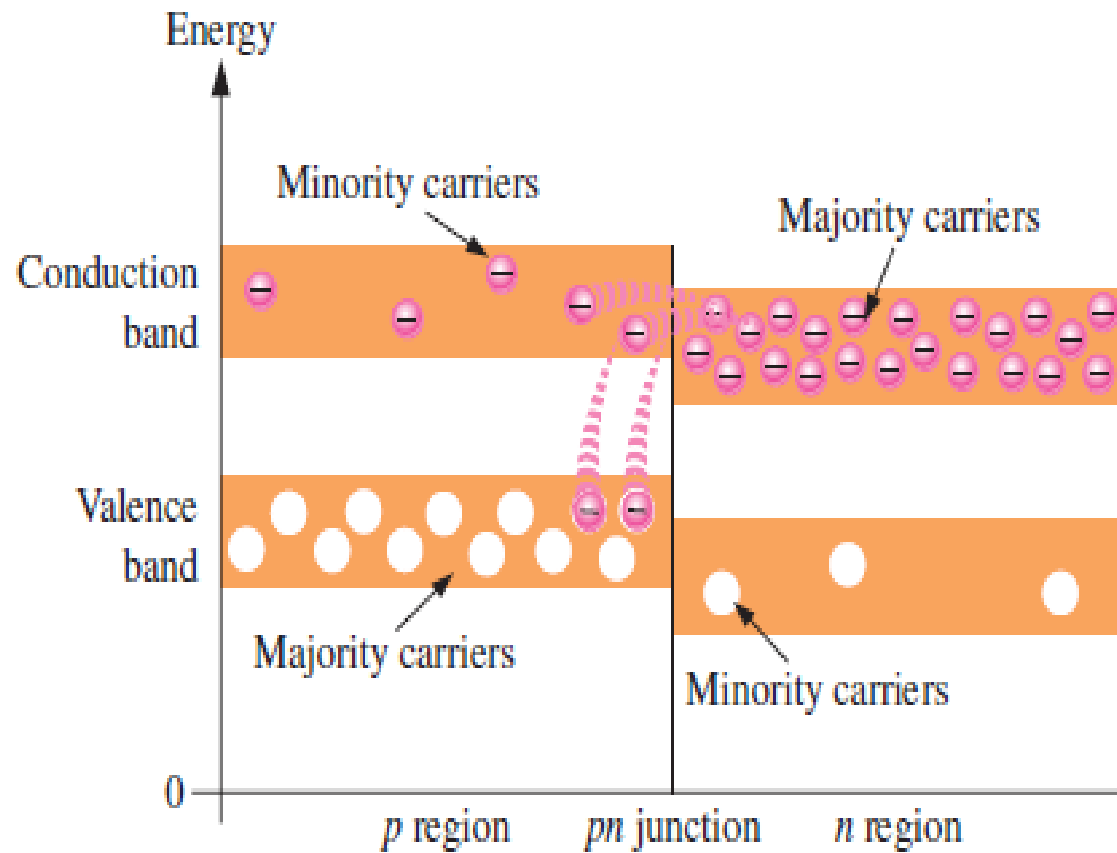
(a) At the instant of junction formation
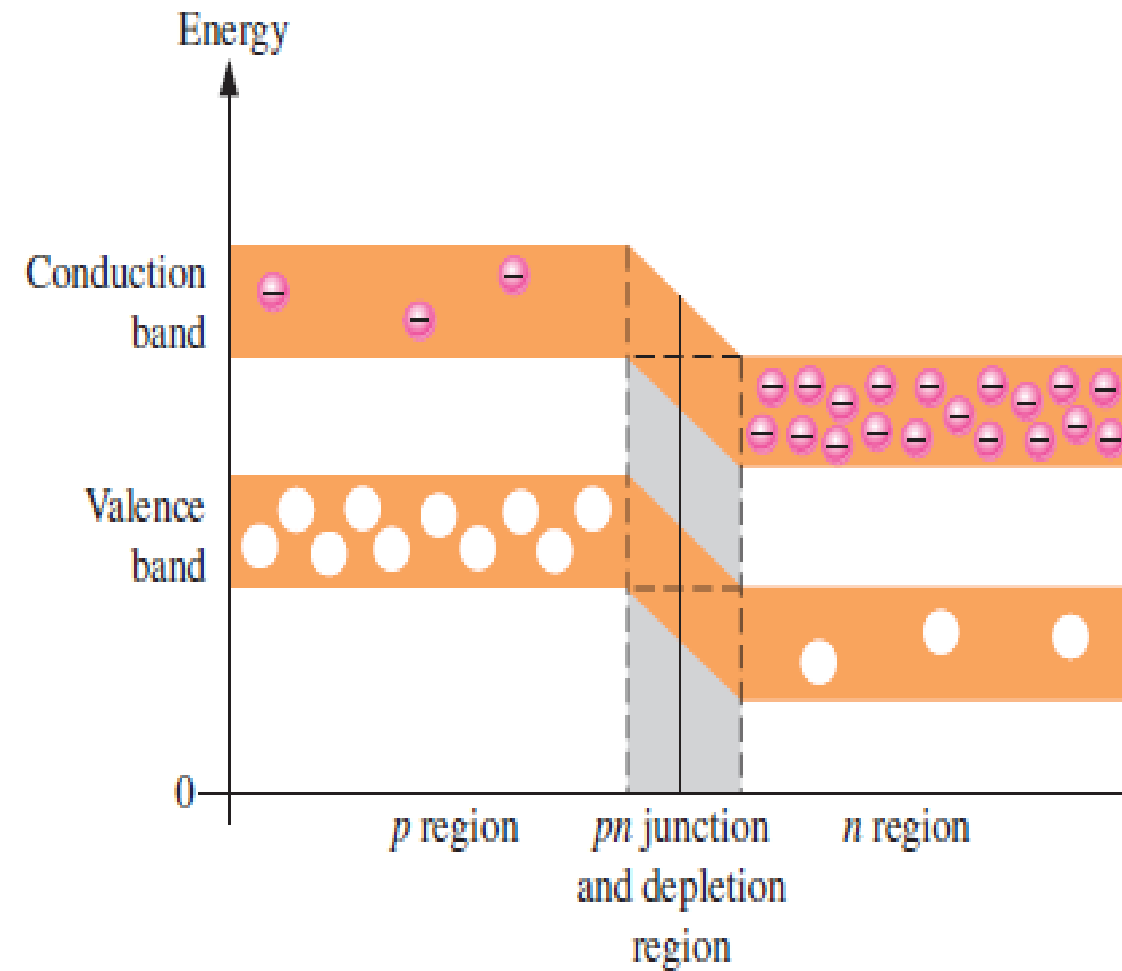
(b) At equilibrium

▲ FIGURE 1–20

Energy diagrams illustrating the formation of the *pn* junction and depletion region.

The decrease in the energy level of the conduction band in the n region is due to the loss of the higher-energy electrons that have diffused across the junction to the p region. Soon, there are no electrons left in the n-region conduction band with enough energy to get across the junction to the p-region conduction band, as indicated by the alignment of the top of the n-region conduction band and the bottom of the p-region conduction band in Figure 1–20(b).

At this point, the junction is at equilibrium; and the depletion region is complete because diffusion has ceased. There is an energy gradient across the depletion region which acts as an "energy hill" that an *n*-region electron must climb to get to the *p* region. Notice that as the energy level of the *n*-region conduction band has shifted downward, the energy level of the valence band has also shifted downward. It still takes the same amount of energy for a valence electron to become a free electron. In other words, the energy gap between the valence band and the conduction band remains the same.

DIODE

A diode is a 2 lead semiconductor that acts as a one way gate to electron flow. – Diode allows current to pass in only one direction.

A pn-junction diode is formed by joining together n-type and p-type silicon. In practice, as the n-type Si crystal is being grown, the process is abruptly altered to grow p-type Si crystal. Finally, a glass or plastic coating is placed around the joined crystal. The p-side is called anode and the n-side is called cathode. When the anode and cathode of a pn-junction diode are connected to external voltage such that the potential at anode is higher than the potential at cathode, the diode is said to be forward biased.

–In a forward-biased diode current is allowed to flow through the device. When potential at anode is smaller than the potential at cathode, the diode is said to be reverse biased. In a reverse-biased diode current is blocked.

DIODE
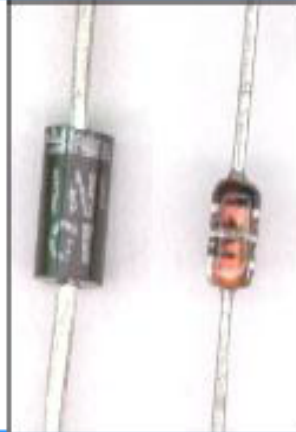
Symbol for DIODE
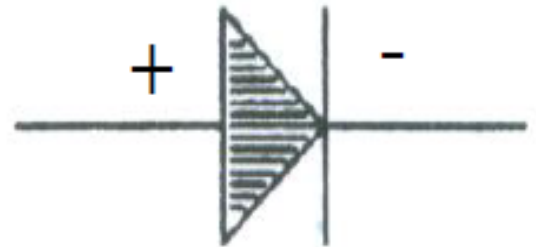
DIODE: HOW IT WORKS –1

When a diode is connected to a battery as shown, electrons from the n-side and holes from the p-side are forced toward the center by the electrical field supplied by the battery. The electrons and holes combine causing the current to pass through the diode. When a diode is arranged in this way, it is said to be forward biased.
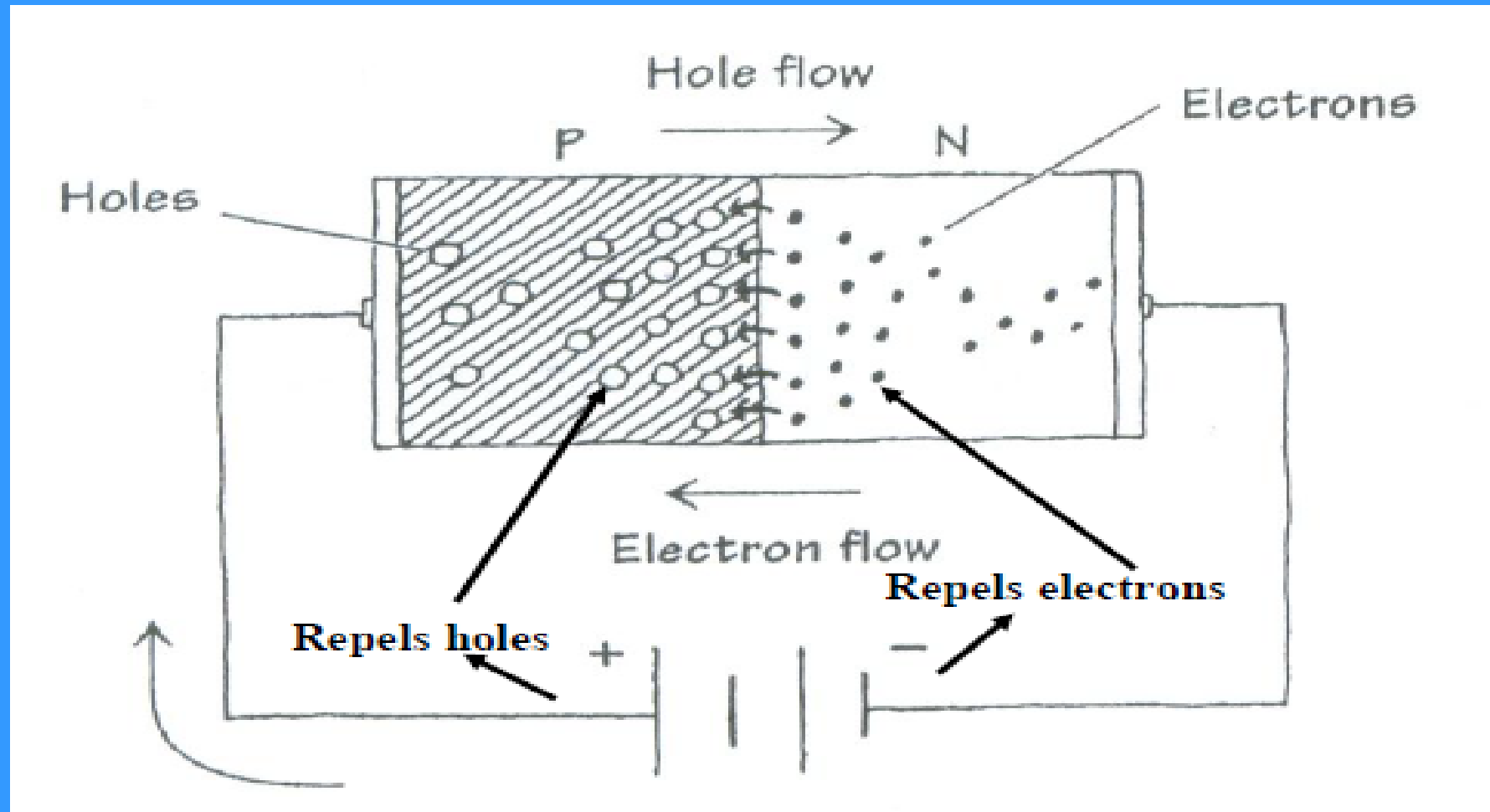
Forward-biased ("open door")

DIODE: HOW IT WORKS—II

A diode's one-way gate feature does not work all the time. Typically for silicon diodes, an applied voltage of 0.6V or greater is needed, otherwise, the diode will not conduct. This feature is useful in forming a voltage-sensitive switch. I-V characteristics for silicon and germanium diodes is shown below.
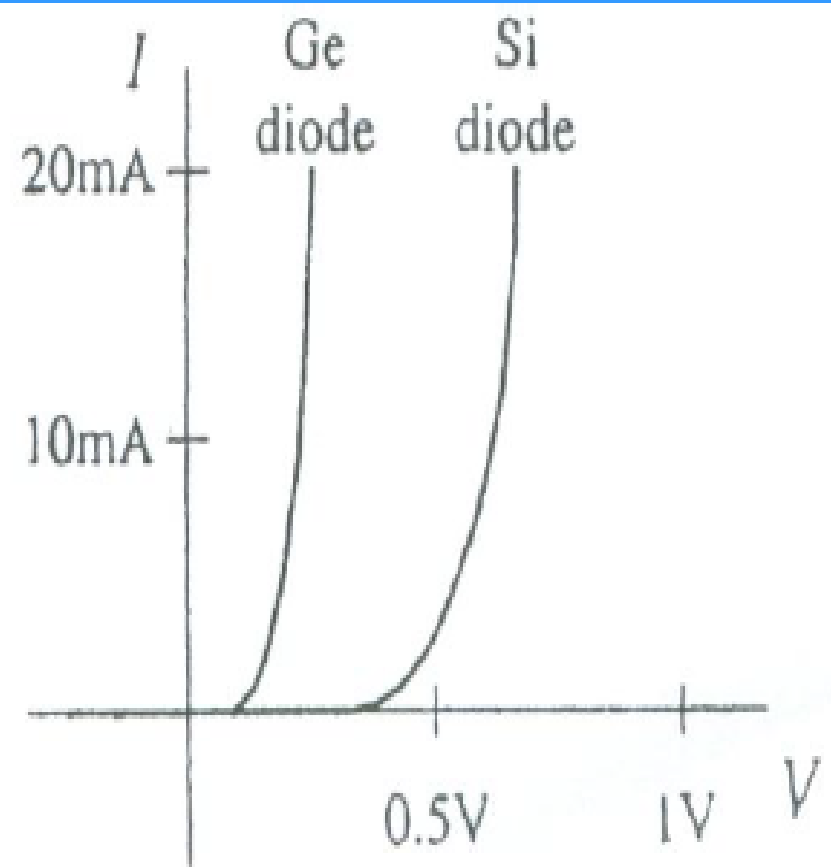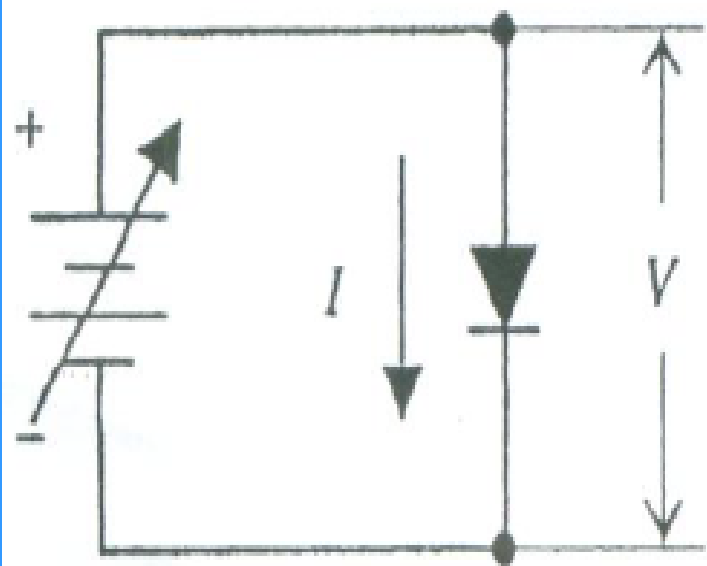
DIODE: HOW IT DOESN'T WORK

When a diode is connected to a battery as shown, holes in the *n* side are forced to the left while electrons in the *p*-side are forced to the right. This results in an empty zone around the pn- junction that is free of charge carries creating a **depletion region**. This depletion region acts as an insulator preventing current from flowing through the diode. When a diode is arranged in this way, it is said to be reverse biased

Reverse-biased ("closed door")

INTRODUCTION

The reader is familiar with the current conduction (*i.e.* flow of electrons) through a conductor. The valence electrons of the conductor atoms are loosely bound to the atomic nuclei. At room temperature, the thermal energy in the conductor is adequate to break the bonds of the valence electrons and leave them unattached to any one nucleus. These unbound electrons move at random within the conductor and are known as *free electrons*. If an electric field is applied across the conductor, these free electrons move through the conductor in an orderly manner, thus constituting electric current. This is how these free electrons move through the conductor or electric current flows through a wire.

Many electronic devices depend for their operation on the movement of electrons in an evacuated space. For this purpose, the free electrons must be ejected from the surface of metallic con-
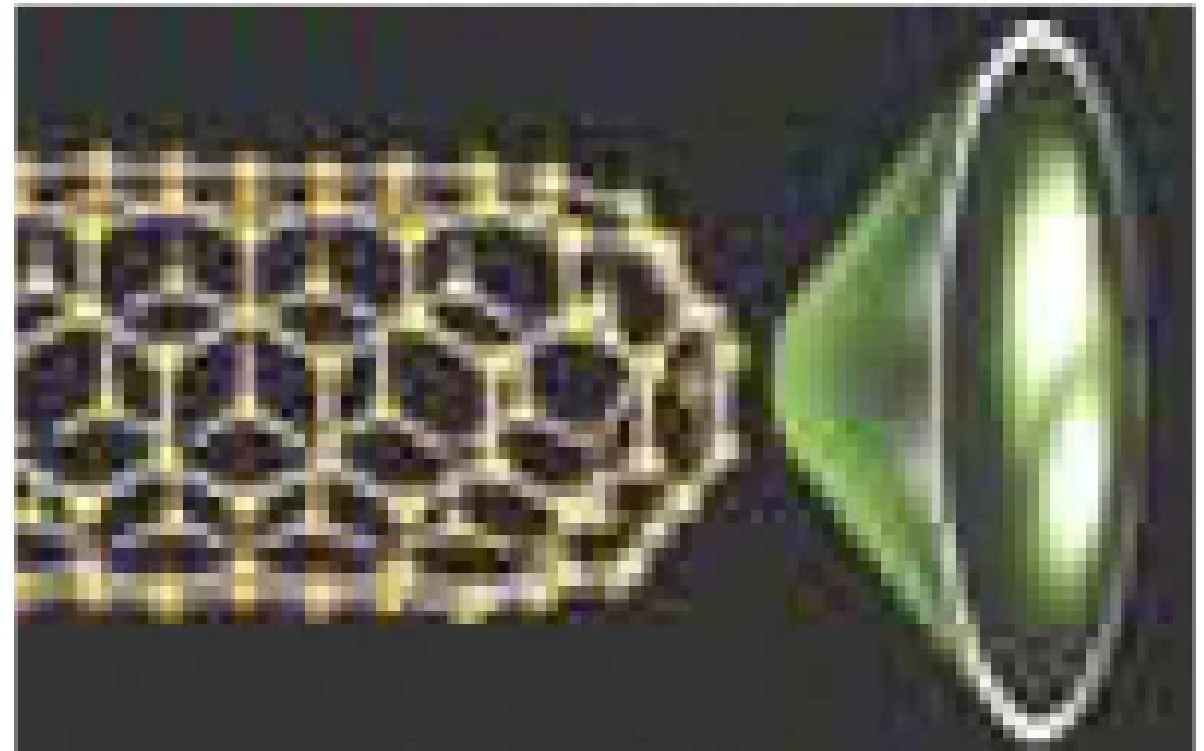
Electron Emission

ductor by supplying sufficient energy from some external source. This is known as *electron emission*. The emitted electrons can be made to move in vacuum under the influence of an electric field, thus constituting electric current in vacuum. In this chapter, we shall confine our attention to the various aspects of electron emission.

2.1   Electron Emission

The liberation of electrons from the surface of a substance is known as **electron emission.**

For electron emission, metals are used because they have many free electrons. If a piece of metal is investigated at room temperature, the random motion of free electrons is as shown in Fig. 2.1. However, these electrons are free only to the extent that they may transfer from one atom to another within the metal but they cannot leave the metal surface to provide electron emission. It is because the free electrons that start at the surface of metal find behind them positive nuclei pulling them back and none pulling forward. Thus at the surface of a metal, a free electron encounters forces that prevent it to leave the metal. In other words, the metallic surface offers a barrier to free electrons and is known as *surface barrier*.
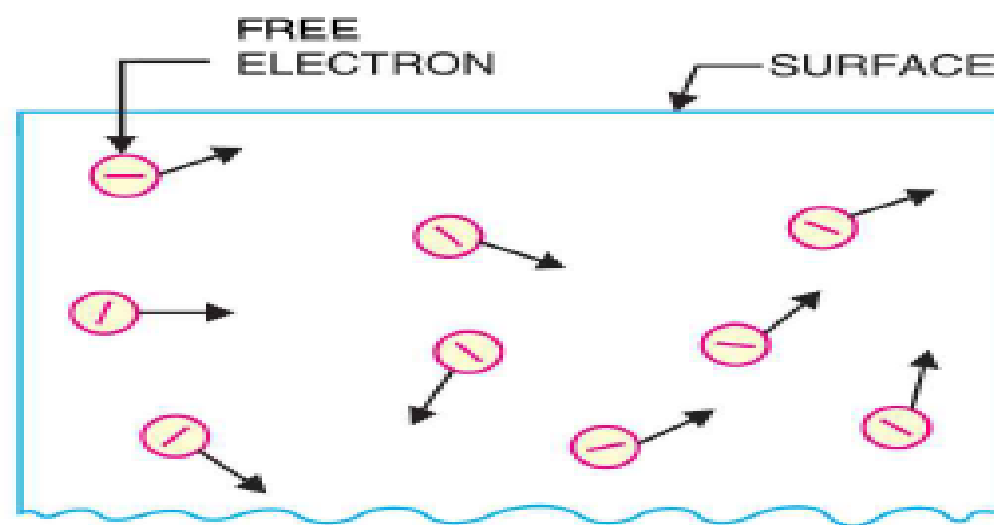
Fig. 2.1

However, if sufficient external energy is given to the free electron, its kinetic energy is increased and thus electron will cross over the surface barrier to leave the metal. This additional energy required by an electron to overcome the surface barrier of the metal is called *work function* of the metal.

The amount of additional energy required to emit an electron from a metallic surface is known as **work function** *of that metal.*

Thus, if the total energy required to liberate an electron from a metal is 4 eV* and the energy already possessed by the electron is 0.5 eV, then additional energy required (*i.e.*, work function) is 4.0 – 0.5 = 3.5 eV. The work function of pure metals varies roughly from 2 to 6 eV. It depends upon the nature of metal, its purity and the conditions of its surface. It may be noted that it is desirable that metal used for electron emission should have low work function so that a small amount of energy is required to cause emission of electrons.

2.2 Types of Electron Emission

The electron emission from the surface of a metal is possible only if sufficient additional energy (*equal to the work function of the metal*) is supplied from some external source. This external energy may come from a variety of sources such as heat energy, energy stored in electric field, light energy or kinetic energy of the electric charges bombarding the metal surface. Accordingly, there are following four principal methods of obtaining electron emission from the surface of a metal :

* Work function is the additional energy required for the liberation of electrons. Therefore, it should have the conventional unit of energy *i.e.* joules. But this unit is very large for computing electronics work. Therefore, in practice, a smaller unit called *electron volt* (abbreviated as eV) is used.

One electron-volt is the amount of energy acquired by an electron when it is accelerated through a potential difference of 1 V.

Thus, if an electron moves from a point of 0 potential to a point of +10V, then amount of energy acquired by the electron is 10 eV.

Since charge on an electron $= 1.602 \times 10^{-19}$ C and voltage $= 1$ V,

$$\therefore \quad 1 \text{ electron-volt} = Q V = (1.602 \times 10^{-19}) \times 1 \text{ J}$$

$$or \quad\quad\quad 1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$$

(i) **Thermionic emission.** In this method, the metal is heated to sufficient temperature (about 2500°C) to enable the free electrons to leave the metal surface. The number of electrons emitted depends upon the temperature. The higher the temperature, the greater is the emission of electrons. This type of emission is employed in vacuum tubes.

(ii) **Field emission.** In this method, a strong electric field (*i.e.* a high positive voltage) is applied at the metal surface which pulls the free electrons out of metal because of the attraction of positive field. The stronger the electric field, the greater is the electron emission.

(iii) **Photo-electric emission.** In this method, the energy of light falling upon the metal surface is transferred to the free electrons within the metal to enable them to leave the surface. The greater the intensity (*i.e.* brightness) of light beam falling on the metal surface, the greater is the photo-electric emission.

(iv) **Secondary emission.** In this method, a high velocity beam of electrons strikes the metal surface and causes the free electrons of the metal to be knocked out from the surface.

2.3 Thermionic Emission

The process of electron emission from a metal surface by supplying thermal energy to it is known as **thermionic emission**.

At ordinary temperatures, the energy possessed by free electrons in the metal is inadequate to cause them to escape from the surface. When heat is applied to the metal, some of heat energy is converted into kinetic energy, causing accelerated motion of free electrons. When the temperature rises sufficiently, these electrons acquire additional energy equal to the work function of the metal. Consequently, they overcome the restraining surface barrier and leave the metal surface.

Metals *with lower work function* will require less additional energy and, therefore, will emit electrons at lower temperatures. The commonly used materials for electron emission are *tungsten, thoriated tungsten* and *metallic oxides of barium and strontium*. It may be added here that high temperatures are necessary to cause thermionic emission. For example, pure tungsten must be heated to about 2300°C to get electron emission. However, oxide coated emitters need only 750°C to cause thermionic emission.

Richardson-Dushman equation. The amount of thermionic emission increases rapidly as the emitter temperature is raised. The emission current density is given by Richardson-Dushman equation given below :

$$J_s = A T^2 e^{-\frac{b}{T}} \text{ amp/m}^2 \qquad \qquad ...(i)$$

where J_s = emission current density *i.e.* current per square metre of the emitting surface

  T = absolute temperature of emitter in K

  A = constant, depending upon the type of emitter and is measured in amp/m^2/K^2

  b = a constant for the emitter

  e = natural logarithmic base

The value of b is constant for a metal and is given by :

$$b = \frac{\phi e}{k}$$

where ϕ = work function of emitter

  e = electron charge = 1.602×10^{-19} coulomb

  k = Boltzmann's constant = 1.38×10^{-23} J/K

$\therefore \qquad \qquad b = \dfrac{\phi \times 1.602 \times 10^{-19}}{1.38 \times 10^{-23}} = 11600 \, \phi \text{ K}$

Putting the value of b in exp. (i), we get,

$$J_s = A T^2 e^{-\frac{11600\,\phi}{T}} \qquad \qquad ...(ii)$$

The following points may be noted from eqn. (ii) :

(i) The emission is markedly affected by temperature changes. Doubling the temperature of an emitter may increase electron emission by more than 10^7 times. For instance, emission from pure tungsten metal is about 10^{-6} ampere per sq. cm. at 1300°C but rises to enormous value of about 100 amperes when temperature is raised to 2900°C.

(ii) Small changes in the work function of the emitter can produce enormous effects on emission. Halving the work function has exactly the same effect as doubling the temperature.

Example 2.1. *A tungsten filament consists of a cylindrical cathode 5 cm long and 0.01 cm in diameter. If the operating temperature is 2500 K, find the emission current. Given that $A = 60.2 \times 10^4$ A/m^2/K^2, $\phi = 4.517$ eV.*

Solution. $\qquad\qquad A = 60.2 \times 10^4$ amp/m^2/K^2, $T = 2500$ K, $\phi = 4.517$ eV

$\therefore \qquad\qquad\qquad b = 11600\,\phi\,\mathrm{K} = 11600 \times 4.517\ \mathrm{K} = 52400$ K

Using Richardson-Dushman equation, emission current density is given by :

$$J_s = A T^2 e^{-\frac{b}{T}} \ \text{amp/m}^2 = 60.2 \times 10^4 \times (2500)^2 \times (2.718)^{-\frac{52400}{2500}}$$

$$= 0.3 \times 10^4 \ \text{amp/m}^2$$

Surface area of cathode, $a = \pi d l = 3.146 \times 0.01 \times 5 = 0.157$ cm$^2 = 0.157 \times 10^{-4}$ m^2

$\therefore \qquad$ Emission current $= J_s \times a = (0.3 \times 10^4) \times (0.157 \times 10^{-4}) = \mathbf{0.047\ A}$

2.4 Thermionic Emitter

The substance used for electron emission is known as an *emitter* or *cathode*. The cathode is heated in an evacuated space to emit electrons. If the cathode were heated to the required temperature in open air, it would burn up because of the presence of oxygen in the air. A cathode should have the following properties:

(*i*) **Low work function.** The substance selected as cathode should have low work function so that electron emission takes place by applying small amount of heat energy *i.e.* at low temperatures.

(*ii*) **High melting point.** As electron emission takes place at very high temperatures (>1500°C), therefore, the substance used as a cathode should have high melting point. For a material such as copper, which has the advantage of a low work function, it is seen that it cannot be used as a cathode because it melts at 810°C. Consequently, it will vaporise before it begins to emit electrons.

Thermionic Emitter

(*iii*) **High mechanical strength.** The emitter should have high mechanical strength to withstand the bombardment of positive ions. In any vacuum tube, no matter how careful the evacuation, there are always present some gas molecules which may form ions by impact with electrons when current flows. Under the influence of electric field, the positive ions strike the cathode. If high voltages are used, the cathode is subjected to considerable bombardment and may be damaged.

2.6 Cathode Construction

As cathode is sealed in vacuum, therefore, the most convenient way to heat it is electrically. On this basis, the thermionic cathodes are divided into two types viz directly heated cathode and indirectly heated cathode.

(*i*) **Directly heated cathode.** In this type, the cathode consists of oxide-coated nickel ribbon, called the *filament. The heating current is directly passed through this ribbon which emits the electrons. Fig. 2.2 (*i*) shows the structure of directly heated cathode whereas Fig. 2.2 (*ii*) shows its symbol.

FILAMENT
GLASS STEM
DIRECTLY HEATED CATHODE
(*i*)

GLASS ENVELOPE
FILAMENT
SYMBOL
(*ii*)
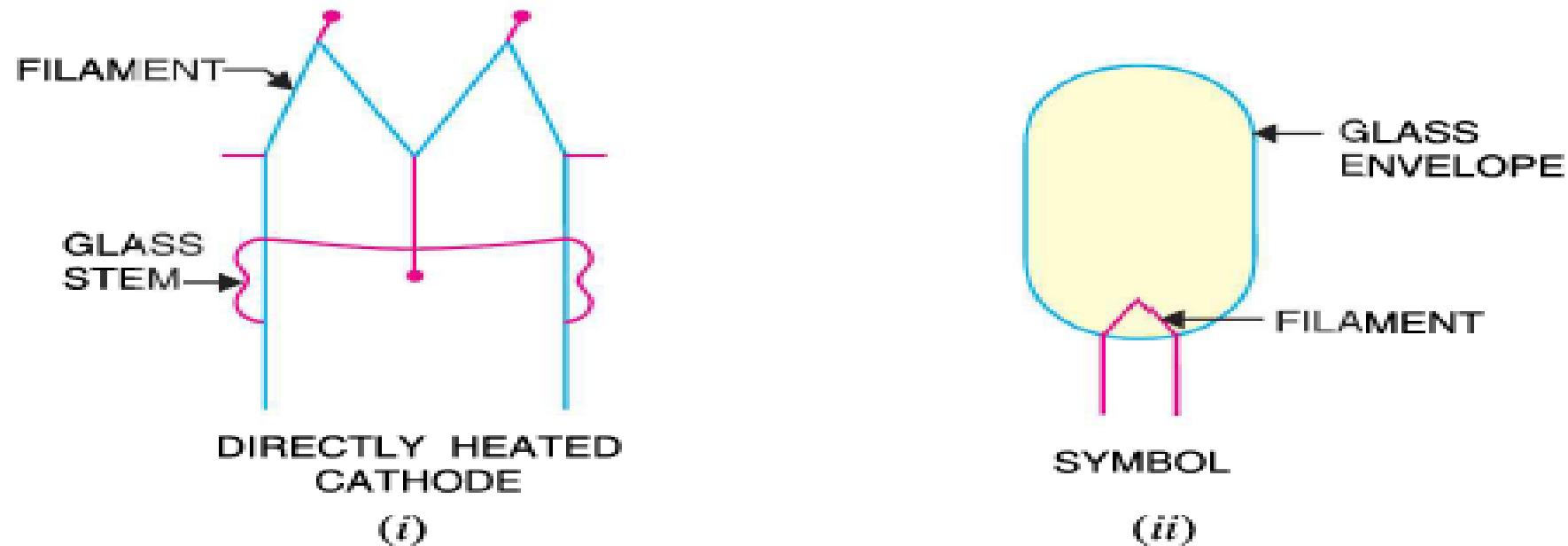
Fig. 2.2

The directly heated cathode is more efficient in converting heating power into thermionic emission. Therefore, it is generally used in power tubes that need large amounts of emission and in small tubes operated from batteries where efficiency and quick heating are important. The principal limitation of this type of cathode is that any variation in heater voltage affects the electron emission and thus produces *hum* in the circuit.

(*ii*) **Indirectly heated cathode.** In this type, the cathode consists of a thin metal sleeve coated with barium and strontium oxides. A filament or heater is enclosed within the sleeve and insulated from it. There is no electrical connection between the heater and the cathode. The heating current is passed through the heater and the cathode is heated indirectly through heat transfer from the heater element. Fig. 2.3 (*i*) shows the structure of indirectly heated cathode whereas Fig. 2.3 (*ii*) shows its symbol.

- -

* *Filament.* The term filament (literally means a thin wire) denotes the element through which the cathode heating current flows. In case of directly heated, cathode is itself the filament. If indirectly heated, heater is the filament.
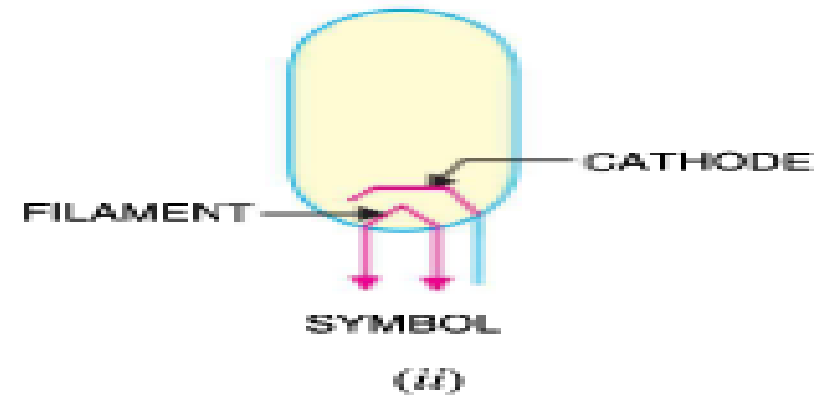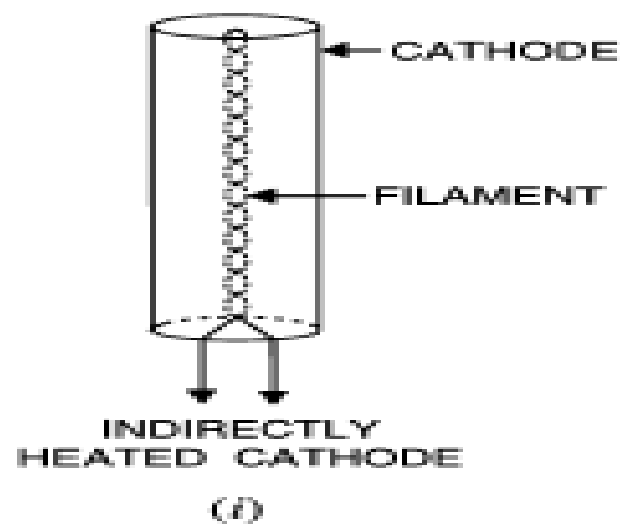
Fig. 2.3

Indirectly heated cathode has many advantages. As cathode is completely separated from the heating circuit, therefore, it can be readily connected to any desired potential as needed, independent of the heater potential. Furthermore, because of relatively large mass of cylindrical cathode, it takes time to heat or cool and as such does not introduce hum due to heater voltage fluctuations. Finally, a.c. can be used in the heater circuit to simplify the power requirements. Almost all modern receiving tubes use this type of cathode.

2.7 Field Emission

The process of electron emission by the application of strong electric field at the surface of a metal is known as **field emission.**

When a metal surface is placed close to a high voltage conductor which is positive w.r.t. the metal surface, the electric field exerts attractive force on the free electrons in the metal. If the positive potential is great enough, it succeeds in overcoming the restraining forces of the metal surface and the free electrons will be emitted from the metal surface as shown in Fig. 2.4.
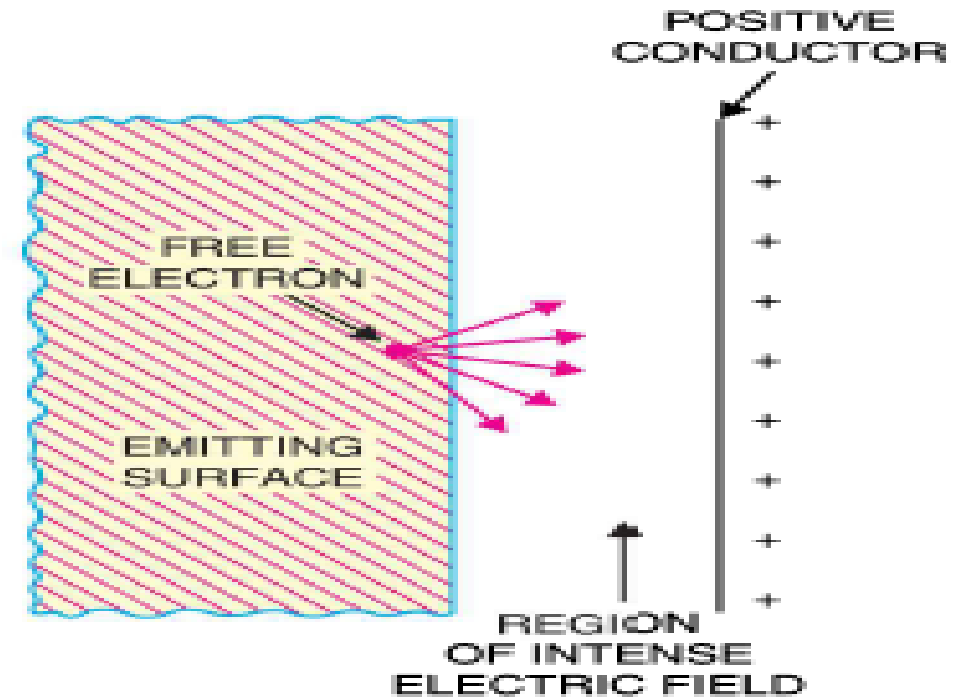
Fig. 2.4

Very intense electric field is required to produce field emission. Usually, a voltage of the order of a million volts per centimetre distance between the emitting surface and the positive conductor is necessary to cause field emission. Field emission can be obtained at temperatures much lower (*e.g.* room temperature) than required for thermionic emission and, therefore, it is also sometimes called *cold cathode emission* or *auto- electronic emission.*

2.8 Secondary Emission

Electron emission from a metallic surface by the bombardment of high-speed electrons or other particles is known as **secondary emission.**

When high-speed electrons suddenly strike a metallic surface, they may give some or all of their kinetic energy to the free electrons in the metal. If the energy of the striking electrons is sufficient, it may cause free electrons to escape from the metal surface. This phenomenon is called *secondary emission*. The electrons that strike the metal are called *primary electrons* while the emitted electrons are known as *secondary electrons*. The intensity of secondary emission depends upon the emitter material, mass and energy of the bombarding particles.

The principle of secondary emission is illustrated in Fig. 2.5. An evacuated glass envelope contains an emitting surface E, the collecting anode A and a source of primary electrons S. The anode is maintained at positive potential w.r.t. the emitting surface by battery B. When the primary electrons strike the emitting surface E, they knock out secondary electrons which are attracted to the anode and constitute a flow of current. This current may be measured by connecting a sensitive galvanometer G in the anode circuit.

The effects of secondary emission are very undesirable in many electronic devices. For example, in a tetrode valve, secondary emission is responsible for the negative resistance. In some electronic devices, however, secondary emission effects are utilised e.g. *electron multiplier, cathode ray tube etc.

2.9 Photo Electric Emission

Electron emission from a metallic surface by the application of light is known as **photo electric emission.**

When a beam of light strikes the surface of certain metals (*e.g.* potassium, sodium, cesium), the energy of photons of light is transferred to the free electrons within the metal. If the energy of the
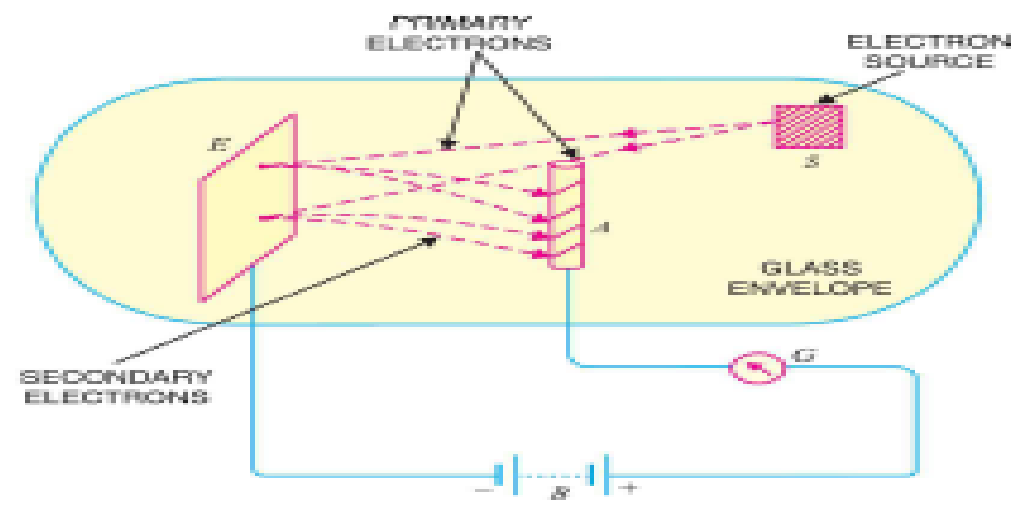
Fig. 2.5
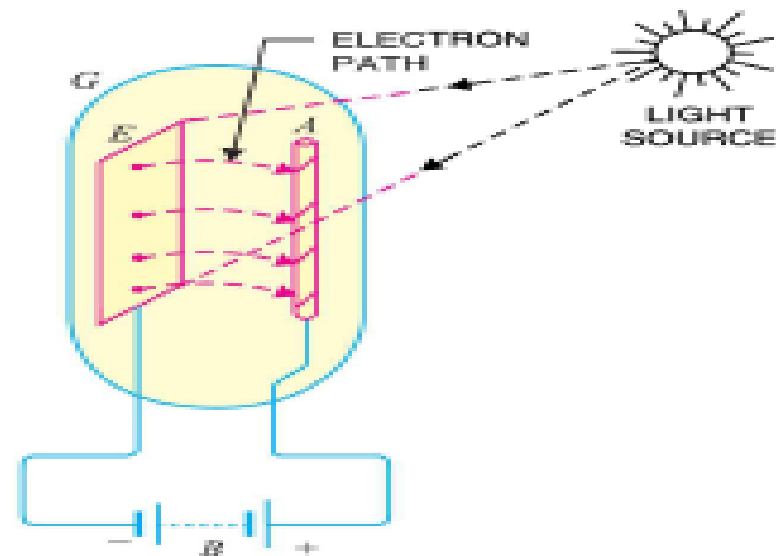
Fig. 2.6

---

* An interesting aspect of secondary emission is that a high-speed bombarding electron may liberate so many as 10 "secondary electrons". This amounts to a multiplication of electron flow by a ratio as great as 10 and is utilised in current multiplier devices.

striking photons is greater than the work function of the metal, then free electrons will be knocked out from the surface of the metal. The emitted electrons are known as *photo electrons* and the phenomenon is known as *photoelectric emission*. The amount of photoelectric emission depends upon the intensity of light falling upon the emitter and frequency of radiations. The greater the intensity and frequency of radiations, the greater is the photo electric emission. Photo-electric emission is utilised in photo tubes which form the basis of television and sound films.
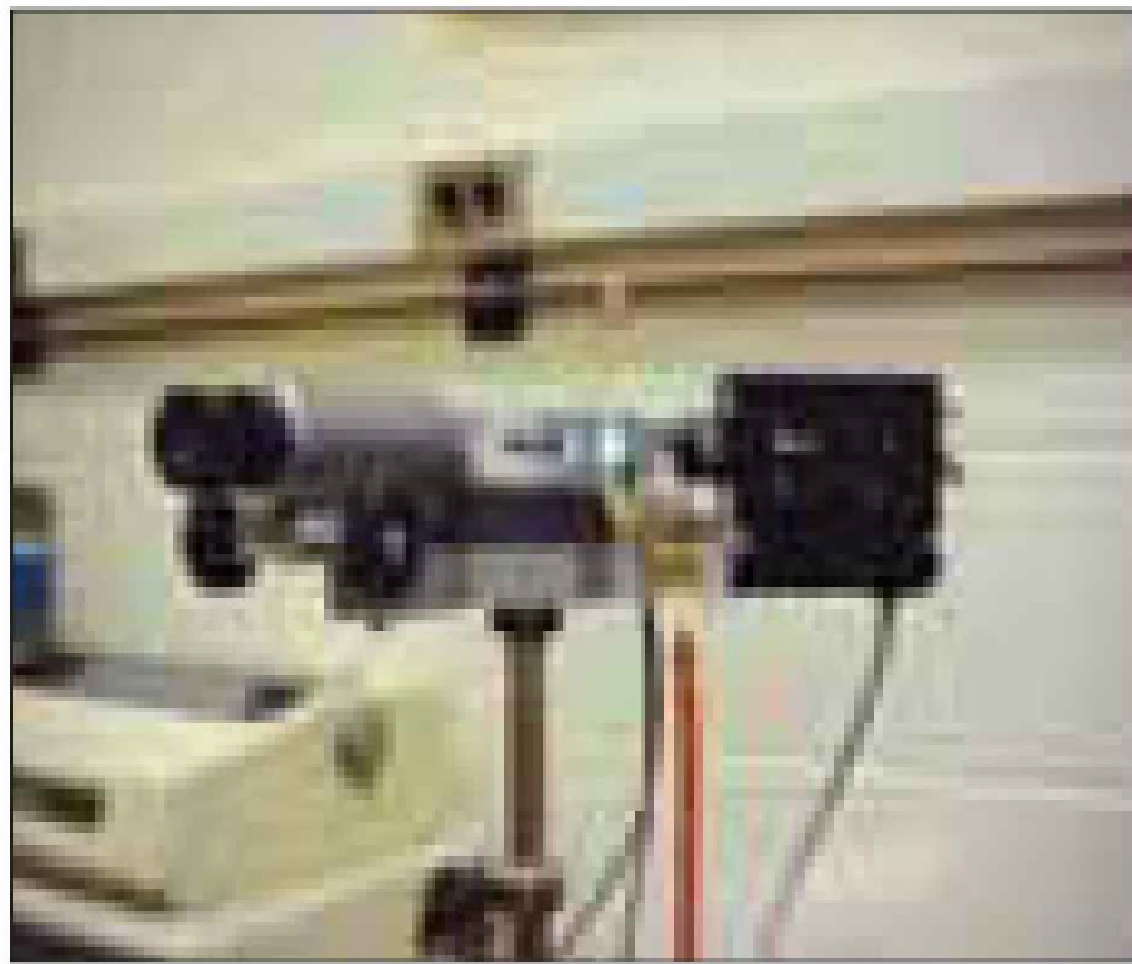
Photo Tube

Fig. 2.6 illustrates the phenomenon of photoelectric emission. The emitter *E* and anode *A* are enclosed in an evacuated glass envelope *G*. A battery *B* maintains the anode at positive potential w.r.t. emitter. When light of suitable intensity and frequency falls on the emitter, electrons are ejected from its surface. These electrons are attracted by the positive anode to constitute current in the circuit. It may be noted that current will exist in the circuit so long as illumination is maintained.

PHOTORESISTORS

- Light sensitive variable resistors.
- Its resistance depends on the intensity of light incident upon it.
  - Under dark condition, resistance is quite high (MΩ: called dark resistance).
  - Under bright condition, resistance is lowered (few hundred Ω).
- Response time:
  - When a photoresistor is exposed to light, it takes a few milliseconds, before it lowers its resistance.
  - When a photoresistor experiences removal of light, it may take a few seconds to return to its dark resistance.

Symbol
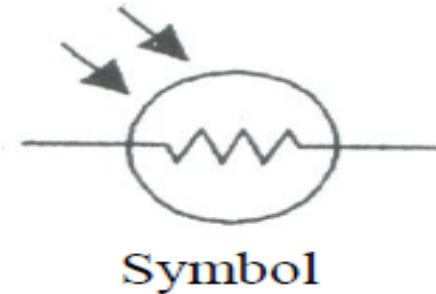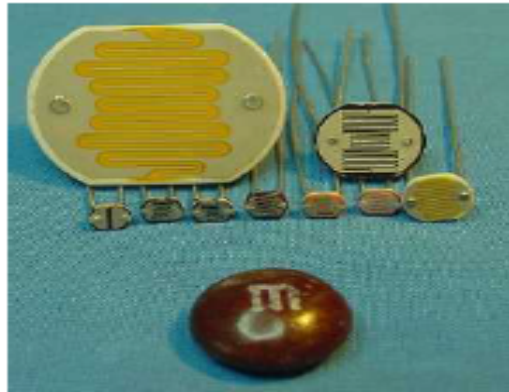
Photoresistors —II

- Some photoresistors respond better to light that contains photons within a particular wavelength of spectrum.
    - Example: Cadmium-sulfide photoresistos respond to light within 400-800nm range.
    - Example: Lead-sulfide photoresistos respond to infrared light.

How Photoresistor Works

- Special semiconductor crystal, such as cadmium sulfide or lead sulfide is used to make photoresistors.

- When this semiconductor is placed in dark, electrons within its structure resist flow through the resistor because they are too strongly bound to the crystal's atoms.

- When this semiconductor is illuminated, incoming photons of light collide with the bound electrons, stripping them from the binding atom, thus creating holes in the process.

- Liberated electrons contribute to the current flowing through the device.
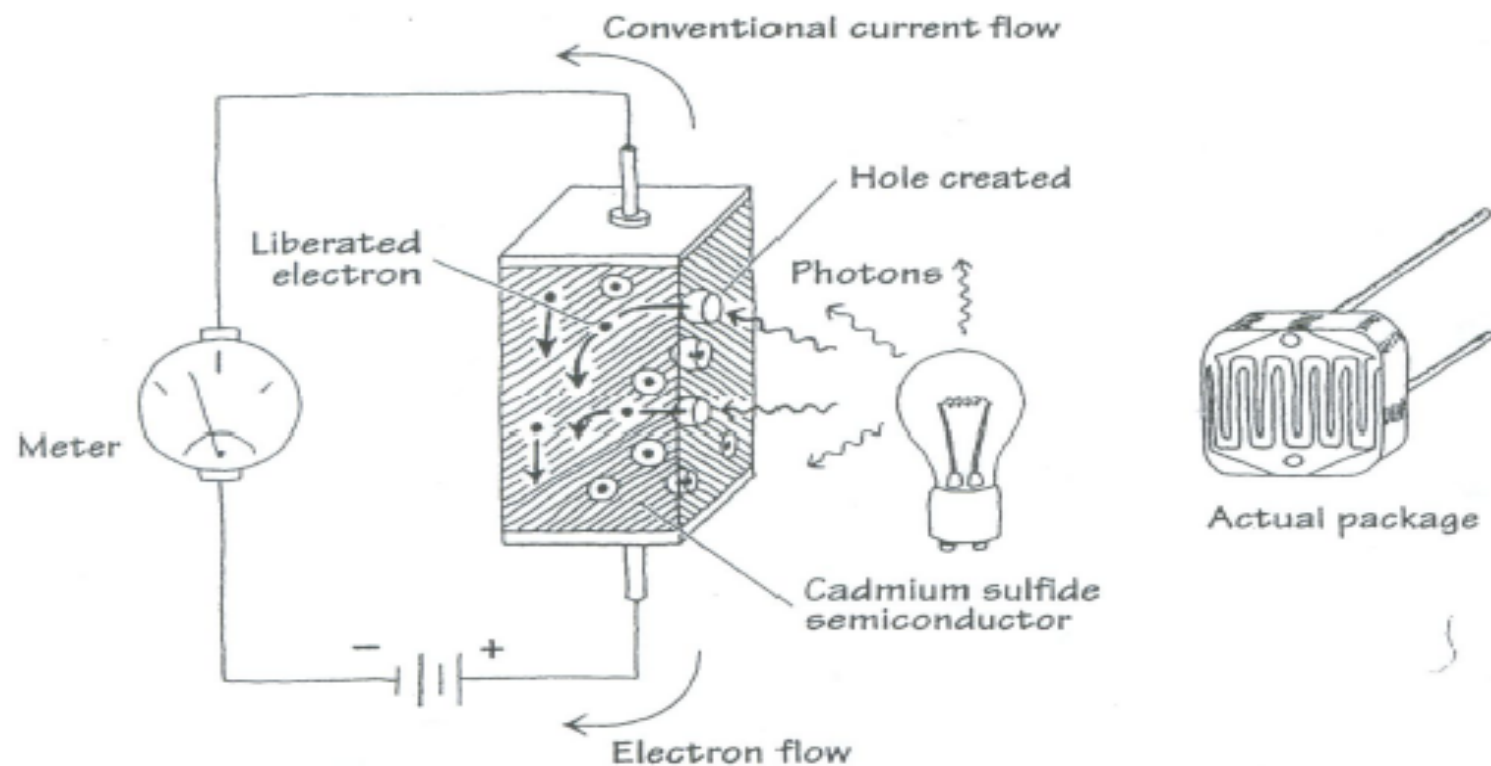
Photoresistor Application —Light Activated Relay

- Light-sensitive voltage divider is being used to trip a relay whenever the light intensity change.
- Light-activated circuit:
  - When the photoresistor is exposed to light, its resistance decreases.
  - $\rightarrow$ Transistor's base current and voltage increase and if the base current and voltage are large enough, the collector-emitter pair of the transistor conducts triggering the relay.
- The value of R_1 in the light-activated circuit should be around 1 KΩ but may have to be adjusted.
- Dark-activated relay works in a similar but opposite manner.
- R_1 in the dark-activated circuit (100KΩ) may also have to be adjusted.
- A 6 to 9-V relay with a 500Ω coil can be used in either circuit.

Light-Activated Relay
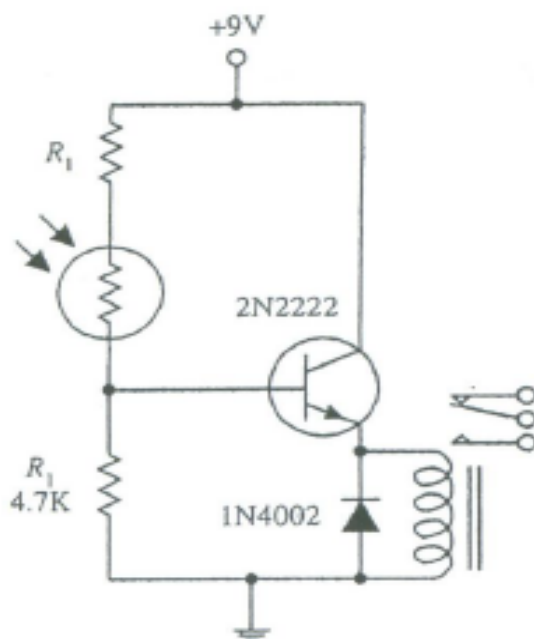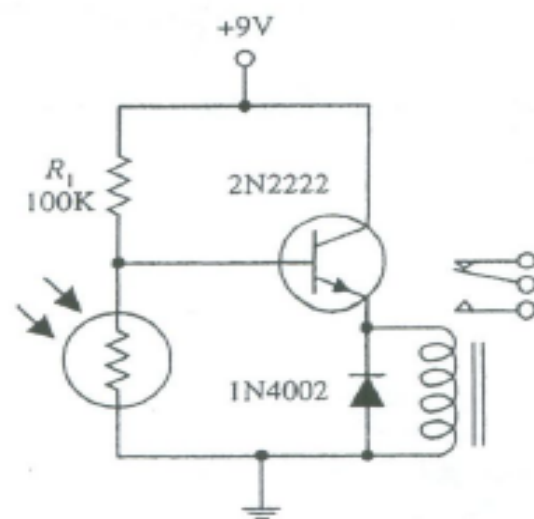
Light activated relay   Dark activated relay

PHOTODIODE

- Photodiode is a 2 lead semiconductor device that transforms light energy to electric current.
- Suppose anode and cathode of a photodiode are wired to a current meter.
  - When photodiode is placed in dark, the current meter displays zero current flow.
  - When the photodiode is expose to light, it acts a a current source, causing current flow from cathode to anode of photodiode through the current meter.
- Photodiodes have very linear light v/s current characteristics.
  - Commonly used as light meters in cameras.
- Photodiodes often have built-in lenses and optical filters.
- Response time of a photodiode slows with increasing surface area.
- Photodiodes are more sensitive than photoresistor.
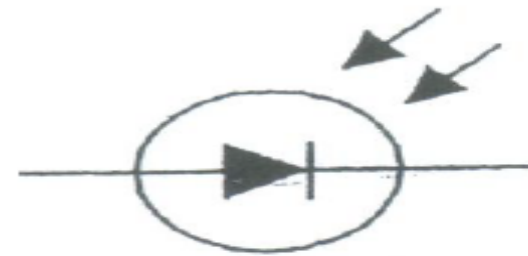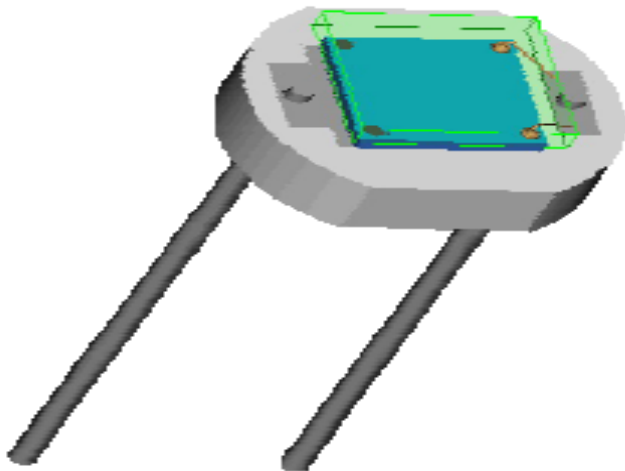
Symbol

How Photodiode Works

- Photodiode: A thin n-type semiconductor sandwiched with a thicker p-type semiconductor.

- N-side is cathode, p-side is anode.

- Upon illumination, a # of photons pass from the n-side and into the p-side of photodiode.

  - Some photons making it into p-side collide with bound electrons within p-semiconductor, ejecting them and creating holes.

  - If these collisions are close to the pn-interface, the ejected electrons cross the junction, yielding extra electrons on the n-side and extra holes on the p-side.

  - Segregation of +ve and -ve charges leads to a potential difference across the pn-junction.

  - When a wire is connected between the cathode and anode, a conventionally positive current flow from the anode to cathode

Very thin layer

Conventional current flow

Cathode   Anode

Electron flow
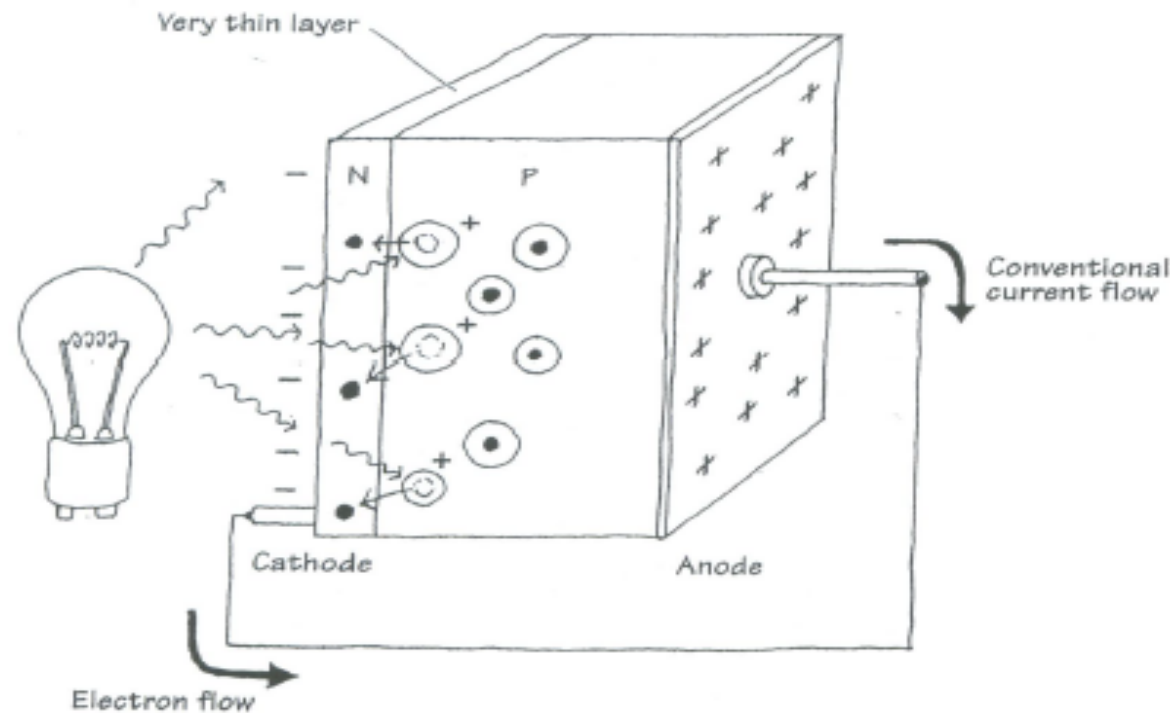
Photodiode Applications—Photovoltaic Current Source

- Photodiode converts light energy directly into electric current that can be measured with meter.
- The input intensity of light and the output current are nearly linear.
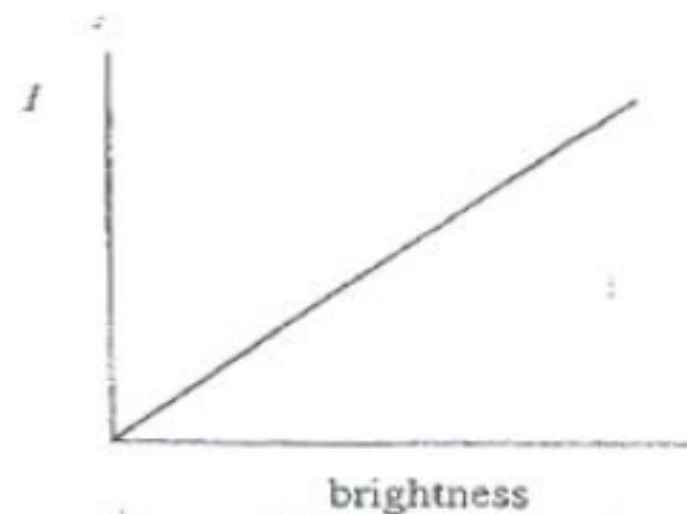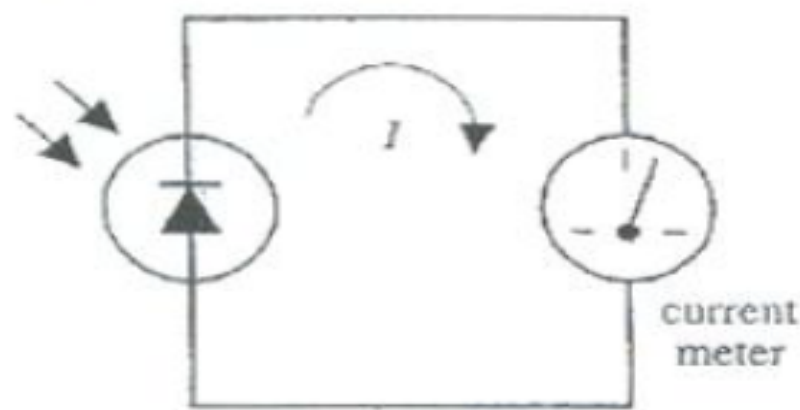
Photodiodes are used for detecting light. There are several types, and they are typically used in reverse bias. They are widely employed in communications applications.

LIGHT EMITTING DIODE (LED)

A light emitting diode (LED) is known to be one of the best optoelectronic devices. The device is capable of emitting a fairly narrow bandwidth of visible or invisible light when its internal diode junction attains a forward electric current or voltage. That is if you forward-biased an LED, current flows and light is produced. The visible lights that an LED emits are usually *orange, red, yellow, or green.* The invisible light includes the infrared light.

There are a variety of colours available. Developments in LED technology are fast finding their way into lighting and display applications, such as the backlighting of television screens. We know that a P-N junction can connect the absorbed light energy into its proportional electric current. The same process is reversed here. That is, the P-N junction emits light when energy is applied on it. Laser diodes also produce light, but a major difference is that the light produced by laser diodes is coherent. They are used in DVD and CD players and laser pointers.

This phenomenon is generally called electro luminance, which can be defined as the emission of light from a semi-conductor under the influence of an electric field. The charge carriers recombine in a forward P-N junction as the electrons cross from the N-region and recombine with the holes existing in the P-region. Free electrons are in the conduction band of energy levels, while holes are in the valence energy band. Thus the energy level of the holes will be less than the energy levels of the electrons. Some part of the energy must be dissipated in order to recombine the electrons and the holes. This energy is emitted in the form of heat and light.

LED Circuit Symbol

The circuit symbol of LED consists of two arrow marks which indicate the radiation emitted by the diode.

Anode   Cathode

The constructional diagram of a LED is shown below.

LED Construction

Emitted Light

Charge Carrier
Recombination

Diffused
P-type

Epitaxial
N-type

Gold Film Cathode
Connection

$+$
V_s
$-$
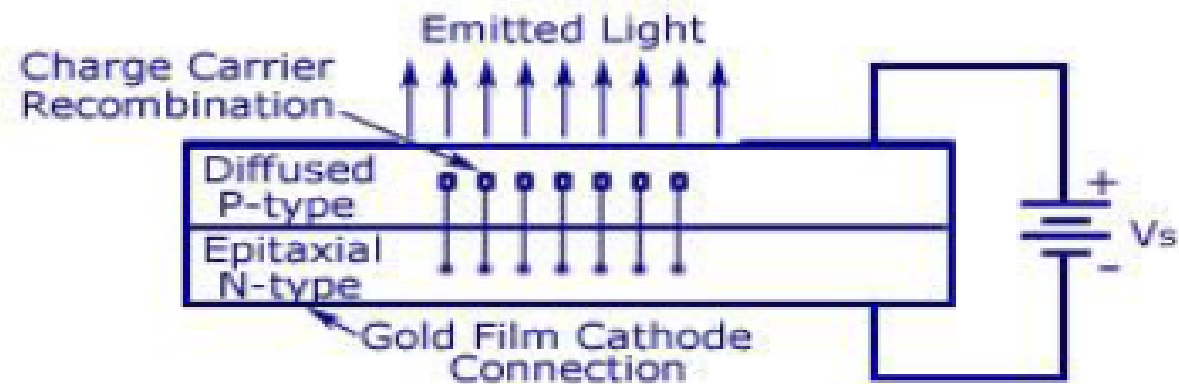
Fig. LED Construction

LED Characteristics

LED

Forward Bias V-I Curve
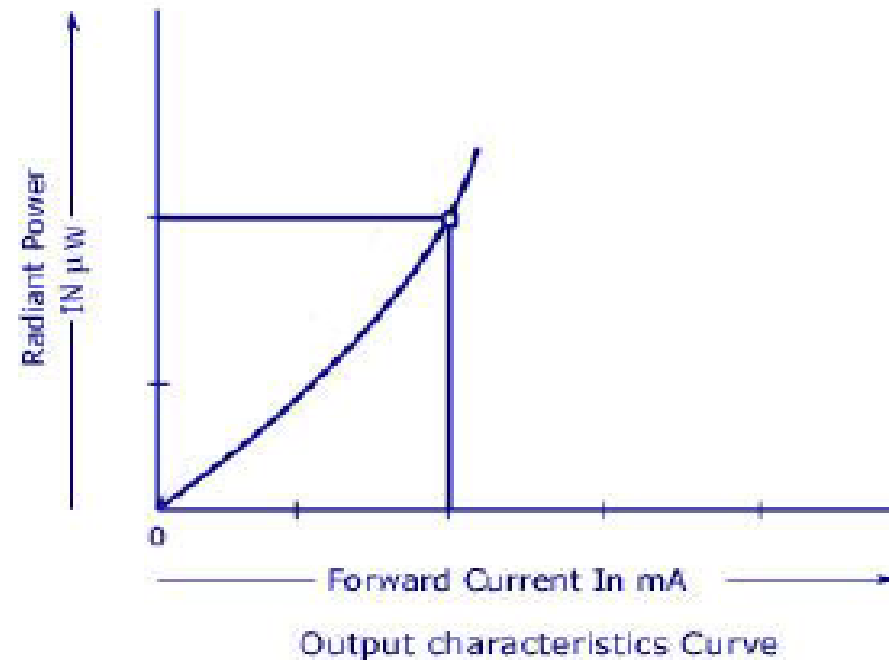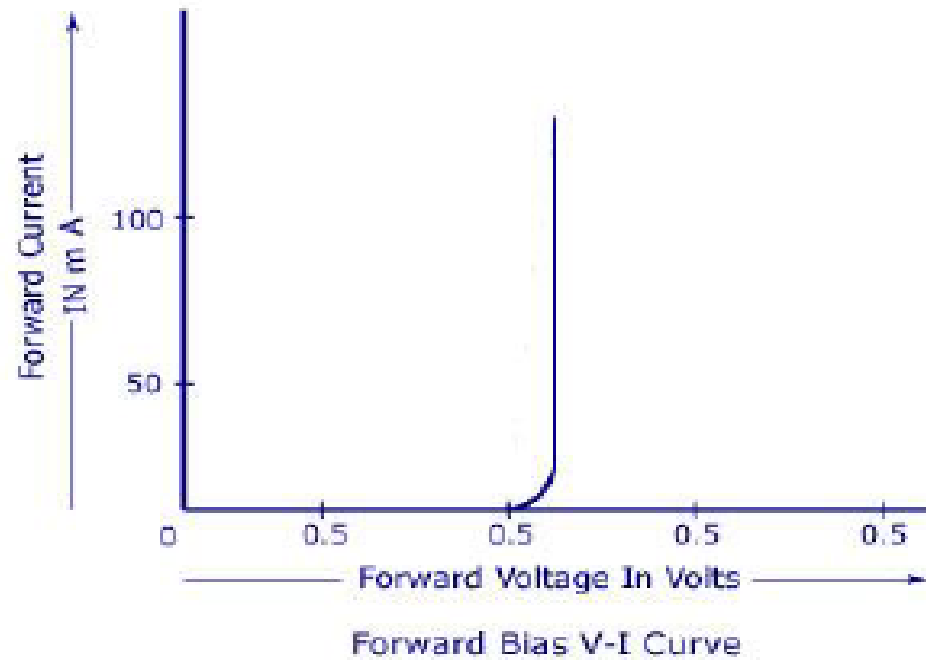
Output characteristics Curve

LED Characteristics

The forward bias Voltage-Current (V-I) curve and the output characteristics curve is shown in the figure above.

LED as an Indicator

The circuit shown below is one of the main applications of LED. The circuit is designed by wiring it in inverse parallel with a normal diode, to prevent the device from being reverse biased. The value of the series resistance should be half, relative to that o f a DC circuit.

LED As An Indicator
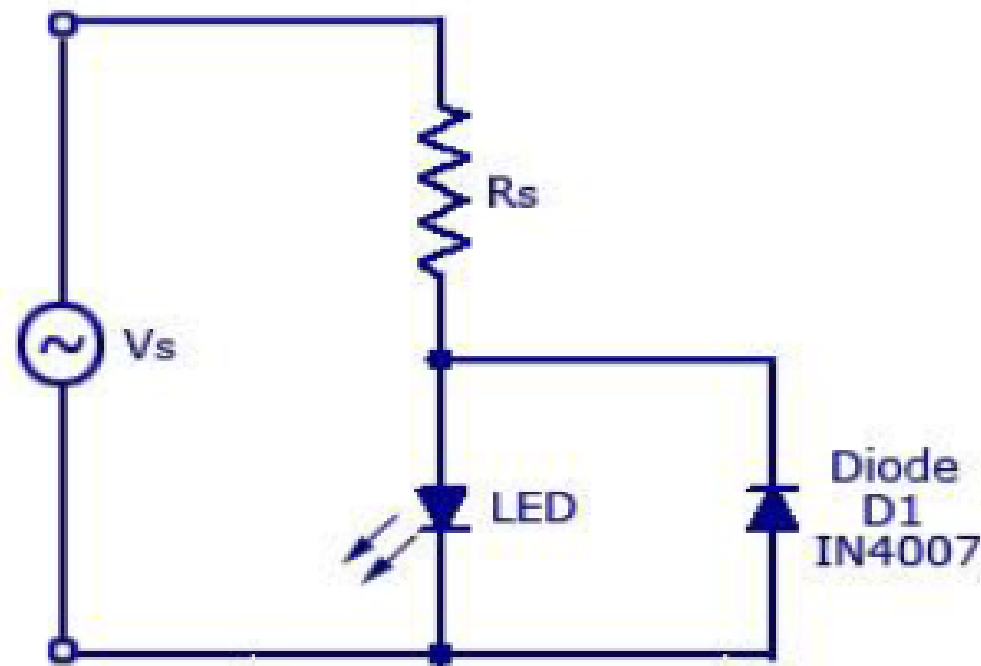
Advantages of LED

- Very low voltage and current are enough to drive the LED.
- Voltage range – 1 to 2 volts.
- Current – 5 to 20 milli amperes.
- Total power output will be less than 150 milli watts.
- The response time is very less – only about 10 nanoseconds.
- The device does not need any heating and warm up time.
- Miniature in size and hence light weight.
- Have a rugged construction and hence can withstand shock and vibrations.
- An LED has a life span of more than 20 years.

Disadvantages

- A slight excess in voltage or current can damage the device.
- The device is known to have a much wider bandwidth compared to the laser.
- The temperature depends on the radiant output power and wavelength.

PHOTODIODES

A **photodiode** is a semiconductor device that converts light into current. The current is generated when photons are absorbed in the photodiode. A small amount of current is also produced when no light is present. A *photodiode* is a diode optimized to produce an electron current flow in response to irradiation by ultraviolet, visible, or infrared light. Silicon is most often used to fabricate photodiodes; though, germanium and gallium arsenide can be used. The junction through which light enters the semiconductor must be thin enough to pass most of the light on to the active region (depletion region) where light is converted to electron hole pairs. In Figure 15 shallow P-type diffusion into an N-type wafer produces a PN junction near the surface of the wafer. The P-type layer needs to be thin to pass as much light as possible. A heavy N+ diffusion on the back of the wafer makes contact with metallization. The top metallization may be a fine grid of metallic fingers on the top of the wafer for large cells. In small photodiodes, the top contact might be a sole bond wire contacting the bare P-type silicon top

Symbol & Construction:

top metal contact

p diffusion

depletion region

n type

n+ contact region

bottom metal contact
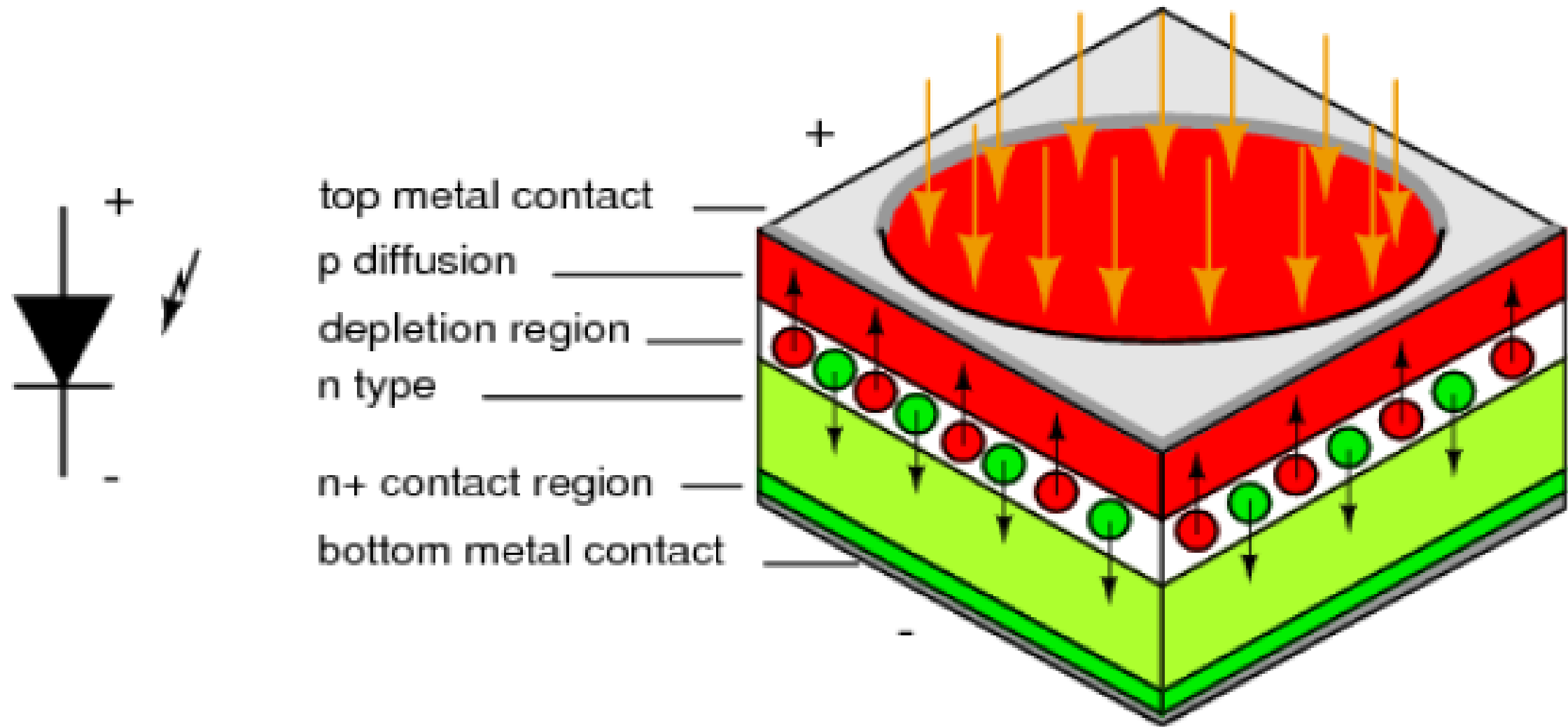
Fig 15 :Photodiode: Schematic symbol and cross section (below Photo diodes)

Applications:

Photodiodes are used in consumer electronics devices such as compact disc players, smoke detectors, and the receivers for infrared remote control devices used to control equipment from televisions to air conditioners. For many applications either photodiodes or photoconductors may be used. Either type of photosensor may be used for light measurement, as in camera light meters, or to respond to light levels, as in switching on street lighting after dark.

Photodiodes may contain optical filters, built-in lenses, and may have large or small surface areas. Photodiodes usually have a slower response time as its surface area increases. The common, traditional solar cell used to generate electric solar power is a large area photodiode.

Infrared LED

IR LED is used in this circuit to transmit infrared light.

An Infrared light-emitting diode (IR LED) is a type of electronic device that emits infrared light not visible to the naked eye.

The wavelength and colour of the light produced depend on the material used in the diode. Infrared LEDs use material that produces light in the infrared part of the spectrum, that is, just below what the human eye can see. Different infrared LEDs may produce infrared light of differing wavelengths, just like different LEDs produce light of different colours.

Since the human eye cannot see the infrared radiations, it is not possible for a person to identify whether the IR LED is working or not, unlike a common LED. To overcome this problem, the camera on a cell phone can be used. The camera can show us the IR rays being emanated from the IR LED in a circuit.

Symbol of IR LED:
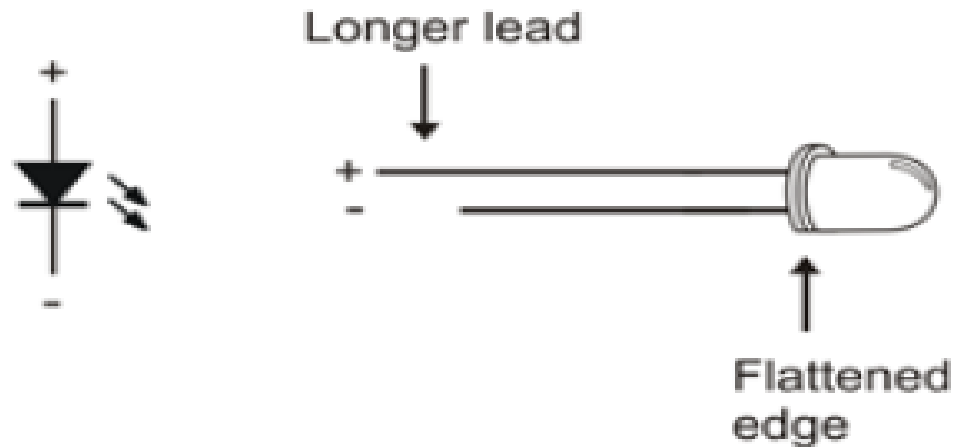
Infrared LEDs can be divided into the following three types: small power one (1mW-10mW), medium power LED (10mW-50mW) and large power LED (50mW-100mW and above). The modulated light can be generated by adding pulse voltage with specific frequency on the driving diode.

Infrared emitting diode infrared light, the principle and characteristics are as follows : the matrix of infrared light-emitting diode light . Infrared emission diode by the infrared radiation efficiency of the material (commonly gallium arsenide GaAs) made of a PN junction , applied to the PN junction forward bias injection current excitation infrared light. Infrared diode current is too small.

Applications:

Infrared LED chips with different wavelengths can be applied in extensive devices, for example:
1. Free air transmission system
2. Optoelectronic switch
3. Floppy disk drive
4. Infrared applied system
5. Smoke detector
6. video surveillance camera
7. Infrared LED chip with wavelength of 940nm: suitable to be used in remote controller, such as remote controllers for household appliances.
8. 808nm: suitable to be used in medical treatment appliances, space optical communication, infrared illumination and the pumping sources of the solid-state lasers.
9. 830nm: suitable to be used in the automated card reader system in freeway.
10. 840nm: suitable to be used in colored zoom infrared waterproof video camera.
11. 850nm: suitable to be used in video cameras that are applied in digital photography, monitoring system, door phone, theft proof alarm and so on.
12. 870nm: suitable to be used in video cameras in marketplace and crossroad.

PHOTOTRANSISTOR

- Phototransistor is a light sensitive transistor.
- In one common type of phototransistor, the base lead of a BJT is replaced by a light sensitive surface.
- When the light sensitive surface @ the base is kept in darkness, the collector-emitter pair of the BJT does not conduct.
- When the light sensitive surface @ the base is exposed to light, a small amount of current flows from the base to the emitter. The small base-emitter current controls the larger collector-emitter current.
- Alternatively, one can also use a field-effect phototransistor (Photo FET).
  - In a photo FET, the light exposure generates a gate voltage which controls a drain-source current.
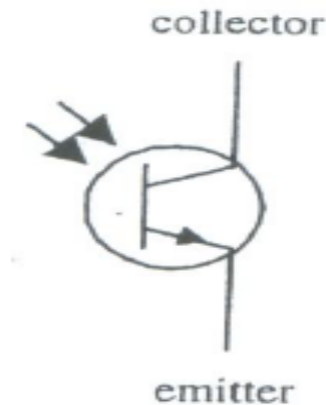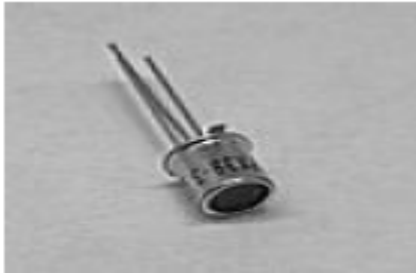
Phototransistor
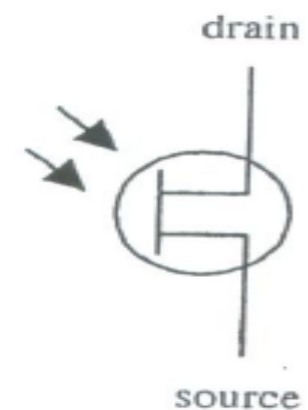
Photo FET

HOW PHOTOTRANSISTOR WORKS

- The bipolar phototransistor resembles a bipolar transistor that has extra large p-type semiconductor region that is open for light exposure.

- When photons from a light source collide with electrons within the p-type semiconductor, they gain enough energy to jump across the pn-junction energy barrier- provided the photons are of the right frequency/energy.

- As electrons jump from the p-region into the lower n-region, holes are created in the p-type semiconductor.

- The extra electrons injected into the lower n-type slab are drawn toward the positive terminal of the battery, while electrons from the negative terminal of the battery are draw into the upper n-type semiconductor and across the np junction, where they combine with the holes, the net result is an electrons current that flows from the emitter to the collector.
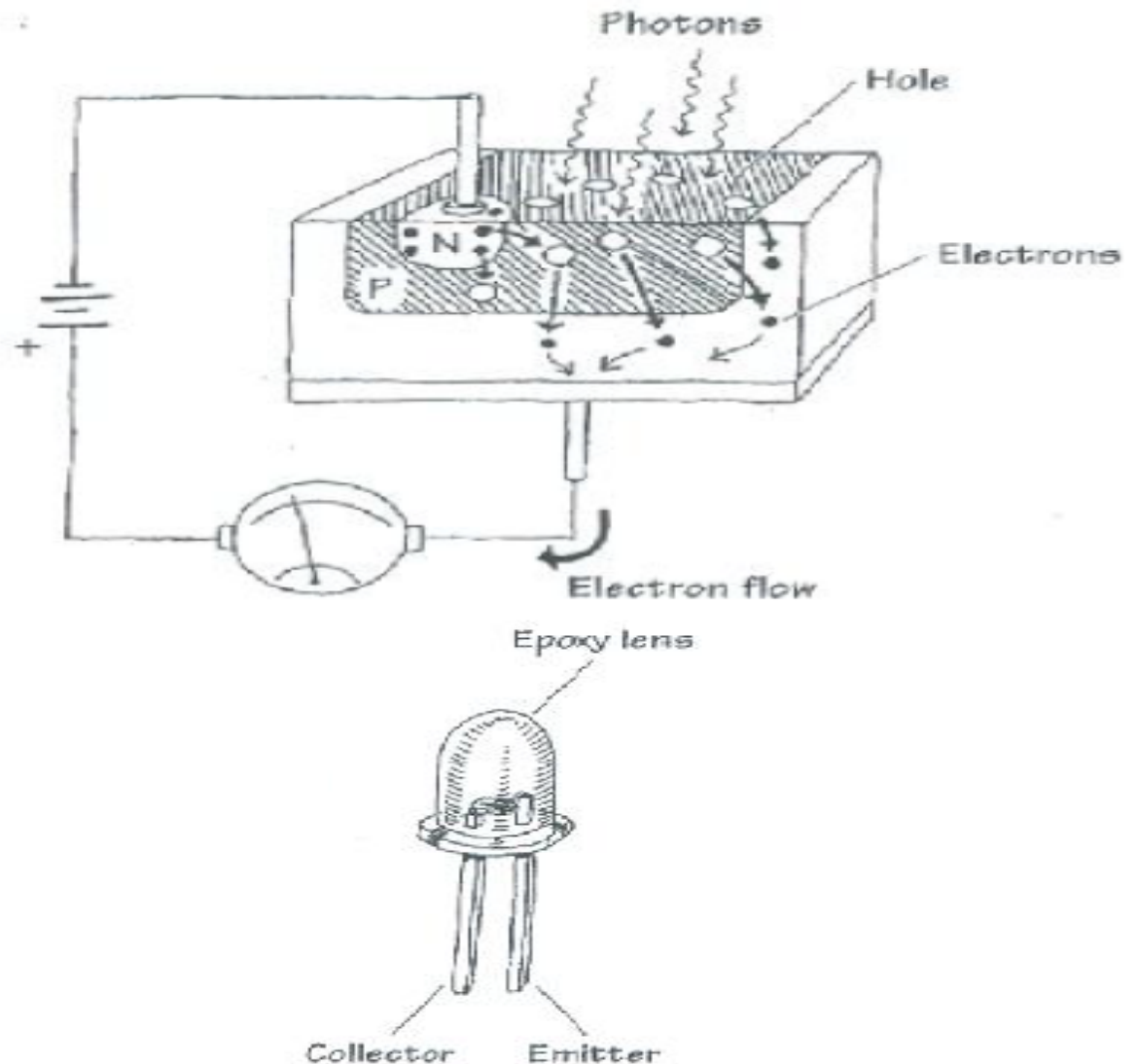
# PHOTOTRANSISTOR APPLICATIONS—LIGHT ACTIVATED RELAY

- A phototransistor is used to control the base current supplied to a power-switching transistor that is used to supply current to a relay.
- When light comes in contact with the phototransistor, the phototransistor turns on, allowing current to pass from the supply into the base of the power-switching transistor.
- This allows the power-switching transistor to turns on, and current flows through the relay, triggering it to switch states.
- The 100K pot is used to adjust the sensitivity of device by controlling current flow through the phototransistor
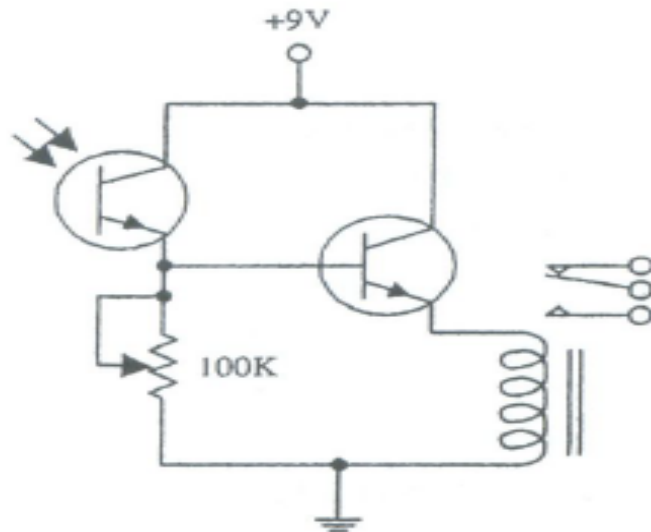
PHOTOTRANSISTOR APPLICATIONS—DARK ACTIVATED RELAY

- A phototransistor is used to control the base current supplied to a power-switching transistor that is used to supply current to a relay.
- When light is removed from the phototransistor, the phototransistor turns off, allowing more current to enter into the base of the power-switching transistor.
- This allows the power-switching transistor to turns on, and current flows through the relay, triggering it to switch states.
- The 100K pot is used to adjust the sensitivity of device by controlling current flow through the phototransistor.
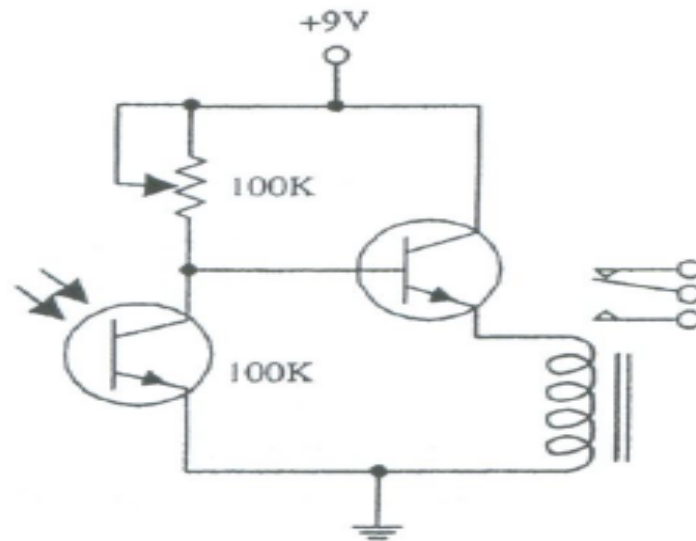
PHOTOTRANSISTOR APPLICATIONS—TACHOMETER

- A phototransistor is being used as a frequency counter or tachometer.
- A rotating disk is connected to a rotating shaft. The rotating disk has one hole in it.
- For the given setup, the disk will allow light to pass through the hole once every revolution.
- The light passing through the disk triggers the phototransistor into conduction.
- A frequency counter is used to count the number of electrical pulses generated.

light source rotating disk with hole

V_{bat} R to frequency counter
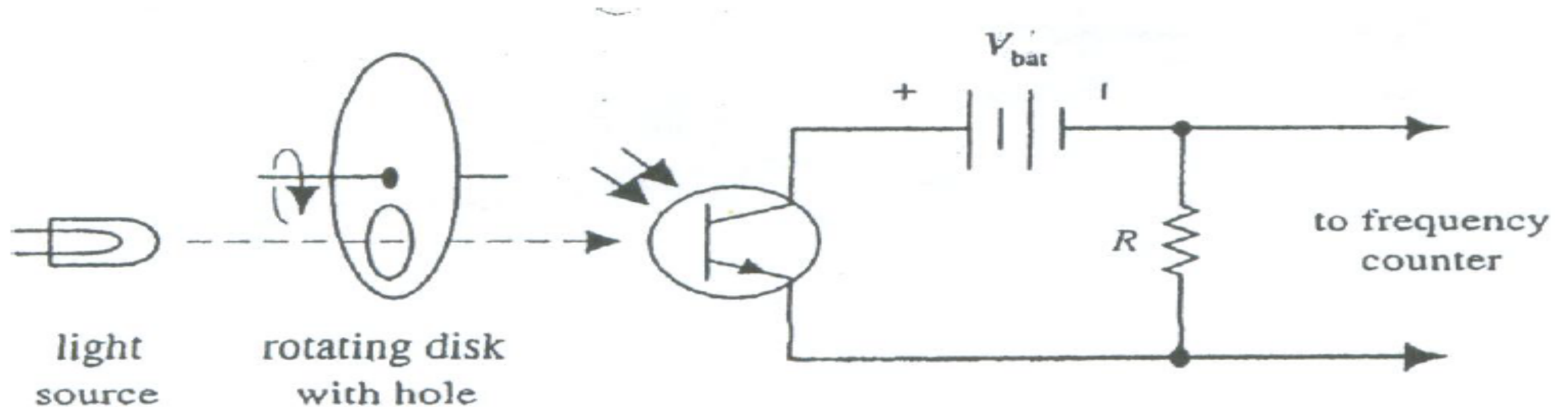
CLASSIFICATION of INTEGRATED CIRCUIT FABRICATION

Integrated circuits offer a wide range of applications and could be broadly classified as: Digital ICs and Linear ICs

Based upon the above requirements, two distinctly different IC technology namely, Monolithic technology and Hybrid technology have been developed.

In monolithic integrated circuits, all circuit components, both active and passive elements and their interconnections are manufactured into or on top of a single chip of silicon.

The monolithic circuit is ideal for applications where identical circuits are required in very large quantities and hence provides lowest per-unit cost and highest order of reliability. In hybrid circuits, separate component parts are attached to a ceramic substrate and interconnected by means of either metallization pattern or wire bonds. This technology is more adaptable to small quantity custom circuits. Based upon the active devices used, ICs can be classified as bipolar (using BJT) and unipolar (using FET). Bipolar and unipolar ICs may further be classified depending upon the isolation technique or type of FET used as in Fig. 1.1.
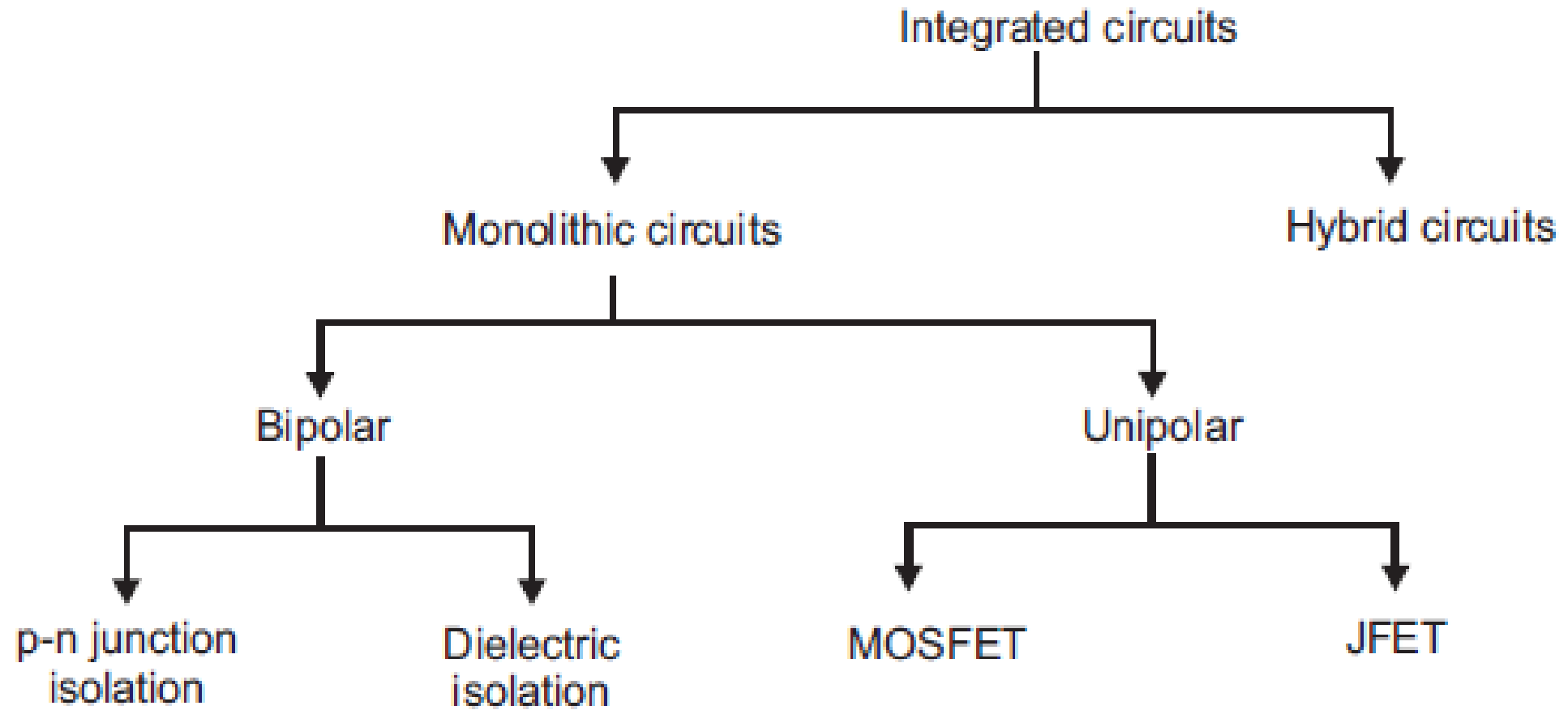
Fig. 1.1 Classification of ICs

FUNDAMENTAL OF MONOLITHIC IC TECHNOLOGY

A monolithic circuit, literally speaking, means a circuit fabricated from a single stone or a single crystal. The origin of the word 'monolithic' is from the Greek word *monos* meaning 'single' and *lithos* meaning 'stone'. So monolithic integrated circuits are, in fact, made in a single piece of single crystal silicon. The most significant advantage of integrated circuit of reducing the cost of production of electronic circuits due to batch production can be easily visualized by a simple example. A standard 10 cm diameter wafer can be divided into approximately 8000 rectangular chips of sides 1 mm. Each IC chip may contain as few as tens of components to several thousand components.

And if 10 such wafers are processed in one batch, we can make 80,000 ICs simultaneously. Many chips so produced will be faulty due to imperfection in the manufacturing process. Even if the yield (percentage of fault free chips/wafer) is only 20 per cent, it can be seen that 16,000 good chips are produced in a single batch. The fabrication of discrete devices such as transistor, diode or an integrated circuit in general can be done by the same technology. The various processes usually take place through a single plane and therefore, the technology is referred to as planar technology. A simple circuit of Fig. 1.3 when fabricated by silicon planar technology will have the cross-sectional view shown in Fig. 1.4.
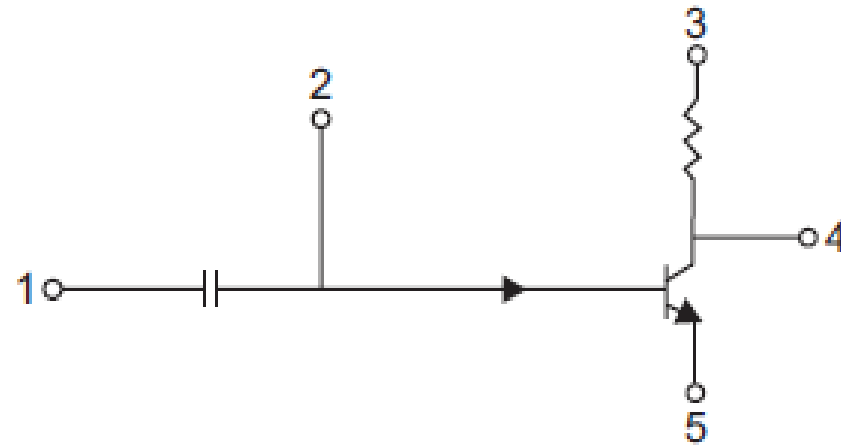
Fig. 1.3 A typical circuit
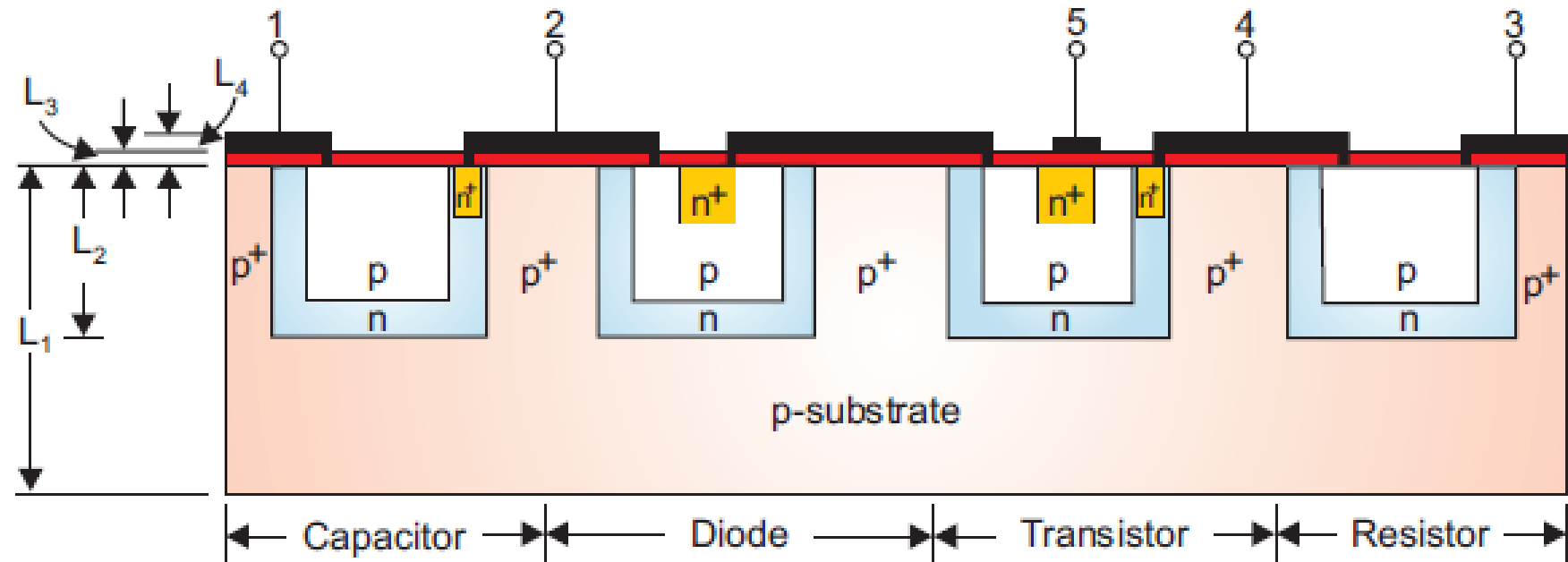
Fig. 1.4 Complete cross-sectional view of the circuit in Fig. 1.3 when transformed into monolithic form

23.1 Integrated Circuit

An *integrated circuit* is one in which circuit components such as *transistors, diodes, resistors, capacitors etc. are automatically part of a small semiconductor chip.*

An integrated circuit consists of a number of circuit components (*e.g.* transistors, diodes, resistors etc.) and their inter connections in a single small package to perform a complete electronic function. These components are formed and connected within a small chip of semiconductor material. The following points are worth noting about integrated circuits :

Fig. 23.1

(*i*) In an *IC*, the various components are automatically part of a small semi-conductor chip and the individual components cannot be removed or replaced. This is in contrast to discrete assembly in which individual components can be removed or replaced if necessary.

(*ii*) The size of an *IC is extremely small. In fact, *ICs* are so small that you normally need a microscope to see the connections between the components. Fig. 23.1 shows a typical semi-conductor chip having dimensions 0.2 mm × 0.2 mm × 0.001 mm. It is possible to produce circuits containing many transistors, diodes, resistors etc. on the surface of this small chip.

(*iii*) No components of an *IC* are seen to project above the surface of the chip. This is because all the components are formed within the chip.
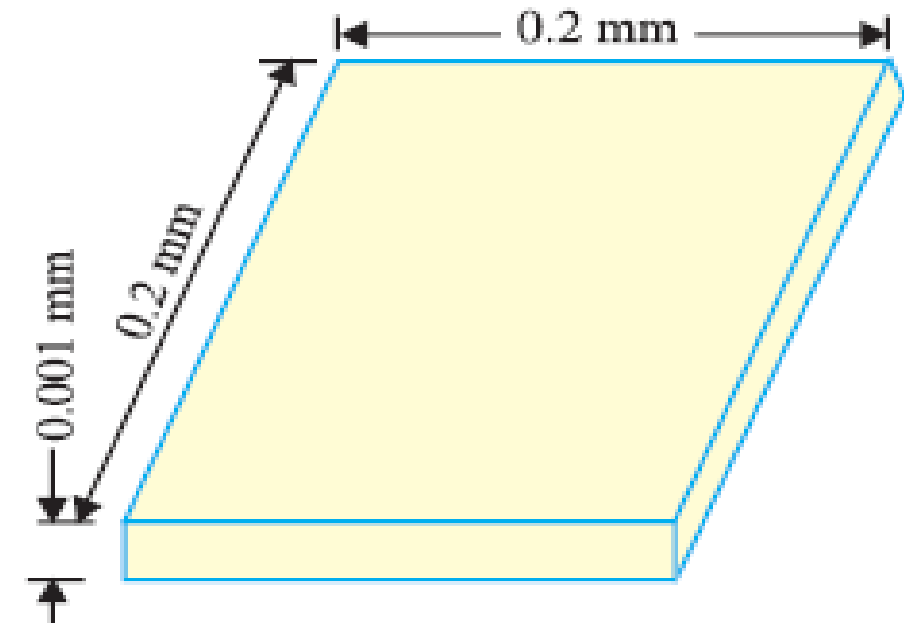
23.2 Advantages and Disadvantages of Integrated Circuits

Integrated circuits free the equipment designer from the need to construct circuits with individual discrete components such as transistors, diodes and resistors. With the exception of a few very simple circuits, the availability of a large number of low-cost integrated circuits have largely rendered discrete circuitry obsolete. It is, therefore, desirable to mention the significant advantages of integrated circuits over discrete circuits. However, integrated circuits have some disadvantages and continuous efforts are on to overcome them.

Advantages : Integrated circuits possess the following advantages over discrete circuits :

    (*i*) Increased reliability due to lesser number of connections.

    (*ii*) Extremely small size due to the fabrication of various circuit elements in a single chip of semi-conductor material.

    (*iii*) Lesser weight and **space requirement due to miniaturized circuit.
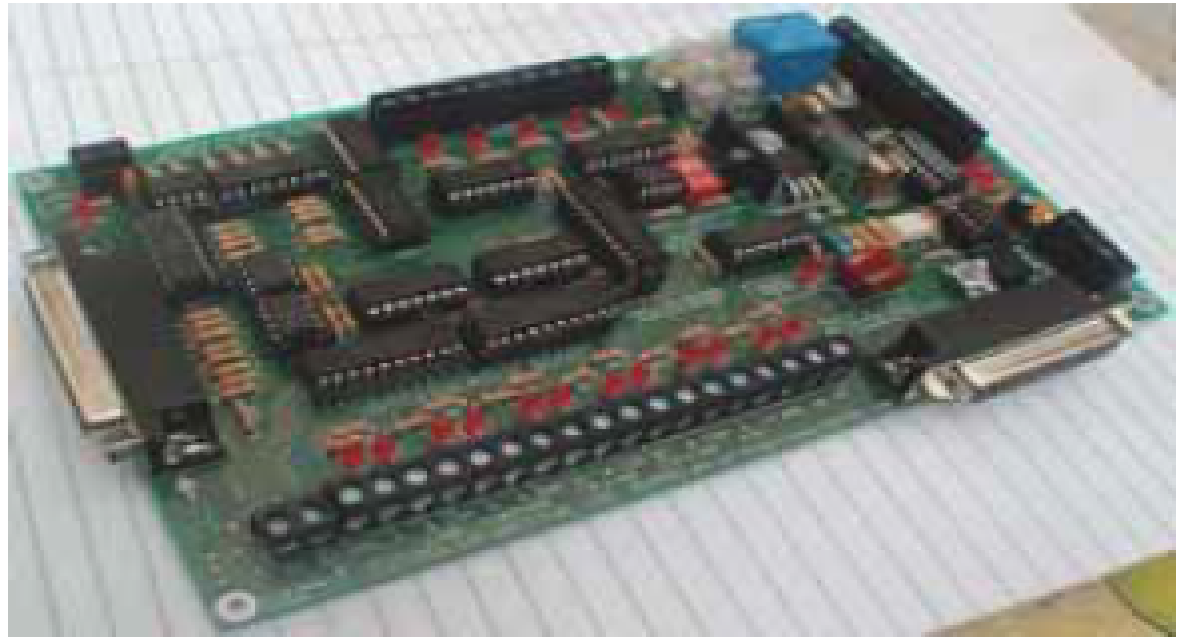
Integrated circuits

---

 *   Since it combines both active (*e.g.*, transistors, diodes etc.) and passive elements (*e.g.*, resistors, capacitors etc.) in a monolithic structure, the complete unit is called an integrated circuit.

**   Typically, this is about 10% of the space required by comparable discrete assembly.

(*iv*) Low power requirements.

(*v*) Greater ability to operate at extreme values of temperature.

(*vi*) Low cost because of simultaneous production of hundreds of alike circuits on a small semi-conductor wafer.

(*vii*) The circuit lay out is greatly simplified because integrated circuits are constrained to use minimum number of external connections.

Disadvantages : The disadvantages of integrated circuits are :

(*i*) If any component in an *IC* goes out of order, the whole *IC* has to be replaced by the new one.

(*ii*) In an *IC*, it is neither convenient nor economical to fabricate capacitances exceeding 30 *pF*. Therefore, for high values of capacitance, discrete components exterior to *IC* chip are connected.

(*iii*) It is not possible to fabricate inductors and transformers on the surface of semi-conductor chip. Therefore, these components are connected exterior to the semi-conductor chip.

(*iv*) It is not possible to produce high power *ICs* (greater than 10 W).

(*v*) There is a lack of flexibility in an *IC* *i.e.*, it is generally not possible to modify the parameters within which an integrated circuit will operate.

23.3 Inside an IC Package

The *IC* units are fast replacing the discrete components in all electronic equipment. These are similar to the discrete circuits that they replaced. However, there are some points to be noted. An integrated circuit (*IC*) usually contains only transistors, diodes and resistors. It is usually very difficult to form inductors in an *IC*. Also, only very small capacitors, in the picofarad range, can be included. When inductors and large values of *C* are needed, they are connected externally to an *IC*. The various components in an *IC* are so small that they cannot be seen with a naked eye. Therefore, individual components cannot be removed or replaced. If a single component within an *IC* fails, the complete *IC* is replaced. When studying circuits using *IC*'s, we are more concerned with the external connections to the *ICs* than with what is actually going on inside. We cannot get into an *IC* to repair its internal circuitry.

23.4 IC Classifications

Four basic types of constructions are employed in the manufacture of integrated circuits, namely :

(*i*) mono-lithic (*ii*) thin-film (*iii*) thick-film (*iv*) hybrid.

Monolithic *ICs* are by far the most common type used in practice. Therefore, in this chapter we shall confine our attention to the construction of this type of *ICs* only. It may be worthwhile to mention here that regardless of the type of method used to fabricate active and passive components, the basic characteristics and circuit operation of an *IC* are the same as for any of their counterparts in a similar circuit using separate circuit components.

23.5 Making Monolithic IC

A **monolithic IC** *is one in which all circuit components and their inter-connections are formed on a single thin wafer called the substrate.*

The word monolithic is from Greek and means "one stone." The word is appropriate because all the components are part of one chip. Although we are mainly interested in using *ICs*, yet it is profitable to know something about their fabrication. The basic production processes for the monolithic *ICs* are as follow :

(*i*) **p-Substrate.** This is the first step in the making of an *IC*. A cylindrical *p*-type *silicon crystal is grown having typical dimensions 25 cm long and 2.5 cm diameter [See Fig. 23.2 (*i*)]. The crystal is then cut by a diamond saw into many thin wafers like Fig. 23.2 (*ii*), the typical thickness of

---

*   Since silicon possesses characteristics which are best suited to *IC* manufacturing processes.

the wafer being 200 µm. One side of wafer is polished to get rid of surface imperfections. This wafer is called the substrate. The *ICs* are produced on this wafer.
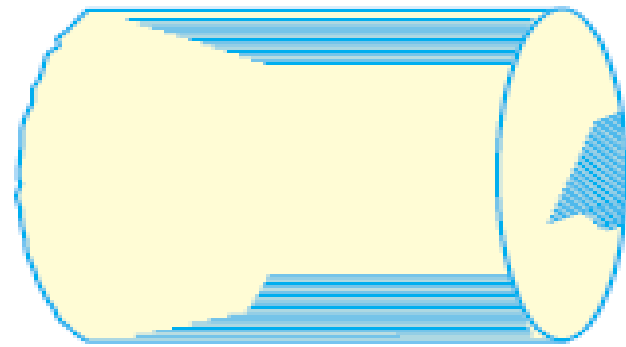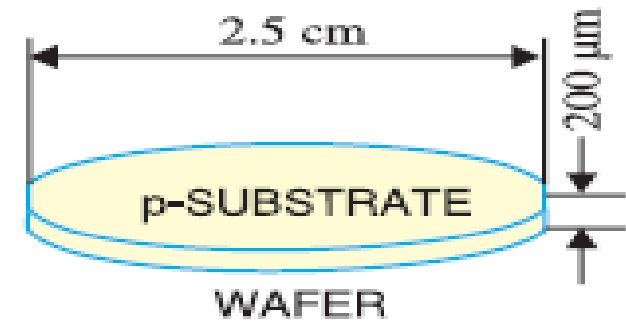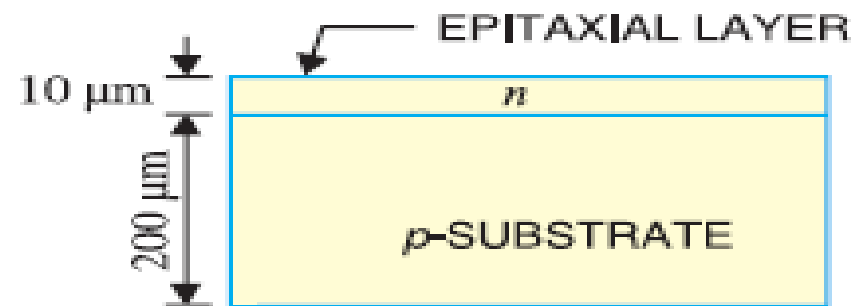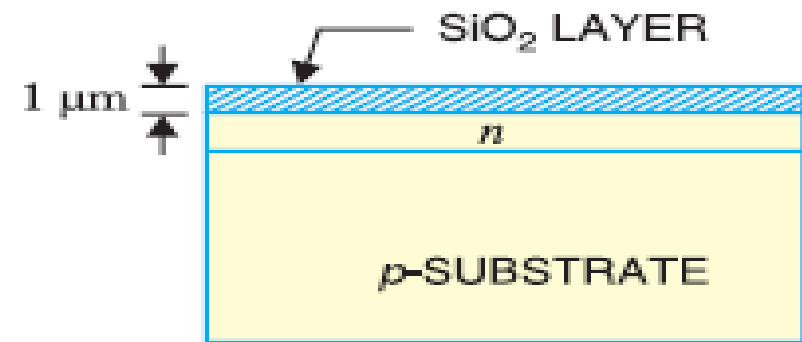

Fig. 23.2

(*ii*) **Epitaxial n layer.** The next step is to put the wafers in a diffusion furnace. A gas mixture of silicon atoms and pentavalent atoms is passed over the wafers. This forms a thin layer of *n*-type
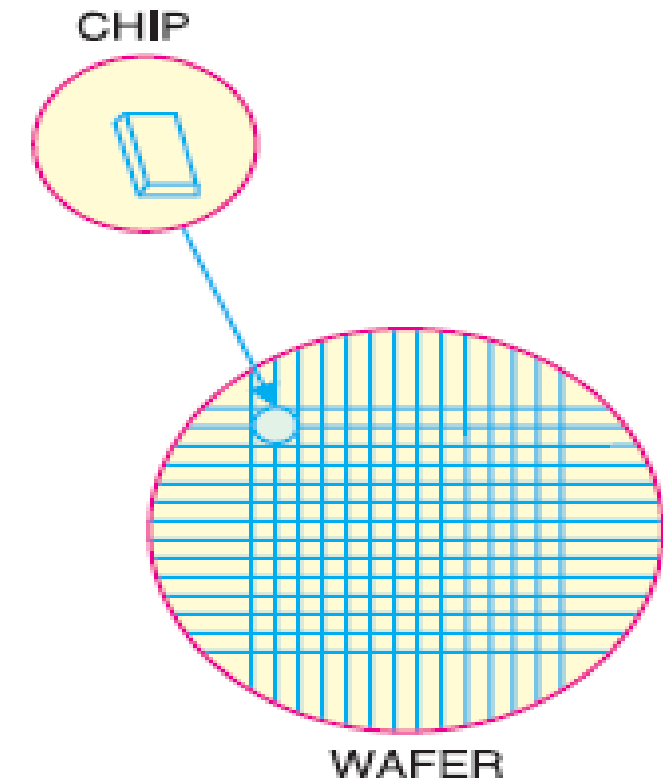

Fig. 23.3

semi-conductor on the heated surface of substrate [See Fig. 23.3 (*i*)]. This thin layer is called the *epitaxial layer* and is about 10 µm thick. It is in this layer that the whole integrated circuit is formed.

(*iii*) **Insulating layer.** In order to prevent the contamination of the epitaxial layer, a thin SiO_2 layer about 1 μm thick is deposited over the entire surface as shown in Fig. 23.3 (*ii*). This is achieved by passing pure oxygen over the epitaxial layer. The oxygen atoms combine with silicon atoms to form a layer of silicon dioxide (SiO_2).

(*iv*) **Producing components.** By the process of **diffusion, appropriate materials are added to the substrate at specific locations to produce diodes, transistors, resistors and capacitors. The production of these components on the wafer is discussed in Art 23.6.

(*v*) **Etching.** Before any impurity is added to the substrate, the oxide layer (*i.e. SiO_2* layer) is etched. The process of etching exposes the epitaxial layer and permits the production of desired components. The terminals are processed by etching the oxide layer at the desired locations.

(*vi*) **Chips.** In practice, the wafer shown in Fig. 23.4 is divided into a large number of areas. Each of these areas will be a separate chip. The manufacturer produces hundreds of alike *ICs* on the wafer over each area. To separate the individual *ICs*, the

CHIP

WAFER

Fig. 23.4

wafer is divided into small chips by a process similar to glass cutting. This is illustrated in Fig. 23.4. It may be seen that hundreds of alike *ICs* can be produced from a small wafer. This simultaneous mass production is the reason for the low cost of integrated circuits.

After the chip is cut, it is bonded to its mounting and connections are made between the *IC* and external leads. The *IC* is then encapsulated to prevent it from becoming contaminated by the surrounding atmosphere.

23.6 Fabrication of Components on Monolithic IC

The notable feature of an *IC* is that it comprises a number of circuit elements inseparably associated in a single small package to perform a complete electronic function. This differs from discrete assembly where separately manufactured components are joined by wires. We shall now see how various circuit elements (*e.g.* diodes, transistors, resistors etc.) can be constructed in an *IC* form.
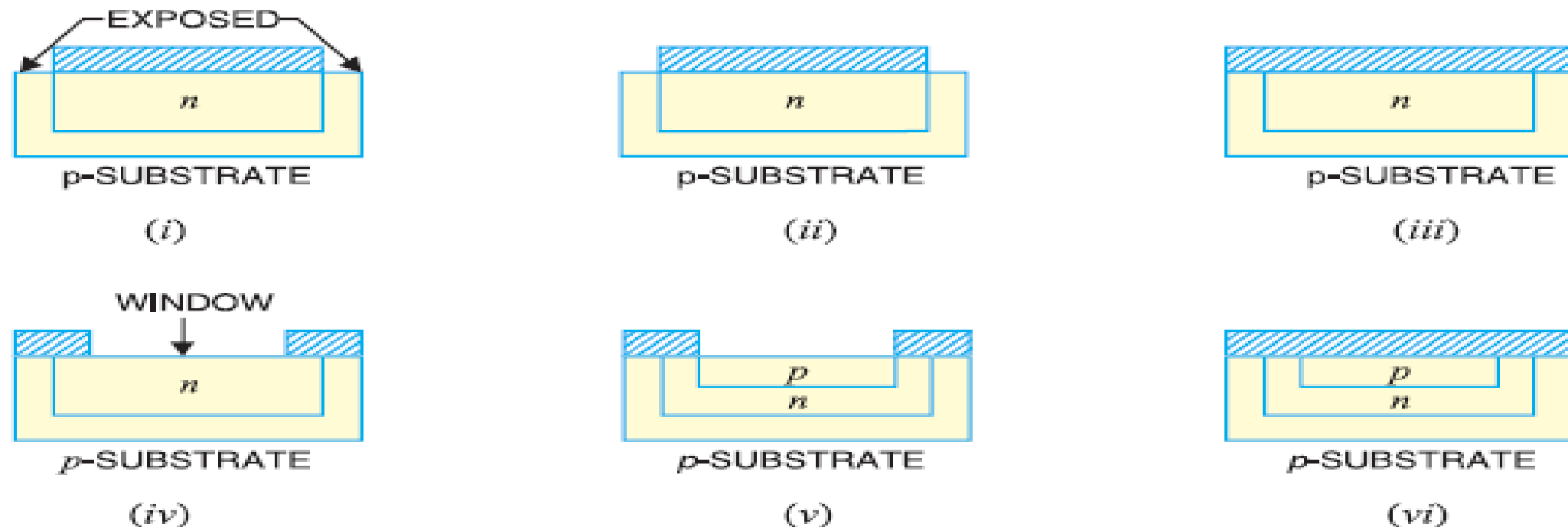
Fig. 23.5

(i) Diodes. One or more diodes are formed by diffusing one or more small n-type deposits at appropriate locations on the substrate. Fig. 23.5 shows how a diode is formed on a portion of substrate of a monolithic IC. Part of SiO_2 layer is etched off, exposing the epitaxial layer as shown in Fig. 23.5 (*i*). The wafer is then put into a furnace and trivalent atoms are diffused into the epitaxial layer. The trivalent atoms change the exposed epitaxial layer from n-type semi-conductor to p-type. Thus we get an island of n-type material under the SiO_2 layer as shown in Fig. 23.5 (*ii*).

Next pure oxygen is passed over the wafer to form a complete SiO_2 layer as shown in Fig. 23.5 (*iii*). A hole is then etched at the centre of this layer ; thus exposing the n-epitaxial layer [See Fig. 23.5 (*iv*)]. This hole in SiO_2 layer is called a **window**. Now we pass trivalent atoms through the window. The trivalent atoms diffuse into the epitaxial layer to form an island of p-type material as shown in Fig. 23.5 (*v*). The SiO_2 layer is again formed on the wafer by blowing pure oxygen over the wafer [See Fig. 23.5 (*vi*)]. Thus a p-n junction diode is formed on the substrate.

The last step is to attach the terminals. For this purpose, we etch the SiO_2 layer at the desired locations as shown in Fig 23.6 (*i*). By depositing metal at these locations, we make electrical contact with the anode and cathode of the integrated diode. Fig. 23.6 (*ii*) shows the electrical circuit of the diode.
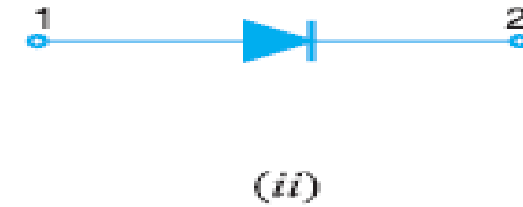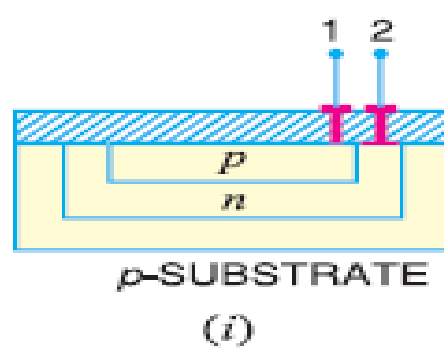
Fig. 23.6

(ii) **Transistors.** Transistors are formed by using the same principle as for diodes. Fig. 23.7 shows how a transistor is formed on a portion of the substrate of a monolithic *IC*. For this purpose, the steps used for fabricating the diode are carried out upto the point where *p* island has been formed and sealed off [See Fig. 23.5 (*vi*) above]. This Fig. is repeated as Fig. 23.7 (*i*) and shall be taken as the starting point in order to avoid repetition.
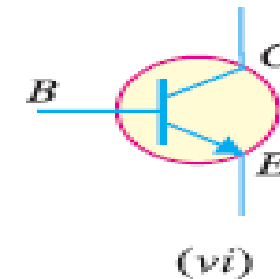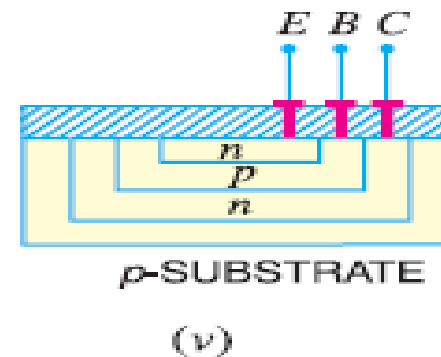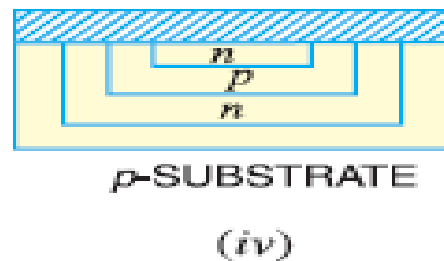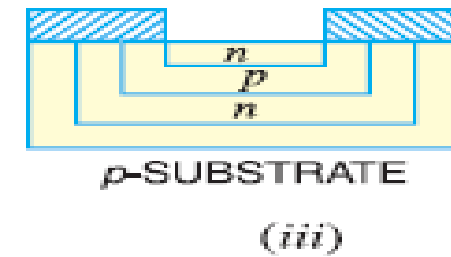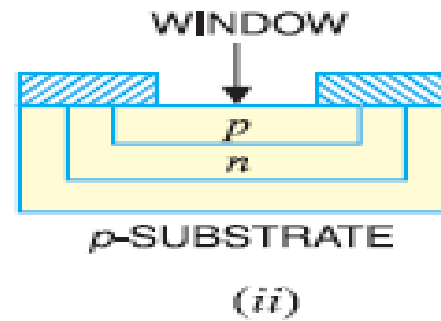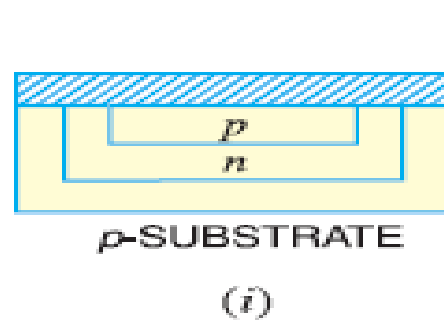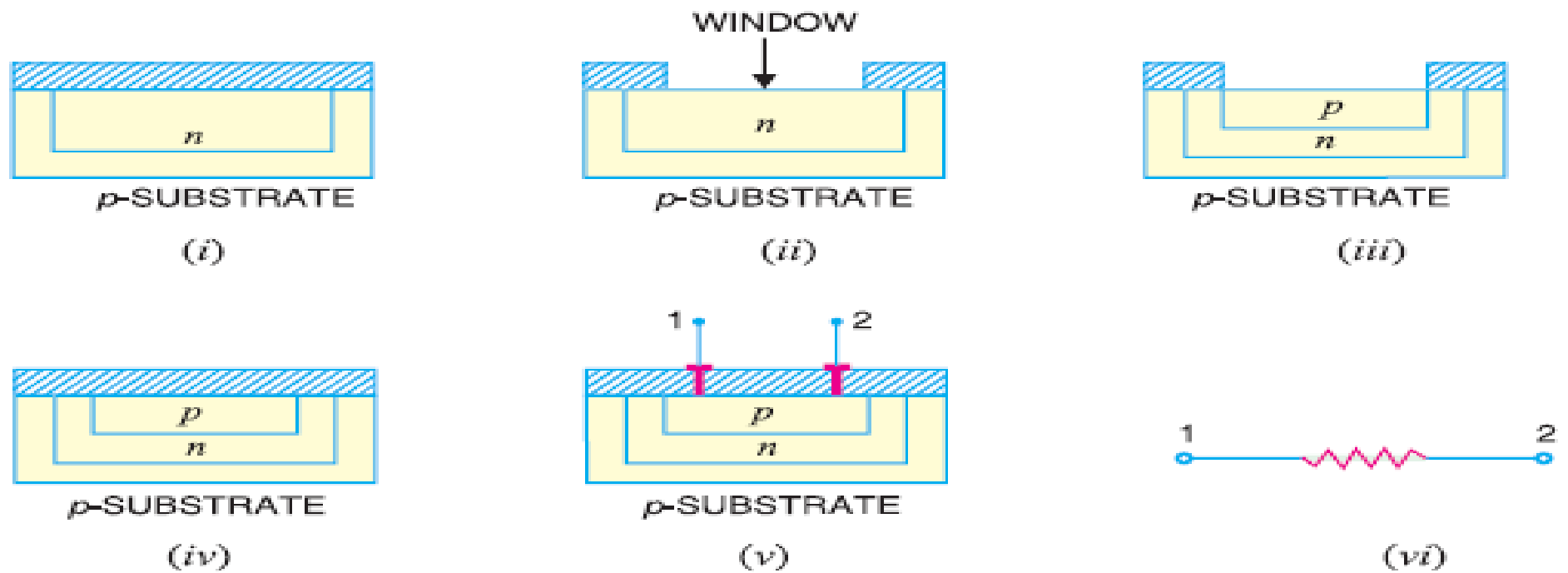

Fig. 23.7

A window is now formed at the centre of SiO_2 layer, thus exposing the p-epitaxial layer as shown in Fig. 23.7(ii). Then we pass pentavalent atoms through the window. The pentavalent atoms diffuse into the epitaxial layer to form an island of n-type material as shown in Fig. 23.7 (iii). The SiO_2 layer is re-formed over the wafer by passing pure oxygen [See Fig. 23.7 (iv)]. The terminals are processed by etching the SiO_2 layer at appropriate locations and depositing the metal at these locations as shown in Fig. 23.7 (v). In this way, we get the integrated transistor. Fig. 23.7 (vi) shows the electrical circuit of a transistor.

(*iii*) **Resistors.** Fig. 23.8 shows how a resistor is formed on a portion of the substrate of a monolithic IC. For this purpose, the steps used for fabricating diode are carried out upto the point where n island has been formed and sealed off [Refer back to Fig. 23.5 (iii)]. This figure is repeated as Fig. 23.8 (i) and shall be taken as the starting point in order to avoid repetition.

A window is now formed at the centre of SiO_2 layer, thus exposing the n-epitaxial layer as shown in Fig. 23.8 (ii). Then we diffuse a p-type material into the n-type area as shown in Fig. 23.8 (iii). The SiO_2 layer is re-formed over the wafer by passing pure oxygen [See Fig. 23.8 (iv)]. The terminals are processed by etching SiO_2 layer at two points above the p island and depositing the metal at these locations [See Fig. 23.8 (v)]. In this way, we get an integrated resistor. Fig. 23.8 (vi) shows the electrical circuit of a resistor.
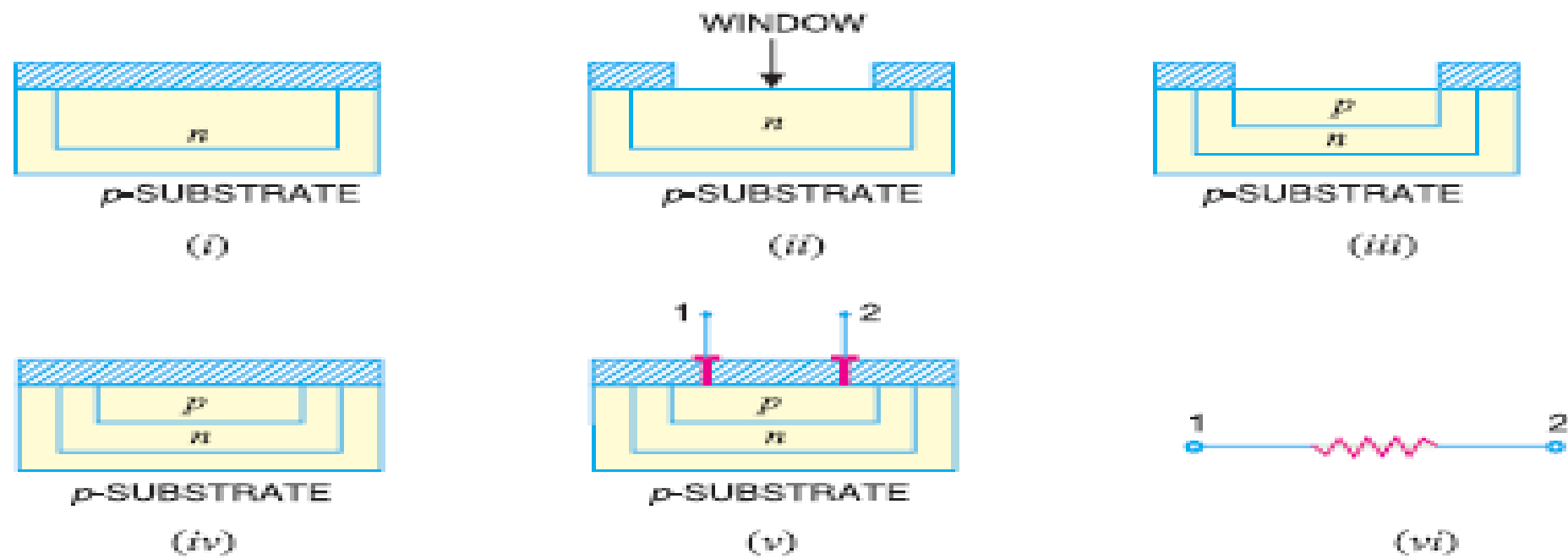
Fig. 23.8

The value of resistor is determined by the material, its length and area of cross-section. The high-resistance resistors are long and narrow while low-resistance resistors are short and of greater cross-section.
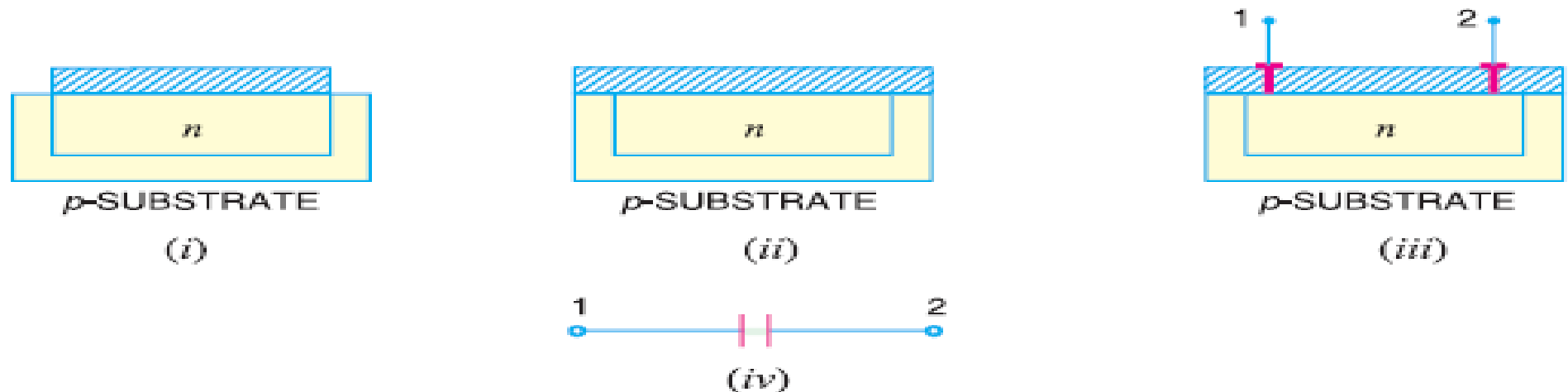
Fig. 23.8

The value of resistor is determined by the material, its length and area of cross-section. The high-resistance resistors are long and narrow while low-resistance resistors are short and of greater cross-section.
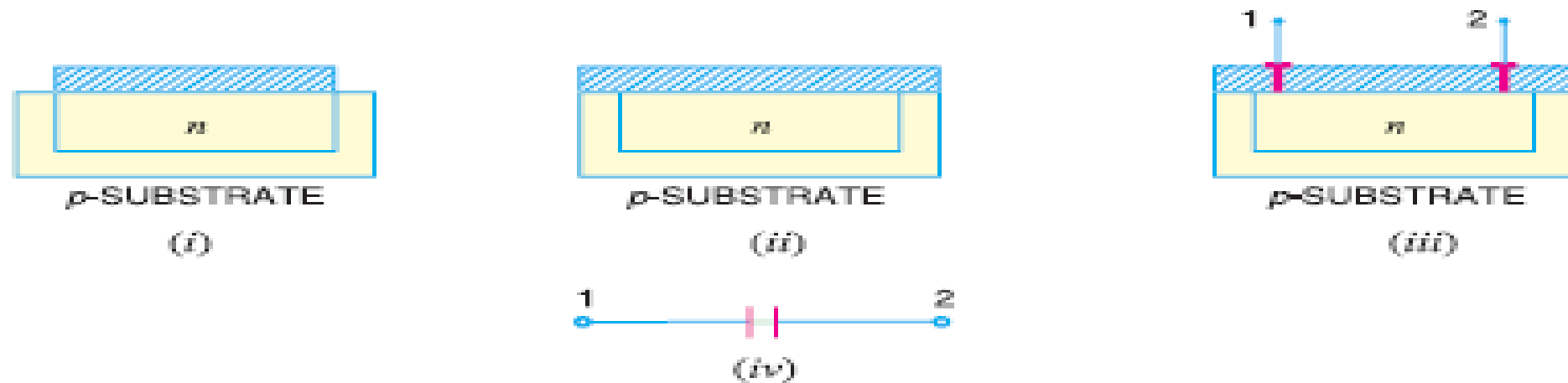

Fig. 23.9

(iv) Capacitors. Fig. 23.9 shows the process of fabricating a capacitor in the monolithic *IC*. The first step is to diffuse an *n*-type material into the substrate which forms one plate of the capacitor as shown in Fig. 23.9 (*i*). Then SiO_2 layer is re-formed over the wafer by passing pure oxygen as shown in Fig. 23.9 (*ii*).

The SiO_2 layer formed acts as the dielectric of the capacitor. The oxide layer is etched and terminal 1 is added as shown in Fig. 23.9 (*iii*). Next a large (compared to the electrode at terminal 1) metallic electrode is deposited on the SiO_2 layer and forms the second plate of the capacitor. The oxide layer is etched and terminal 2 is added. This gives an integrated capacitor. The value of capacitor formed depends upon the dielectric constant of SiO_2 layer, thickness of SiO_2 layer and the area of cross-section of the smaller of the two electrodes.