# UNIT II – SEMICONDUCTOR Material

## 2.1 Introduction

Devices in which a controlled flow of electrons can be obtained are the basic building blocks of all the electronic circuits. The seed of the development of modern solid-state semiconductor electronics goes back to 1930's when it was realized that some solid-state semiconductors and their junctions offer the possibility of controlling the number and the direction of flow of charge carriers through them. Simple excitations like light, heat or small applied voltage can change the number of mobile charges in a semiconductor. They are small in size, consume low power, operate at low voltages and have long life and high reliability.

Semiconductors are the materials which have a conductivity between conductors (generally metals) and non-conductors or insulators (such ceramics). Semiconductors can be compounds such as gallium arsenide or pure elements, such as germanium or silicon. Physics explains the theories, properties and mathematical approach governing semiconductors. Examples of Semiconductors: Gallium arsenide, germanium, and silicon are some of the most commonly used semiconductors. Silicon is used in electronic circuit fabrication, and gallium arsenide is used in solar cells, laser diodes, etc.

Holes and electrons are the types of charge carriers accountable for the flow of current in semiconductors. Holes are the positively charged electric charge carrier whereas electrons are the negatively charged particles. Both electrons and holes are equal in magnitude but opposite in polarity. In a semiconductor, the mobility of electrons is higher than that of the holes. It is mainly because of their different band structures and scattering mechanisms.

Electrons travel in the conduction band whereas holes travel in the valence band. When an electric field is applied, holes cannot move as freely as electrons due to their restricted movent. The elevation of electrons from their inner shells to higher shells results in the creation of holes in semiconductors. Since the holes experience stronger atomic force by the nucleus than electrons, holes have lower mobility.

| **Intrinsic Semiconductor** | **Extrinsic Semiconductor** |
|---|---|
| 1. It is a pure semiconductor with all tetravalent atoms and impurities are completely absent. For example: Silicon, Germanium and Gallium Arsenide (GaAs). | 1. When a small amount of impurity (trivalent or tetravalent) is added (doped) in a pure semiconductor, it becomes an extrinsic semiconductor. |
| 2. Number of electrons in the conduction band is equal to the number of holes in the valance band i.e. $\mathbf{N_e=N_h}$ | 2. In case of N-type semiconductor, $\mathbf{N_e>>N_h}$and in case of P-type, $\mathbf{N_h>> N_e}$ |
| 3. Fermi level in an intrinsic semiconductor lies at midway in the forbidden gap. | 3. In N-type semiconductor, Fermi level lies slightly below the conduction band and in P-type semiconductor, Fermi level lies slightly above the valance band. |

| | |
|---|---|
| 4. The electrical conductivity depends only on temperature. | 4. The electrical conductivity depends on temperature as well as on the concentration of impurity atoms. |
| 5. It is a bad conductor at room temperature. | 5. It is a good conductor at room temperature. |

## 2.2 Intrinsic and Extrinsic Semiconductors

## 2.3 Intrinsic semiconductors:

The semiconductor in which generation of free charge carriers in the form of electrons in the conduction band and holes in the valance band takes place purely by thermal excitation are called intrinsic semiconductors. It means that this effect is temperature dependent and produces equal number of holes and free electrons. The electrons and holes are called intrinsic charge carriers and the resulting conductivity is called intrinsic conductivity.

## 2.4 Concentration of electrons in conduction band:

For the conduction band, the total concentration of electrons, ***n*** is the integral of number of occupied energy states at a given energy per unit volume. If $E_c$ represents the energy at the bottom of the conduction band, number density of electrons in the conduction band that can have energies lying between $E_c$ and $\infty$ is given by,

$$n = \int_{E_C}^{\infty} D(E)\, f(E)\, dE \qquad (2.1)$$

where D(E)= density of states at the bottom of the conduction band per unit volume

$$D(E) = \frac{4\pi}{h^3}\left(2m_e^*\right)^{3/2}\left(E - E_C\right)^{1/2} \qquad (2.2)$$

Here, $m_e$* is the effective mass of electrons. The Fermi Dirac distribution function is given as

$$f(E) = \frac{1}{1 + e^{(E - E_F)/kT}} \qquad (2.3)$$

Substituting (2.2) and (2.3) in (2.1), we get

$$n = \frac{4\pi}{h^3}\left(2m_e^*\right)^{3/2}\int_{E_C}^{\infty}\frac{(E-E_C)^{1/2}}{1+e^{(E-E_F)/kT}}\,dE \tag{2.4}$$

For $E \geq E_C$ and E - E$_F$ >> KT (which is always satisfied at all temperatures because E$_F$ lies at tye center of forbidden band in intrinsic semiconductors), so 1 in the denominator can be neglected in the above equation (4), which reduces to

$$n = \frac{4\pi}{h^3}\left(2m_e^*\right)^{3/2}\int_{E_C}^{\infty}(E-E_C)^{1/2}e^{(E_F-E)/kT}\,dE$$

$$n = \frac{4\pi}{h^3}\left(2m_e^*\right)^{3/2}\int_{E_C}^{\infty}(E-E_C)^{1/2}e^{(E_F-E+E_C-E_C)/kT}\,dE$$

$$n = \frac{4\pi}{h^3}\left(2m_e^*\right)^{3/2}e^{(E_F-E_C)/kT}\int_{E_C}^{\infty}(E-E_C)^{1/2}e^{(E_C-E)/kT}\,dE \tag{2.5}$$

Putting $(E-E_C)/kT = x$ so that $dE = kTdx$ and changing the limits of integration as $x \to 0$ When $E \to E_C$ and $x \to \infty$ when $E \to \infty$

Expression (2.5) now becomes

$$n = \frac{4\pi}{h^3}\left(2m_e^*\right)^{3/2}e^{(E_F-E_C)/kT}\int_{0}^{\infty}x^{1/2}(kT)^{1/2}e^{-x}kTdx$$

i.e. $$n = \frac{4\pi}{h^3}\left(2m_e^*\right)^{3/2}(kT)^{3/2}e^{(E_F-E_C)/kT}\int_{0}^{\infty}x^{1/2}e^{-x}dx$$

Using, $\int_0^\infty x^{1/2}e^{-x}dx = \frac{\sqrt{\pi}}{2}$, we get

$$n = \frac{4\pi}{h^3}\left(2m_e^*\right)^{3/2}(kT)^{3/2}e^{(E_F-E_C)/kT}\frac{\sqrt{\pi}}{2}$$

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

This is an expression of concentration (or density) of electrons in the conduction band of an intrinsic semiconductor. This can be written as

$$n = N_C e^{(E_F-E_C)/kT}$$

Where

$N_C = 2\left[\frac{2\pi m_e^* kT}{h^2}\right]^{3/2}$ is called effective density of states in the conduction band. Note that ***n*** depends only on temperature.

## 2.5 Concentration of holes in valance band

For the valance band, the total concentration of holes, ***p*** is the integral of number of occupied energy states at a given energy per unit volume. Since a hole signifies a vacancy created by removal of an electron in the valance band, the Fermi function or occupational probability for a hole is given by

$$1 - f(E) = 1 - \frac{1}{1+e^{(E-E_F)/kT}} = 1 - \left[1 + e^{(E-E_F)/kT}\right]^{-1}$$

i.e. $$1 - f(E) = 1 - \left[1 - e^{(E-E_F)/kT}\right]$$ (for $(E-E_F)/kT)<<1$)

i.e. $$1 - f(E) = e^{(E-E_F)/kT} \qquad (2.6)$$

If $E_v$ represents energy at the top of the valance band, number density of holes in the valance band that can have energies lying between -∞ and $E_V$ will be

$$p = \int_{-\infty}^{E_V} D(E)[1 - f(E)]dE \qquad (2.7)$$

where D(E)= density of states at the top of the valance band per unit volume

$$D(E) = \frac{4\pi}{h^3}\left(2m_h^*\right)^{3/2}\left(E_V - E\right)^{1/2} \qquad (2.8)$$

Here, $m_h$* is the effective mass of holes.

Substituting (2.6) and (2.8) in (2.7), we get

$$p = \frac{4\pi}{h^3}\left(2m_h^*\right)^{3/2}\int_{-\infty}^{E_V}(E_V - E)^{1/2}e^{(E-E_F)/kT}dE \qquad (2.9)$$

$$p = \frac{4\pi}{h^3}\left(2m_h^*\right)^{3/2}\int_{-\infty}^{E_V}(E_V - E)^{1/2}e^{(E-E_F+E_V-E_V)/kT}dE$$

$$p = \frac{4\pi}{h^3}\left(2m_h^*\right)^{3/2}e^{(E_V-E_F)/kT}\int_{-\infty}^{E_V}(E_V - E)^{1/2}e^{(E-E_V)/kT}dE \qquad (2.10)$$

Putting $(E_V - E)/kT = x$ so that $dE = -kTdx$ and changing the limits of integration as

$x \to 0$ when $E \to E_V$ and $x \to \infty$ when $E \to -\infty$

Expression (5) now becomes

$$p = \frac{4\pi}{h^3}\left(2m_h^*\right)^{3/2}e^{(E_V-E_F)/kT}\int_{\infty}^{0}x^{1/2}(kT)^{1/2}e^{-x}(-kTdx)$$

i.e.

$$p = \frac{4\pi}{h^3}\left(2m_h^*\right)^{3/2}(kT)^{3/2}e^{(E_V-E_F)/kT}\int_{0}^{\infty}x^{1/2}e^{-x}dx$$

Using $\int_0^\infty x^{1/2}e^{-x}dx = \frac{\sqrt{\pi}}{2}$, we get

$$p = \frac{4\pi}{h^3}\left(2m_h^*\right)^{3/2}(kT)^{3/2}e^{(E_V-E_F)/kT}\,\frac{\sqrt{\pi}}{2}$$

i.e. $$p = 2\left[\frac{2\pi m_h^* kT}{h^2}\right]^{3/2} e^{(E_V - E_F)/kT}$$

This is an expression of concentration (or density) of holes in the valance band of an intrinsic semiconductor. This can be written as

$$p = N_V e^{(E_V - E_F)/kT}$$

Where

$N_V = 2\left[\frac{2\pi m_h^* kT}{h^2}\right]^{3/2}$ is called effective density of states in the valance band. Note that ***p*** depends only on temperature.

## 2.6 Intrinsic Concentration of Charge Carriers

In an intrinsic semiconductor, thermal energy produces equal number of free electrons and holes. As the two charge carrier concentrations are equal, they are denoted by a common symbol $\boldsymbol{n_i}$, which is called intrinsic density or intrinsic concentration. Thus,

$$n = p = n_i$$

and $\boldsymbol{n_i^2 = np}$.

**This is called Law of Mass Action**

Substituting the value of *n* and p, we get

$$n_i^2 = 2\left[\frac{2\pi m_e^* kT}{h^2}\right]^{3/2} e^{(E_F - E_C)/kT} \times 2\left[\frac{2\pi m_h^* kT}{h^2}\right]^{3/2} e^{(E_v - E_F)/kT}$$

$$n_i^2 = 4\left[\frac{2\pi kT}{h^2}\right]^{3} (m_e^* m_h^*)^{3/2} e^{(E_V - E_C)/kT}$$

$$n_i^2 = 4\left[\frac{2\pi kT}{h^2}\right]^{3} (m_e^* m_h^*)^{3/2} e^{-E_g/kT} \qquad (\because E_C - E_V = E_g)$$

$$n_i = 2\left[\frac{2\pi kT}{h^2}\right]^{3/2} (m_e^* m_h^*)^{3/4} e^{-E_g/2kT}$$

Note that the intrinsic density of charge carriers depends only on temperature and $E_g$ but is independent of E<sub>F</sub>.

## 2.7 Fermi level in an Intrinsic Semiconductor

We know that in an intrinsic semiconductor, *n*= *p*

i.e.
$$2\left[\frac{2\pi m_e^* kT}{h^2}\right]^{3/2} e^{(E_F-E_C)/kT} = 2\left[\frac{2\pi m_h^* kT}{h^2}\right]^{3/2} e^{(E_v-E_F)/kT}$$

$$e^{(2E_F-E_C-E_V)/kT} = \left(\frac{m_h^*}{m_e^*}\right)^{3/2}$$

Taking logarithm on both sides, we get

$$\frac{2E_F-E_C-E_V}{kT} = \frac{3}{2}\ln\left(\frac{m_h^*}{m_e^*}\right)$$

i.e.
$$2E_F-E_C-E_V = \frac{3}{2}kT\ln\left(\frac{m_h^*}{m_e^*}\right)$$

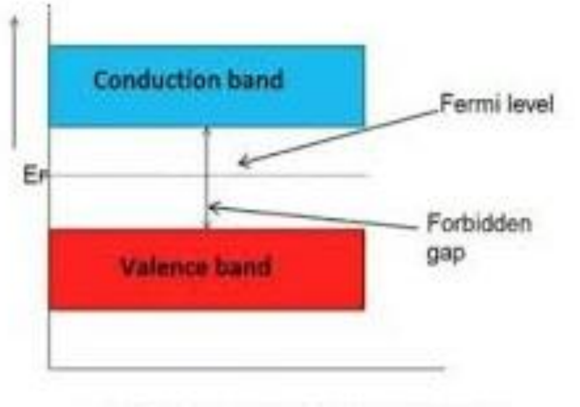
$$2E_F = E_C+E_V+\frac{3}{2}kT\ln\left(\frac{m_h^*}{m_e^*}\right)$$

$$E_F = \frac{E_C+E_V}{2}+\frac{3}{4}kT\ln\left(\frac{m_h^*}{m_e^*}\right)$$

Note that if $m_e^* = m_h^*$, then we get

$$E_F = \frac{E_C+E_V}{2}$$

Thus, the Fermi level in an intrinsic semiconductor lies exactly half way between the top of the valance band and bottom of the conduction band i.e. at the center of forbidden energy gap as shown in Fig. below.



### Limitations of intrinsic semiconductors:

The conductivity of an intrinsic semiconductor depends on its temperature, but at room temperature its conductivity is very low. As such, no important electronic devices can be developed using these semiconductors. For example Ge has a conductivity of 1.67 S/m at room temperature, which is $10^7$ times smaller than that of copper at same temperature.

## 2.8 Extrinsic Semiconductors:

One way of increasing the conductivity of a semiconductor crystal is to add a controlled amount of certain impurities. An intentionally or deliberately introduction of controlled amount of impurity into the intrinsic semiconductor is called doping. The impurity added is called dopant. A semiconductor doped with impurity atoms is called an extrinsic semiconductor. The conduction that occurs then is called impurity conduction or extrinsic conduction and is of paramount importance in the operation of semiconductor devices.

Typical doping levels range from $10^{20}$ to $10^{27}$ impurity atoms per $m^3$. Pentavalent elements from group 5 (Phosphorus, Arsenic and Antimony) or trivalent elements from group 3 (Boron, Gallium and Indium) are generally used as Si or Ge atoms and easily substitute themselves in place of some of Si or Ge atoms in the semiconductor crystal without distorting it. Depending on the two different (pentavalent or trivalent) types of doping, two types of extrinsic semiconductors are named as n-type semiconductors (with pentavalent doping) and p-type semiconductors (with trivalent doping).

## Applications of Semiconductors

Let us now understand the uses of semiconductors in daily life. Semiconductors are used in almost all electronic devices. Without them, our life would be much different.

Their reliability, compactness, low cost and controlled conduction of electricity make them ideal to be used for various purposes in a wide range of components and devices, transistors, diodes, photo-sensors, microcontrollers, integrated chips and much more are made up of semiconductors.

## Uses of Semiconductors in Everyday life

1. Temperature sensors are made with semiconductor devices.
2. They are used in 3D printing machines
3. Used in microchips and self-driving cars
4. Used in calculators, solar plates, computers and other electronic devices.
5. Transistor and MOSFET used as a switch in Electrical Circuits are manufactured using the semiconductors.

## Industrial Uses of Semiconductors

The physical and chemical properties of semiconductors make them capable of designing technological wonders like microchips, transistors, LEDs, solar cells, etc.The microprocessor used for controlling the operation of space vehicles, trains, robots, etc is made up of transistors and other controlling devices which are manufactured by semiconductor materials.

## Importance of Semiconductors

Here we have discussed some advantages of semiconductors which makes them highly useful everywhere.

1. They are highly portable due to the smaller size
2. They have high conductivity.
3. Conductivity can be tailored to the desired value by controlling the doping concentration.
4. Conductivity is temperature independent at ambient temperatures.
5. They require less input power
6. Semiconductor devices are shockproof
7. They have a longer lifespan
8. They are noise-free while operating

## P-Type semiconductor:

The extrinsic p-TypeSemiconductor is formed when a trivalentimpurity is added to a pure semiconductor in a small amount, and as a result, a large number of holes are created in it. A large number of holes are provided in the semiconductor material by the addition of trivalent impurities likeGallium andIndium**.** Such type of impurities which produces p-type semiconductor are known as an AcceptorImpurities because each atom of them create one hole which can accept one electron. A trivalent impurity like boron, having three valence electrons is added to silicon crystal in a small amount. Each atom of the impurity fits in the silicon crystal in such a way that its three valence electrons form covalent bonds with the three surrounding silicon atoms.

## N- Type semiconductor:

When a small amount of Pentavalentimpurity is added to a pure semiconductor providing a large number of free electrons in it, the extrinsic semiconductor thus formed is known as n-TypeSemiconductor**.** The conduction in the n-type semiconductor is because of the free electrons denoted by the pentavalent impurity atoms. These electrons are the excess free electrons with regards to the number of free electrons required to fill the covalent bonds in the semiconductors. Silicon of Group IV has four valence electrons and phosphorus of Group V has five valence electrons. If a small amount of phosphorus is added to a pure silicon crystal, one of the valence electrons of phosphorus becomes free to move around (free electron) as a surplus electron. When this free electron is attracted to the "+" electrode and moves, current flows.

## 2.9 Fermi level in n-type semiconductor

At zero kelvin, all the donor levels are filled with electrons, when the temperature is slightly increased, a small fraction of donors are ionized. The position of Fermi level at various temperatures can be obtained as follows:

**a) At very low temperature:**

$n = N_D F(E_c)$

Here, n is number density of free electrons in conduction

$N_D$ is number density of donor atoms

$F(E_C)$ is occupation probability for $E=E_C$

$$\therefore n = \frac{N_D}{1+\exp\left[\frac{E_C-E_F}{kT}\right]} \approx N_D xp\left[\frac{E_F-E_C}{kT}\right] \qquad (2.11)$$

Likewise, $N_D^+ = N_D[1-f(E_D)]$

Where, $N_D^+$ is number density of donor ions and $[1-f(E_D)]$ is probability that electron is missing in state $E = E_D$

$$\therefore \quad N_D^+ = N_D\left[1-\left(1+e^{\frac{E_D-E_F}{kT}}\right)^{-1}\right]$$

$$\text{i.e.} N_D^+ = N_D e^{\frac{E_D-E_F}{kT}} \qquad (2.12)$$

Now, clearly there are as many electrons in the conduction band, after excitation, as the number of donor ions.

i.e. $n = N_D^+$

Using, (2.11) and (2.12), we obtain, $E_F - E_C = E_D - E_F$

i.e.$E_F = \frac{E_C + E_D}{2}$

Therefore, at very low temperature, $E_F$ lies in between midway between $E_C$ and $E_D$ in n-type semiconductors.

b) **At moderate and normal temperatures**,

All the electrons in the donor states get excited to the conduction band, so

$n = N_D$

as, $n = 2\left[\frac{2\pi m_e kT}{\hbar^2}\right]^{3/2} e^{(E_F - E_C)/kT}$

or, $n = A_C e^{(E_F - E_C)/kT}$

Where, $A_C = 2\left[\frac{2\pi m_e kT}{\hbar^2}\right]^{3/2} = \text{constant}$

$\therefore \frac{A_C}{n} = \frac{A_C}{N_D} = e^{-(E_F - E_C)/kT}$

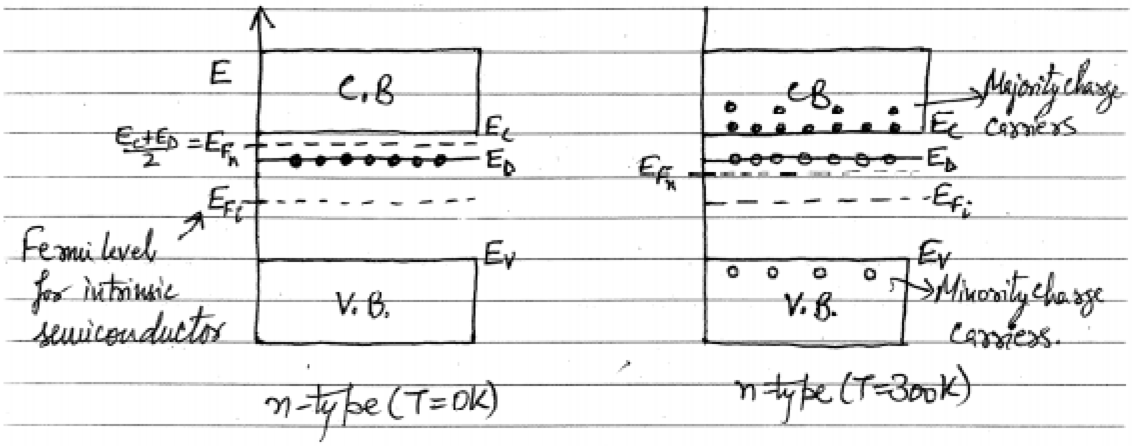
Taking logarithm on the both sides, we get,

or, $\ln\left(\frac{A_C}{N_D}\right) = -\frac{(E_F - E_C)}{kT}$

Hence,

$E_F = E_C - k_B T\,(A_{C/}N_D)$

Which shows that, $E_F$lies slightly below the bottom of the conduction band in-between $E_C$ and $E_D$ and with increase in temperature, it shifts further down.



## 2.10 Fermi level in p-type semiconductor

At zero kelvin, all the acceptor levels are filled with holes and the valance band is completely filled with electrons.When the temperature is slightly increased, a small fraction of holesmove from the acceptor level to the valnce band and a small fraction of acceptor atoms are ionized. The position of Fermi level at various temperatures can be obtained as follows:

**a) At very low temperature:**

$p = N_A F(E_V)$

Here, p is number density of free holes in valance band

$N_A$ is number density of acceptor atoms

$F(E_V)$ is occupation probability (for hole) for $E=E_V$ in valance band

$$\therefore p = \frac{N_A}{1+\exp\left[\frac{E_V-E_F}{kT}\right]} \approx N_A \exp\left[\frac{E_F-E_V}{kT}\right] \tag{2.13}$$

Likewise, $N_A^- = N_A[1 - F(E_A)]$

Where, $N_A^-$ is number density of acceptor ions and $[1 - F(E_A)]$ is probability that hole is missing in state $E = E_A$

$$\therefore \quad N_A^- = N_A\left[1 - \left(1 + e^{\frac{E_A-E_F}{kT}}\right)^{-1}\right]$$

$$\text{i.e.} \quad N_A^- = N_A e^{\frac{E_A-E_F}{kT}} \tag{2.14}$$

Now, clearly there are as many holes in the valance band, as the number of acceptor ions.

i.e. $p = \quad N_A^-$

Using, (2.13) and (2.14), we obtain, $E_F$-$E_C$ = $E_D$-$E_F$

i.e.$E_F = \frac{E_A+E_V}{2}$

Therefore, at very low temperature, $E_F$ lies in between midway between $E_V$ and $E_A$ in p-type semiconductors.

b) **At moderate and normal temperatures**, all the holes in the acceptor states jump to the valance band, so

$p = N_A$

as, $p = 2\left[\frac{2\pi m_h kT}{ђ^2}\right]^{3/2} e^{(E_V-E_F)/kT}$

or, $p = A_V e^{(E_V-E_F)/kT}$

Where, $A_V = 2\left[\frac{2\pi m_h kT}{ђ^2}\right]^{3/2} = \text{constant}$

$\therefore\ \frac{A_V}{p} = \frac{A_V}{N_A} = e^{(E_F-E_V)/kT}$

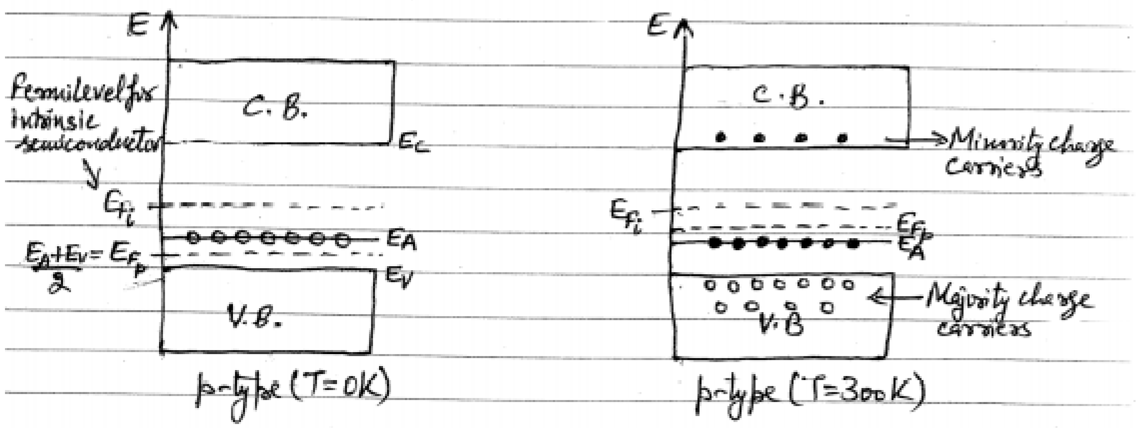
Taking logarithm on the both sides, we get,

or, $\ln\left(\frac{A_V}{N_A}\right) = \frac{(E_F-E_V)}{kT}$

Hence,

$E_F = E_V + kT\ (A_{V/}N_A)$

Which shows that, $E_F$ lies slightly above the top of the valance band in-between $E_A$ and $E_V$ and with increase in temperature, it shifts further up.



With increase in doping concentration in a p-type semiconductor, the acceptor levels broader into acceptor band, which ultimately overlaps on the valance band. The Fermi level moves down closer to the valance band and finally at very high doping concentration it will shift into the valance band. In the process, the forbidden gap between the bands also decreases with increase in doping.

## 2.11 PN-junction:

Consider one side of a sample of silicon is doped with pentavalent impurity making it n-type and other side with trivalent impurity making it p-type, as shown in figure 4. We assume that there is an abrupt discontinuity between the p- and n-region, which we call the metallurgical junction and label as ‘M’. The fixed (immobile) ionized donors and free electrons (in the conduction band) in the n-region along with the fixed ionized acceptors and holes (in the valance band) in the p-region are also shown in figure 4. When a p-type semiconductor is brought into close contact with an n-type semiconductor, the assembly formed is called pn-junction.The portion of contact between the two types of semiconductors is called junction.

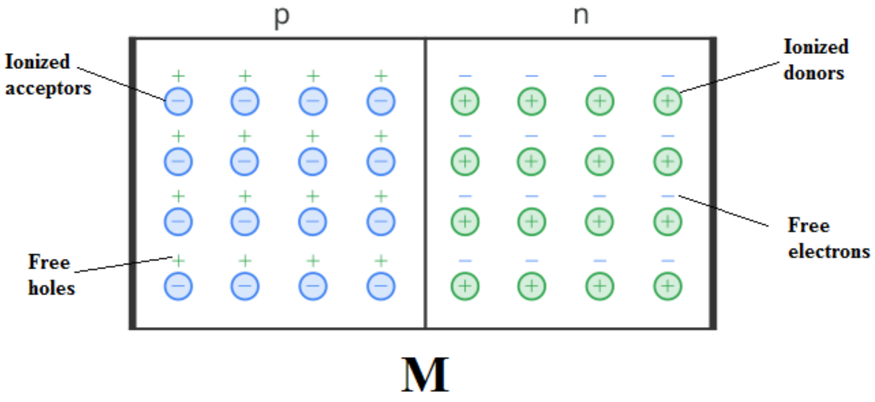


**Fig. 4: Basic pn-junction development**

## 2.12 Unbiased PN-junction:

There are large number of free electrons in n-type semiconductor and a large number of holes in the p-type semiconductor. At the junction, electrons diffuse into p-region from n-region and are neutralized by an equal number of holes. Similarly, holes from the p-region enter into the n-region through the junction where they are neutralized by an equal number of electrons, as shown in figure 5(a). This charge transfer of electrons and holes across the pn-junction is known as **diffusion**.The space between the dotted lines in the figure 5 represents the width of junction.

However, because the electrons have moved across the pn- junction from the n-type to the p-type, they leave behind positively charged donor ions ($N_D$) on the negative side and now the holes from the acceptor impurity migrate across the junction in the opposite direction into the region where there are large numbers of free electrons. As a result, the charge density of the p-type along the junction is filled with negatively charged acceptor ions ($N_A$), and the charge density of the n-type along the junction becomes positive, as shown in figure 5 (b) & (c). The fig. 5 (b) shows that there are no electrons and holes in the depletion region whereas; (c) represents the variation of concentration of uncovered (or un-neutralized) charge in the depletion region around the junction. The positively and negatively charged ions are called un-neutralized charges. This process continues back and forth until the number of electrons which have crossed the junction have a large enough electrical charge to repel or prevent any more charge carriers from crossing over the junction. Eventually a state of equilibrium (electrically neutral situation) will occur producing a “potential barrier” zone around the area of the junction as the donor atoms repel the holes and the acceptor atoms repel the electrons.

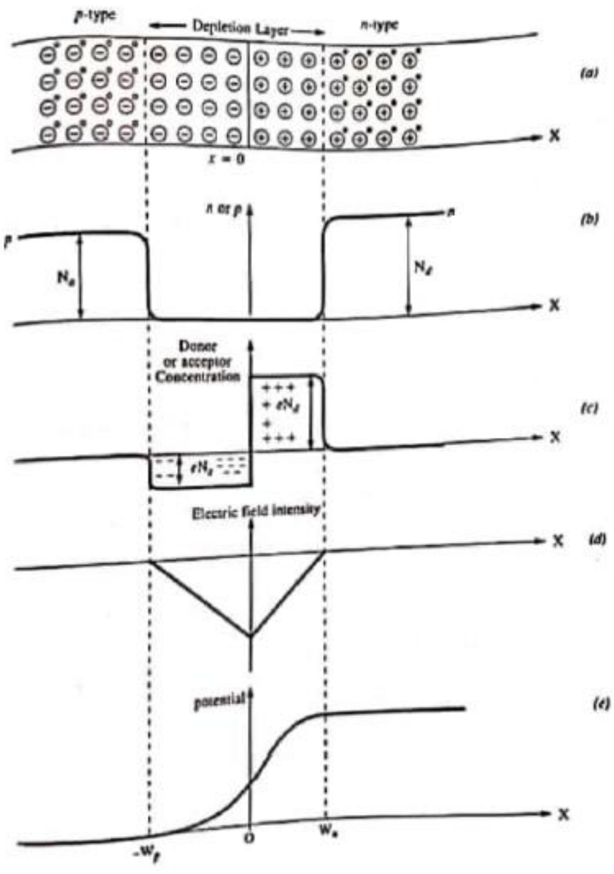


**Fig. 5: Un-biased pn-junction**

As a result of the production of un-neutralized charges, p-type semiconductor develops a negative potential and n-type semiconductor develops a positive potential. The potential difference is called the barrier potential or built-in potential and is denoted by $V_B$. The figure 5 (e) shows the variation of potential on the both side of the junction. Since no free charge carriers can rest in a position where there is a potential barrier, the regions on either sides of the junction now become completely depleted of any more free carriers in comparison to the n and p-type materials further away from the junction. The width of these p and n layers depends on how heavily each side is doped with acceptor density $N_A$, and donor density $N_D$, respectively. This area around the **pn-Junction** (the region in which no free charge carrier is available) is called depletion region. The thickness of this region may be a fraction of micron and barrier field is approximately $10^7$ volt per meter. The depletion region is also termed as space charge region or width. There is a difference between the energies of an electron in two regions and is called barrier energy $E_B$. Therefore, $E_B = e\ V_B$. The figure 5 (d) gives the variation of electric field intensity on the either side of junction.

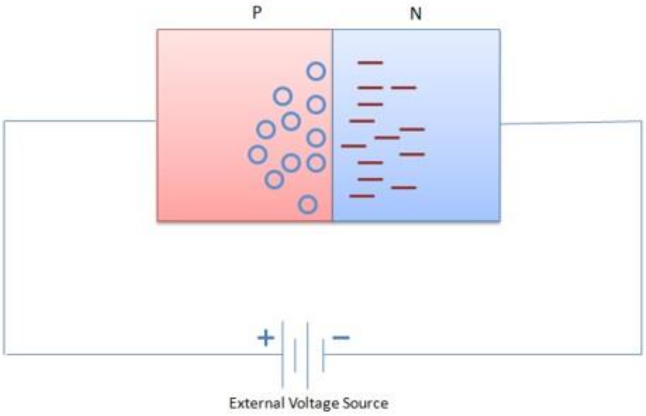
## 2.13 Biased PN-junction:

A pn-junction is said to be biased, if a direct potential difference is applied across the n and p-region of the pn-junction. The term bias refers to the application of DC voltage to set up certain operating conditions. Or when an external source of energy is applied to a pn-junction it is called a bias voltage or simply biasing. This method either increases or decreases the barrier potential of the junction. As a result, the reduction of the barrier potential causes current carriers to return to the depletion region. Following two bias conditions are applied with respect to pn-junctions.

- **Forward Biasing** − An external voltage is added of the same polarity to the barrier potential, which causes an increase in the width of the depletion region.
- **Reverse Biasing** − A PN junction is biased in such a way that the application of external voltage action prevents current carriers from entering the depletion region.

### (a) Forward Biasing

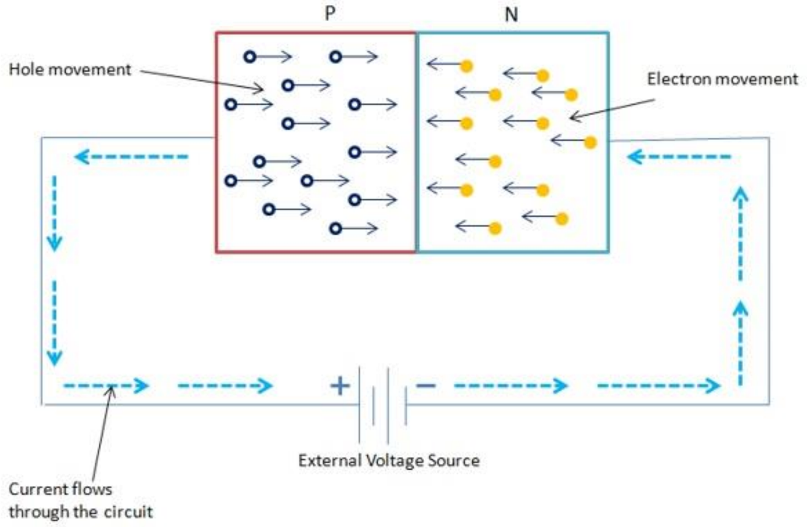
The following figure shows a forward biased PN junction diode with external voltage applied. You can see that the positive terminal of the battery is connected to the P material and the negative terminal of the battery is connected to the N material.



Following are the observations −

- This bias voltage repels the majority current carriers of each P and N type material. As a result, large number of holes and electrons start appearing at the junction.
- At the N-side of the junction, electrons move in to neutralize the positive ions in the depletion region.
- On the P-side material, electrons are dragged from negative ions, which cause them to become neutral again. This means that forward biasing collapses the depletion region and hence the barrier potential too. It means that when P-N junction is forward biased, it will allow a continuous current flow.

The following figure shows the flow of current carriers of a forward-biased diode. A constant supply of electrons is available due to an external voltage source connected to the diode. The flow and direction of the current is shown by large arrows outside the diode in the diagram. Note that the electron flow and the current flow refer to the same thing.

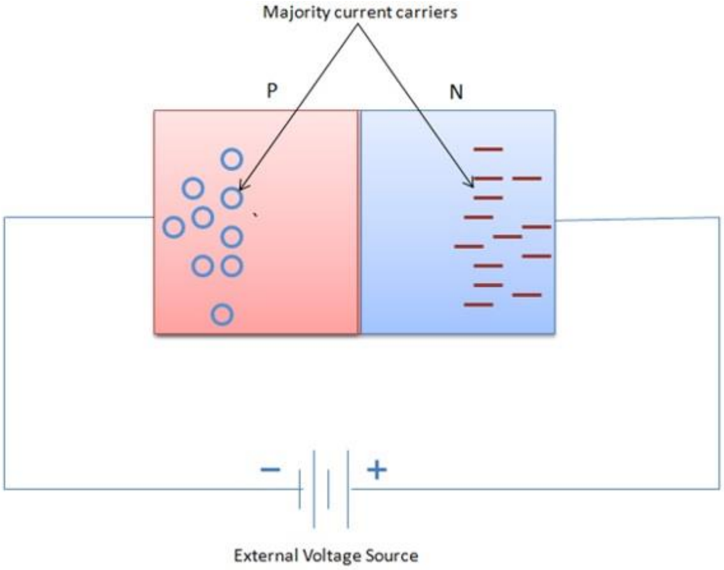


Following are the observations −

- Suppose electrons flow through a wire from the negative battery terminal to the N material. Upon entering this material, they flow immediately to the junction.
- Similarly, on the other side an equal number of electrons are pulled from P side and are returned to the positive battery terminal. This action creates new holes and causes them to move toward the junction.
- When these holes and electrons reach the junction they join together and effectively disappear. As a result, new holes and electrons emerge at the outer ends of the diode. These majority carriers are created on a continuous basis. This action continues as long as the external voltage source is applied.
- When diode is forward biased it can be noticed that electrons flow through the entire structure of diode. This is common in N type material, whereas in the P material holes are the moving current carriers. Notice that the hole movement in one direction must begin by electron movement in the opposite direction. Therefore, the total current flow is the addition of holes and electrons flow through a diode.

### (b) Reverse Biasing

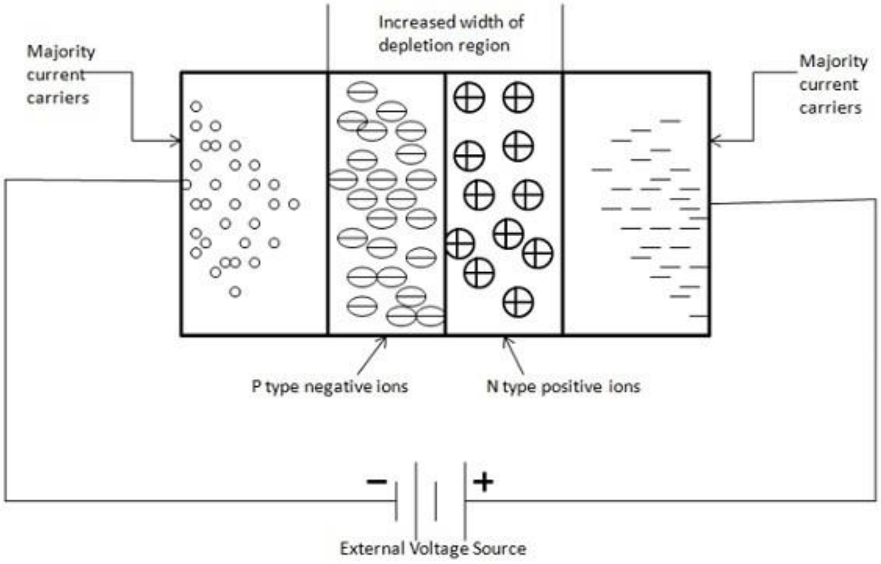
The following figure shows reverse biased PN junction diode with external voltage applied. You can see that the positive terminal of the battery is connected to the N material and the negative terminal of the battery is connected to the P material. Note that in such an arrangement, battery polarity is to oppose the material polarity of the diode so that dissimilar charges attract. Hence, majority charge carriers of each material are dragged away from the junction. Reverse biasing causes the diode to be nonconductive.



The following figure shows the arrangement of the majority current carriers in a reverse biased diode.

- Due to circuit action electrons of the N material are pulled toward the positive battery terminal.
- Each electron that moves or departs the diode causes a positive ion to emerge in its place. As a result, this causes an equivalent increase in the width of the depletion region on the N side of the junction.

- The P side of the diode has a similar effect alike the N side. In this action, a number of electrons leave the negative battery terminal and enter the P type material.
- These electrons then straight away move in and fill a number of holes. Each occupied hole then becomes a negative ion. These ions in turn are then repelled by the negative battery terminal and driven toward the junction. Due to this, there is an increase in the width of the depletion region on the P side of the junction.



The overall width of the depletion region directly depends on an external voltage source of a reverse-biased diode. In this case, the diode cannot efficiently support the current flow through the wide depletion region. As a result, the potential charge starts developing across the junction and increases until the barrier potential equals the external bias voltage. After this, the diode behaves as a nonconductor.

## 2.14 Fermi level in PN-junction:

In p-type semiconductor, Fermi level lies close to the top of valance band. In n-type semiconductors, it lies near the bottom of the conduction band. When a contact between both semiconductors is made, the Fermi level of two semiconductors attains same value. If VB is the barrier height, then the shift in the conduction band of p-type semiconductor eVB with respect to the conduction band of n-type semiconductor. Energy of electron of p-type is more than the energy of electron of n-type. Therefore, there is no restriction on electron of p-type to jump from n-type by crossing the depletion region. The same alignment of Fermi level in pn-junction is shown in figure 6.

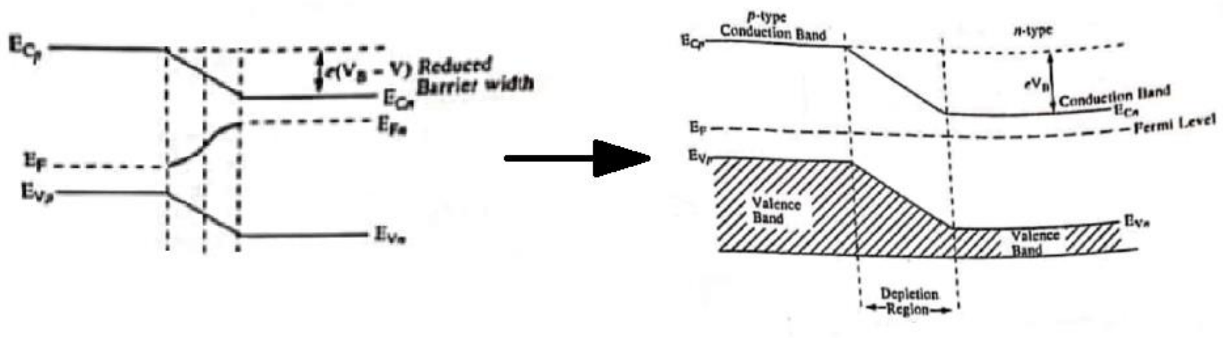


**Fig. 6: Energy band diagram of pn-junction**

# Zener Diode

A Zener diode can be defined as a heavily doped semiconductor device that is designed to operate the electric circuit in the reverse direction. It is also called a breakdown diode.

Working Principle of 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.

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

**(i)** Forward Characteristics

**(ii)** Reverse Characteristics

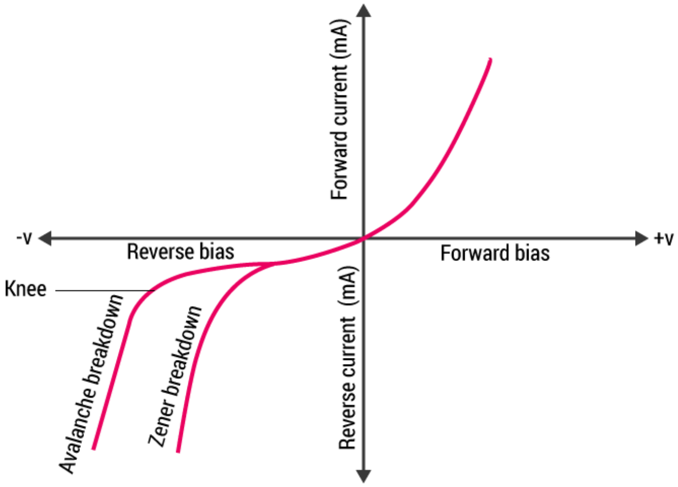
### Forward Characteristics

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.

### Reverse Characteristics

As we increase the reverse voltage, initially a small reverse saturation current Io. 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 Vz.

he value of Vz 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 Vz. 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.



## 2.15 Zener and avalanche breakdown:

**(a) Zener breakdown:** This type of breakdown occurs when the pn-junction is heavily doped, depletion region is reduced. If voltage is increased, a large number of electrons and holes are produced and resistance is reduced. Zener mechanism is useful only at low values of reverse voltage.

**(b) Avalanche breakdown:** This type of breakdown occurs when the pn-junction is lightly doped, depletion region become wide. When reverse voltage is applied, the new electron hole pair (EHP) accelerated by electric field. This lead to avalanche of charge carriers and large current is set up. Avalanche breakdown occurs when the electric field across the depletion region of the order of $2 \times 10^7$ Vm$^{-1}$.

### (c)Differentiation between Zener and Avalanche breakdown:

| Zener breakdown | | Avalanche breakdown | |
|---|---|---|---|
| 1 | The process in which the electrons move across the barrier from the valence band of p-type material to the conduction band of n-type material is known as Zener breakdown. | 1 | The process of applying high voltage and increasing the free electrons or electric current in semiconductors and insulating materials is called an avalanche breakdown. |
| 2 | This is observed in Zener diodes having a Zener breakdown voltageof 5 to 8 volts. | 2 | This is observed in Zener diode having a Zener breakdown voltage is greater than 8 volts. |
| 3 | The valence electrons are pulled into conduction due to the high electric field in the narrow depletion region. | 3 | The valence electrons are pushed to conduction due to the energy imparted by accelerated electrons, which gain |

| | | | |
|---|---|---|---|
| | | | their velocity due to their collision with other atoms. |
| 4 | The increase in temperature decreases the breakdown voltage. | 4 | The increase in temperature increases the breakdown voltage. |
| 5 | The VI characteristics of a Zener breakdown have a sharp curve. | 5 | The VI characteristic curve of the avalanche breakdown is not as sharp as the Zener breakdown. |
| 6 | It occurs in diodes that are highly doped. | 6 | It occurs in diodes that are lightly doped. |

# Photodiode

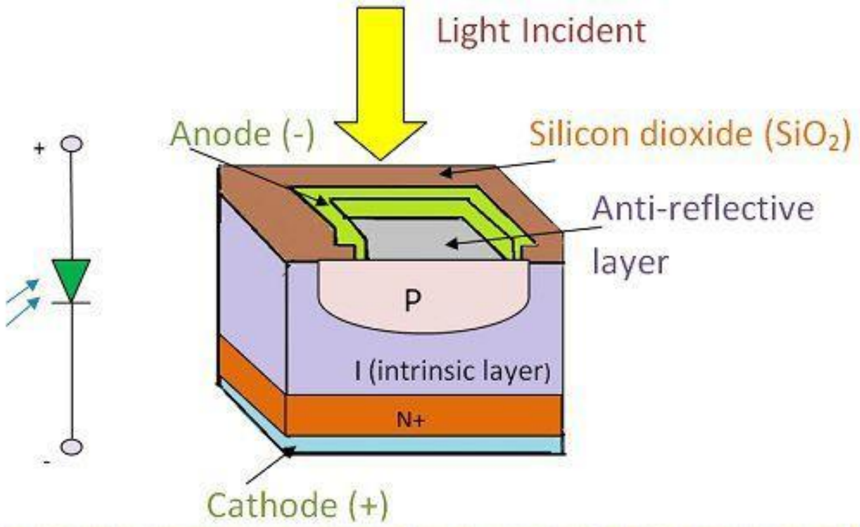
**Definition: Photodiode** is a two terminal electronic device which, when exposed to light the current starts flowing in the diode. It is operated in reverse biased mode only. It converts **light energy** into **electrical energy.** P-N junction of the photodiode is illuminated by light and light energy dislodges valence electrons and the diode starts conducting.

## Construction of Photodiode

The photodiode is made up of two layers of P-type and N-type semiconductor. In this, the P-type material is formed from diffusion of the lightly doped P-type substrate. Thus, the layer of P+ ions is formed due to the diffusion process. And N-type epitaxial layer is grown on N-type substrate. The P+ diffusion layer is developed on N-type heavily doped epitaxial layer. The contacts are made up of metals to form two terminal cathode and anode.

The front area of the diode is divided into two types that are active surface and non-active surface. The non-active surface is made up of **$SiO_2$ (Silicon di Oxide)** and the active surface is coated with **anti-reflection material**. The active surface is called so because the light rays are incident on it.

While on the non-active surface the light rays do not strike. The active layer is coated with anti-reflection material so that the light energy is not lost and the maximum of it can be converted into current. The entire unit has dimensions of the order of **2.5 mm.**



**Photodiode Symbol & Construction of Photodiode**

Electronics Coach

### Working Principle of Photodiode

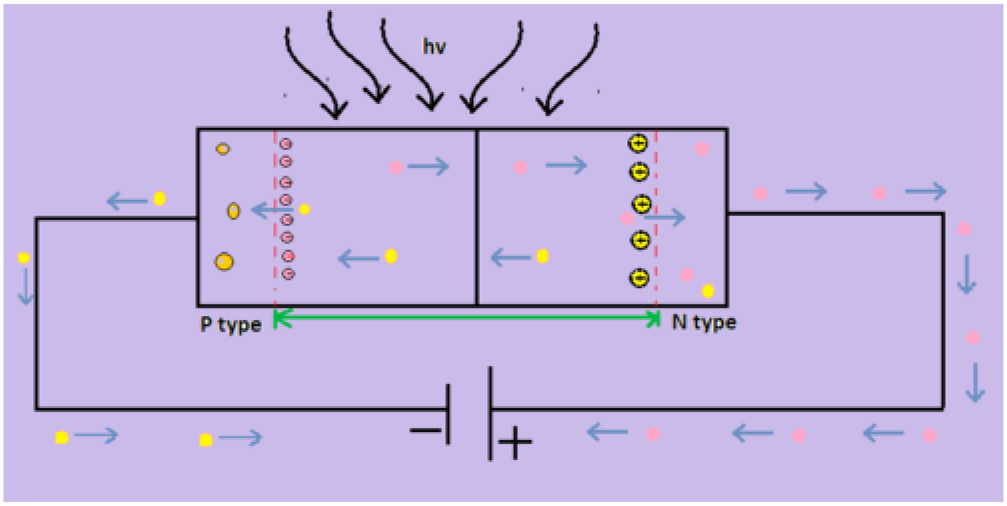
Generally, when a light is made to illuminate the PN junction, covalent bonds are ionized. This generates hole and electron pairs. Photocurrents are produced due to generation of electron-hole pairs. Electron hole pairs are formed when photons of energy more than 1.1eV hits the diode. When the photon enters the depletion region of diode, it hits the atom with high energy. This results in release of electron from atom structure. After the electron release, free electrons and

hole are produced.

In general, an electron will have a negative charge and holes will have a positive charge. The depletion energy will have built-in electric field. Due to that electric field, electron-hole pairs move away from the junction. Hence, holes move to anode and electrons move to the cathode to produce photocurrent.

The photon absorption intensity and photon energy are directly proportional to each other. When energy of photos is less, the absorption will be more. This entire process is known as Inner Photoelectric Effect.

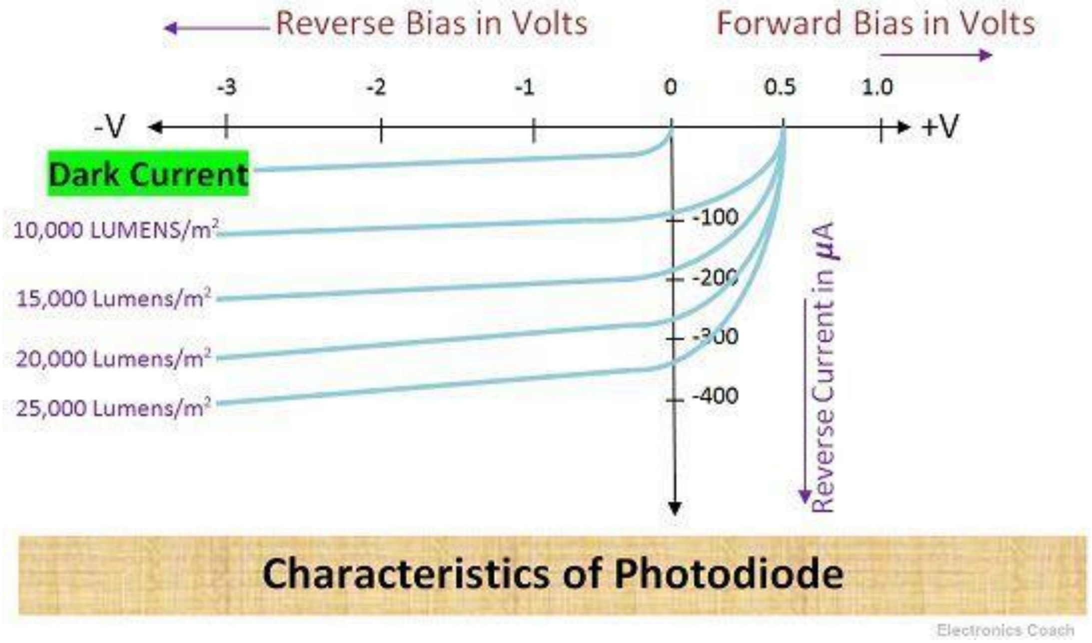
Intrinsic Excitations and Extrinsic Excitations are the two methods via which the photon excitation happens. The process of intrinsic excitation happens, when an electron in the valence band is excited by photon to conduction band.



In this way, the photodiode converts light energy into electrical energy. The current which flows in photodiode before light rays are incident on it is called **dark current.** As leakage current flows in the conventional diode, similarly the **dark current** flows in the photodiode.

## V-I Characteristics of Photodiode

The characteristics curve of the photodiode can be understood with the help of the below diagram. The characteristics are shown in the negative region because the photodiode can be operated in reverse biased mode only.



The reverse saturation current in the photodiode is denoted by $I_0$. It varies linearly with the intensity of photons striking the diode surface. The current under large reverse bias is the summation of reverse saturation current and short circuit current.

$$\mathbf{I = I_{sc} + I_0 (1 - e^{V/\eta V_t})}$$

Where Isc is the short circuit current, V is positive for forward voltage and negative for reverse bias, Vt is volt equivalent for temperature, η is unity for germanium and 2 for silicon.

Advantages of Photodiodes

1. The reverse current is low in the tens of microamperes.
2. The rise and fall times in case of photodiodes is very small making it suitable for high-speed counting and switching applications.

Disadvantages of Photodiodes

Photodiodes have lower light sensitivity than cadmium sulphide LDRs (Light dependent resistors), thus they CdS LDRs are considered more suitable for some applications.

Applications of Photodiodes

1. It is used for detection of both visible as well as invisible light rays.
2. Photodiodes are used for the communication system for encoding & demodulation purpose.
3. It is also used for digital and logic circuits which require fast switching and high-speed operation.
4. These diodes also find application in character recognition techniques and IR remote control circuits.

Photodiodes are considered as one of the significant optoelectronics devices which are extensively used in the optical fiber communication system.

## LED

The lighting emitting diode is a p-n junction diode. It is a specially doped diode and made up of a special type of semiconductors. When the light emits in the forward biased, then it is called a light-emitting diode.

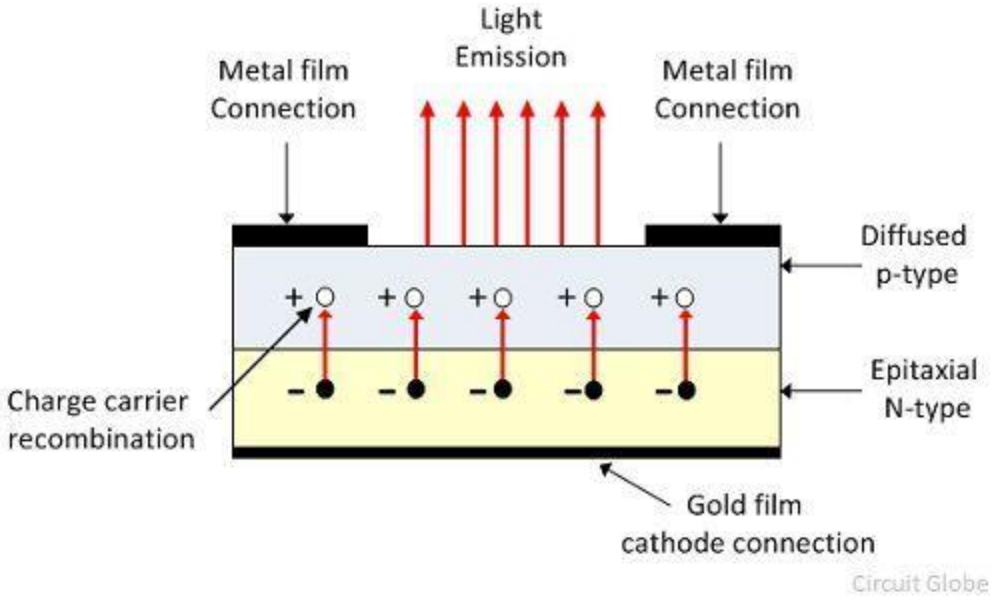
### Construction of LED

### Working Principle of LED

The LED is a PN-junction diode which emits light when an electric current passes through it in the forward direction. In the LED, the recombination of charge carrier takes place. The electron from the N-side and the hole from the P-side are combined and gives the energy in the form of heat and light. The LED is made of semiconductor material which is colourless, and the light is radiated through the junction of the diode.

Construction of LED

The recombination of the charge carrier occurs in the P-type material, and hence P-material is the surface of the LED. For the maximum emission of light, the anode is deposited at the edge of the P-type material. The cathode is made of gold film, and it is usually placed at the bottom of the N-region. This gold layer of cathode helps in reflecting the light to the surface.

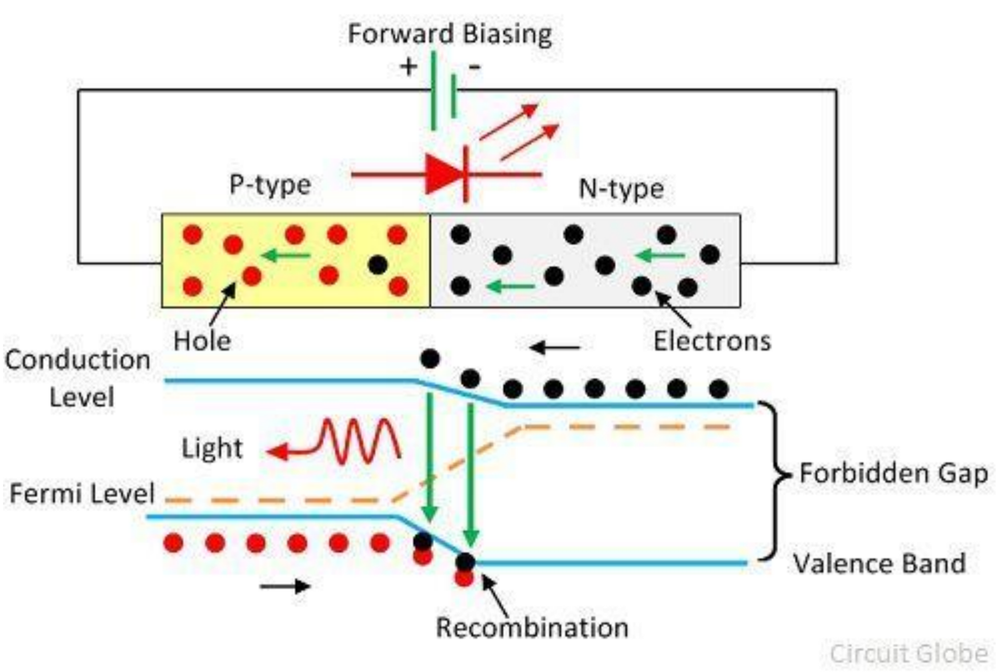


the higher and The gallium arsenide phosphide is used for the manufacturing of LED which emits red or yellow light for emission. The LED are also available in green, yellow amber and red in colour.

**Working of LED**

The working of the LED depends on the quantum theory. The quantum theory states that when the energy of electrons decreases from the higher level to lower level, it emits energy in the form of photons. The energy of the photons is equal to the gap between lower level. el.

The LED is connected in the forward biased, which allows the current to flows in the forward direction. The flow of current is because of the movement of electrons in the opposite direction. The recombination shows that the electrons move from the conduction band to valence band and they emits electromagnetic energy in the form of photons. The energy of photons is equal to the gap between the valence and the conduction



For example, let us consider the quantum theory, the energy of the photon is the product of both the Planck constant and frequency of electromagnetic radiation. The mathematical equation is shown

$$Eq = hf$$

Where his known as a Planck constant, and the velocity of electromagnetic radiation is equal to the speed of light i.e c. The frequency radiation is related to the velocity of light as an $f= c / \lambda$. $\lambda$ is denoted as a wavelength of electromagnetic **radiation and the above equation will become as a**

$$Eq = he / \lambda$$

From the above equation, we can say that the wavelength of electromagnetic radiation is inversely proportional to the forbidden gap.

The advantages of light-emitting diode include the following.

The cost of LED’s is less and they are tiny.

By using the LED’s electricity is controlled.

The intensity of the LED differs with the help of the microcontroller.

Long Lifetime

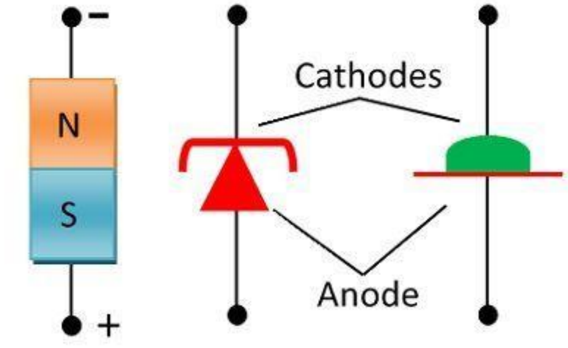
# Tunnel Diode

**Definition:** The tunnel diode is a highly conductive, heavily doped PN-junction diode in which the current induces because of the tunneling. The tunneling is the phenomenon of conduction in the semiconductor material in which the charge carrier punches the barrier instead of climbing through

The tunnel diode is a heavily doped PN-junction diode. The concentration of impurity in the normal PN-junction diode is about 1 part in $10^8$. And in the tunnel diode, the concentration of the impurity is about 1 part in $10^3$. Because of the heavy doping, the diode conducts current both in the forward as well as in the reverse direction. It is a fast switching device; thereby it is used in high-frequency oscillators, computers and amplifiers.

Symbol of Tunnel Diode

The symbol of tunnel diode is shown in the figure below. The cathode and anode are the two terminals of semiconductor material. The p-type material attracts the electrons and hence it is called anode while the n-type material emits the electrons and it is named as the cathode.



Tunnel Diode Symbols

Circuit Globe

Construction of Tunnel Diode

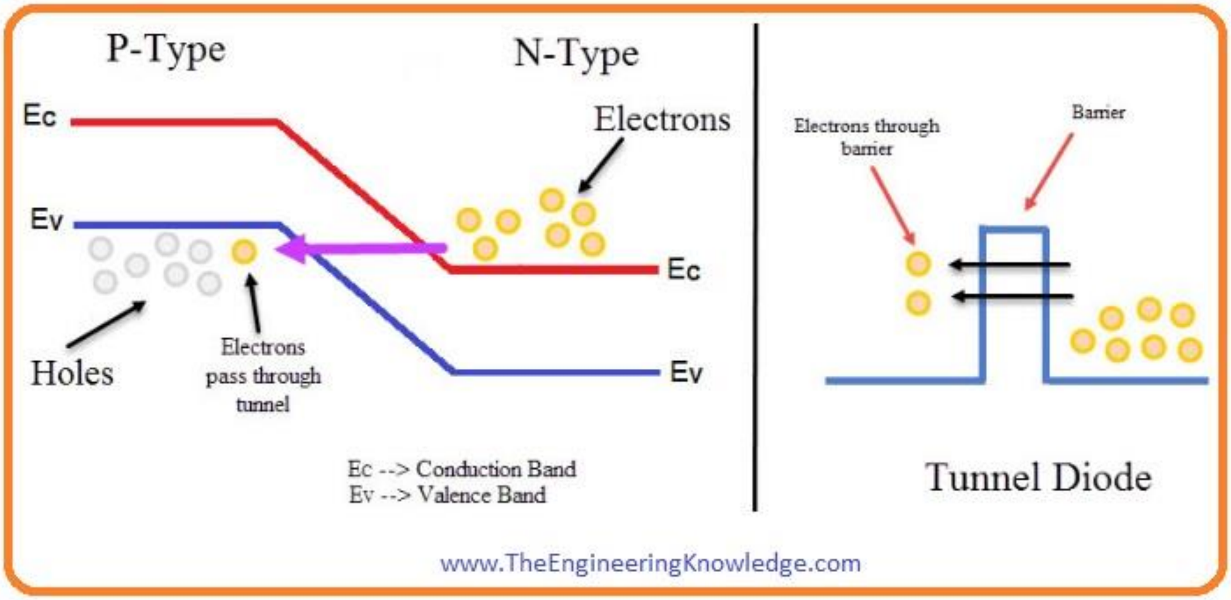
The device is constructed by using the two terminals namely anode and cathode. The p-type semiconductor acts as an anode, and the n-type semiconductor material acts as a cathode. The gallium arsenide, germanium and gallium antimonite are used for manufacturing the tunnel diode.

The ratio of the peak value of the forward current to the value of the valley current is maximum in case of germanium and less in silicon. Hence silicon is not used for fabricating the tunnel diode. The doping density of the tunnel diode is 1000 times higher than that of the ordinary diode.

## TUNNEL DIODE WORKING PRINCIPLE

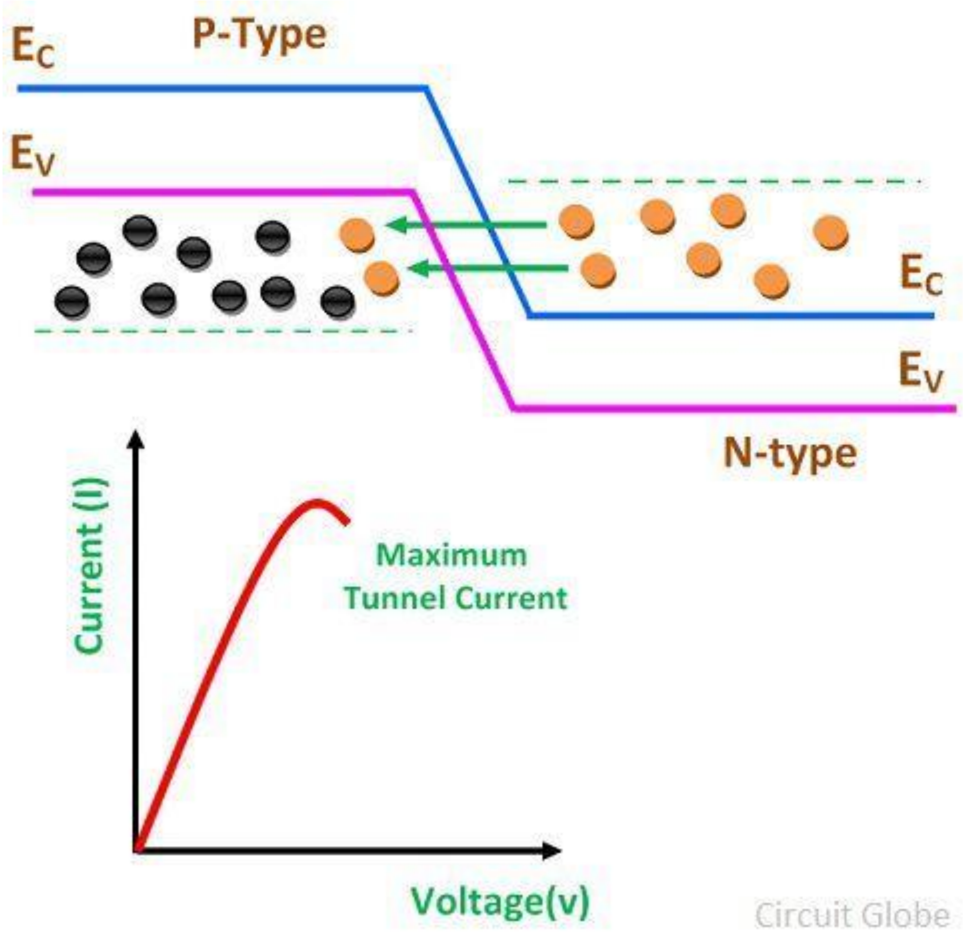
### UNBIASED DIODE

When the tunnel diode is unbiased it means no voltage appear across the diode. In that case, the conduction band of the entire semiconductor material overlaps with the valence band of the p-type material. And this happens because of the heavy doping. So, due to this overlapping when the temperature raises the electrons tunnel from the conduction band of the n region to the valence band of the p region and similarly the holes tunnel from the Valance band of the p region to the conduction band of the n region. the note for the point is that equal number of electrons from n region and holes from p region will flow in opposite direction so because of this equal number of electrons and holes the net current across the diode will be zero so we can say that "zero" current flows through this tunnel diode when it is in unbiased condition.

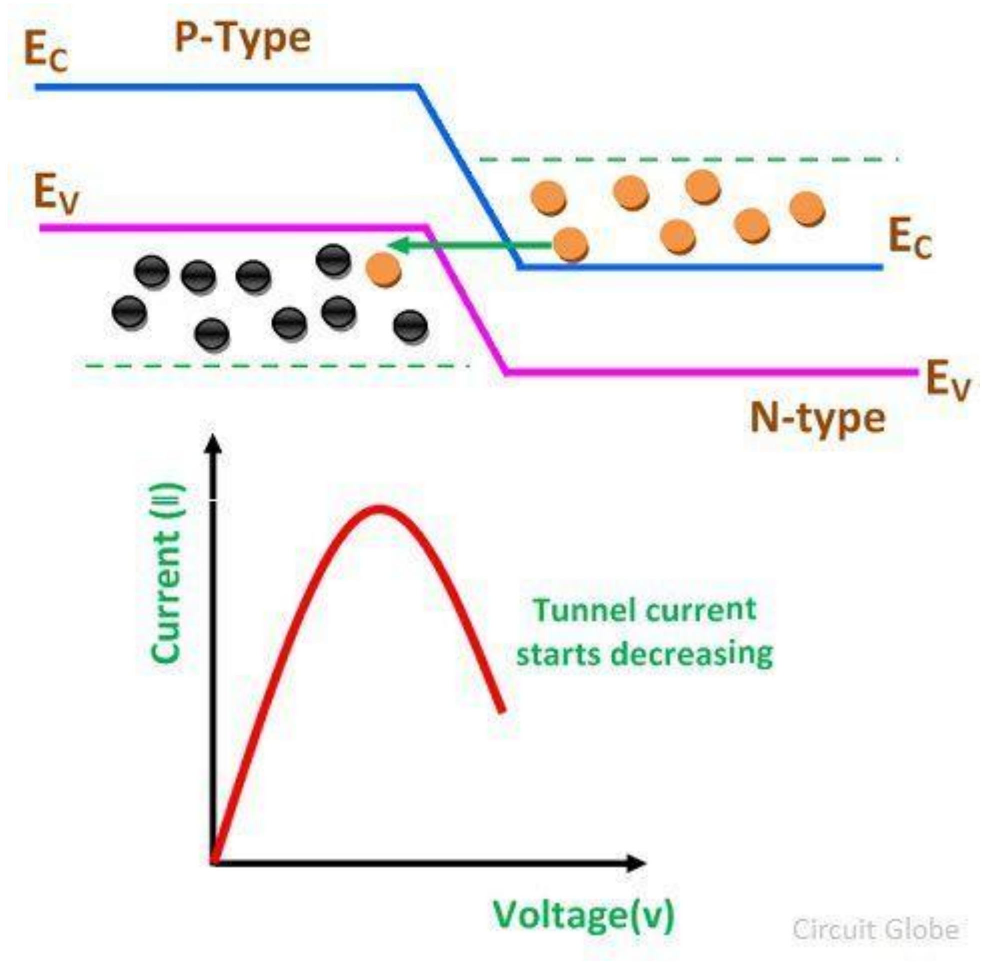


When we applied small voltage across tunnel diode and the magnitude is less than built-in voltage of the depletion region. So there are no forward current flows to the junction. And this is because no electrons can cross this depletion region. It means zero current flows through the diode. But however few electrons from n region of the conduction band are tunneled into the p region of the balance band and because of this tunneling of electrons small forward current flows through the diode. So as you can see below the VI characteristic graph of the tunnel diode when small voltage is applied then small tunnel current will flow through the diode.

When the voltage applied to the tunnel is slightly increased then due to this increase in voltage as you can see here overlapping of the conduction band and balance band is increased or in simple words the energy level of an inside conduction band becomes exactly equal to the energy level of a p side balance band. And as a result of this overlapping maximum tunnel current flows through the diode. So see here this is the VI characteristic graph of tunnel. So when the voltage is slightly increased so because of this overlapping here maximum tunnel current flows through the diode



Circuit Globe

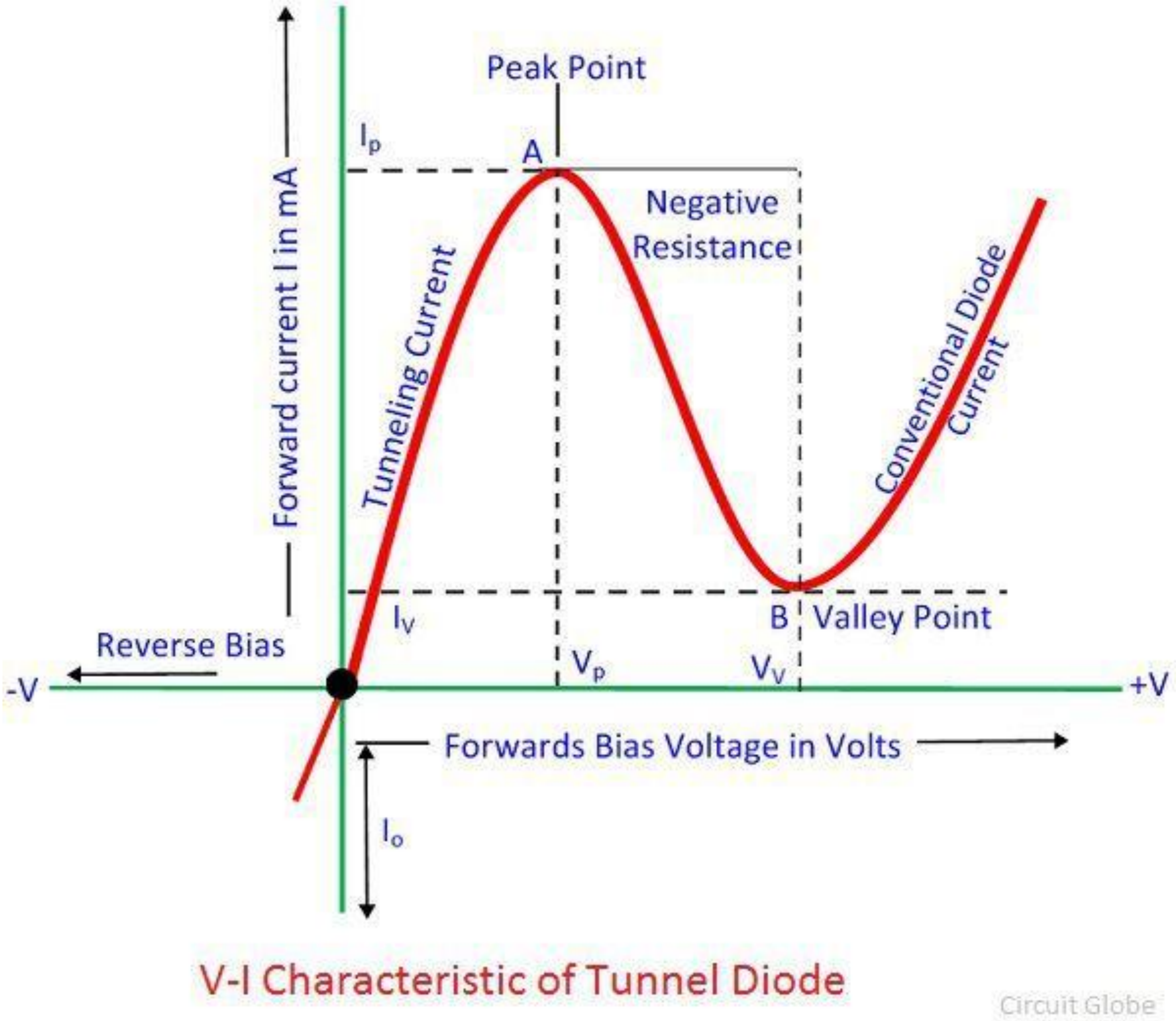


if the applied voltage is further increased then as you can see here a slight missile end of the conduction band and the balance band takes place but however here the conduction band of the n-type material and the valence band of the p-type material are still overlap. Because of this overlapping the electrons will tunnel from the conduction band of n region to the balance band of p region. This movement of electrons the tunneling current starts decreasing. If we go on increasing the voltage then the tunneling current drops to 0. At this particular point the conduction band and valence band are no longer overlap. Then it will operate in the same manner as a conventional PN junction semiconductor in which electrons are passing from conduction band of n type semiconductor to conduction band of p type semiconductor.

## V-I CHARACTERISTICS OF TUNNEL DIODE

Now let’s see the VI characteristic graph of the tunnel diode. If we go increasing the forward biasing voltage then the current increases to its peak point value which I represented as Ip. When we go increasing the voltage after this peak point voltage then the diode current starts decreasing till it reaches its minimum value called valley point. At this particular point the conduction band of n-type and variance band of p-type are no longer overlap. So when we go increasing the voltage after this valley voltage then the current starts increasing again as an ordinary PN junction diode. The region between point A and point B is called negative resistance region. In this negative resistance region as we discussed earlier current decreases with increase in applied voltage.

.



V-I Characteristic of Tunnel Diode