PROGRAMME: B.Sc. PHYSICS
General

SEMESTER - VI

Course: Digital & Analog Circuits and Instrumentation (DSE)

Lecture Note: 1

Semiconductor junction Diode

Semiconductor Junction Diodes

5.1 INTRODUCTION

A junction between a *p*-type and an *n*-type semiconductor is most important in electronic applications and in understanding other semiconductor devices like transistors. We shall discuss in this Chapter the physical principle and the operation of a *p-n* junction. An important group of devices formed with *p-n* junctions is also given.

5.2 *p-n* JUNCTIONS

When one side of a single-crystal semiconductor is doped with acceptors and the other side is doped with donors, a *p-n* junction is produced (Fig. 5.1). It is a two-terminal device and is referred to as a *junction diode*. It appeared in the 1950s.

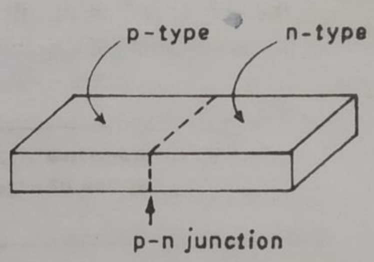

Fig. 5.1. A *p-n* junction

A *p-n* junction can be step-graded or linearly graded. In a *step-graded* or *abrupt p-n junction*, the acceptor and the donor concentrations in the semiconductor are constants up to the junction. To produce this kind of junctions, a small pellet of a trivalent impurity like aluminium is placed on a wafer of *n*-type silicon and the system is heated to a high temperature for a short period. Aluminium dissolves in silicon in the process, so that over a portion silicon changes from *n*-type to *p*-type. A *p-n* junction is thus formed.

In a *linearly-graded junction*, the impurity concentration varies almost linearly with distance from the junction. This kind of junctions can be obtained by drawing a single crystal from a melt of silicon containing the impurity of one type at the beginning. As the crystal is drawn, the impurity of the other type is added at a controlled rate to change the semiconductor type.

A precise control of the impurity profile is obtained by the solid state diffusion method. In the diffused mesa junction method, *p*-type impurities are diffused into an *n*-type single crystal, called the *substrate*. A *p-n* junction is thus developed.

5.3 THE UNBIASED *p-n* JUNCTION

A schematic diagram of a *p-n* junction is depicted in Fig. 5.2. The positively charged donor ions are represented by encircled plus signs. The negatively charged acceptor ions are similarly shown by encircled minus signs.

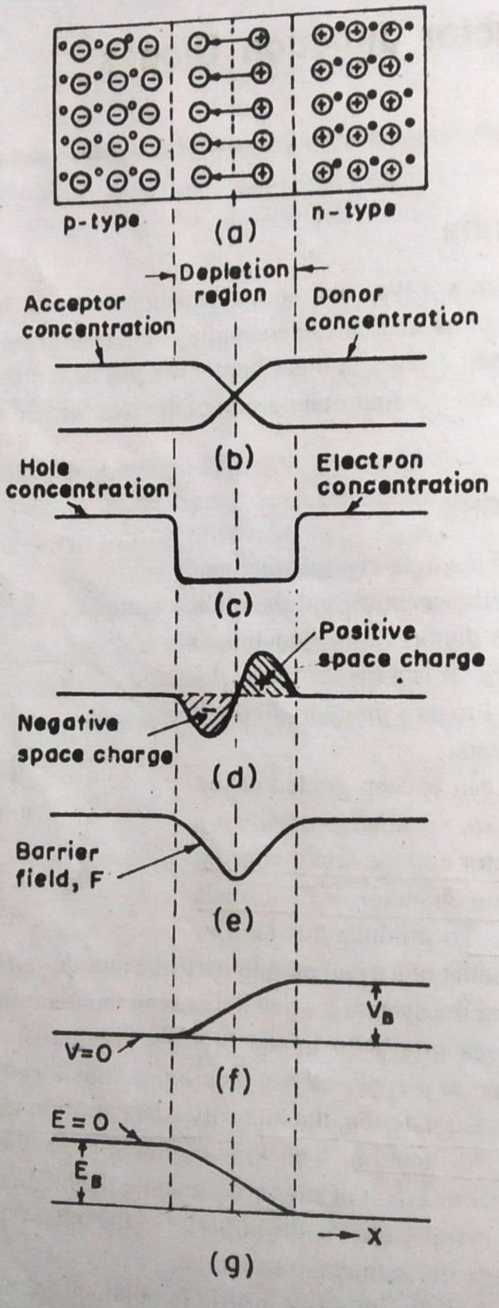

Fig. 5.2 A schematic diagram of a *p-n* junction showing (a) mobile and immobile charges, (b) variation of donor and acceptor ion concentrations, (c) variation of electron and hole concentrations, (d) concentration of uncovered charges, (e) electric field variation, (f) variation of electrostatic potential, and (g) electron energy variation.

For simplicity, we assume that there are only holes in the p-side and electrons in the n-side. Because of the density gradient across the junction, electrons will initially diffuse through the junction to the left, and holes to the right. In the process, the electrons and the holes combine with each other. As a result, the acceptor ions near the junction in the p-side and the donor ions near the junction in the n-side are left unneutralized. The unneutralized ions in the vicinity of the junction are termed *uncovered charges*. The positive and the negative uncovered charges generate an electric field across the junction directed from the n-side to the p-side. The direction of this field, called the *barrier field*, is shown by the arrowheads in Fig. 5.2(a). The barrier field opposes the diffusion of electrons and holes through the junction. Equilibrium is established when the barrier field is just sufficient to prevent further diffusion of electrons and holes. At this situation, there is no movement of charge carriers across the junction. Since the neighbourhood of the junction is depleted of mobile charges, it is referred to as the *depletion region*, the *space-charge region*, or the *transition region*. The thickness of the depletion region is about 0.5 μm.

The variation of the concentrations of donors and acceptors in the system is shown in Fig. 5.2(b). The mobile charge concentration is zero in the depletion region, as shown in Fig. 5.2(c). To the left of this region, the hole concentration equals the acceptor concentration, and to its right the electron concentration equals the donor concentration. The uncovered charge density ρ is zero at tne junction, positive to its right and negative to its left in the depletion region [see Fig. 5.2(d)].

The barrier field F can be found from the Poisson equation (in SI units)

$$\frac{d^2 V}{dx^2} = -\frac{\rho}{\varepsilon}\,, \tag{5.1}$$

where V is the electrostatic potential at a distance x from a suitably chosen origin, and ε is the permittivity of the semiconductor. Integrating Eq. (5.1) yields

$$F = -\frac{dV}{dx} = \int_{x_0}^{x} \frac{\rho}{\varepsilon}\, dx\,, \tag{5.2}$$

where the field F is zero at $x = x_0$. The variation of the barrier field across the junction is shown in Fig. 5.2(e).

The electrostatic potential at a point is given by

$$V(x) = -\int_{x_0}^{x} F\, dx \tag{5.3}$$

The variation of $V(x)$ in the depletion region is depicted in Fig. 5.2(f). It rises in the n-side to give a *potential barrier* V_B which prevents further diffusion of carriers across the junction.

The barrier field across the junction causes the expenditure of an energy E_B $(= e\, V_B)$ to transfer an electron from the n-side to the p-side of the junction. E_B is termed the *barrier energy*. An equal amount of energy is to be spent to transfer a hole from the *p*-side to the *n*-side. The value of E_B depends on the semiconductor, the doping, and the width of the depletion region.

Since the holes carry positive charges, the potential energy for a hole follows the variation of V shown in Fig. 5.2(f). The charge of an electron being negative, the plot of the potential energy E of an electron is obtained by multiplying V by $-e$ and is shown in Fig. 5.2(g).

In actual practice, the *p*-side contains electrons as minority carriers and the *n*-side has holes as minority carriers. When the *p-n* junction is unbiased, the barrier field does not prevent the electrons from the *p*-side and the holes from the *n*-side to cross the junction. The current caused by this movement of the minority carriers across the junction is exactly balanced in the equilibrium condition by the current generated by an equal and opposite flow of the majority carriers having sufficient thermal energy to surmount the barrier energy.

Energy-Band Diagram

For an *n*-type semiconductor, the Fermi level E_F lies near the conduction band edge E_C while for a *p*-type semiconductor it lies near the valence band edge E_v. When a *p-n* junction is produced, under equilibrium condition, the Fermi energy E_F attains a constant value throughout the system, as shown in Fig. 5.3. In this situation, the conduction band edge E_{cp} of the *p*-side will be at a higher level than the conduction band edge E_{cn} of the *n*-side. Similarly, the valence band edge E_{vp} of the *p*-side will lie higher than the valence band edge E_{vn} of the *n*-side. The barrier energy is

$$E_B = E_{cp} - E_{cn} = E_{vp} - E_{vn} . \tag{5.4}$$

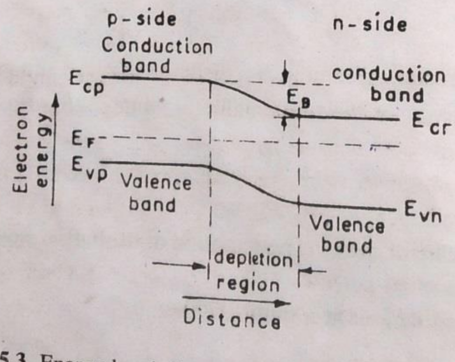

Fig. 5.3 Energy-band diagram of an unbiased *p-n* junction.

Observation

Although a barrier potential V_B exists across a *p-n* junction, this voltage does not drive a current in a metal wire externally connected to the terminals of a *p-n* junction diode. For, if a current would have flown, it would heat the connecting wire. As no external source of energy is present, this heating of the wire must be accompanied by a simultaneous cooling of the *p-n* junction. This contradicts the attainment of the thermal equilibrium. Therefore the current in the circuit must be zero. This can be accounted for, if the net driving voltage is zero, i.e. if the barrier voltage across the *p-n* junction is balanced by the contact potentials at the metal-semiconductor contacts at the ends of the diode. The current being zero, the connecting wire can be cut and the voltage across the cut would be zero. Therefore, a voltmeter connected across the terminals of a *p-n* junction gives a zero voltage reading; *it does not record the barrier voltage* V_B.

5.4 THE BIASED *p-n* JUNCTION

A *p-n* junction is said to be *forward-biased* when the positive pole of a battery is connected to the *p*-side and the negative pole to the *n*-side of the junction. In this situation, the barrier energy is reduced, allowing a large current to flow across the junction. On the other hand, if the positive pole of a battery is connected to the *n*-side and the negative pole to the *p*-side of a *p-n* junction, the junction is said to be *reverse-biased.* Now the barrier energy is enhanced and a small current flows across the junction. Thus, a *p-n* junction is not a bilateral device and it can be used as a rectifier (see Chapter 6).

A. Forward-biased *p-n* Junction

Figures 5.4(a) and (b) show a forward-biased *p-n* junction and its circuit representation, respectively. The diode is represented in Fig. 5.4(b) by an arrowhead pointing to the direction of the current flow in the forward bias state. The forward biasing voltage *V* supplied by the battery *B* exerts a force on the holes of the *p*-side and on the electrons of the *n*-side of the junction, driving them towards the junction. As a result, the width of the depletion layer and the barrier energy are diminished. The reduction in the barrier energy is eV, where e is the electronic charge and V is the battery voltage. The decrease in the barrier height increases the current flow due to the majority carriers. That is, the flow of holes from the *p*-side to the *n*-side and that of the electrons from the *n*-side to the *p*-side increase. Since electrons and holes are oppositely charged, both these flows cause a current from the *p*-side to the *n*-side across the junction. The forward bias, however, does not affect the minority carrier current that flows in the direction opposite to the direction of the majority carrier current. This is due to the fact that the minority carrier current is determined by the temperature only.

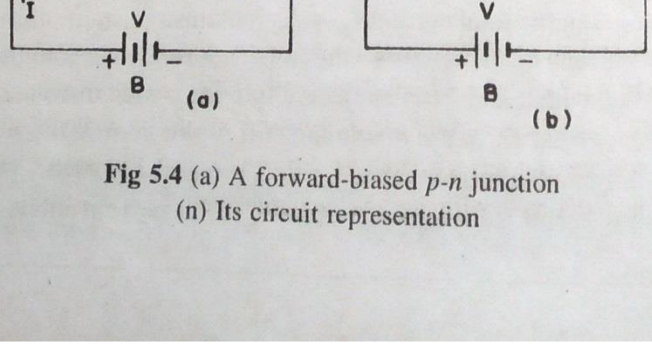

Fig 5.4 (a) A forward-biased *p-n* junction
(n) Its circuit representation

B. Reverse-biased *p-n* Junction

Figures 5.5(a) and (b) show, respectively, a reverse-biased *p-n* junction and its circuit representation. Here the battery voltage *V* pulls the holes in the *p*-side and the electrons in the *n*-side away from the junction. The width of the depletion layer and the height of the barrier increase accordingly. The increase in the barrier energy is *eV*. The rise of the barrier height reduces the current to a negligible value since very few majority carriers can now cross the junction by surmounting the enhanced barrier. But the number of minority carriers moving down the potential barrier is left unaffected since this number depends on the temperature only. As a result, a small current, called the *reverse saturation current* (I_s), flows. The reverse saturation current is practically independent of the applied reverse bias, but it increases with a rise of temperature of the junction diode. This is because, the minority carrier density contributing to I_s increases with a rise of temperature.

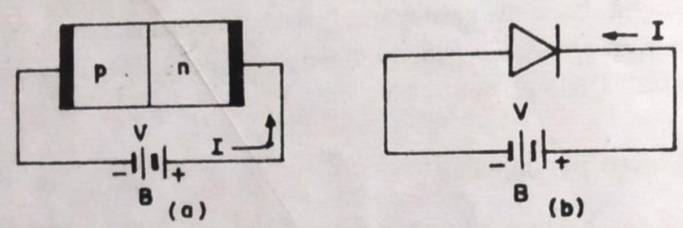

Fig. 5.5 (a) A reverse-biased *p-n* junction (b) Its circuit representation

As the depletion region is devoid of free carriers, this region has a much higher resistivity than any part of the semiconductor. So, most of the battery voltage *V* appears across the depletion region for both forward and reverse-bias. The potential energy diagrams of forward-biased and reverse-biased junctions relative to the unbiased junction are shown in Fig. 5.6(a) and (b), respectively.

Observation

In a forward-biased *p-n* junction, holes are injected into the *n*-side and electrons into the *p*-side. The injected holes are the minority carries in the *n*-region, so are the injected electrons in the *p*-region. As one moves away from the junction, the injected minority carrier density falls off exponentially due to electron-hole recombination. The injected hole current I_{pn} in the *n*-region thus decreases away from the junction into the *n*-region. Deep in the *n*-side the current is carried entirely by the electrons. As the junction is approached, some electrons recombine with the holes injected from the *p*-side. So, the electron current I_{nn} decreases toward the junction, but the total current $I_{pn} + I_{nn}$ remains constant. Similarly, in the *p*-side the hole current I_{pp} decreases as the holes approach the junction and some of them recombine with the electrons injected into the *p*-side from the *n*-side. The injected electron current I_{np} in the *p*-side falls off as we go into the *p*-region from the junction, but again the total current $I_{np} + I_{pp}$ remains constant. Thus, the current flowing in the circuit is $I = I_{pn} + I_{nn} = I_{np} + I_{pp}$. The current enters the *p*-region as a hole current and leaves the *n*-region as an electron current of the same magnitude. Note that the *p-n* diode is a *bipolar* device since the current is carried by both holes and electrons.

Junction Capacitance

The space-charge region around a *p-n* junction consists of immobile positively charged donor ions in the *n*-side and immobile negatively charged acceptor ions in the *p*-side of the junction. Also, the application of a reverse bias to a *p-n* junction increases the width of the space charge region. This voltage-dependent positive and negative charges, spatially separated from each other, endows the *p-n* junction with a capacitance, termed the *transition, space-charge*, or *depletion-region capacitance*. It can be shown that the junction capacitance decreases as the reverse voltage increases. This feature of the junction capacitance is utilized in *varactor diodes* (see Sec. 5.6).

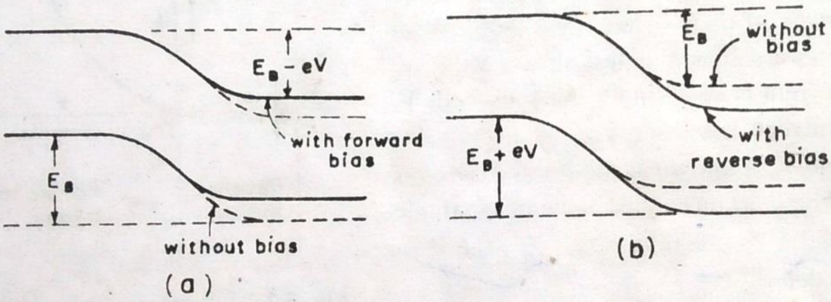

Fig 5.6 Potential energy diagrams for a (a) forward-biased and (b) reverse-biased *p-n* junction

A forward-biased *p-n* junction has a more important source of capacitance. Under such a bias, holes are injected into the *n*-side where they become minority carries, and electrons are injected into the *p*-side where they become the minority carriers. The injected minority carrier concentration decreases as we move away from the junction. This injected charge on either side of the junction gives rise to a capacitance, termed the *diffusion* or *storage capacitance*.

5.5 CURRENT-VOLTAGE CHARACTERISTIC OF A *p-n* JUNCTION

When a voltage V is applied across a *p-n* junction, the total current I flowing through the junction can be shown to be given by

$$I = I_s \left[\exp\left(\frac{eV}{\eta k_B T} \right) - 1 \right], \tag{5.5}$$

where I_s is the reverse saturation current, e is the electronic charge, k_B is the Boltzmann constant, T is the absolute temperature of the junction, and η is a dimensionless number depending on the semiconductor. For germanium, η is unity while for silicon, η is approximately 2.

Substituting $e = 1.6 \times 10^{-19}$ C, $k_B = 1.38 \times 10^{-23}$ J/K, and $T = 300$ K (room temperature), we have from Eq. (5.5)

$$I = I_s \left[\exp\left(\frac{39\,V}{\eta} \right) - 1 \right] \tag{5.6}$$

If V is positive, the *p-n* junction is forward biased, and if V is negative, the junction is reverse biased. Neglecting the voltage drops in the bulk *p*-type and *n*-type regions, the voltage V is approximately the voltage applied across the diode terminals.

Figure 5.7 gives the plot of the current I versus the voltage V, as described by Eq. (5.6). This plot is termed the *current-voltage characteristic* or the *volt-ampere characteristic* of the *p-n* junction diode. When the forward bias V is less than a value V_γ, the current is very small. As V exceeds V_γ, the current increases very sharply. The voltage V_γ is known as the *break-point*, *offset*, *threshold* or *cutin* voltage of the diode. $V_\gamma \approx 0.2$ V for Ge diodes and $V_\gamma \approx 0.6$ V for Si diodes.

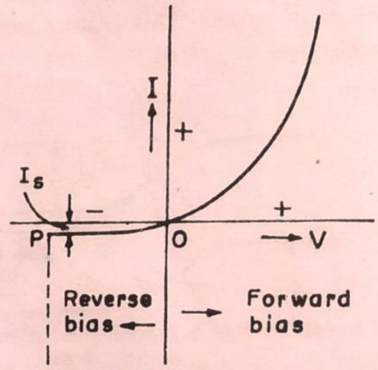

Fig. 5.7 Volt-ampere characteristic of a *p-n* diode.

When V is positive and sufficiently high, the term unity in the parentheses of Eq. (5.6) can be neglected, so that the current I increases exponentially with the voltage V. If the diode is reverse-biased and $|V|$ is a few times $\eta/39$, we have $I' - I_s$. Thus, the reverse current is independent of the applied bias and equals the *reverse saturation current*. For Ge diodes, I_s is a few microampere and for Si diodes, I_s is a few nanoampere at room temperature. Ordinarily, the range of forward current for the diode operation is many orders of magnitude greater than the reverse saturation current. Typically, the forward current is in the range of milliampere while the reverse current is in the range of microampere or less.

Since $\eta = 2$ for Si diodes and $\eta = 1$ for Ge diodes, for a given forward bias (larger than V_γ) the current for a Si diode is less than that for a Ge diode. However, Si diodes can be used over a larger range of temperatures than Ge diodes.

At a reverse voltage corresponding to the point P in Fig. 5.7, the reverse current increases abruptly, as indicated by the dashed line, almost parallel to the current axis. The dashed portion of the characteristic is called the *breakdown region*. A *p-n* diode operated in the breakdown region, is known as *breakdown diode* which is discussed in more detail in Sec. 5.6.

Diode Resistance

The ratio between the voltage V across the junction and the current I flowing through the junction is called the *static* or the *dc resistance* r_{dc} of the diode. That is,

$$r_{dc} = \frac{V}{I} \quad (5.7)$$

As Ohm's law is not valid for the *p-n* junction, r_{dc} is not a constant but varies with the applied voltage.

The inverse of the slope of the volt-ampere characteristic of the *p-n* diode defines the *dynamic* or the *ac resistance* (r_{ac}) of the diode. Thus,

$$r_{ac} = \frac{dV}{dI} \quad (5.8)$$

From Eq. (5.6), we obtain for the dynamic resistance at room temperature

$$r_{ac} = \frac{dV}{dI} = \frac{\eta}{39(I + I_s)} \quad (5.9)$$

Clearly, r_{ac} is not a constant and is determined by the operating voltage. For a reverse bias such that $|39\ V/\eta| >> 1$, the dynamic resistance is extremely large. For a forward bias (greater than V_γ), $I >> I_s$ and $r_{ac} = \eta/(39\ I)$. For a Ge diode ($\eta = 1$) at room temperature, the forward ac resistance is $r_{ac} = 26/I$, where I is in mA and r_{ac} is in ohm. Thus, for a forward current of 26 mA, the dynamic resistance of a Ge diode is 1 ohm.

The nonohmic behaviour of a junction diode gives a variable resistance that is exploited in a device, called *varistor* (variable resistor). To achieve the forward *I-V* characteristic in either direction, a varistor is obtained by connecting two identical diodes in parallel, oppositely poled (Fig. 5.8).

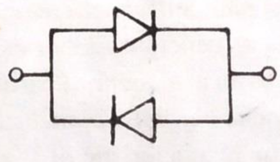

Fig. 5.8 Two oppositely poled diodes in parallel

Observation

(1) The temperature dependance of the volt-ampere characteristic of a *p-n* junction given by Eq. (5.5) is contained in the terms I_s and $eV/(\eta\ k_BT)$. The reverse saturation current I_s nearly doubles for every 10°C increase in temperature. If $I_s = I_{s0}$ at $T = T_0$, then I_s at another temperature T is obtained from $I_s(T) = I_{s0} \times 2^{(T-T_0)/10}$ for either Ge or Si. With increasing temperature, the decrease of the term $eV/(\eta\ k_BT)$ tends to reduce the current I, but for ordinary voltages the rise in I_s predominates, so that the current I increases at a fixed voltage as the temperature is raised. The current can be brought back to its initial value by reducing the voltage.

(2) A *p-n* junction diode conducts when the applied forward voltage V exceeds the cutin voltage V_g. The forward characteristic of the diode can be roughly ap-

proximated by a straight line (Fig. 5.8A) of slope $1/r$ for $V > V_\gamma$ where r (also denoted by R_f) is the diode forward resistance. Under this condition, the equivalent circuit of the diode consists of a battery of voltage V_γ in series with a resistance r (Fig. 5.8B). The diode current is $i = (V - V_\gamma)/r$. Typically, r is a few ohm. For $V < V_\gamma$ and also for a reverse bias below breakdown, the diode acts as an open circuit. The characteristic is thus piecewise linear.

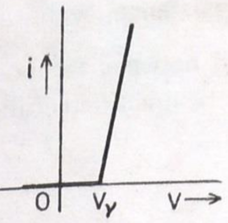

Fig. 5.8A Piecewise linear approximation of the *p-n* diode characteristic.

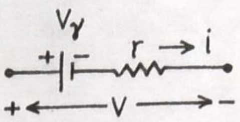

Fig. 5.8B Equivalent circuit of a conducting junction diode.

5.6 SPECIAL DIODES

(a) Breakdown Diode

Breakdown diodes are *p-n* junction diodes operated in the breakdown region of the reverse-voltage characteristic (Fig. 5.7). These diodes are designed with sufficient power-dissipation capacities to work in the breakdown region. The following two mechanisms can cause breakdown in a junction diode:

(i) **Avalanche Breakdown:** With increasing reverse-bias voltage, the electric field across the junction of a *p-n* diode increases. At a certain reverse bias, the electric field imparts a sufficiently high energy to a thermally generated carrier crossing the junction. This carrier, on colliding with a crystal ion on its way, disrupts a covalent bond and produces a new electron-hole pair. These carriers can also gain sufficient energy from the applied field and collide with other crystal ions to generate further electron-hole pairs. The process is cumulative and produces an avalanche of carriers in a very short time. This mechanism, known as *avalanche multiplication*, causes a large reverse current, as shown by the dashed part of the characteristic of Fig. 5.7. The diode is said to work in the region of *avalanche breakdown*.

(ii) **Zener Breakdown:** Zener breakdown occurs when the reverse bias voltage across a *p-n* junction is sufficiently high, so that the resulting electric field at the junction exerts a large force on a bound electron to tear it out of its covalent bond. Thus, a direct rupture of the covalent bonds produces a large number of electron-hole pairs, thereby increasing the reverse current. This process is referred to as *Zener breakdown*. Unlike avalanche breakdown, Zener breakdown does not originate from the collision of carriers with the crystal ions.

For heavily doped diodes, Zener breakdown occurs at voltages below 6 V. For lightly doped diodes the breakdown voltage is greater, and avalanche multiplication is predominant. Despite the two different mechanisms, breakdown diodes are commonly referred to as *Zener diodes*.

Zener breakdown occurs in a junction having a narrow depletion region. With a rise of temperature, the energy of the valence electron increases so that a lower applied voltage can pull these electrons out of their covalent bonds. So, the Zener breakdown voltage (which is less than 6 V) decreases with increasing temperature. That is, the temperature coefficient of the Zener breakdown voltage is *negative*. On the other hand, avalanche breakdown takes place in a junction having a wide depletion layer. With a rise of temperature, the crystal ions vibrate with a greater amplitude, increasing the possibility of collisions of the carriers and the crystal ions. The carriers thus have less opportunity to pick up enough energy between collisions to trigger the avalanche process. Consequently, the avalanche breakdown voltage (which is higher than 6 V) increases with a rise of temperature, Thus the temperature coefficient of the avalanche breakdown voltage is *positive*. The change of the breakdown voltage is about ± 0.1 percent per centigrade degree change of the diode temperature.

The volt-ampere characteristic of a Zener diode showing the breakdown region is given in Fig. 5.9. In such diodes, the doping is controlled and the surface imperfections are avoided, so that the breakdown voltage is sharp and distinct. The minimum diode current for the Zener diode to fire is I_{Z0} just below the knee of the characteristic curve. If V_Z is the operating voltage of a Zener diode and I_Z is the corresponding Zener current, the *dc resistance* of the diode is V_Z/I_Z. The *dynamic resistance* r_Z of the Zener diode is given by $r_Z = \Delta V_Z/\Delta I_Z$, where a change ΔI_Z in the operating current produces a change ΔV_Z of the operating Zener voltage Thus, r_Z is the reciprocal of the slope of the volt-ampere characteristic. In the ideal situation, $r_Z = 0$, where the Zener characteristic is parallel to the current axis. In practice, r_Z varies from a few ohm to about a few hundred ohm. Both the dc and the dynamic resistance of a Zener diode vary with the operating point.

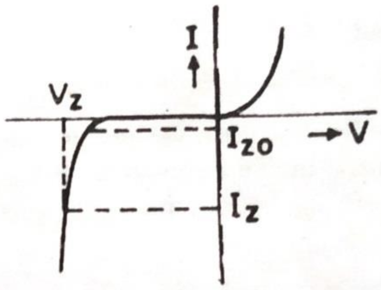

Fig. 5.9 Current-voltage characteristic of a Zener diode.

The circuit symbol for a Zener diode is shown in Fig. 5.10, Since r_Z is small, the voltage across the Zener diode changes vary little for a large change in the Zener current. Hence the voltage across a Zener diode can serve as a reference and the diode is then said to work as a *reference diode*. Fig. 5.11 depicts the application of a Zener diode as a reference diode. The supply voltage V and the series resistance R are so chosen that the diode current I_Z is within the specified range and the diode operates in the breakdown region. The voltage V_Z across the load resistance R_L remains fairly constant although the supply voltage V and the load current I_L can vary. The upper limit of the diode current depends on the power dissipation rating of the diode. If W is the power dissipation rating, the maximum allowable diode current is $I_{Zm} = W/V_Z$. If

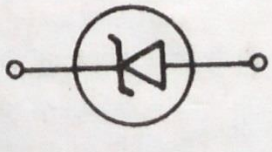
Fig. 5.10 Circuit symbol of a Zener diode.

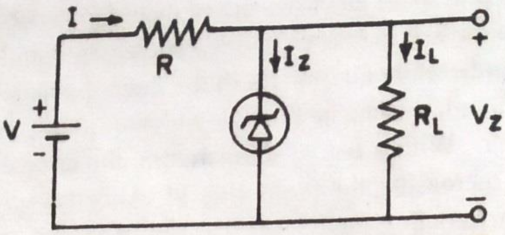

Fig. 5.11 Application of a Zener diode as a reference diode.

the current supplied by the voltage source in Fig. 5.11 is I, we have from Kirchhoff's current and voltage laws

$$I = I_Z + I_L, \tag{5.10}$$

and $$V_Z = V - IR, \tag{5.11}$$

Also, $$V_Z = I_L R_L, \tag{5.12}$$

In the ideal case, V_Z remains constant. Let the supply voltage V be kept constant and the load current be changed by varying the load resistance R_L Since $\Delta V_Z = 0$ and $\Delta V = 0$, Eq. (5.11) yields $\Delta I = 0$. Hence Eq. (5.10) gives

$$\Delta I = \Delta I_Z + \Delta I_L = 0$$

or, $$\Delta I_Z = -\Delta I_L.$$

So, when the load resistance is decreased but the supply voltage remains constant, the load current I_L increases and the Zener current I_Z decreases equally such that the supply current I remains fixed.

If, on the other hand, the load resistance R_L is kept constant and the supply voltage V is changed, we find from Eq. (5.11)

$$\Delta V = R\,\Delta I,$$

since $\Delta V_Z = 0$. Equation (5.12) gives $\Delta I_L = 0$, and Eq. (5.10) finally yields $\Delta I = \Delta I_Z$ Thus, when the supply voltage is changed and the load resistance is kept constant, the supply current I and the Zener current I_Z change equally to maintain the load current I_L at a constant value.

Observation

Two identical Zener diodes can be arranged in parallel by connecting their p-sides together and the n-sides together. The breakdown voltage of the equivalent Zener diode remains the same, but its dynamic resistance becomes half that of each diode in the combination. The current carried by the equivalent Zener diode is two times the current passing through each diode. The current-handling capacity and the wattage rating is thus increased by such parallel connection. However, since it is difficult to get identical Zener diodes, such parallel connections are rarely used.

When two Zener diodes of breakdown voltages V_1 and V_2 are connected in series, the breakdown voltage of the equivalent Zener diode is $(V_1 + V_2)$ and its dynamic resistance is the sum of the dynamic resistances of the individual diodes. Such a series connection can be used in situations where more than one reference voltage, e.g. V_1 (or V_2) and $V_1 + V_2$, is required.

(b) Varactor Diode

The name 'varactor' originates from the terms *var*iable re*actor*. It is a device, the reactance of which can be changed in a controlled manner by a bias voltage. Since a *p-n* junction diode has a voltage-dependent depletion-region capacitance, it finds application as a varactor diode, also called a *varicap*.

The depletion-region capacitance C of an abrupt *p-n* junction varies inversely as the square root of the reverse voltage. For a linearly graded junction, C varies inversely as the one-third of the reverse voltage. Therefore, the voltage sensitivity of the junction capacitance is higher for an abrupt junction than for a linearly graded junction. Special growth techniques can cause the junction capacitance to vary inversely as the square of the reverse voltage. If this junction capacitance C is used with an inductance L in a resonant circuit, the resonant frequency $1/\sqrt{LC}$ will vary linearly with the applied reverse voltage. Such varactors are employed in frequency modulation circuits.

Varactor diodes can be used in the tuning stage of a radio receiver or in a TV tuner in place of bulky variable capacitors. The circuit thereby becomes compact and its performance is improved. Varactors also find applications in harmonic generators, microwave frequency multipliers, self-balancing bridge circuits, active filters, and parametric amplifiers.

(c) Light Emitting Diode (LED)

When a *p-n* diode is forward biased, the carriers injected across the junction enhance the carrier concentrations above the thermal equilibrium values. These excess carriers recombine and emit electromagnetic energy corresponding to the band gap. For GaAs, the electromagnetic energy is released as infrared radiation. For the alloy semiconductor $GaAs_{1-x}P_x$, the energy appears as visible light. The *p-n* junction diode formed from such semiconductors is termed the *light-emitting diode* (abbreviated LED). The biasing electrical energy applied to the diode is responsible for the emitted optical energy. The electrons injected to the *p*-side of the junction drops down from the conduction band to the valence band to recombine with the holes there. The process releases photons with energy $h\nu = E_g$, where E_g is the band gap, h is Planck's constant, and ν is the frequency of the emitted radiation. Similarly, the holes injected in the *n*-side of the junction recombine with the electrons to emit photons. The electron-hole recombination and the resulting emission of light in a *p-n* diode are schematically depicted in Fig. 5.12.

The wavelength of the emitted radiation is given by

$$\lambda = \frac{hc}{E_g},$$

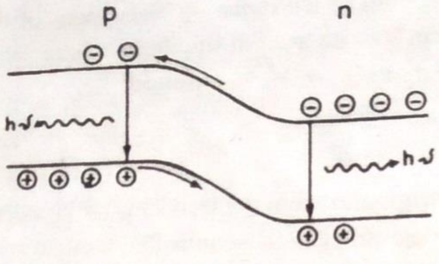

Fig. 5.12 Electron-hole recombination and consequent emission of light in a *p-n* diode.

where c is the velocity of light in free space. Substituting $h = 6.62 \times 10^{-34}$ J.s, $c = 3 \times 10^8$ m/s, and expressing E_g in eV, we obtain from the above expression

$$\lambda = \frac{1240}{E_g} \text{ nm.}$$

The light is concentrated in the vicinity of the junction because most of the carriers are concentrated there. The efficiency of light generation increases with the injected current and with a drop of temperature.

$GaAs_{1-x}P_x$ diodes have lower cost and are easily fabricated. Hence they are most commercially available. The brightness of the radiation from such diodes is a maximum when $x = 0.4$. Then $E_g = 1.9$ eV and $\lambda = 650$ nm. The emitted radiation is thus red. With $x = 0.65, 0.85$ and 1, respectively, orange, yellow, and green radiations can be generated but with less brightness and efficiency. LEDs using $GaAs_{1-x}P_x$ find applications in signal lamps and displays. The infrared LED can serve as a source for optical communication systems. The first LED was fabricated in the late 1960s.

Note that though silicon is used in many semiconductor devices, it is not an LED material. Silicon being an indirect-gap semiconductor, the recombining electrons and holes have different momenta, and so quanta of lattice vibrations or phonons must participate in the electron-hole recombination to conserve momentum. The emitted radiation would be infrared. The probability for such phonon-assisted recombination is *much smaller* than that for recombination in direct-gap semiconductor like GaAs. The recombining electrons and holes in direct-gap semiconductors have equal momentum, and so momentum is conserved without the participation of phonons. The recombination process in direct-gap semiconductors is thus very efficient.

Under certain circumstances, the emitted light can be coherent. The diode then operates as an *injection junction laser*, which is discussed below.

(d) *p-n* Junction Laser

Laser is an acronym of **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation. Basically a *p-n* junction laser is a LED, where the emitted radiation itself triggers or stimulates emission. In the LED, photons are emitted spontaneously, requiring no external stimulation by photons. However, if photons of frequency ν, where $h\nu = E_g$, are present, they can stimulate emission analogous to the forced oscillation of a classical oscillator. This phenomenon is known as *stimulated emission*. The important characteristic of stimulated emission is that, as the process is started, it is synchronized with the stimulating radiation. The emitted radiation is thus coherent.

A *p-n* junction laser has two parallel polished planes perpendicular to the junction plane (Fig. 5.13). These polished planes reflect most of the incident light. The reflected light serves as the stimulator. A small fraction of the light comes out from one end. The other surfaces of the diode are roughened to prevent stimulating radiation from other directions.

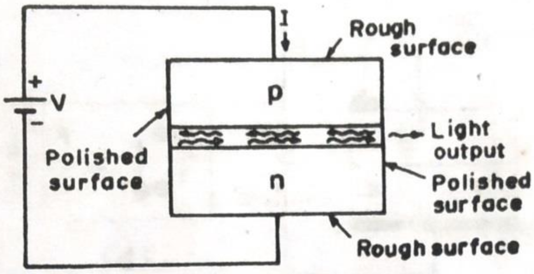

Fig. 5.13 *p-n* junction laser

Lasers are used as coherent light sources in research work, medical appliances, industry, and optical communication.

For a fuller discussion on lasers, refer to Chapter 22.

(e) Photo Diode

When a reverse-based *p-n* junction diode is illuminated with light, extra electron-hole pairs are generated in *p*-and *n*-regions. As a result, the minority carrier concentration changes significantly but the majority carrier concentration does not change much. The generated minority carriers increase the reverse current. The diode current is found to vary almost linearly with the light flux. The diodes so operated are referred to as *photo diodes*. These diodes are used in fast reading of tapes and film sound tracks, light-operated switches, and light-detection systems.

(f) Solar Cell

The solar cell is a *p-n* junction diode that transforms sunlight to electricity with a large conversion efficiency. Figure 5.14 (a) shows schematically a *p-n* junction solar cell with a load resistance R_L. When light is allowed to fall on the diode, photons are absorbed and electron-hole pairs are produced in *p*- and *n*- regions. The electrons and the holes that are generated over a short distance from the junction can reach the depletion region of width W by diffusion [Fig. 5.14 (b)].

The electron-hole pairs are then separated under the action of the strong barrier field existing across the depletion region. The electrons of the *p*-side can slide down the barrier potential to reach the *n*-side whereas the holes in the *n*-side can move towards the *p*-side [Fig. 5.14 (c)].

When the diode is open-circuited, the collection of electrons and holes on the two sides of the junction produces an *open-circuit voltage* V_{oc} [Fig. 5.14 (d)]. In the presence of a load resistance R_L, a current I flows in the circuit [Fig. 5.14(a)]. The maximum current flows when the diode terminals are externally short-circuited. This maximum current is known as the *short-circuit current* I_{sc}. The current varies as the intensity of sunlight and continues to flow as long as the diode is exposed to sunlight.

Solar cells are made of silicon, cadmium sulphide, gallium arsenide etc. They serve as long-duration power supplies in satellites and space vehicles.

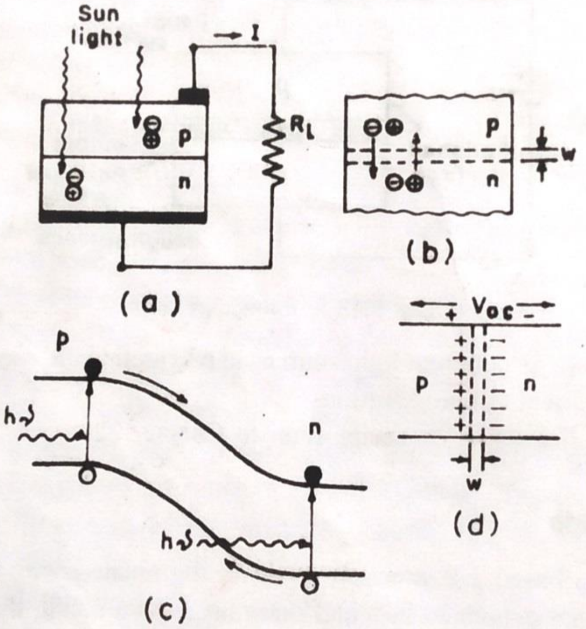

Fig. 5.14 (a) *p-n* junction solar cell having a load resistance R_L
(b) Electron and hole diffusion (c) Energy band diagram for (b), and
(d) Creation of the open-circuit voltage V_{oc}

(g) Tunnel Diode

When the concentration of the impurity atoms is very high (about 10^{24} or 10^{25} m^{-3}) in both *p*- and *n*- regions and the junction is abrupt, the depletion layer width of a *p-n* junction diode becomes very small (≈10 nm). The typical current-voltage characteristic [Fig. 5.15(a)] reveals a region of negative slope when the diode is forward biased. The occurrence of this negative slope is explained by *quantum mechanical tunneling*. Hence the diode is termed the tunnel diode.

The mechanism of the tunnel diode is discussed below. Since the doping is high, the Fermi levels lie within the bands of the semiconductor at absolute zero of temperature. In the absence of any bias voltage, the Fermi level (E_F) is constant across the junction at thermal equilibrium [(Fig. 5.15(b)]. Since E_F is the highest occupied energy level, no current can flow through the junction.

When the junction is reverse biased, the Fermi level in the *p*-side goes up relative to that in the *n*-side [Fig. 5.15(c)]. The electrons in the valence band of the *p*-side (shown by a tiny circle) are separated by a thin barrier represented by the depletion region from the available states in the conduction band of the *n*-side. Hence the electrons can tunnel from the valence band of the *p*-side to the conduction band of the *n*-side.* The greater the reverse bias, the larger the number of electrons that find available unoccupied states at the energy on the other side of the thin barrier. Consequently, the tunnel current increases with the reverse bias. The portion *OA* of the characteristic of Fig. 5.15(a) is thus explained.

When the junction is forward biased, the Fermi level in the *n*-side moves up relative to that in the *p*-side [Fig. 5.15(d)]. The electrons in the conduction band of the *n*-side now find empty allowed energy states in the valence band of the *p*-side across the barrier. The electrons can thus tunnel from the *n*-side to the *p*-side. This tunnel current is much greater than the normal current across the junction.

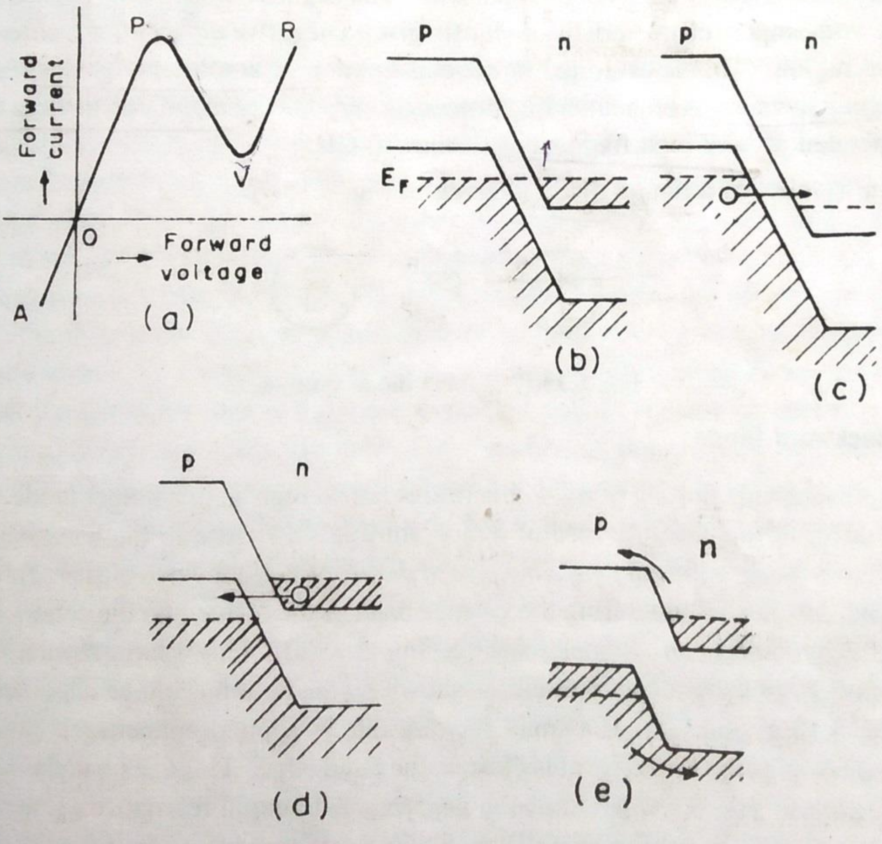

Fig. 5.15 (a) A typical tunnel diode characteristic; (b - e) Energy-band diagram of the tunnel diode; (b) zero bias, (c) reverse-bias, (d) and (e) forward bias

---

* For details of quantum mechanical tunneling, refer to the book 'Quantum Mechanics, Statistical Mechanics and Solid State Physics – An Introduction' by D. Chattopadhyay and P.C. Rakshit (S. Chand & Co., New Delhi).

5.7 SOLVED PROBLEMS

1. The reverse saturation current at 300 K of a *p-n* junction Ge diode is 5 μA. Find the voltage to be applied across the junction to obtain a forward current of 50 mA.

Ans. The current I for an applied voltage V is given for a Ge *p-n* diode by

$$I = I_s \left[\exp\left(\frac{eV}{k_B T}\right) - 1 \right]$$

Here $I = 50 \text{ mA} = 50 \times 10^{-3}$ A, $I_s = 5\ \mu\text{A} = 5 \times 10^{-6}$ A, and $T = 300$ K. So,

$$\exp\left(\frac{eV}{k_B T}\right) = \frac{I}{I_s} + 1 = \frac{50 \times 10^{-3}}{5 \times 10^{-6}} + 1 = 10^4 .$$

or

$$V = \frac{k_B T}{e} \ln 10^4 = \frac{1.38 \times 10^{-23} \times 300}{1.6 \times 10^{-19}} \times 2.303 \times 4$$

$$= 0.238 \text{ volt} .$$

2. Calculate the ratio of the current for a forward bias of 0.06 V to the current for the same value of reverse bias applied to a Ge *p-n* diode at 27°C.

Ans. The current I and the bias voltage V for a Ge diode are related by

$$I = I_s \left[\exp\left(\frac{eV}{k_B T}\right) - 1 \right],$$

where I_s is the reverse saturation current at the absolute temperature T, e is the electronic charge, and k_B is Boltzmann's constant. For a forward bias $V = 0.06$ volt, we have for the current

$$I_1 = I_s \left[\exp\left(\frac{e \times 0.06}{k_B T} \right) - 1 \right] \quad \text{(i)}$$

For a reverse bias of $V = -0.06$ volt, the current is

$$I_2 = I_s \left[\exp\left(\frac{-e \times 0.06}{k_B T} \right) - 1 \right] \quad \text{(ii)}$$

From (i) and (ii), we get

$$\frac{I_1}{I_2} = \frac{\exp\left(\frac{0.06\, e}{k_B T} \right) - 1}{\exp\left(\frac{-0.06\, e}{k_B T} \right) - 1} \quad \text{(iii)}$$

Now, $$\frac{0.06\, e}{k_B T} = \frac{0.06 \times 1.6 \times 10^{-19}}{1.38 \times 10^{-23} \times (27 + 273)} = 2.319$$

Hence $$\frac{I_1}{|I_2|} = \frac{e^{2.319} - 1}{|e^{-2.319} - 1|} = 10.16$$

3. The current flowing through a *p-n* junction Si diode is 60 mA for a forward bias of 0.9 volt at 300 K. Determine the static and dynamic resistance of the diode.

Ans. The current through the silicon diode is $I = 60$ mA $= 60 \times 10^{-3}$ A. The forward bias is $V = 0.9$ volt. So, the static resistance is

$$r_{dc} = \frac{V}{I} = \frac{0.9}{60 \times 10^{-3}} = 15 \text{ ohm}.$$

The dynamic resistance is

$$r_{ac} = \frac{26\,\eta}{I},$$

where $\eta = 2$ for Si and I is in mA.

Therefore,

$$r_{ac} = \frac{26 \times 2}{60} = 0.87 \text{ ohm}.$$

4. In the circuit of Fig. 5.11, the supply voltage is $V = 15$ volt. The 12 V, 0.36 W Zener diode operates at a minimum diode current of 2 mA. Calculate the series resistance R and the range over which the load resistance R_L can be varied.

Ans. The maximum allowable Zener current is $\frac{0.36 \text{ W}}{12 \text{ V}} = 0.03 \text{ A} = 30 \text{ mA}$. If the load resistance R_L is infinite, the Zener current I_z attains its maximum value of $I =$ 30 mA. The voltage drop across the series resistance R is $V_R = V - V_z = 15 - 12 = 3$ volt. Since the current through R is $I = 30 \text{ mA} = 30 \times 10^{-3}$ A, we have

$$R = \frac{V_R}{I} = \frac{3}{30 \times 10^{-3}} = 100 \text{ ohm}.$$

As R_L is decreased, the current I_L through it increases and the Zener current I_Z decreases correspondingly, since $I_L + I_Z = I = 30$ mA.

For the minimum value R_{Lm} of R_L, the Zener current I_Z attains is minimum value of 2 mA.

Hence, $$I_L = 30 - 2 = 28\text{mA}.$$

Therefore,

$$R_{Lm} = \frac{12 \text{ volt}}{28 \text{ mA}} = 428.6 \text{ ohm}.$$

Hence the allowable range of variation of R_L is $428.6 \ \Omega \le R_L < \infty$.

5. In the circuit of Fig.5.11, the Zener diode has the same specification as in problem 4. In the circuit, $R = 100 \ \Omega$ and $R_L = 1 \text{ k}\Omega$. Determine the limits between which the supply voltage V can vary without loss of regulation in the circuit.

(cf. C.U. 2001)

Ans. When V is a minimum, the Zener current I_z attains its minimum value of 2 mA. The load current is $I_L = 12 \text{ volt}/1 \text{ k}\Omega = 12$ mA. Hence the minimum allowable value of I is $12 + 2 = 14$ mA. The corresponding voltage drop across R is $V_R = IR = 14 \times 10^{-3} \times 100 = 1.4$ volt. Thus the minimum value of V is

$$V_{min} = V_Z + V_R = 12 + 1.4 = 13.4 \text{ volt}.$$

At the maximum value V_{max} of V, the Zener current becomes a maximum, i.e. 30 mA. Now $I = 30 + 12 = 42$ mA and $V_R = 42 \times 10^{-3} \times 100 = 4.2$ volt. Therefore,

$$V_{max} = 4.2 + 12 = 16.2 \text{ volt}.$$

So, V can vary between 13.4 volt and 16.2 volt.

REVIEW QUESTIONS

1. What is a *p-n* junction diode? How does a barrier field appear across a *p-n* junction?
2. For an unbiased *p-n* junction, sketch the variation of the space charge, electric field, electrostatic potential, and electron energy as a function of distance across the junction.
3. What is an abrupt and a linearly graded *p-n* junction?