

## **UNIT-II: Semiconductor Physics**

Intrinsic and Extrinsic semiconductors, Dependence of Fermi level on carrier-concentration and temperature, Carrier generation and recombination, Carrier transport: diffusion and drift, Hall effect, p-n junction diode, Zener diode and their V-I Characteristics, Bipolar Junction Transistor (BJT): Construction, Principle of operation.

### **Unit II Semiconductor Physics**

#### **INTRODUCTION**

A Semiconductor is a solid which behaves as an insulator at absolute zero whose resistivity lies between  $10^{-4}$  to  $0.5 \Omega\text{-m}$ . They have negative temperature coefficient of resistance in which the current conduction is carried out by electrons and holes.

There are two types of semiconductors.

1. Intrinsic Semiconductor
2. Extrinsic Semiconductor

Semiconductors are materials whose electronic properties are intermediate between those of conductors and insulators. These electrical properties of a solid depend on its band structure. A semiconductor has two bands of importance (neglecting bound electrons as they play no part in the conduction process) the valence and the conduction bands. They are separated by a forbidden energy gap. At OK the valence band is full and the conduction band is empty, the semiconductor behaves as an insulator. Semiconductor has both positive (hole) and negative (electron) carriers of electricity whose densities can be controlled by doping the pure semiconductor with chemical impurities during crystal growth.

At higher temperatures, electrons are transferred across the gap into the conduction band leaving vacant levels in the valence band. It is this property that makes the semiconductor a material with special properties of electrical conduction.

Generally there are two types of semiconductors. Those in which electrons and holes are produced by thermal activation in pure Ge and Si are called intrinsic semiconductors. In other type the current carriers, holes or free electrons are produced by the addition of small quantities of elements of group III

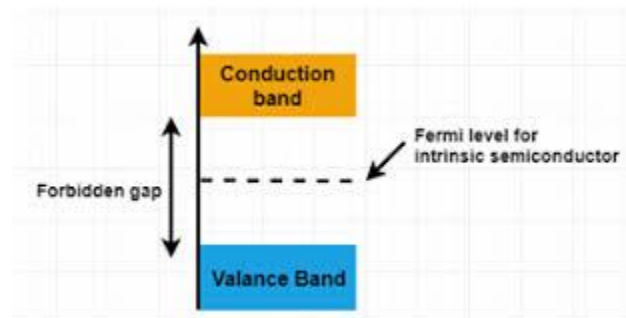
or V of the periodic table, and are known as extrinsic semiconductors. The elements added are called the impurities or dopants.

## Intrinsic semiconductors:

A pure semiconductor which is not doped is termed as intrinsic semiconductor. In Si crystal, the four valence electrons of each Si atom are shared by the four surrounding Si atoms. An electron which may break away from the bond leaves deficiency of one electron in the bond. The vacancy created in a bond due to the departure of an electron is called a hole. The vacancy may get filled by an electron from the neighboring bond, but the hole then shifts to the neighboring bond which in turn may get filled by electron from another bond to whose place the hole shifts, and so on thus in effect the hole also undergoes displacement inside a crystal. Since the hole is associated with deficiency of one electron, it is equivalent for a positive charge of unit magnitude. Hence in a semiconductor, both the electron and the hole act as charge carriers. In an intrinsic semiconductor, for every electron freed from the bond, there will be one hole created. It means that, the no of conduction electrons is equal to the no of holes at any given temperature. Therefore there is no predominance of one over the other to be particularly designated as charge carriers.

### Intrinsic Semiconductor

A semiconductor in which charge carriers are created only by increasing temperature is known as an intrinsic semiconductor.



### Expression for density of electrons in the conduction band

The number of electrons per unit volume in the conduction band at an equilibrium temperature is called density (concentration) of electrons.

Let  $dn$  be the number of electrons in the conduction band in the energy interval  $dE$  between  $E$  and  $E+dE$  in the conduction band at any temperature  $T$ .

$$dn = Z(E) F(E) dE$$

Where  $Z(E) dE$  is the density of energy states in the energy interval  $dE$  and  $F(E)$  is the probability of occupation of an electron in the energy state  $E$ .

If  $E_c$  is the bottom of the conduction band, then the density of electrons in the conduction band is

$$n = \int_{E_c}^{\infty} Z(E) F(E) dE$$

Now,  $Z(E)dE = \frac{4\pi}{h^3} (2m)^{3/2} E^{1/2} dE$  can be written as

$$Z(E)dE = \frac{4\pi}{h^3} (2m_e^*)^{3/2} E^{1/2} dE, \text{ where } m^* \text{ is the effective mass of electron.}$$

Since  $E_c$  is the bottom of the conduction band,

$$Z(E)dE = \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_c)^{1/2} dE$$

$$\text{And } F(E) = \frac{1}{1 + \exp(\frac{E - E_F}{kT})} = \exp(-\frac{E - E_F}{kT}) = \exp(\frac{E_F - E}{kT}) \quad (\text{since } E - E_F \gg kT)$$

$$n = \int_{E_c}^{\infty} \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_c)^{1/2} \exp(\frac{E_F - E}{kT}) dE$$

$$n = \frac{4\pi}{h^3} (2m_e^*)^{3/2} \int_{E_c}^{\infty} (E - E_c)^{1/2} \exp(\frac{E_F - E}{kT}) dE$$

$$n = \frac{4\pi}{h^3} (2m_e^*)^{3/2} \exp(\frac{E_F}{kT}) \int_{E_c}^{\infty} (E - E_c)^{1/2} \exp(\frac{-E}{kT}) dE$$

Let  $E - E_c = x \Rightarrow E = E_c + x$  and  $dE = dx$

$$n = \frac{4\pi}{h^3} (2m_e^*)^{3/2} \exp(\frac{E_F}{kT}) \int_0^{\infty} x^{1/2} \exp(-\frac{E_c + x}{kT}) dx$$

$$n = \frac{4\pi}{h^3} (2m_e^*)^{3/2} \exp(\frac{E_F - E_c}{kT}) \int_0^{\infty} x^{1/2} \exp(-\frac{x}{kT}) dx$$

By gamma function  $\int_0^{\infty} x^{1/2} \exp(-\frac{x}{kT}) dx = (kT)^{3/2} \frac{\pi^{1/2}}{2}$

$$n = \frac{4\pi}{h^3} (2m_e^*)^{3/2} \exp(\frac{E_F - E_c}{kT}) (kT)^{3/2} \frac{\pi^{1/2}}{2}$$

$$\boxed{\therefore n = 2 \left( \frac{2\pi m_e^* kT}{h^2} \right)^{3/2} \exp(\frac{E_F - E_c}{kT})}$$

### Expression for density of holes in the valence band

The number of holes per unit volume in the valence band at an equilibrium temperature is called density of holes. Let  $dp$  be the number of holes in the energy interval  $dE$  between  $E$  and  $E + dE$  in the valence band at any temperature  $T$ .

$$dp = Z(E)[1 - F(E)] dE$$

where  $Z(E) dE$  is the density of energy states in the energy interval  $dE$  in the valence band and  $[1 - F(E)]$  is the probability of presence of a hole (absence of an electron) in the energy state  $E$ . If  $E_v$  is the energy corresponding to the top of the valence band, then the density of holes in the valence band is

$$p = \int_{-\infty}^{E_v} Z(E)[1 - F(E)] dE$$

Now,  $Z(E)dE = \frac{4\pi}{h^3} (2m)^{3/2} E^{1/2} dE$  can be written as

$$Z(E)dE = \frac{4\pi}{h^3} (2m_h^*)^{3/2} E^{1/2} dE, \text{ where } m^* \text{ is the effective mass of hole.}$$

Since  $E_v$  is the top of the valence band,  $Z(E)dE = \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_v - E)^{1/2} dE$

$$\text{And } [1 - F(E)] = 1 - \frac{1}{1 + \exp(\frac{E - E_F}{kT})} = \exp(\frac{E - E_F}{kT})$$

$$p = \int_{-\infty}^{E_v} \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_v - E)^{1/2} \exp(\frac{E - E_F}{kT}) dE$$

$$p = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \int_{-\infty}^{E_v} (E_v - E)^{1/2} \exp(\frac{E - E_F}{kT}) dE$$

$$p = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \exp(\frac{-E_F}{kT}) \int_{-\infty}^{E_v} (E_v - E)^{1/2} \exp(\frac{E}{kT}) dE$$

Let  $E_v - E = x \Rightarrow E = E_v - x$  and  $dE = -dx$

$$p = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \exp(\frac{E_F}{kT}) \int_0^{\infty} x^{1/2} \exp(\frac{E_v - x}{kT}) dx$$

$$p = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \exp(\frac{E_v - E_F}{kT}) \int_0^{\infty} x^{1/2} \exp(-\frac{x}{kT}) dx$$

By gamma function  $\int_0^{\infty} x^{1/2} \exp(-\frac{x}{kT}) dx = (kT)^{3/2} \frac{\pi^{1/2}}{2}$

$$p = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \exp(\frac{E_v - E_F}{kT}) (kT)^{3/2} \frac{\pi^{1/2}}{2}$$

$$\therefore p = 2 \left( \frac{2\pi m_h^* kT}{h^2} \right)^{3/2} \exp(\frac{E_v - E_F}{kT})$$

### Intrinsic Carrier Concentration ( $n_i$ )

The number of charge carriers per unit volume in an intrinsic semiconductor is known as intrinsic carrier concentration. It is numerically equal to the square root of the product of densities of the electrons in the conduction band and holes in the valence band of an intrinsic semiconductor.

In an intrinsic semiconductor  $n=p=n_i \Rightarrow n_i^2=n \cdot p$

$$\begin{aligned} \text{i.e., } n_i^2 &= 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} \exp\left(\frac{E_F - E_c}{kT}\right) \cdot 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} \exp\left(\frac{E_v - E_F}{kT}\right) \\ \Rightarrow n_i^2 &= 4\left(\frac{2\pi kT}{h^2}\right)^3 (m_e^* m_h^*)^{3/2} \exp\left(\frac{E_v - E_c}{kT}\right) \\ \Rightarrow n_i &= 2\left(\frac{2\pi kT}{h^2}\right)^{3/2} (m_e^* m_h^*)^{3/4} \exp\left(\frac{E_v - E_c}{2kT}\right) \\ \therefore n_i &= 2\left(\frac{2\pi kT}{h^2}\right)^{3/2} (m_e^* m_h^*)^{3/4} \exp\left(\frac{-E_g}{2kT}\right), \text{ where } E_g = E_c - E_v = \text{band gap} \end{aligned}$$

### Fermi Level in an intrinsic semiconductor

For an intrinsic semiconductor, **density of electrons in CB ( $n$ ) = density of holes in VB ( $p$ )**

$$\text{i.e., } 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} \exp\left(\frac{E_F - E_c}{kT}\right) = 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} \exp\left(\frac{E_v - E_F}{kT}\right) \quad \rightarrow [1]$$

$$\Rightarrow \exp\left[\frac{2E_F - (E_v + E_c)}{kT}\right] = \left(\frac{m_h^*}{m_e^*}\right)^{3/2} \quad \rightarrow [2]$$

$$\Rightarrow \frac{2E_F}{kT} = \frac{(E_v + E_c)}{kT} + \frac{3}{2} \log\left(\frac{m_h^*}{m_e^*}\right)$$

$$\Rightarrow E_F = \frac{(E_v + E_c)}{2} + \frac{3kT}{4} \log\left(\frac{m_h^*}{m_e^*}\right) \quad \rightarrow [3]$$

$$\text{When } m_e^* = m_h^*, \quad E_F = \frac{E_v + E_c}{2}$$

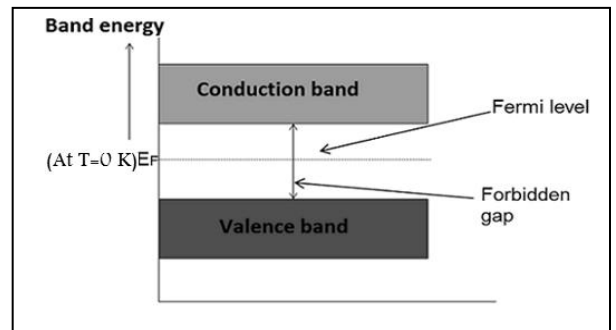


fig 2.1

It is clear from [3] that if the temperature is increased above 0 K, the Fermi level moves towards the bottom of the conduction band.

### Extrinsic Semiconductors

A semiconductor, in which charge carriers are created not only by increasing temperature but also by adding impurities to it, is known as an extrinsic semiconductor.

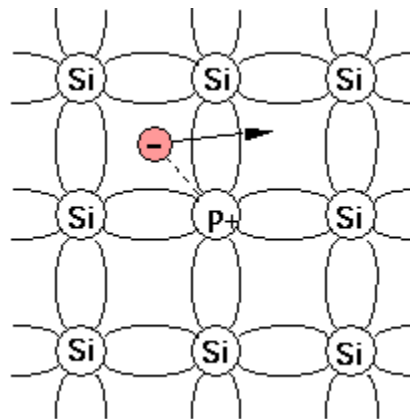
#### EXTRINSIC SEMICONDUCTORS:

Intrinsic Semiconductors are rarely used in semiconductor devices as their conductivity is not sufficiently high. The electrical conductivity is extremely sensitive to certain types of impurity. It is the ability to modify electrical characteristics of the material by adding chosen impurities that make extrinsic semiconductors important and interesting.

Addition of appropriate quantities of chosen impurities is called doping, usually, only minute quantities of dopants (1 part in  $10^3$  to  $10^{10}$ ) are required. Extrinsic or doped semiconductors are classified into main two main types according to the type of charge carries that predominate. They are the n-type and the p-type.

### **N-TYPE SEMICONDUCTORS :**

Doping with a pentavalent impurity like phosphorous, arsenic or antimony the semiconductor becomes rich in conduction electrons. It is called n-type the bond structure of an n-type semiconductor is shown in Fig below.

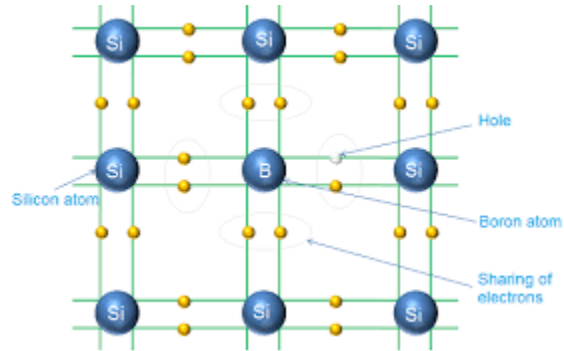


Even at room temperature, nearly all impurity atoms lose an electron into the conduction band by thermal ionization. The additional electrons contribute to the conductivity in the same way as those excited thermally from the valence bond. The essential difference beam

### **P-TYPE SEMICONDUCTORS:-**

p-type semiconductors have holes as majority charge carries. They are produced by doping an intrinsic semiconductor with trivalent impurities.(e.g. boron, aluminium, gallium, or indium). These dopants have

three valence electrons in their outer shell. Each impurity is short of one electron for covalent bonding. The vacancy thus created is bound to the atom at OK. It is not a hole. But at some higher temperature an electron from a neighbouring atom can fill the vacancy leaving a hole in the valence bond for conduction. It behaves as a positively charge particle of effective mass  $m_h^*$ . The bond structure of a ptype semiconductor is shown in Fig below.



### Carrier concentration of n-type semiconductor

N-type semiconductor is obtained by doping an intrinsic semiconductor with pentavalent impurity atoms like phosphorus, arsenic, antimony etc. In this type of semiconductor electrons are the majority charge carriers.

#### Expression for carrier concentration:

The energy level diagram of a N-type semiconductor is shown in fig. Density of electrons in the conduction band is given as

$$n = 2 \left( \frac{2\pi m_e^* kT}{h^2} \right)^{3/2} \exp\left(\frac{E_F - E_C}{kT}\right) \rightarrow [1]$$

Let  $N_d$  be the donor concentration and  $E_d$  be the donor energy level. At very low temperatures all donor levels are filled with electrons. With increase in temperature the ionization of donor impurity atoms takes place and density of electrons in the conduction band increases.

$$\text{Number of ionised impurities} = N_d [1 - F(E_d)] \approx N_d \exp\left(\frac{E_d - E_F}{kT}\right)$$

At low temperatures,

The no. of ionised donor impurities = no. of electrons in the conduction band  $\rightarrow [2]$

$$2 \left( \frac{2\pi m_e^* kT}{h^2} \right)^{3/2} \exp\left(\frac{E_F - E_C}{kT}\right) = N_d \exp\left(\frac{E_d - E_F}{kT}\right) \rightarrow [3]$$

Taking logarithm and rearranging, we get

$$\left(\frac{E_F - E_C}{kT}\right) - \left(\frac{E_d - E_F}{kT}\right) = \log N_d - \log 2 \left( \frac{2\pi m_e^* kT}{h^2} \right)^{3/2}$$

$$2E_F - (E_d + E_C) = kT \log \frac{N_d}{2 \left( \frac{2\pi m_e^* kT}{h^2} \right)^{3/2}}$$

$$E_F = \frac{(E_d + E_C)}{2} + \frac{kT}{2} \log \frac{N_d}{2 \left( \frac{2\pi m_e * kT}{h^2} \right)^{3/2}} \rightarrow [4]$$

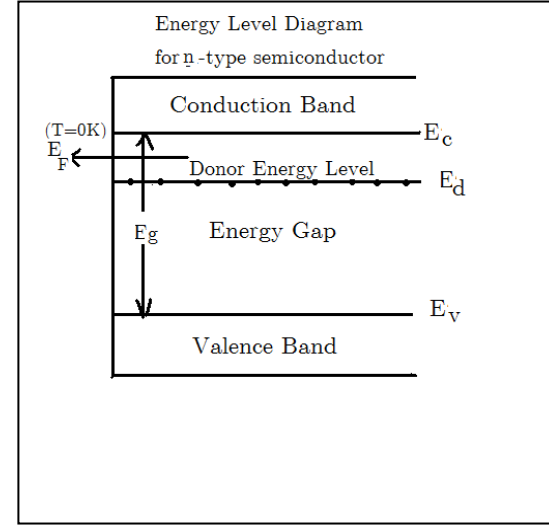
$$\text{At } T=0 \text{ K, } E_F = \frac{(E_d + E_C)}{2}$$

At 0 K, Fermi level lies exactly at the middle of the donor level  $E_d$  and the bottom of the conduction band  $E_c$  as shown in fig. Now, substituting  $E_F$  from equation [4] in equation [1],

$$\exp\left(\frac{E_F - E_C}{kT}\right) = \exp\left\{\left(\frac{E_d - E_C}{2kT}\right) + \frac{1}{2} \log \frac{N_d}{2 \left( \frac{2\pi m_e * kT}{h^2} \right)^{3/2}}\right\} \rightarrow [5]$$

$$= \exp\left\{\left(\frac{E_d - E_C}{2kT}\right) + \log \frac{N_d^{1/2}}{\left[2 \left( \frac{2\pi m_e * kT}{h^2} \right)^{3/2}\right]^{1/2}}\right\}$$

$$= \left[ \exp\left(\frac{E_d - E_C}{2kT}\right) \frac{N_d^{1/2}}{\left[2 \left( \frac{2\pi m_e * kT}{h^2} \right)^{3/2}\right]^{1/2}} \right]$$



**Fig. 2.2**

$$\therefore n = 2 \left( \frac{2\pi m_e * kT}{h^2} \right)^{3/2} \frac{N_d^{1/2}}{\left[2 \left( \frac{2\pi m_e * kT}{h^2} \right)^{3/2}\right]^{1/2}} \exp\left(\frac{E_d - E_C}{2kT}\right) \rightarrow [6]$$

$$\therefore n = (2N_d)^{1/2} \left( \frac{2\pi m_e * kT}{h^2} \right)^{3/4} \exp\left(\frac{E_d - E_C}{2kT}\right) \rightarrow [7]$$

Thus it is evident from equation [7], at low temperatures; the density of electrons in the conduction band is proportional to the square root of donor concentration in an N-type semiconductor.

### Carrier concentration of p-type semiconductor

P-type semiconductor is obtained by doping an intrinsic semiconductor with trivalent impurity atoms like aluminum, gallium, indium etc. In this type of semiconductor holes are the majority charge carriers.



### Expression for carrier concentration:

The energy level diagram of a P-type semiconductor is shown in fig. Density of holes in the valence band is given as

$$p = 2 \left( \frac{2\pi m_h^* kT}{h^2} \right)^{3/2} \exp\left(\frac{E_v - E_F}{kT}\right) \rightarrow [1]$$

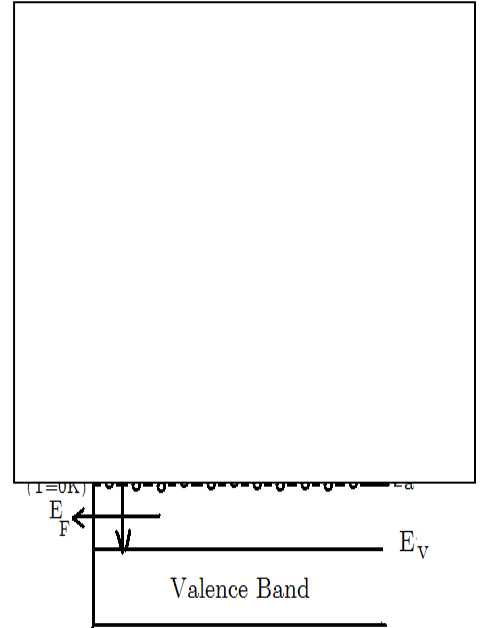
Let  $N_a$  be the acceptor concentration and  $E_a$  be the acceptor energy level. At very low temperatures all acceptor levels are

empty. With increase in temperature the ionization of acceptor impurity atoms takes place i.e., the electrons move from valence band to and occupy the vacant sites in the acceptor level by leaving holes in the valence band.

$$\text{Number of ionised impurities} = N_a F(E_a) \approx N_a \exp\left(\frac{E_F - E_a}{kT}\right)$$

At low temperatures, the no. of ionized acceptor impurities = no. of holes in the valence band  $\rightarrow [2]$

$$2 \left( \frac{2\pi m_h^* kT}{h^2} \right)^{3/2} \exp\left(\frac{E_v - E_F}{kT}\right) = N_a \exp\left(\frac{E_F - E_a}{kT}\right) \rightarrow [3]$$



Taking logarithm and rearranging, we get

$$\left( \frac{E_v + E_a - 2E_F}{kT} \right) = \log \frac{N_a}{2 \left( \frac{2\pi m_h^* kT}{h^2} \right)^{3/2}}$$

$$E_F = \frac{(E_v + E_a)}{2} - \frac{kT}{2} \log \frac{N_a}{2 \left( \frac{2\pi m_h^* kT}{h^2} \right)^{3/2}} \rightarrow [4]$$

$$\text{At } T=0 \text{ K, } E_F = \frac{(E_v + E_a)}{2}$$

Fig.5.2

At 0 K, Fermi level lies exactly at the middle of the acceptor level  $E_a$  and the top of the valence band  $E_v$  as shown in fig. Now, substituting  $E_F$  from equation [4] in equation [1],

$$\exp\left(\frac{E_v - E_F}{kT}\right) = \exp\left\{\left(\frac{E_v - E_a}{2kT}\right) + \log \frac{N_a^{1/2}}{\left[2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}\right]^{1/2}}\right\} \rightarrow [5]$$

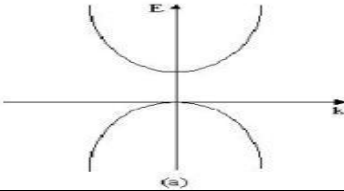
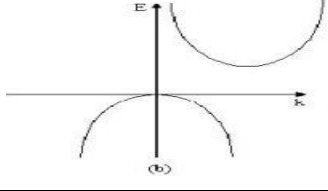
$$= \left[ \exp\left(\frac{E_v - E_a}{kT}\right) \frac{N_a^{1/2}}{\left[2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}\right]^{1/2}} \right]$$

$$\therefore p = 2 \left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} \frac{N_a^{1/2}}{\left[2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2}\right]^{1/2}} \exp\left(\frac{E_v - E_a}{2kT}\right) \rightarrow [6]$$

$$\therefore p = (2N_a)^{1/2} \left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/4} \exp\left(\frac{E_v - E_a}{2kT}\right) \rightarrow [7]$$

Thus it is evident from equation [7], at low temperatures; the density of holes in the valence band is proportional to the square root of acceptor concentration in a P-type semiconductor.

### Direct and Indirect Band gap semiconductors

| S. No | Direct Band gap semiconductors  | Indirect Band gap semiconductors   |
|-------|---|--|
| 1.    |    |    |
| 2.    | The semiconductors in which the maximum of the valence band coincides with the minimum of the conduction band for the same value of the propagation constant $K$ is known as a direct band gap semiconductor. | The semiconductors in which the maximum of the valence band does not coincide with the minimum of the conduction band for the same value of the propagation constant $K$ is known as an indirect band gap semiconductor. |
| 3.    | In these semiconductors the electron-hole recombination takes place directly from conduction band to valence band.  | In these semiconductors the electron-hole recombination does not take place directly from conduction band to valence band but via a trap centre in the forbidden band.   |
| 4.    | Life time of the charge carriers is less due to direct recombination  | Life time of the charge carriers is more due to indirect recombination   |
| 5.    | Current amplification is less   | Current amplification is more  |

|    |   |  |
|----|---|--|
| 6. | Light is produced due to recombinations | Heat is produced due to recombinations |
|----|---|--|

**Radiative and non radiative recombination mechanism in semiconductors**

Like other solids, semiconductor materials have an [electronic band structure](#) determined by the crystal properties of the material. Energy distribution among electrons is described by the [Fermi level](#) and the [temperature](#) of the electrons.

At [absolute zero](#) temperature, all of the electrons have energy below the Fermi level; but at non-zero temperatures the energy levels are filled following a Boltzmann distribution.

In undoped semiconductors the Fermi level lies in the middle of a *forbidden band* or [band gap](#) between two *allowed bands* called the [valence band](#) and the [conduction band](#). The valence band, immediately below the forbidden band, is normally very nearly completely occupied. The conduction band, above the Fermi level, is normally nearly completely empty. Because the valence band is so nearly full, its electrons are not mobile, and cannot flow as electric current.

However, if an electron in the valence band acquires enough energy to reach the conduction band, it can flow freely among the nearly empty conduction band energy states. Furthermore, it will also leave behind a hole that can flow as current exactly like a physical charged particle.

Carrier generation describes processes by which electrons gain energy and move from the valence band to the conduction band, producing two mobile carriers; while recombination describes processes by which a conduction band electron loses energy and re-occupies the energy state of an electron hole in the valence band.

.

Recombination and generation are always happening in semiconductors, both optically and thermally, a material at [thermal equilibrium](#) will have generation and recombination rates that are balanced so that the net [charge carrier](#) density remains constant..

## Radiative Recombination

### Radiative recombination

*Band-to-band recombination* is the name for the process of electrons jumping down from the conduction band to the valence band in a radioactive manner. During band-to-band recombination, a form of [spontaneous emission](#), the energy absorbed by a material is released in the form of [photons](#). Generally these photons contain the same or [less](#) energy than those initially absorbed. This effect

is how [LEDs](#) create light. Because the photon carries relatively little [momentum](#), radioactive recombination is significant only in [direct band gap](#) materials. This process is also known as *bimolecular recombination*<sup>[3]</sup>.

### Non-radiative recombination

Non-radiative recombination is a process in [phosphors](#) and [semiconductors](#), whereby [charge carriers](#) recombine with releasing [phonon](#) instead of [photons](#). Non-radiative recombination in optoelectronics and phosphors is an unwanted process, lowering the light generation efficiency and increasing heat losses.

Non-radiative life time is the average time before an [electron](#) in the [conduction band](#) of a [semiconductor](#) recombines with a [hole](#). It is an important parameter in [optoelectronics](#) where [radiative recombination](#) is required to produce a [photon](#); if the non-radiative life time is shorter than the radiative, a carrier is more likely to recombine non-radiatively. This results in low internal [quantum efficiency](#).

### **DRIFT CURRENT**

In an electric field  $E$ , the drift velocity  $V_d$  of carriers superposes on the thermal velocity  $V_{th}$ . But the flow of charge carriers results in an electric current, known as the drift current. Let a field  $E$  be applied, in the positive direction creating drift currents  $J_{nd}$  and  $J_{pd}$  of electrons and holes respectively.

Without  $E$ , the carriers move randomly with rms velocity  $V_{th}$ . Their mean velocity is zero. The current density will be zero. But the field  $E$  applied, the electrons have the velocity  $V_{de}$  and the holes  $V_{dh}$ .

Consider free electrons in a Semiconductor moving with uniform velocity  $V_{de}$  in the

negative  $x$  direction due to an electric field  $E$ . Consider a smaller rectangular block of  $AB$  of

length  $V_{de}$  inside the Semiconductor. Let the area of the side faces each be unity. The total

charge  $Q$  in the elements  $AB$  is

$Q = \text{Volume of the element} \times \text{density of particles} \times \text{charge on each particle}$

$$= (V_{de} \times 1 \times 1) \times n \times -q$$

$$\text{Thus } Q = -qnV_{de}$$

Where  $n$  is the number density of electrons. The entire charge of the block will cross

the face  $B$ , in unit time. Thus the drift current density  $J_{nd}$  due to free electrons at the face  $B$

will be.

$$J_{nd} = -q n V_{de}$$

Similarly for holes  $J_{pd} = q n V_{dh}$

but  $V_{de} = -E$

and  $V_{dh} = \cdot_p E$

hence  $J_{nd} = n q \cdot_n E$

and  $J_{pd} = p q \cdot_p E$

The total drift current due to both electrons and holes  $J_d$  is

$$J_d = J_{nd} + J_{pd} = (nq \cdot_n + pq \cdot_p) E$$

Even though electrons and holes move in opposite direction the effective direction of current

flow, is the same for both and hence they get added up. Ohm's Law can be written in terms

of electrical conductivity, as

$$J_d = \cdot E$$

Equating the RHS of eq we have

$$\cdot = nq \cdot_n + pq \cdot_p = \cdot_n + \cdot_p$$

For an intrinsic Semiconductor  $n = p = n_i$

$$\cdot_i = n_i q (\cdot_n + \cdot_p)$$

## DIFFUSION CURRENTS:

1. Diffusion Current: Electric current is Setup by the directed movement of charge carriers. The movements of charge carriers could be due to either drift or diffusion. Non uniform

Concentration of carriers gives rise to diffusion. The first law of diffusion by Fick

States that the flux  $F$ , i.e., the particle current is proportional and is directed to opposite to the

Concentration gradient of particles. It can be written mathematically, in terms of Concentration  $N$ , as

$$F = -D \nabla N$$

Where  $D$  stands for diffusion constant.

In one dimension it is written as

$$F = -D \cdot \nabla N$$

$\cdot x$



In terms of  $J_e$  and  $J_p$  the flux densities of electrons holes and their densities  $n$  and  $p$  respectively.

We get  $J_e = -D_n \cdot n \cdot x$

and  $J_p = -D_p \cdot p$

$\cdot x$

Where  $D_n$  and  $D_p$  are the electron and hole diffusion constant constants respectively.

Then the diffusion current densities become

$J_{n \text{ diff}} = q D_n \cdot n$

$\cdot x$

$J_{p \text{ diff}} = - q D_p \cdot p \cdot x$

**HALL EFFECT:**

When a material carrying current is subjected to a magnetic field in a direction perpendicular to the direction of current, an electric field is developed across the material in a

direction perpendicular to both the direction of the magnetic field and the current direction.

This phenomenon is called Hall Effect.

Hall Effect finds important application in studying the electron properties of semiconductor, such as determination of carrier concentration and carrier mobility. It also used to

determine whether a semiconductor is n-type, or p-type.

#### THEORY:

Consider a rectangular slab of an n-type Semiconductor carrying current in the positive x-direction. The magnetic field  $B$  is acting in the positive direction as indicated in

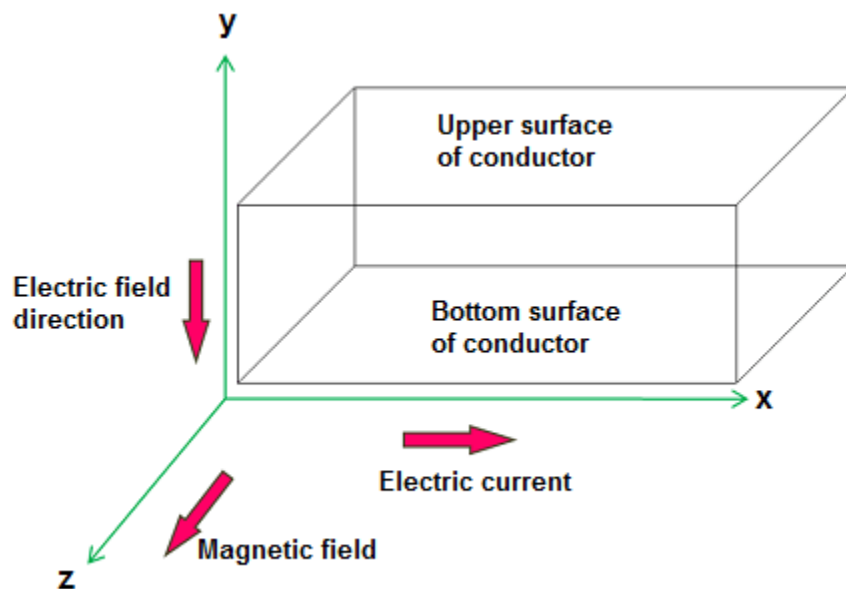
fig above. Under the influence of the magnetic field, electrons experience a force  $F_L$  given

by

$$F_L = - Bev \text{ ----- (1)}$$

Where  $e$  = magnitude of the charge of the electron

$v$  = drift velocity



Applying the Fleming's Left Hand Rule, it indicates a force  $F_H$  acting on the electrons in the

negative  $y$ -direction and electron are deflected down wards. As a consequence the lower face

of the specimen gets negatively charged (due to increases of electrons) and the upper face

positively charged (due to loss of electrons).

Hence a potential  $V_H$ , called the Hall voltage appears between the top and bottom faces of the specimen, which establishes an electric field

$E_H$ , called the Hall field across the conductor in negative  $y$ -direction. The field  $E_H$  exerts an

upward force  $F_H$  on the electrons. It is given by

$$F_H = -eE_H \text{-----(2)}$$

upward direction. The two opposing forces  $F_L$  and  $F_H$  establish an equilibrium under which

$$|F_L| = F_H$$

using eqns 1 and 2

$$-Bev = -eE_H$$

$$E_H = Bv \text{-----(3)}$$

If 'd' is the thickness of the Specimen

$$E_H = V_h/d$$

$$v_h = E_H d = Bvd \text{ from eqn (3)----- 4}$$

If  $\omega$  is the width of the specimen in z- direction.

The current density

$$J = I / \omega d$$

But  $J = nev = \rho v$  ----- 5

Where  $n$  = electron concentration

And  $\rho$  charge density

$$\rho v = I / \omega d$$

$$\text{Or } v = I / \rho \omega d \text{-----6}$$

Substitution for  $V$ , FROM 6 AND 4

$$V_H = BI / \rho \omega$$

$$\text{or } \rho = BI / V_H \omega$$

**Thus, by measuring  $V_H$ ,  $I$ , and  $\omega$  and by knowing  $B$ , the charge density  $\rho$  can be determined.**

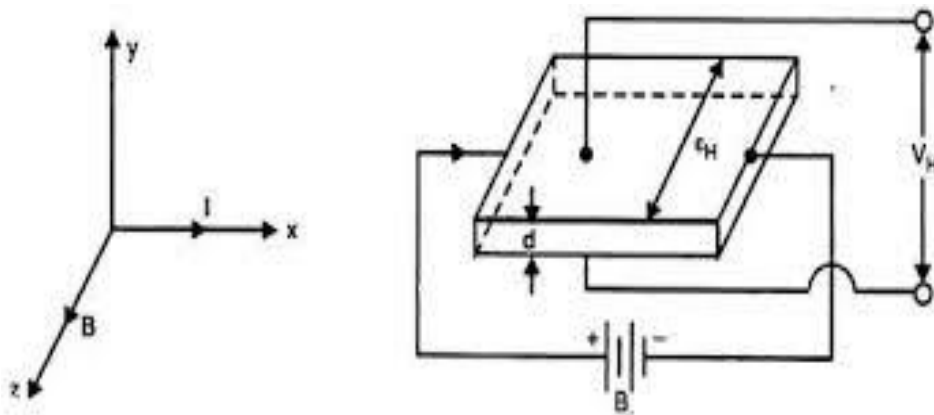


Fig. 7.22. Set up for the measurement of Hall voltage.

The Hall field  $E_H$ , for a given material depends on the current density  $J$ , and the

applied field B

i.e.,  $E_H \propto JB$  (propositional to)

$$E_H = R_H JB$$

Where  $R_H$  is called the Hall Coefficient

BI

Since  $V_H = BI / \rho \omega$

$$E_H = V_H / d \omega$$

$$J = I / d \omega$$

$$BI / \rho \omega J = R_H BI / d \omega$$

**This leads to  $R_H = I / \rho$**

Mobility of charge carriers:

The mobility  $\mu$  is given by  $\mu = V/E$

$$\text{But } J = \sigma E = nev = \rho v$$

Therefore  $\sigma E = \rho v$

Or  $E = \rho v / \sigma$

$$\mu = \sigma / \rho = \sigma R_H \quad (1/\rho = R_H)$$

$\sigma$  is the conductivity of the semiconductor.

### (C) Applications

Determination of the type of Semiconductor:

The Hall Coefficient  $R_H$  is negative for an n-type Semiconductor and positive for a p-type material. Thus,

the sign of the Hall coefficient can be utilized to determine whether a given Semiconductor is n or p type.

(b) Determination of Carrier Concentration: Equation relates the Hall Coefficient  $R_H$  and charge density is

$$1 - 1$$

$$R_H = 1 / p = -1 / ne \quad (\text{for n-type})$$

$$R_H = 1/ne \quad (\text{for p-type})$$

$$\text{Thus } n = 1 / e R_H$$

$$\text{And } 1 / e R_H$$

Determination of mobility:

According to equation the mobility of charge carriers is given by

$$\mu = \sigma |R_H|$$

Determination of  $\sigma$  and  $R_H$  leads to a value of mobility of charge carriers.

Measurement of Magnetic Induction (B):-

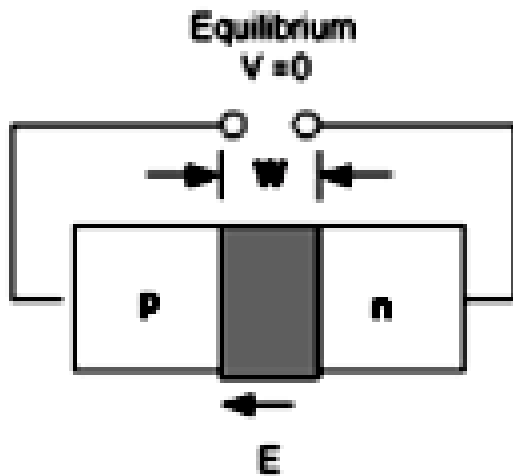
The Hall Voltage is proportional to the flux density B. As such measurement of  $V_H$  can be used to estimate B.

P      P-N Junction Diode

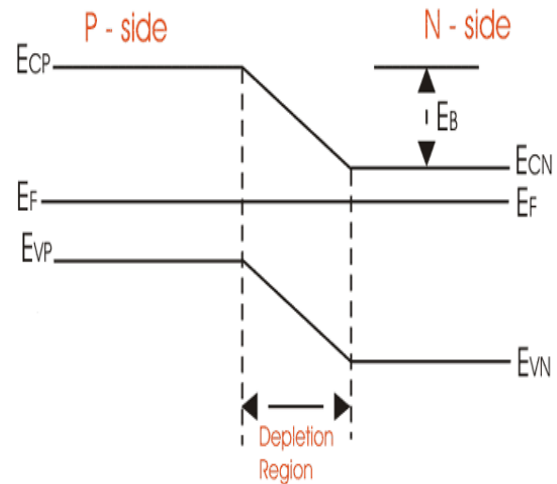
A P-N Junction diode is formed by placing a p-type crystal in contact with n-type crystal and subjecting to high pressure so that it becomes a single piece. The assembly so obtained is called p-n junction or junction diode. The p-type region has holes as majority      n      chargecarriers. Similarly the n-type region has electrons as majority charge carriers.

Formation of P-N junction





Open circuited P-N junction



Energy band diagram of a P-N

Joining n-type material with p-type material causes excess electrons from the n-type material to diffuse to the p-type side and excess holes from the p-type material to diffuse to the n-type side. Movement of electrons to the p-type side exposes positive ion cores in the n-type side while movement of holes to the n-type side exposes negative ion cores in the p-type side, resulting in an electric field at the junction and forming the depletion region. A voltage results from the electric field formed at the junction called as junction potential

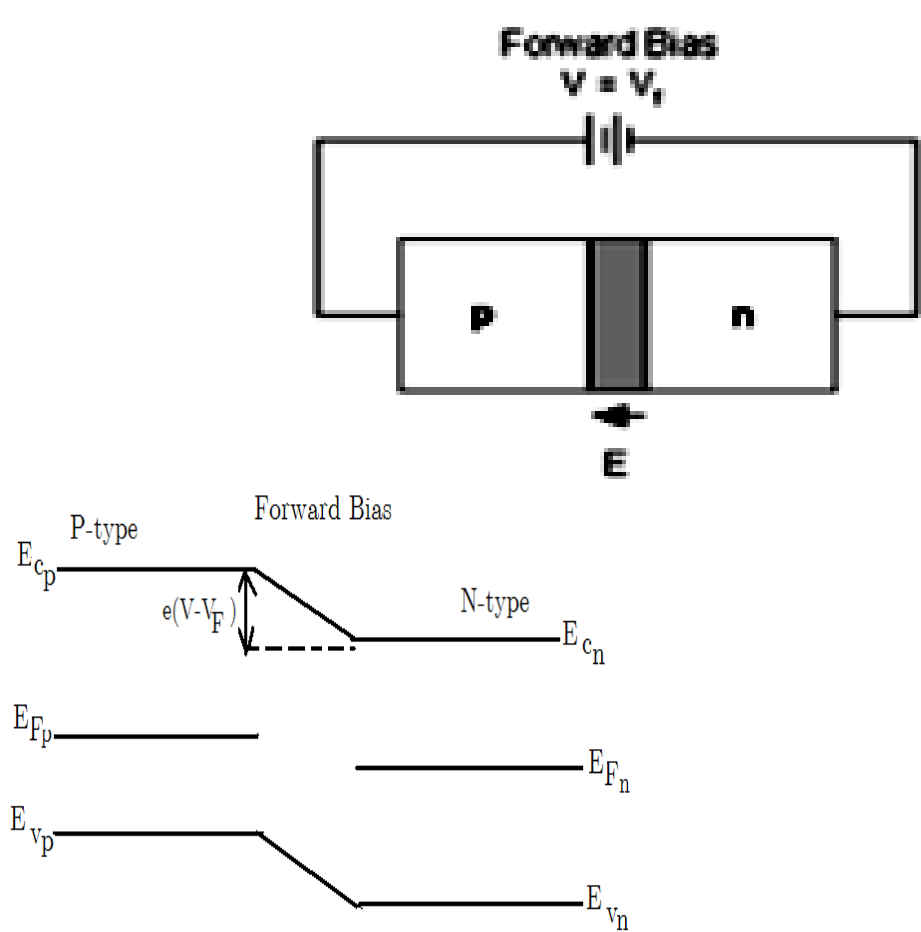
## Biasing

Connecting a p-n junction to an external dc voltage source is called biasing. There are two such connections in which a diode can be operated.

1. Forward bias

2. Reverse bias

Forward bias:



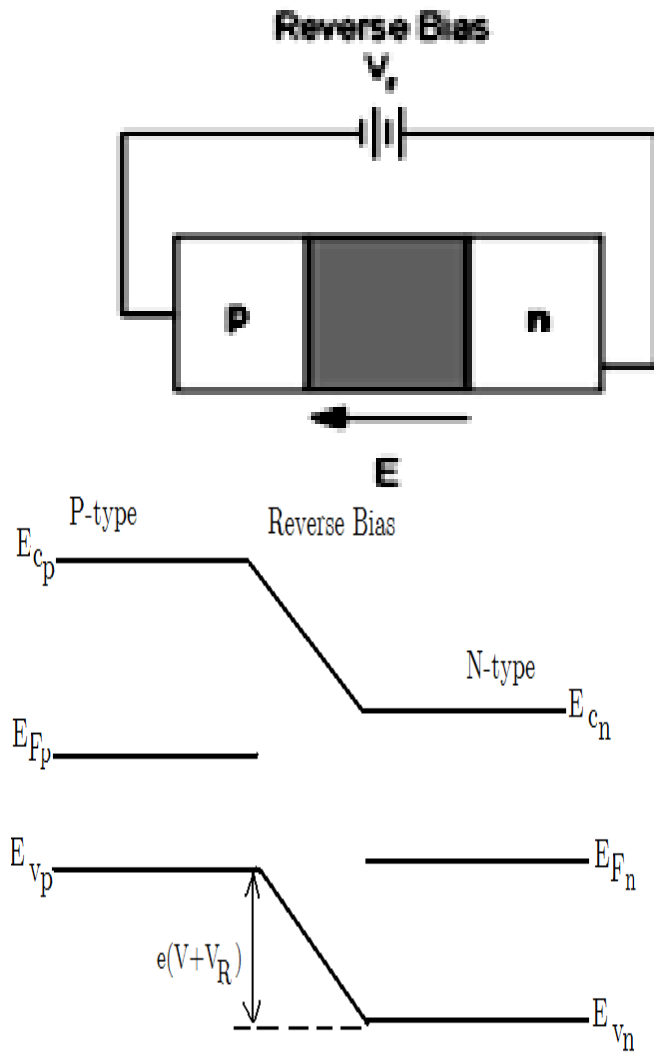
1. Forward biased P-N junction (5.5)  
forward biased P-N junction

2. Energy diagram of

When external voltage applied to the junction is in such a direction that it cancels the potential barrier, thus permitting current flow is called forward biasing. To apply forward bias, the +ve terminal of the battery is connected to p-type material and the -ve terminal to n-type. The applied forward potential establishes the 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.

Since the potential barrier voltage is very small, 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 as in fig 5.5. Therefore a current flows in the circuit. This is current called forward current.

Reverse bias:

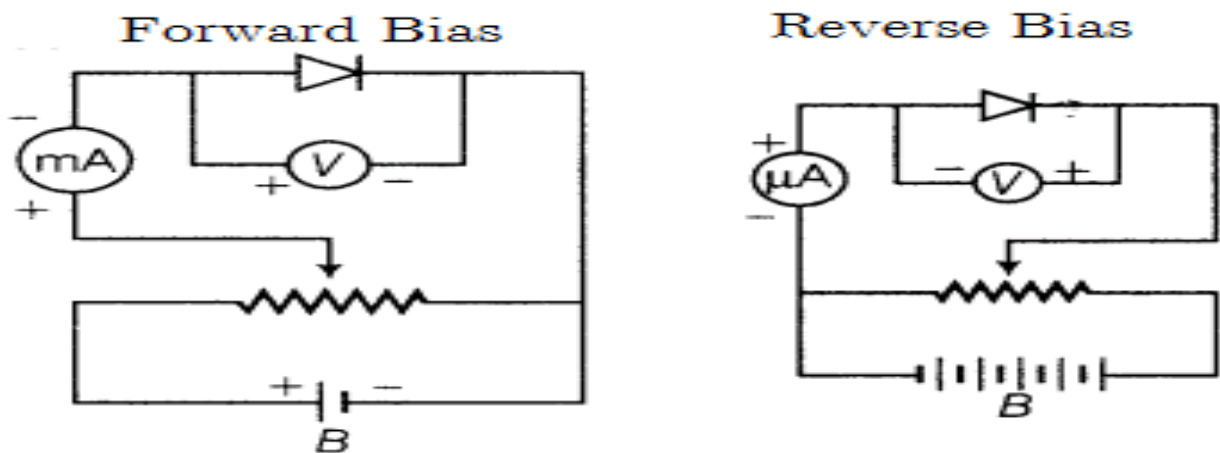


1.Reverse biased P-N junction (5.6)  
Reverse biased P-N junction

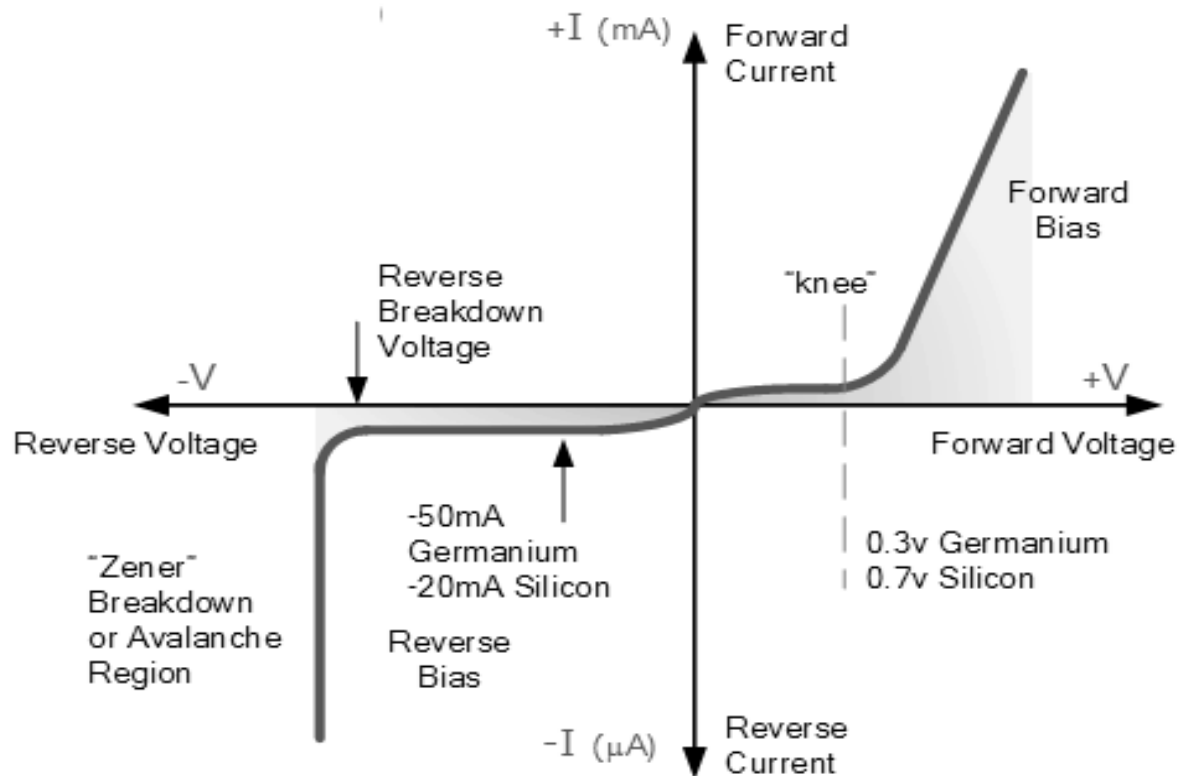
2.Energy diagram of

When the external voltage applied to the junction is in such a direction the potential barrier is increased it is called reverse biasing. To apply reverse bias, –ve terminal of the battery is connected to p-type material and the +ve terminal to n-type. The 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 as in fig 5.6. Thus a high resistance path is established for the entire circuit and hence current does not flow.

### V-I Characteristics of a P-N junction



Circuit diagram of P-N junction(5.7(a))



Junction diode characteristics(5.7)(b)

### V-I Characteristics of a P-N junction

The V-I characteristics of a semiconductor diode can be obtained with the help of the circuit shown in fig. The supply voltage  $V$  is a regulated power supply, the diode is forward biased in the circuit shown. The resistor  $R$  is a current limiting resistor. The voltage across the diode is measured with the help of voltmeter and the current is recorded using an ammeter.

By varying the supply voltage different sets of voltage and currents are obtained. By plotting these values on a graph, the forward characteristics can be obtained. It can be noted from the graph the current remains zero till the diode voltage attains the barrier potential. For silicon diode, the barrier potential is 0.7 V

and for Germanium diode, it is 0.3 V. The barrier potential is also called as knee voltage or cut-in voltage. The reverse characteristics can be obtained by reverse biasing the diode. It can be noted that at a particular reverse voltage, the reverse current increases rapidly. This voltage is called breakdown voltage.

## V-I Characteristics Of Zener Diode

**Zener Diode** is a reverse-biased heavily-doped PN junction diode which operates in the breakdown region. The reverse breakdown of a PN- junction may occur either due to Zener effect or avalanche effect. Zener effect dominates at reverse voltages less than 5 volt whereas avalanche effect dominates above 5 V. Hence, first one should be called Zener diode. But for simplicity, both types are called Zener Diodes. The breakdown voltage of a Zener diode can be set by controlling the doping level. For Zener diodes, silicon is preferred to Ge because of its high temperature and current capability. This post includes explanation of operation of Zener diode and V-I Characteristics of Zener Diode.

### Operation of Zener Diode:

- Zener Diodes are normally used only in the reverse bias direction.
- It means that the anode must be connected to the negative side of the voltage source and the cathode must be connected to the positive side.
- A main difference between Zener diodes and regular silicon diodes is the way they are used in the circuits.
- It is primarily used to regulate the circuit voltage as it has constant  $V_z$ .
- A large change in IR will cause only a small change in  $V_z$ . It means that a zener diode can be used as an alternate current path. The constant  $V_z$  developed across the diode can then be applied to a load.
- Thus the load voltage remains at constant by altering the current flow through the Zener diode.

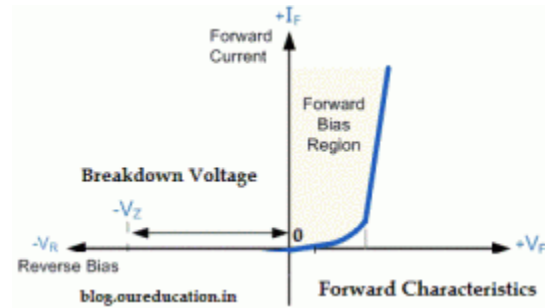
**The V-I Characteristics of a Zener Diode can be divided into two parts**

- (i) Forward Characteristics
- (ii) Reverse Characteristics

### Forward Characteristics



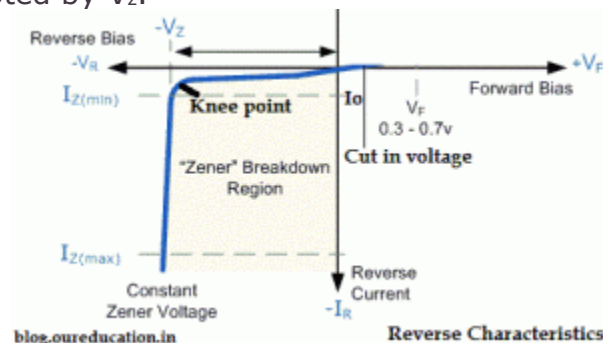
The forward characteristics of a Zener diode is shown in figure. It is almost identical to the forward characteristics of a P-N junction diode.



Forward Characteristics of Zener Diode

## Reverse Characteristics

As we increase the reverse voltage, initially a small reverse saturation current  $I_o$ . Which is in  $A$ , will follow. This current flows due to the thermally generated minority carriers. At a certain value of reverse voltage, the reverse current will increase suddenly and sharply. This is an indication that the breakdown has occurred. This breakdown voltage is called as Zener breakdown voltage or Zener voltage and it is denoted by  $V_Z$ .

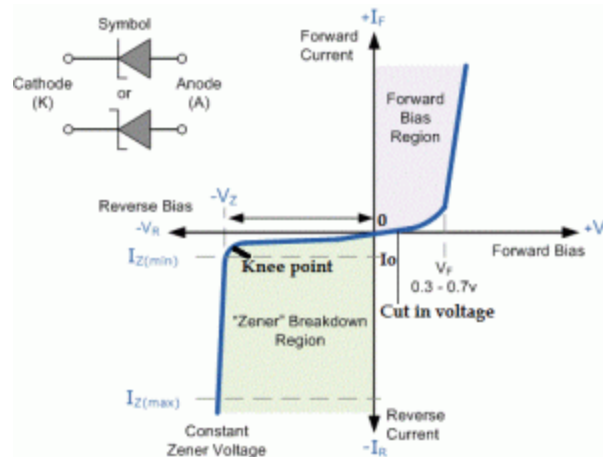


Reverse Characteristics of Zener Diode

The value of  $V_Z$  can be precisely controlled by controlling the doping levels of P and N regions at the time of manufacturing a Zener diode. After breakdown has occurred. The voltage across Zener diode remains constant equal to  $V_Z$ . Any increase in the source voltage will result in the increase in reverse Zener current. The Zener current after the reverse breakdown must be controlled by connecting a resistor R as shown in figure. This is essential to avoid any damage to the device due to excessive heating.

## Zener Region and its importance

Reverse breakdown of the zener diode operates in a region called zener region, as shown in figure. In this region the voltage across zener diode remains constant but current changes depending on the supply voltage. zener diode is operated in this region when it is being used as a voltage regulator. The complete v-i characteristics of zener diode is as shown in figure



V-I Characteristics of Zener Diode

### Application of zener Diode are as follows:

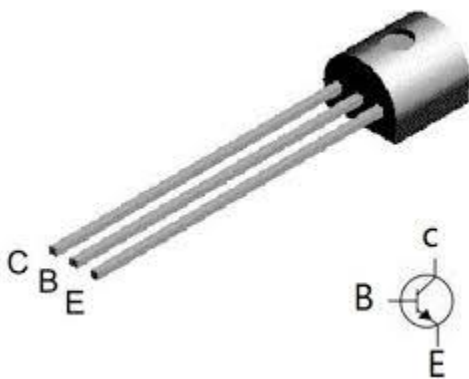
Zener diodes have a large number of application. few of them are

- (i) Zener diode is used as a voltage regulator.
- (ii) Zener diode is used as a peak clipper in wave shaping circuits.
- (iii) Zener diode is used as a fixed reference voltage in transistor biasing circuits.
- (iv) Zener diode is used for meter protection against damage from accidental application of excessive voltages.

## Bipolar Transistor Basics

In the Diode tutorials we saw that simple diodes are made up from two pieces of semiconductor material, either silicon or germanium to form a simple PN-junction and we also learnt about their properties and characteristics. If we now join together two individual signal diodes back-to-back, this will give us two PN-junctions connected together in series that share a common P or N terminal. The fusion of these two diodes produces a three layer, two junction, three terminal device forming the basis of a Bipolar Transistor, or BJT for short.

Transistors are three terminal active devices made from different semiconductor materials that can act as either an insulator or a conductor by the application of a small signal voltage. The transistor's ability to change between these two states enables it to have two basic functions: "switching" (digital electronics) or "amplification" (analogue electronics). Then bipolar transistors have the ability to operate within three different regions:



Typical Bipolar Transistor

- 1. Active Region - the transistor operates as an amplifier and  $I_c = \beta \cdot I_b$  •
- 2. Saturation - the transistor is "fully-ON" operating as a switch and  $I_c = I(\text{saturation})$  •
- 3. Cut-off - the transistor is "fully-OFF" operating as a switch and  $I_c = 0$

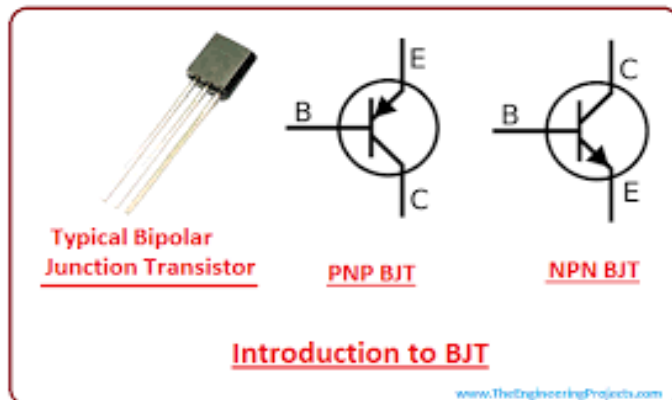
The word Transistor is an acronym, and is a combination of the words Transfer Varistor used to describe their mode of operation way back in their early days of development. There are two basic types of bipolar transistor construction, NPN

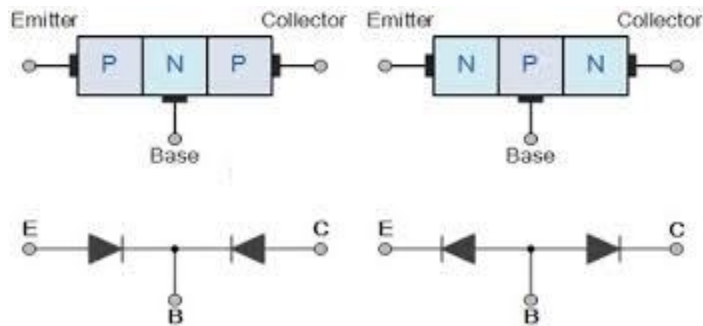
and PNP, which basically describes the physical arrangement of the P-type and N-type semiconductor materials from which they are made.

The Bipolar Transistor basic construction consists of two PN-junctions producing three connecting terminals with each terminal being given a name to identify it from the other two. These three terminals are known and labelled as the Emitter ( E ), the Base ( B ) and the Collector ( C ) respectively.

Bipolar Transistors are current regulating devices that control the amount of current flowing through them in proportion to the amount of biasing voltage applied to their base terminal acting like a current-controlled switch. The principle of operation of the two transistor types NPN and PNP, is exactly the same the only difference being in their biasing and the polarity of the power supply for each type.

## Bipolar Transistor Construction





The construction and circuit symbols for both the NPN and PNP bipolar transistor are given above with the arrow in the circuit symbol always showing the direction of "conventional current flow" between the base terminal and its emitter terminal. The direction of the arrow always points from the positive P-type region to the negative N-type region for both transistor types, exactly the same as for the standard diode symbol.

### Bipolar Transistor Configurations

As the Bipolar Transistor is a three terminal device, there are basically three possible ways to connect it within an electronic circuit with one terminal being common to both the input and output. Each method of connection responding differently to its input signal within a circuit as the static characteristics of the transistor vary with each circuit arrangement.

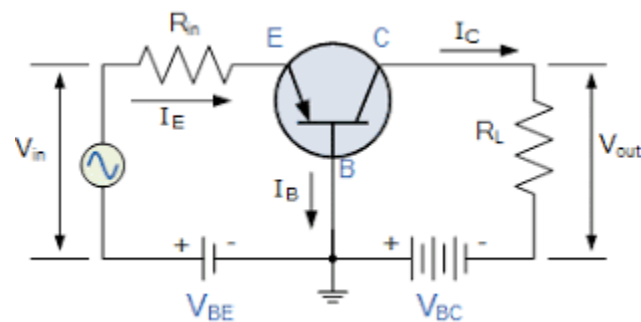
1. Common Base Configuration - has Voltage Gain but no Current Gain.
- • 2. Common Emitter Configuration - has both Current and Voltage Gain.
- • 3. Common Collector Configuration - has Current Gain but no Voltage Gain.

### The Common Base (CB) Configuration

As its name suggests, in the Common Base or grounded base configuration, the BASE connection is common to both the input signal AND the output signal with the input signal being applied between the base and the emitter terminals. The corresponding output signal is taken from between the base and the collector terminals as shown with the base terminal grounded or connected to a fixed reference voltage point. The input current flowing into the emitter is quite large as it is the sum of both the base current and collector current respectively therefore,

the collector current output is less than the emitter current input resulting in a current gain for this type of circuit of "1" (unity) or less, in other words the common base configuration "attenuates" the input signal.

### The Common Base Transistor Circuit



This type of amplifier configuration is a non-inverting voltage amplifier circuit, in that the signal voltages  $V_{in}$  and  $V_{out}$  are in-phase. This type of transistor arrangement is not very common due to its unusually high voltage gain characteristics. Its output characteristics represent that of a forward biased diode while the input characteristics represent that of an illuminated photo-diode. Also this type of bipolar transistor configuration has a high ratio of output to input resistance or more importantly "load" resistance ( $R_L$ ) to "input" resistance ( $R_{in}$ ) giving it a value of "Resistance Gain". Then the voltage gain ( $A_v$  for a common base configuration) is therefore given as:

### Common Base Voltage Gain

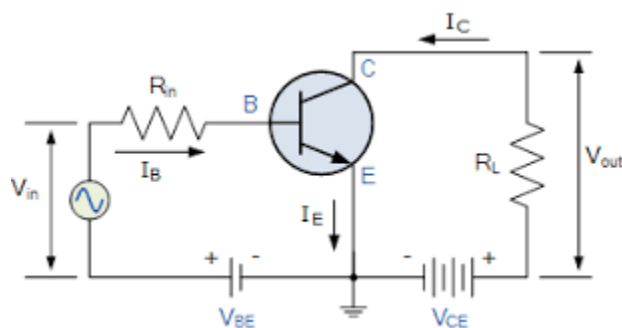
$$A_v = V_{out} / V_{input} = I_c R_L / I_E R_{IN}$$

The common base circuit is generally only used in single stage amplifier circuits such as microphone pre-amplifier or radio frequency (Rf) amplifiers due to its very good high frequency response.

**The Common Emitter (CE) Configuration** In the Common Emitter or grounded emitter configuration, the input signal is applied between the base, while the

output is taken from between the collector and the emitter as shown. This type of configuration is the most commonly used circuit for transistor based amplifiers and which represents the "normal" method of bipolar transistor connection. The common emitter amplifier configuration produces the highest current and power gain of all the three bipolar transistor configurations. This is mainly because the input impedance is LOW as it is connected to a forward-biased PN-junction, while the output impedance is HIGH as it is taken from a reverse-biased PN-junction.

### The Common Emitter Amplifier Circuit



In this type of configuration, the current flowing out of the transistor must be equal to the currents flowing into the transistor as the emitter current is given as  $I_E = I_C + I_B$ . Also, as the load resistance ( $R_L$ ) is connected in series with the collector, the current gain of the common emitter transistor configuration is quite large as it is the ratio of  $I_C/I_B$  and is given the Greek symbol of Beta, ( $\beta$ ). As the emitter current for a common emitter configuration is defined as  $I_E = I_C + I_B$ , the ratio of  $I_C/I_E$  is called Alpha, given the Greek symbol of  $\alpha$ . Note: that the value of Alpha will always be less than unity.

Since the electrical relationship between these three currents,  $I_B$ ,  $I_C$  and  $I_E$  is determined by the physical construction of the transistor itself, any small change in the base current ( $I_B$ ), will result in a much larger change in the collector current ( $I_C$ ). Then, small changes in current flowing in the base will thus control the current in the emitter-collector circuit. Typically, Beta has a value between 20 and 200 for most general purpose transistors.

By combining the expressions for both Alpha,  $\alpha$  and Beta,  $\beta$  the mathematical relationship between these parameters and therefore the current gain of the transistor can be given as:

$$\text{Alpha}(\alpha) = I_c / I_E \quad \text{and} \quad \text{Beta}(\beta) = I_c / I_B$$

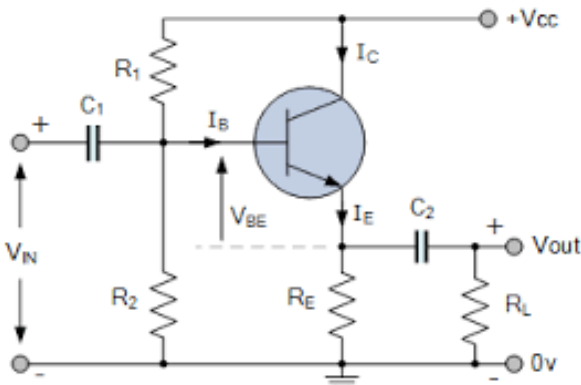
$$\text{Therefore } I_E = I_c + I_B$$

Where: " $I_c$ " is the current flowing into the collector terminal, " $I_b$ " is the current flowing into the base terminal and " $I_e$ " is the current flowing out of the emitter terminal. Then to summarise, this type of bipolar transistor configuration has a greater input impedance, current and power gain than that of the common base configuration but its voltage gain is much lower. The common emitter configuration is an inverting amplifier circuit resulting in the output signal being 180° out-of-phase with the input voltage signal.

### The Common Collector (CC) Configuration

In the Common Collector or grounded collector configuration, the collector is now common through the supply. The input signal is connected directly to the base, while the output is taken from the emitter load as shown. This type of configuration is commonly known as a Voltage Follower or Emitter Follower circuit. The emitter follower configuration is very useful for impedance matching applications because of the very high input impedance, in the region of hundreds of thousands of Ohms while having a relatively low output impedance.

#### The Common Collector Transistor Circuit



The common emitter configuration has a current gain approximately equal to the  $\beta$  value of the transistor itself. In the common collector configuration the load resistance is situated in series with the emitter so its current is equal to that of the emitter current. As the emitter current is the combination of the collector AND the



base current combined, the load resistance in this type of transistor configuration also has both the collector current and the input current of the base flowing through it. Then the current gain of the circuit is given as:

The Common Collector Current Gain

$$A_I = I_E / I_B$$

$$= I_C + I_B / I_B$$

$$= \beta + 1$$

This type of bipolar transistor configuration is a non-inverting circuit in that the signal voltages of  $V_{in}$  and  $V_{out}$  are inphase. It has a voltage gain that is always less than "1" (unity). The load resistance of the common collector transistor receives both the base and collector currents giving a large current gain (as with the common emitter configuration) therefore, providing good current amplification with very little voltage gain.



