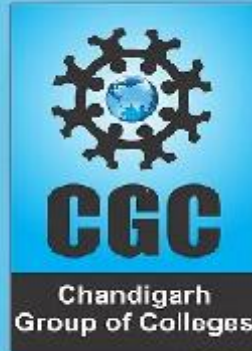


# SEMICONDUCTOR PHYSICS

## STUDY MATERIAL



## UNIT - 2 : SEMICONDUCTORS

BY:

**DR. ARVIND SHARMA**

**&**

**DR. SANGEETA**

**ASSOCIATE PROFESSOR**

**ASSISTANT PROFESSOR**

**DEPARTMENT OF APPLIED SCIENCES**

**CHANDIGARH ENGINEERING COLLEGE, LANDRAN**

# **Table of content**

## **UNIT – II**

- 2.1 Introduction**
- 2.2 Intrinsic and extrinsic semiconductors**
- 2.3 Intrinsic semiconductors**
- 2.4 Concentration of electrons in conduction band**
- 2.5 Concentration of holes in valance band**
- 2.6 Intrinsic concentration of charge carriers**
- 2.7 Fermi level in an intrinsic semiconductor**
- 2.8 Intrinsic conductivity**
- 2.9 Band Gap in an intrinsic semiconductor**
- 2.10 Extrinsic semiconductors**
- 2.11 Fermi level in n-type semiconductor**
- 2.12 Fermi level in p-type semiconductor**
- 2.13 Law of mass action in extrinsic semiconductor**
- 2.14 Charge neutrality condition**
- 2.15 Minority charge concentration**
- 2.16 Drift and diffusion current**
- 2.17 PN-junction**
- 2.18 Unbiased PN-junction**
- 2.19 Biased PN-junction**
- 2.20 Fermi level in PN-junction**
- 2.21 Law of junction**
- 2.22 Built-in Potential**
- 2.23 Diode equation**
- 2.24 Zener and avalanche breakdown**
- 2.25 Electric field**

- 2.26 Electric potential**
- 2.27 Space charge width**
- 2.28 Junction capacitance**
- 2.29 Metal-metal junction**
- 2.30 Metal-semiconductor junction**
- 2.31 Schottky junction**
- 2.32 Ohmic junction**
- 2.33 Differentiation between Schottky and PN-diode**

## **UNIT II - SEMICONDUCTORS**

### **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 movement. 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.

## 2.2 Intrinsic and Extrinsic Semiconductors

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. $N_e = N_h$	2. In case of N-type semiconductor, $N_e \gg N_h$ and in case of P-type, $N_h \gg N_e$
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.

## 2.3 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,

$$n = \int_{E_c}^{\infty} D(E) f(E) dE \quad (2.1)$$

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

$$D(E) = \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_c)^{1/2} \quad (2.2)$$

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

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

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

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

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

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

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

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

Putting  $(E - E_C)/KT = x$  so that  $dE = KTdx$  and changing the limits of integration as

$x \rightarrow 0$  When  $E \rightarrow E_C$  and  $x \rightarrow \infty$  when  $E \rightarrow \infty$

Expression (2.5) now becomes

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

i.e.

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

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

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

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

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

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

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

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

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  $-\infty$  and  $E_v$  will be



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

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

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

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

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

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

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

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

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

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

Expression (5) now becomes

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

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

i.e.

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

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

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

i.e.

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  $n_i$ , which is called intrinsic density or intrinsic concentration. Thus,

$$n = p = n_i$$

and  $n_i^2 = np.$

**This is called Law of Mass Action**

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

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

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

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

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

Note that the intrinsic density of charge carriers depends only on temperature and  $E_g$  but is independent of  $E_F$ .

## 2.7 Fermi level in an Intrinsic Semiconductor

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

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

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

Taking logarithm on both sides, we get

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

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

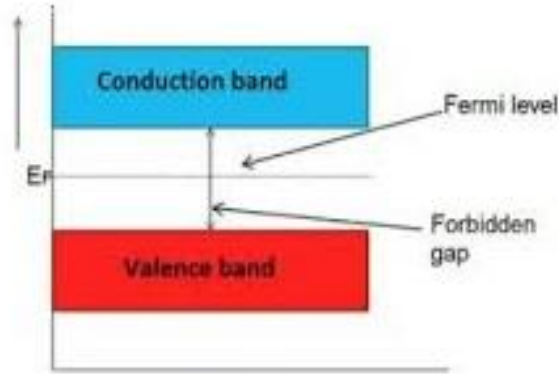
$$2E_F = E_C + E_V + \frac{3}{2} kT \ln \left( \frac{m_h^*}{m_e^*} \right)$$

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

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.



## 2.8 Intrinsic Conductivity

Since, both holes and electrons act as intrinsic charge carriers, so total current density in an intrinsic semiconductor will be

$$J = J_e + J_h = ne\mu_e E + pe\mu_h E = (ne\mu_e + pe\mu_h)E$$

Therefore, intrinsic conductivity will be

$$\sigma = \frac{J}{E} = (ne\mu_e + pe\mu_h) = en_i(\mu_e + \mu_h) \quad (\because n = p = n_i)$$

Substituting,

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

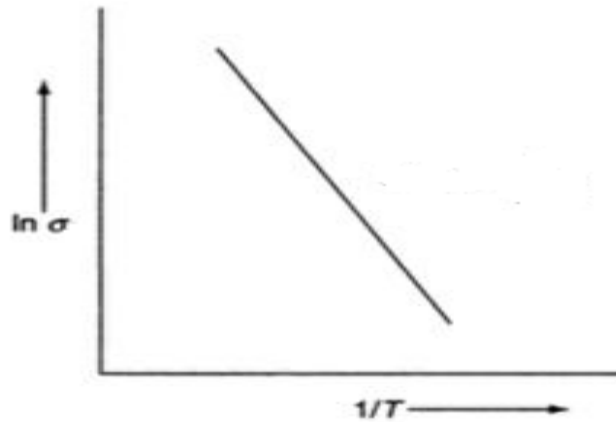
$$\sigma = 2e(\mu_e + \mu_h) \left[ \frac{2\pi kT}{h^2} \right]^{3/2} (m_e^* m_h^*)^{3/4} e^{-E_g/2kT}$$

i.e.  $\sigma = \sigma_0 e^{-E_g/2kT}$ , where  $\sigma_0 = 2e(\mu_e + \mu_h) \left[ \frac{2\pi kT}{h^2} \right]^{3/2} (m_e^* m_h^*)^{3/4}$

Taking logarithm on both sides, we get

$$\ln \sigma = \ln \sigma_0 - \frac{E_g}{2kT}$$

Plot of  $\ln \sigma$  versus  $1/T$  shows the conductivity increases with increase in temperature for intrinsic semiconductors.



## 2.9 Band Gap in an Intrinsic Semiconductor

Since,  $\sigma = \sigma_0 e^{-E_g/2kT}$

Taking reciprocal on both sides, we get

$$\rho = \rho_0 e^{E_g/2kT}$$

Or  $\frac{RA}{l} = \rho_0 e^{E_g/2kT}$   $(\because R = \frac{\rho l}{A})$  or  $(\rho = \frac{RA}{l})$

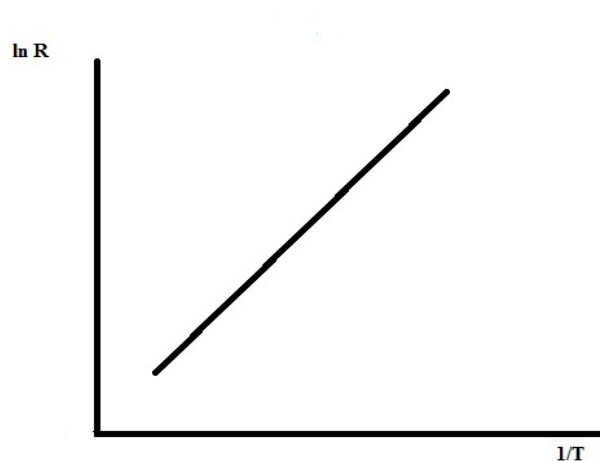
$$\therefore R = C e^{E_g/2kT}, \quad (\text{where } C = \frac{\rho_0 l}{A})$$

Taking logarithm on both sides, we get

$$\ln R = \ln C + \frac{E_g}{2kT}$$

Plot of  $\ln R_i$  versus  $1/T$  gives a straight line whose slope is  $\frac{E_g}{2kT}$ .

i.e.  $E_g = 2k \times \text{slope}$



### **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.

## **2.10 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  $\text{m}^3$ . 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-Type Semiconductor is formed when a trivalent impurity 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 like Gallium and Indium. Such type of impurities which produces p-type semiconductor are known as an Acceptor Impurities 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 Pentavalent impurity is added to a pure semiconductor providing a large number of free electrons in it, the extrinsic semiconductor thus formed is known as n-Type Semiconductor. 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.

## **2.11 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  $E=E_C$

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

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  $E = E_D$

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

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

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

$$\text{i.e. } n = N_D^+$$

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

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

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

$$n = N_D$$

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

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

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

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

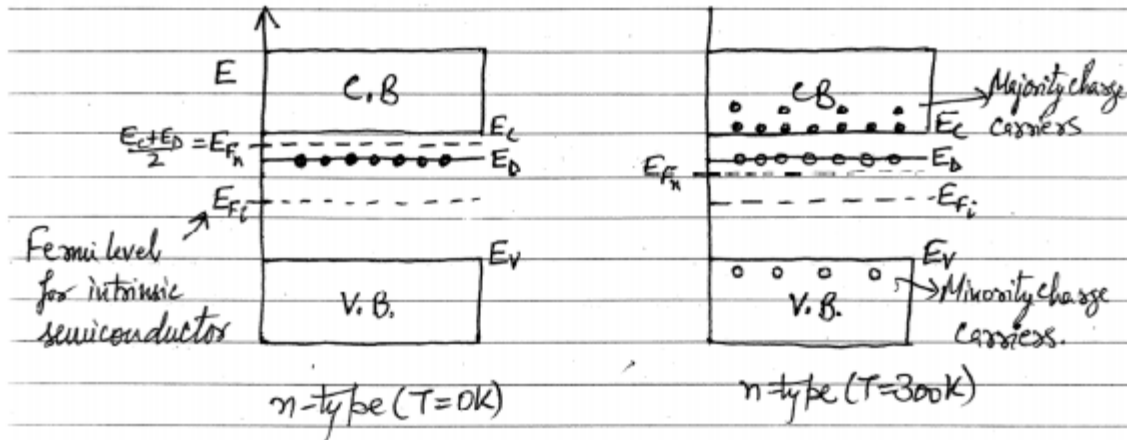
Taking logarithm on the both sides, we get,

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

Hence,

$$E_F = E_C - k_B T (\ln A_C / N_D)$$

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.



## 2.12 Fermi level in p-type semiconductor

At zero kelvin, all the acceptor levels are filled with holes and the valence band is completely filled with electrons. When the temperature is slightly increased, a small fraction of holes move from the acceptor level to the valence 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 valence band

$N_A$  is number density of acceptor atoms

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

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

$$\text{Likewise, } N_A^- = N_A [1 - F(E_A)]$$

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

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

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

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

$$\text{i.e. } p = N_A^-$$

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

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

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

$$p = N_A$$

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

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

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

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

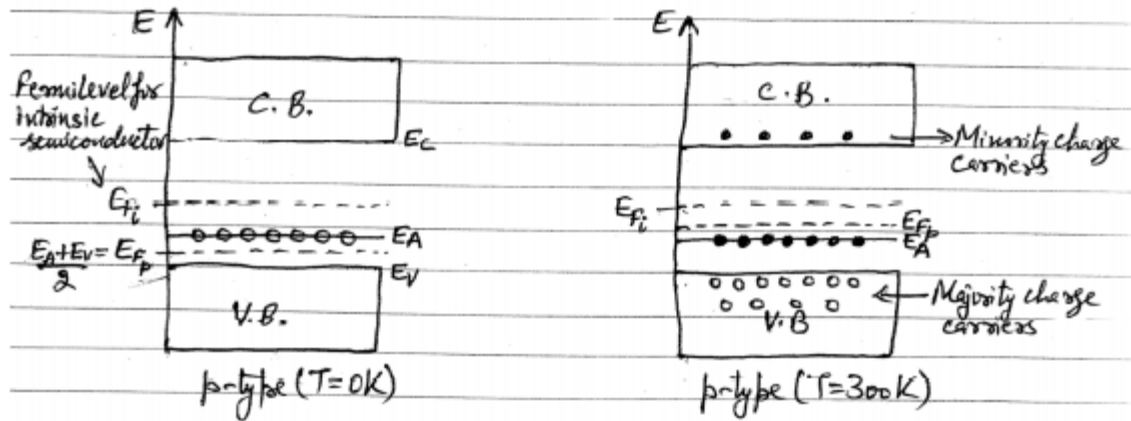
Taking logarithm on the both sides, we get,

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

Hence,

$$E_F = E_V + k T (\ln A_V / N_A)$$

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 broaden 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.

## 2.13 Law of mass action in extrinsic semiconductors:

In case of intrinsic semiconductors, we found that The law of mass action is applied for both intrinsic and extrinsic semiconductors. For extrinsic semiconductor the law of mass action states that the product of majority carriers and minority carriers is constant at fixed temperature and is independent of amount of donor and acceptor impurity added.

### Law of mass action for n-type semiconductor

The law of mass action for n-type semiconductor is mathematically written as

$$n_n p_n = n_i^2 = \text{constant}$$

Where  $n_n$  = number of electrons in n-type semiconductor and  $p_n$  = number of holes in n-type semiconductor.

The electrons are the majority carriers and holes are the minority carriers in n-type semiconductor. In n-type semiconductor, as the number of electrons (majority) in the conduction band increases the number of holes (minority) in the valence band decreases. Therefore, the product of electrons (majority) and holes (minority) remains constant at fixed temperature.

### **Law of mass action for p-type semiconductor**

The law of mass action for p-type semiconductor is mathematically written as

$$p_p n_p = n_i^2 = \text{constant}$$

Where  $p_p$  = number of holes in p-type semiconductor and  $n_p$  = number of electrons in p-type semiconductor. The holes are the majority carriers and electrons are the minority carriers in p-type semiconductor. In p-type semiconductor, as the number of holes (majority) in the valence band increases the number of electrons in the conduction band (minority) decreases. Therefore, the product of holes (majority) and electrons (minority) remains constant at fixed temperature.

## **2.14 Charge neutrality condition:**

Note that an extrinsic semiconductor is an electrically neutral material in its equilibrium condition. The charge neutrality condition applied to the n-type semiconductor implies that the total negative charge of free electrons in the conduction band is equal to the total positive charge of donor ions and holes in the valence band,

$$\text{i.e. } n_n = N_D + p_n$$

Where  $N_D$  is the donor impurity concentration in the saturation region (i.e. when  $N_D^+ = N_D$ ). But, in the n-type semiconductor, electrons are the majority charge carriers i.e.  $n_n \gg p_n$

So that, we get,  $n_n \approx N_D$

Hence the majority carrier concentration in the n-type semiconductor is approximately equal to the donor impurity concentration.

Likewise, the charge neutrality condition applied to the p-type semiconductor implies that the total positive charge of holes in the valance band is equal to the total negative charge of acceptor ions and free electrons in the conduction band, i.e.  $p_p = N_A + n_p$

Where,  $N_A$  is the acceptor impurity concentration in saturation region (i.e.  $N_A^- = N_A$ ). But in the p-type semiconductor, holes are the majority charge carriers, i.e.  $p_p \gg n_p$

So that, we get,  $p_p \approx N_A$

Hence the majority carrier concentration in p-type semiconductor is approximately equal to the acceptor impurity concentration.

## 2.15 Minority charge concentration:

**For n-type semiconductors:**

$$n_n = N_D \quad (\text{from charge neutrality condition})$$

$$n_n p_n = n_i^2 \quad (\text{from law of mass action})$$

$$\text{So, } p_n = n_i^2 / N_D$$

**For p-type semiconductors:**

$$P_p = N_A \quad (\text{from charge neutrality condition})$$

$$P_p n_p = n_i^2 \quad (\text{from law of mass action})$$

$$\text{So, } n_p = n_i^2 / N_A$$

## 2.16 Drift and diffusion current:

Under the condition of thermal equilibrium, the free electrons and holes are uniformly distributed inside the semiconductor (intrinsic as well as extrinsic) crystal and in the absence of an external stimulus their average velocity is zero and no current flows in the crystal.

The thermal equilibrium may be disturbed by an external agent and the random motion of the charge carriers acquire a directional movement leading to the flow of current in the crystal. Electric field and the concentration gradients are examples of such disturbing agents or external stimulus leading to drift current and diffusion current respectively

### a) Drift Current:

When an electric field ( $E$ ) is applied across a semiconductor, the charge carriers acquire a directional motion over and above their random thermal motion and produce a drift current as shown in fig. 1. Therefore, the drift current density due to motion of free electrons in the conduction band is given by,

$$J_n = n q \mu_n E$$

And, the drift current due to motion of free holes in the valance band is given by,

$$J_p = p q \mu_p E$$

Where,  $q$ ,  $n$ ,  $p$ ,  $\mu_n$  and  $\mu_p$  represents the magnitude of charge on the hole/electron, concentration of electrons, concentration of holes, mobility of electrons and mobility of holes respectively.

Therefore, the total drift current density can be written as,

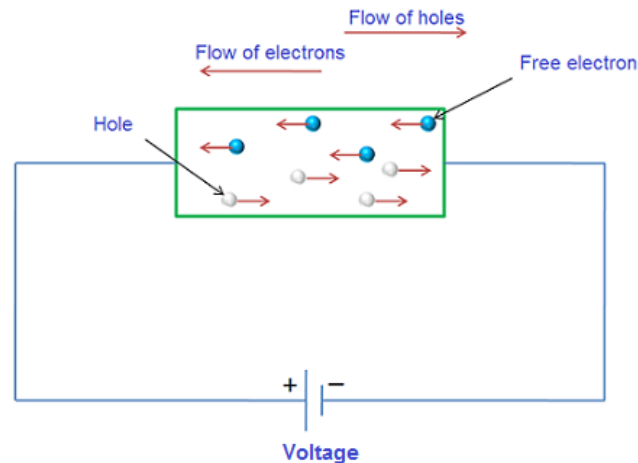
$$J^{drift} = J_n + J_p$$

$$J^{drift} = q E (n \mu_n + p \mu_p)$$

And hence the drift current is given by,

$$I^{drift} = \frac{J^{drift}}{E} = q (n \mu_n + p \mu_p)$$





**Fig. 1: semiconductor material under the applied electric field and the motion of electrons and holes**

Note that, the drift current flows only when an external electric field is applied across the semiconductor. Although electrons and holes move in opposite direction, the direction of conventional current flows due to the both charge carriers in the same direction. The electrons move with more ease compared to holes through the semiconductor I.e. the mobility of the electrons in a semi conducting material is considerably greater than that of holes. Therefore, the velocity of electron is greater than the velocity of holes under the applied electric field and hence semiconductor current is largely due to free electrons. Only a small fraction of current is due to holes.

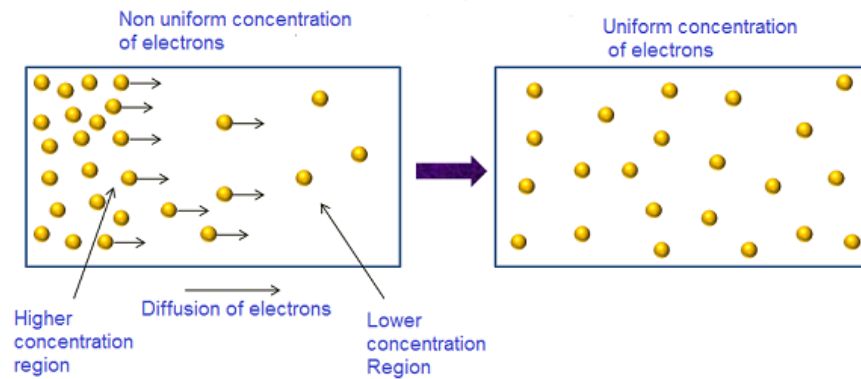
### **b) Diffusion Current:**

In case of semiconductors, current can also flow without the application of an external electric field. Another important component of current exists if there is a spatial variation of charge carrier concentration in a semiconductor material. The charge carriers tend to move from region of high concentration to the region of low concentration as shown in the fig. 2. This current component is called the diffusion current.

The diffusion current can be defined as the flow of charge carriers within a semiconductor travels from a higher concentration region to a lower concentration region. A higher concentration region is nothing but where the number of electrons present in the

semiconductor. Similarly, a lower concentration region is where the less number of electrons present in the semiconductor. The process of diffusion mainly occurs when a semiconductor is doped non-uniformly.

In an N-type semiconductor, when it is doped non-uniformly then a higher concentration region can be formed at the left side whereas the lower concentration region can be formed at the right side. The electrons in the higher concentration region are more in the semiconductor so they will experience a repulsive force from each other.



**Fig. 2: Flow of charge from high to low concentration**

To understand the diffusion process, let us assume an electron density that varies in x-direction as shown in the fig. 3. Here 'l' is the mean free path and  $n(x)$  is the electron density. The semiconductor is at uniform temperature so that the average thermal energy of electron does not vary with x-axis, only the density varies. Therefore,

$$l = v_{th}\tau_c$$

$v_{th}$ : Thermal velocity of electron

$\tau_c$ : Relaxation time

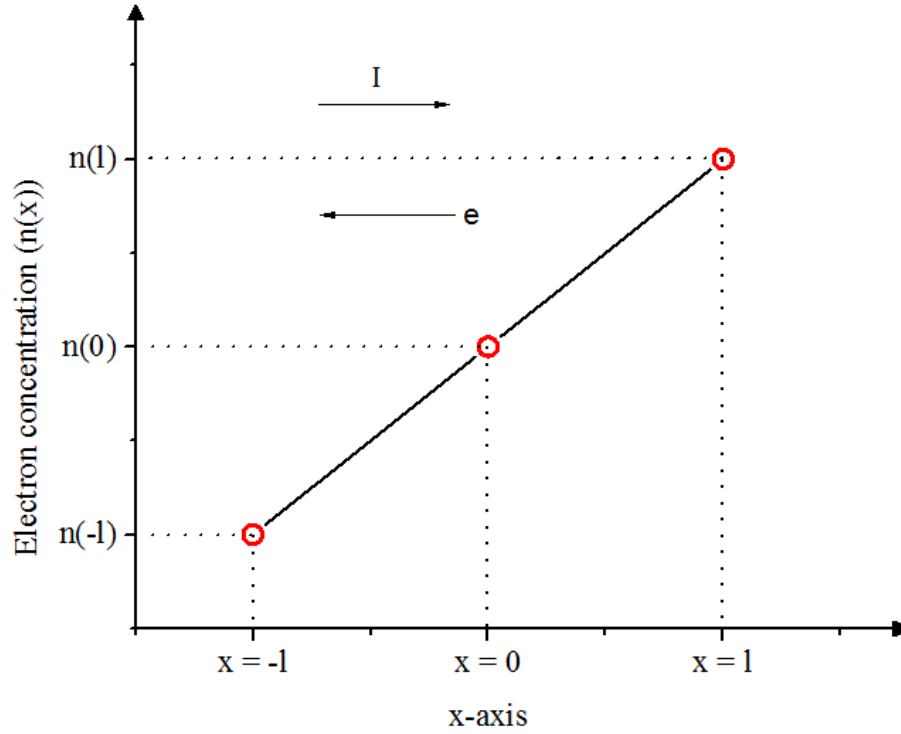
Let us evaluate the net flow of electrons per unit time per unit area crossing the plane at  $x=0$ . The electrons at  $x=-l$  are one mean free path away from the left side have equal chances of moving left or right in a mean free time ( $\tau_c$ ). Therefore, one half of them will move across the plane  $x=0$ . Therefore the average rate of flow of electron per unit area is defined as,

$F_1 = \frac{1}{2} \times \text{average rate of flow of electrons per unit area crossing the plane } x=0 \text{ from the left}$

$F_2 = \frac{1}{2} \times \text{average rate of flow of electrons per unit area crossing the plane } x=0 \text{ from the right.}$

Therefore,  $F_1 = \frac{1}{2} \frac{n(-l)l}{\tau_c}$  and  $F_2 = \frac{1}{2} \frac{n(l)l}{\tau_c}$

Or,  $F_1 = \frac{1}{2} n(-l)v_{th}$  and  $F_2 = \frac{1}{2} n(l)v_{th}$



**Fig. 3: Electron concentration versus spatial distance**

Net rate of flow of charge carriers from left to right is,

$$F = F_1 - F_2$$

$$F = \frac{1}{2} n(-l)v_{th} - \frac{1}{2} n(l)v_{th}$$

$$F = \frac{1}{2} v_{th}[n(-l) - n(l)]$$

Approximating the densities at  $x = \mp l$  by the first two terms of a Taylor series expansion,

$$F = \frac{1}{2} v_{th} \left[ \left( n(0) - l \frac{dn}{dx} \right) - \left( n(0) + l \frac{dn}{dx} \right) \right]$$

$$F = -v_{th} l \frac{dn}{dx}$$

$l \frac{dn}{dx}$ , is the gradient of concentration of charge carriers.

Therefore,  $F = -D_n \frac{dn}{dx}$ , where,  $D_n$  is the diffusion coefficient or diffusivity constant for electrons having unit  $\text{cm}^2/\text{s}$ . The diffusion coefficient mean the amount of a substance that diffuses from one section to another through every unit of the cross-section for each unit of time as the gradient of volume-concentration is unity. Now, because each electron carries a charge of ‘-q’, the carrier flow gives rise to diffusion current and the diffusion current density is defined as,

$$J_n = -q F = q D_n \frac{dn}{dx}$$

Hence, the diffusion current density is proportional to the spatial derivative of electron density. Diffusion current results from the random thermal motion of charge carriers in a concentration gradients. For an electron density that increases with ‘x’ the gradient is positive and the electron will diffuse towards the negative x-direction, i.e. backward and hence the current is positive. The corresponding current density expression for the diffusion of holes is given by,

$$J_p = q F = -q D_p \frac{dp}{dx}$$

Here,  $D_p$  is the diffusion coefficient or diffusivity constant for holes. The total diffusion current density is now given by,

$$J^{diffusion} = J_n + J_p$$

$$J^{diffusion} = q D_n \frac{dn}{dx} - q D_p \frac{dp}{dx}$$

Hence, the total current in a semiconductor is represented by,

$$I = I^{drift} + I^{diffusion} = \frac{(J^{drift} + J^{diffusion})}{E}$$

$$I = q (n \mu_n + p \mu_p) + \frac{1}{E} \left( q D_n \frac{dn}{dx} - q D_p \frac{dp}{dx} \right)$$

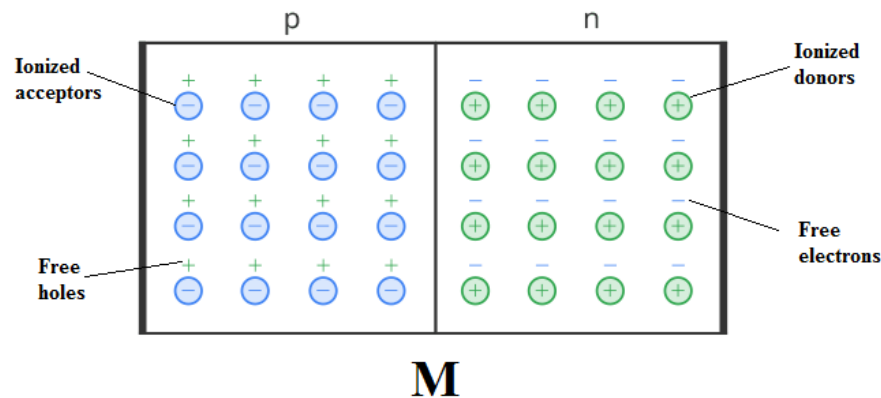
### c) Difference between drift and diffusion current

Drift Current		Diffusion Current	
1	The movement of charge carriers is because of the applied electric field is known as drift current.	1	The diffusion current can be occurred because of the diffusion in charge carriers.
2	It requires electrical energy for the process of drift current.	2	Some amount of external energy is enough for the process of diffusion current.
3	This current obeys Ohm's Law.	3	This current obeys Fick's Law.
4	The direction of charge carriers in the semiconductor is reverse to each other.	4	For charge carriers, the densities of diffusion are reverse in symbol to each other.
5	The direction of the drift current, as well as the electric field, will be the same.	5	The direction of this current can be decided by the concentration of the carrier slope.
6	It depends on the permittivity.	6	It is independent of permittivity.
7	The direction of this current mainly depends on the polarity of the applied electric field.	7	The direction of this current mainly depends on the charge within the concentrations of carrier.

## 2.17 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.



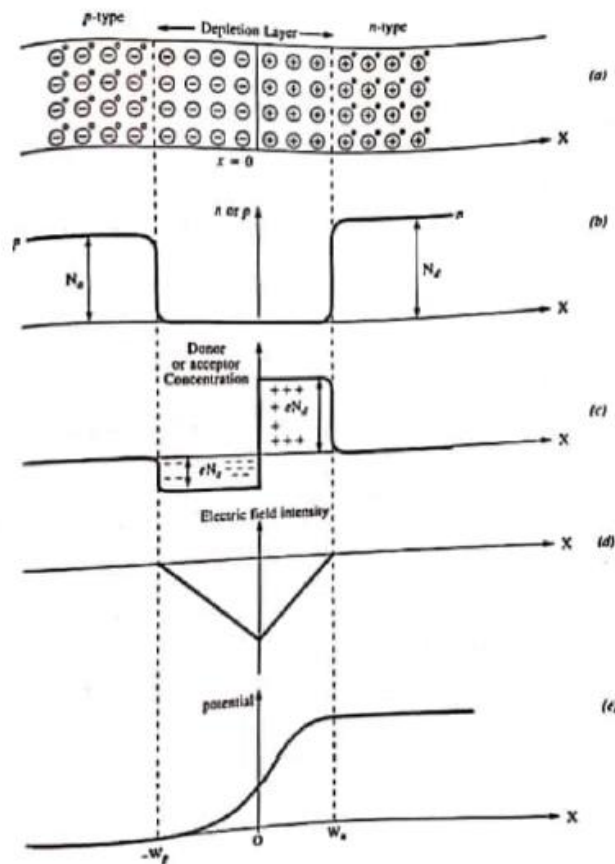
**Fig. 4: Basic pn-junction development**

## **2.18 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.

## 2.19 Biased PN-junction:

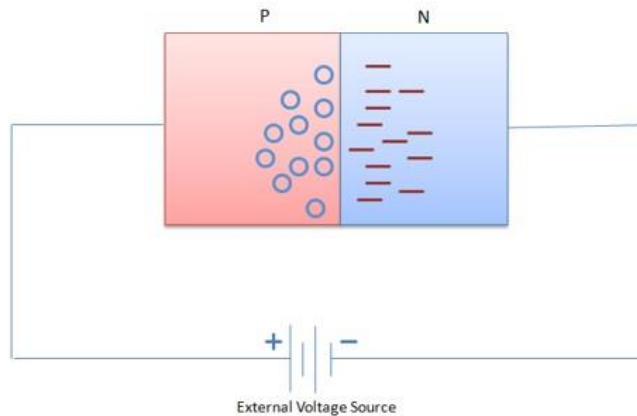
A pn-junction is said to be biased, if a direct potential difference is applied across the n and p-region of the pn-junction. The term bias refers to the application of DC voltage to set up certain operating conditions. Or when an external source of energy is applied to a pn-junction it is called a bias 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

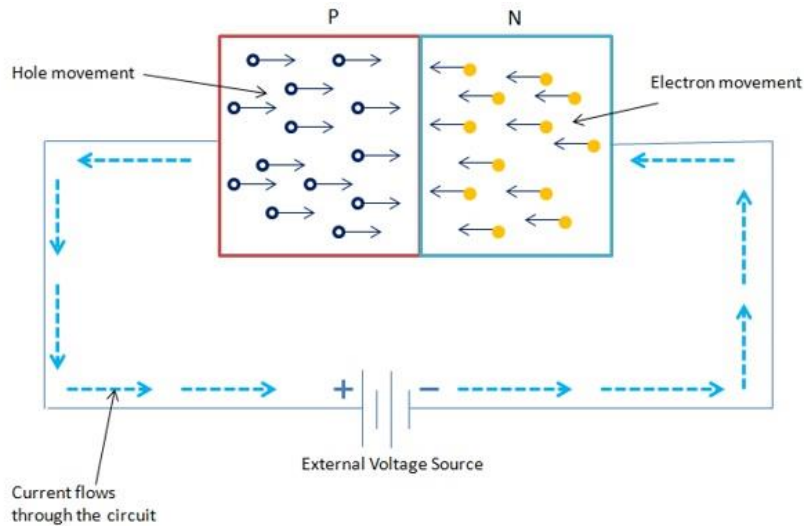
The following figure shows a forward biased PN junction diode with external voltage applied. You can see that the positive terminal of the battery is connected to the P material and the negative terminal of the battery is connected to the N material.



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.

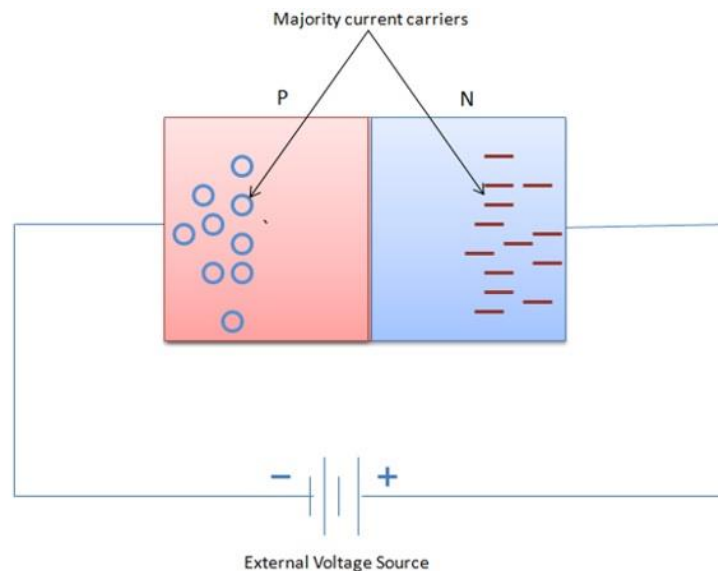


Following are the observations –

- 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

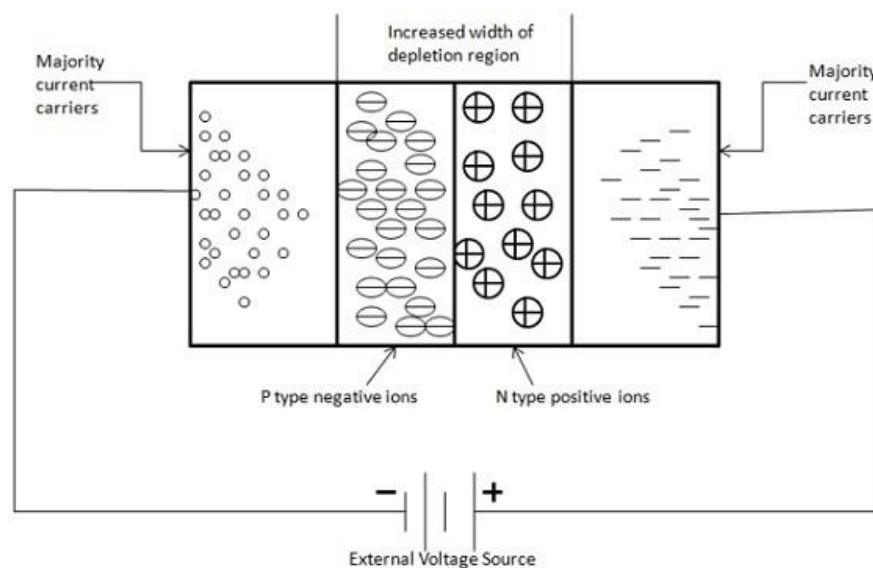
The following figure shows reverse biased PN junction diode with external voltage applied. You can see that the positive terminal of the battery is connected to the N material and the negative terminal of the battery is connected to the P material. Note that in such an arrangement, battery polarity is 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.

## 2.20 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  $V_B$  is the barrier height, then the shift in the conduction band of p-type semiconductor  $eV_B$  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.

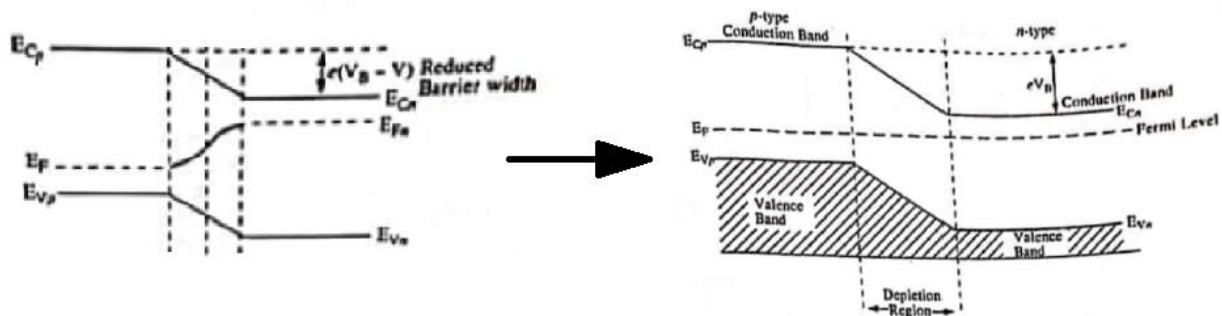


Fig. 6: Energy band diagram of pn-junction

## 2.21 Law of junction:

Note that the majority carriers that cross the junction and arrives in the opposite region become minority carriers in that region. Therefore, due to forward bias, the number of minority charge carriers in that region increases to a very large value. The increase in the minority carrier concentration can be estimated as follows:

Under equilibrium condition, the electron concentration on the p-side of the junction is calculated before, as

$$n_p = n_n \exp \left[ \frac{-eV_o}{k T} \right]$$

Due to the application of forward bias,  $V_o$  reduces to  $V_o - V_F$ , whereas,  $V_F$  is the forward biasing potential. This causes an increase in the value of  $n_p$ . Let  $n_p^*$  be the enhanced concentration of electrons at the junction on the p-sed. Then,

$$n_p^* = n_n \exp \left[ \frac{-e(V_o - V_F)}{k T} \right]$$

$$n_p^* = n_n \exp \left[ \frac{-eV_o}{k T} \right] \exp \left[ \frac{eV_F}{k T} \right]$$

In above equation, as we know,  $n_p = n_n \exp \left[ \frac{-eV_o}{k T} \right]$ , therefore, we get,

$$n_p^* = n_p \exp \left[ \frac{eV_F}{k T} \right]$$

Similarly,

$$p_n^* = p_n \exp \left[ \frac{eV_F}{k T} \right]$$

Hence, upon applying a small forward bias, an exponential increase in the number of electron occurs on p-sed of the junction or the probability of the majority carriers diffusing the junction is increased by a factor of,  $\exp \left[ \frac{eV_F}{k T} \right]$ . This is known as law of junction.

## 2.22 Built-in Potential:

To find the barrier potential, we must keep in mind the following points:

- (a) For a semiconductor doped with different impurity and different concentration, the position of the conduction band and valance band do not change, whereas; the

position of Fermi level changes. Therefore, if  $E_{Cn}$  and  $E_{Vn}$  are the energies of bottom of conduction band and top of valance band then,

$$E_{Cn} = E_{Cp} = E_C$$

$$E_{Fn} \neq E_{Fp}$$

The difference in the Fermi level corresponding to n-type and p-type semiconductor is equal to the energy gained by an electron in jumping over the potential barrier.

$$i.e. E_{Fn} - E_{Fp} = eV_B$$

- (b) The energy band treatment allows a simple way to calculate  $V_B$ . When the junction is formed  $E_{Fn}$  and  $E_{Fp}$  must shift and line up. This shift in the  $E_{Fp}$  and  $E_{Fn}$  to line up is clearly  $\phi_p - \phi_n$ , the work function difference. Thus the potential barrier  $eV_B$  is  $\phi_p - \phi_n$ .

$$eV_B = \phi_p - \phi_n = (E_C - E_{Fp}) - (E_C - E_{Fn})$$

But on the p- and n-side, the electron concentration in thermal equilibrium are given by,

$$n_{po} = N_C \exp \left[ -\frac{(E_C - E_{Fp})}{k T} \right]$$

$$n_{no} = N_C \exp \left[ -\frac{(E_C - E_{Fn})}{k T} \right]$$

From above two equations, we can now substitute the values  $(E_C - E_{Fp})$  and  $(E_C - E_{Fn})$  in the expression for ' $eV_B$ '. Therefore, we obtain,

$$eV_B = kT \ln \left( \frac{n_{no}}{n_{po}} \right)$$

Since,  $n_{po} = n_i^2 / N_a$  and  $n_{no} = N_d$ , we readily obtain the built-in potential as,

$$V_B = \frac{k T}{e} \ln \left[ \frac{N_a N_d}{n_i^2} \right]$$

## 2.23 Diode Equation:

When a forward bias or reverse bias is applied to a pn-junction, a net current starts flowing through the circuit. A relation may be derived between the current flowing through the junction and the potential difference applied. This relation is termed as the diode equation or rectifier equation. In deriving this equation, we make the assumptions that the density of minority carriers injected into a region (p-type or n-type) is negligible as compared to the density of majority carriers and that the depletion region is marked by well-defined boundaries within which there is no generation current.

Let  $n_p$  be the concentration of electrons (minority carriers) in p-region and  $n_n$  be the concentration of electrons (majority carriers) in n-region. Then at equilibrium,  $n_n$  and  $n_p$  are related as,

$$n_p = n_n \exp \left[ \frac{-eV_B}{k T} \right] \quad (1)$$

Similarly, the hole concentration in n-region and the hole concentration in p-region are related as,

$$p_n = p_p \exp \left[ \frac{-eV_B}{k T} \right] \quad (2)$$

When a forward bias is applied, the barrier potential is reduced to  $V_B - V$ . Therefore, electron density in p-region is found using equation (1),

$$n_p + \Delta n_p = n_n \exp \left[ \frac{-e(V_B - V)}{k T} \right] = n_n \exp \left[ \frac{-eV_B}{k T} \right] \exp \left[ \frac{eV}{k T} \right] \quad (3)$$

And the hole density in n-region is found by using equation (2)

$$p_n + \Delta p_n = p_p \exp \left[ \frac{-e(V_B - V)}{k T} \right] = p_p \exp \left[ \frac{-eV_B}{k T} \right] \exp \left[ \frac{eV}{k T} \right] \quad (4)$$



The increase in the electron density  $\Delta n_p$  in the p-region and the increase in hole density in n-region can be found by subtracting equation (1) from equation (3) and equation (2) from equation (4) respectively and is given by

$$\Delta n_p = n_n \exp \left[ \frac{-eV_B}{k T} \right] \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right] \quad (5)$$

$$\Delta p_n = p_p \exp \left[ \frac{-eV_B}{k T} \right] \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right] \quad (6)$$

If 'A' is the area of the junction,  $v_e$  is the drift velocity of electrons and  $v_h$  is the drift velocity of holes, then the current due to electrons is,

$$I_e = e A v_e \Delta n_p$$

On substituting the expression of  $\Delta n_p$  in above equation, we get,

$$I_e = e A v_e n_n \exp \left[ \frac{-eV_B}{k T} \right] \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right]$$

Consider,  $B_e = e A v_e n_n \exp \left[ \frac{-eV_B}{k T} \right]$ , therefore,

$$I_e = B_e \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right] \quad (7)$$

Als, the current due to holes is,

$$I_h = e A v_h \Delta p_n$$

On substituting the expression of  $\Delta p_n$  in above equation, we get,

$$I_h = e A v_h p_p \exp \left[ \frac{-eV_B}{k T} \right] \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right]$$

Consider,  $B_h = e A v_h p_p \exp \left[ \frac{-eV_B}{k T} \right]$ , therefore,

$$I_h = B_h \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right] \quad (8)$$

Therefore, the total current across the junction is,

$$I = I_e + I_h$$

Using equation (7) and (8), we get,

$$\begin{aligned} I &= B_e \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right] + B_h \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right] \\ I &= (B_e + B_h) \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right] \end{aligned} \quad (9)$$

When a reverse potential (-V) is applied, the exponential term becomes very small in comparison to (-1). Therefore, under reverse potential,

$$I = -(B_e + B_h) = -I_s$$

Here,  $I_s$  is a constant called the reverse saturation current, Therefore, equation (9) may be written as,

$$I = I_s \left[ \exp \left[ \frac{eV}{k T} \right] - 1 \right] \quad (10)$$

The equation (10) is called the voltage current equation or rectifier equation or diode equation.

### Special cases:

(a) When the pn-junction is forward bias, then V is large. Therefore,

$$\exp \left( \frac{e V}{k T} \right) \gg 1$$

Therefore, equation (10) becomes,

$$I = I_s \exp \left[ \frac{eV}{kT} \right] \quad (11)$$

Thus the current increases with an increase in the forward bias in an exponential way.

- (b) When the pn-junction is reverse biased, then  $V$  is negative, Therefore,  $\exp \left( \frac{eV}{kT} \right)$  is very small and can be neglected, therefore, from equation (10),

$$I = I_s(0 - 1) = -I_s$$

i.e. the current during reverse biasing is independent of applied voltage.

- (c) At ordinary temperatures, reverse saturation current is quite small due to low rate of generation of minority carriers. But it depends strongly on temperature. Its value is generally doubled for each  $10^\circ\text{C}$  rise in junction temperature and continues to increase up to  $100^\circ\text{C}$  for germanium device and up to  $200^\circ\text{C}$  for silicon devices. The equation (10) is derived by using the term  $\frac{eV}{kT}$  in the exponential. Generally, the equation (10) is given as,

$$I = I_s \left[ \exp \left[ \frac{eV}{\eta kT} \right] - 1 \right] \quad (12)$$

Here,  $\eta = 1$  for germanium and 2 for silicon.

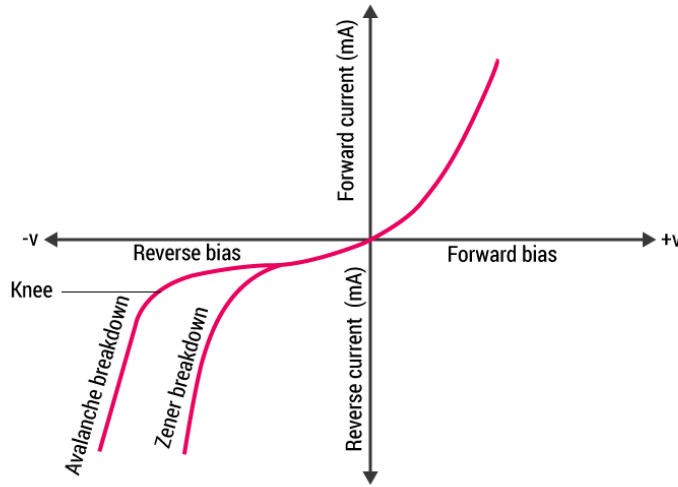
## **2.24 Zener and avalanche breakdown:**

- (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  $2 \times 10^7 \text{Vm}^{-1}$ .

**(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 voltage of 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.



## 2.25 Electric field:

An electric field is created in the depletion region by the separation of positive and negative space charge densities. Figure 7.4 shows the volume charge density distribution in the pn junction assuming uniform doping and assuming an abrupt junction approximation. We will assume that the space charge region abruptly ends in the n region at  $x = +x_n$  and abruptly ends in the p region at  $x = -x_p$  ( $x_p$  is a positive quantity).

The electric field is determined from Poisson's equation, which, for a one-dimensional analysis, is

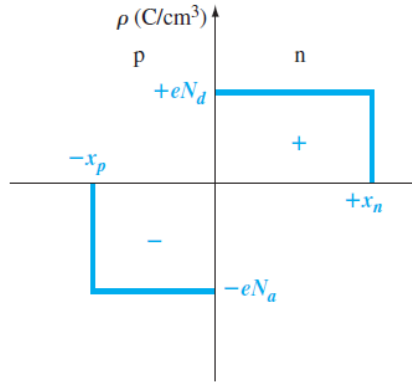
$$\frac{d^2\phi(x)}{dx^2} = \frac{-\rho(x)}{\epsilon_s} = \frac{-dE(x)}{dx} \quad (1)$$

where  $\phi(x)$  is the electric potential,  $E(x)$  is the electric field,  $\rho(x)$  is the volume charge density, and  $\epsilon_s$  is the permittivity of the semiconductor. From Figure 6, the charge densities are,

$$\rho(x) = -e N_a; -x_p < x < 0 \quad (2a)$$

And,

$$\rho(x) = -e N_d; 0 < x < x_n \quad (2b)$$



**Fig. 7: The space charge density in a uniformly doped pn junction assuming the abrupt junction approximation.**

The electric field in the p-region is found by integrating equation (1). We have,

$$E = \int \frac{\rho(x)}{\epsilon_s} dx = - \int \frac{e N_a}{\epsilon_s} dx = - \frac{e N_a}{\epsilon_s} x + C_1 \quad (3)$$

where  $C_1$  is a constant of integration. The electric field is assumed to be zero in the neutral p-region for  $x < -x_p$  since the currents are zero in thermal equilibrium. Since there are no surface charge densities within the pn junction structure, the electric field is a continuous function. The constant of integration is determined by setting  $E = 0$  at  $x = -x_p$ .

$$C_1 = - \frac{e N_a}{\epsilon_s} x_p$$

The electric field in the p region is then given by

$$E = - \frac{e N_a}{\epsilon_s} (x + x_p); \quad -x_p < x < 0 \quad (4)$$

In the n-region, the electric field is determined from,

$$E = \int \frac{\rho(x)}{\epsilon_s} dx = \int \frac{e N_d}{\epsilon_s} dx = \frac{e N_d}{\epsilon_s} x + C_2 \quad (5)$$

where  $C_2$  is again a constant of integration and is determined by setting  $E = 0$  at  $x = x_n$ ,

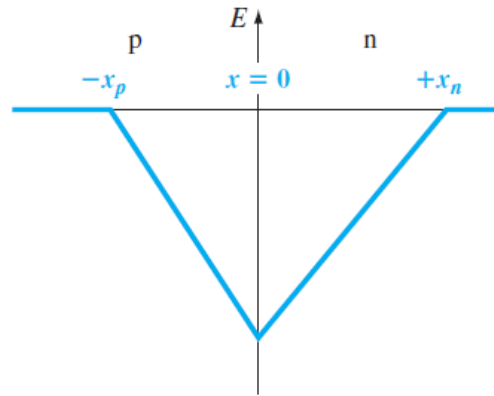
$$C_2 = -\frac{e N_d}{\epsilon_s} x_n$$

Since the E-field is assumed to be zero in the n region and is a continuous function. Then,

$$E = -\frac{e N_d}{\epsilon_s} (x_n - x); \quad 0 < x < x_n \quad (6)$$

The electric field is also continuous at the metallurgical junction, or at  $x = 0$ . Setting Equations (4) and (6) equal to each other at  $x = 0$  gives,

$$N_d x_n = N_a x_p \quad (7)$$



**Fig. 8: Electric field in the space charge region of a uniformly doped pn junction.**

Equation (7) states that the number of negative charges per unit area in the p region is equal to the number of positive charges per unit area in the n region. Figure 8 is a plot of the electric field in the depletion region. The electric field direction is from the n to the p region, or in the negative  $x$  direction for this geometry. For the uniformly doped pn junction, the E-field is a linear function of distance through the junction, and the maximum (magnitude) electric field occurs at the metallurgical junction. An electric field exists in the depletion region even when no voltage is applied between the p and n regions.

The magnitude of the electric field in the depletion region increases with an applied reverse-biased voltage. The electric field is still given by Equations (4) and (6) and is still a linear function of distance through the space charge region. Since  $x_n$  and  $x_p$  increase with reverse-biased voltage, the magnitude of the electric field also increases. The maximum electric field still occurs at the metallurgical junction. The maximum electric field at the metallurgical junction, from Equations (4) and (6), is

$$E = -\frac{e N_d x_n}{\epsilon_s} = -\frac{e N_a x_p}{\epsilon_s}$$

## 2.26 Electric potential:

The potential in the junction is found by integrating the electric field. In the p-region then, we have,

$$\phi(x) = -\int E(x)dx = \int \frac{e N_a}{\epsilon_s}(x + x_p) dx \quad (8)$$

$$\phi(x) = \frac{e N_a}{\epsilon_s} \left( \frac{x^2}{2} + x_p \cdot x \right) + C'_1 \quad (9)$$

where  $C'_1$  is again a constant of integration. The potential difference through the pn-junction is the important parameter, rather than the absolute potential, so we may arbitrarily set the potential equal to zero at  $x = -x_p$ . The constant of integration is then found as,

$$C'_1 = \frac{e N_a}{2\epsilon_s} x_p^2 \quad (10)$$

So that the potential in the p-region can now be written as,

$$\phi(x) = \frac{e N_a}{2\epsilon_s} (x + x_p)^2; -x_p < x < 0 \quad (11)$$

The potential in the n region is determined by integrating the electric field in the n region, or

$$\phi(x) = -\int E(x)dx = \int \frac{e N_d}{\epsilon_s} (x_n - x) dx \quad (12)$$

$$\phi(x) = \frac{e N_d}{\epsilon_s} \left( x_n \cdot x - \frac{x^2}{2} \right) + C'_2 \quad (13)$$



where  $C'_2$  is another constant of integration. The potential is a continuous function, so setting Equation (11) equal to Equation (13) at the metallurgical junction or at  $x = 0$ , gives,

$$C'_2 = \frac{e N_a}{2\epsilon_s} x_p^2 \quad (14)$$

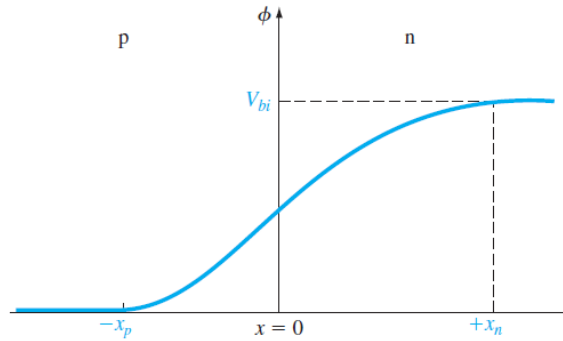
The potential in the n region can thus be written as,

$$\phi(x) = \frac{e N_d}{\epsilon_s} \left( x_n \cdot x - \frac{x^2}{2} \right) + \frac{e N_a}{2\epsilon_s} x_p^2; \quad 0 < x < x_n \quad (15)$$

Figure 9 is a plot of the potential through the junction and shows the quadratic dependence on distance. The magnitude of the potential at  $x = x_n$  is equal to the built-in potential barrier. Then from Equation (15), we have

$$V_B = |\phi(x = x_n)| = \frac{e}{2\epsilon_s} (N_d x_n^2 + N_a x_p^2) \quad (16)$$

The potential energy of an electron is given by  $E = -e\phi$ , which means that the electron potential energy also varies as a quadratic function of distance through the space charge region.



**Fig. 9: Electric potential through the space charge region of a uniformly doped pn-junction.**

## 2.27 Space charge width

We can determine the distance that the space charge region extends into the p and n regions from the metallurgical junction. This distance is known as the space charge width. From Equation (7), we may write, for example

$$x_p = \frac{N_d x_n}{N_a} \quad (17)$$

Then, substituting Equation (17) into Equation (16) and solving for  $x_n$ , we obtain,

$$x_n = \left[ \frac{2\epsilon_s V_B}{e} \left( \frac{N_a}{N_d} \right) \left( \frac{1}{N_a + N_d} \right) \right]^{1/2} \quad (18)$$

Equation (18) gives the space charge width, or the width of the depletion region,  $x_n$  extending into the n-type region for the case of zero applied voltage. Similarly, if we solve for  $x_p$  from Equation (7) and substitute into Equation (16), we find

$$x_p = \left[ \frac{2\epsilon_s V_B}{e} \left( \frac{N_d}{N_a} \right) \left( \frac{1}{N_a + N_d} \right) \right]^{1/2} \quad (19)$$

where  $x_p$  is the width of the depletion region extending into the p region for the case of zero applied voltage. The total depletion or space charge width  $W$  is the sum of the two components, or

$$W = x_n + x_p \quad (20)$$

Using Equations (18) and (19), we obtain

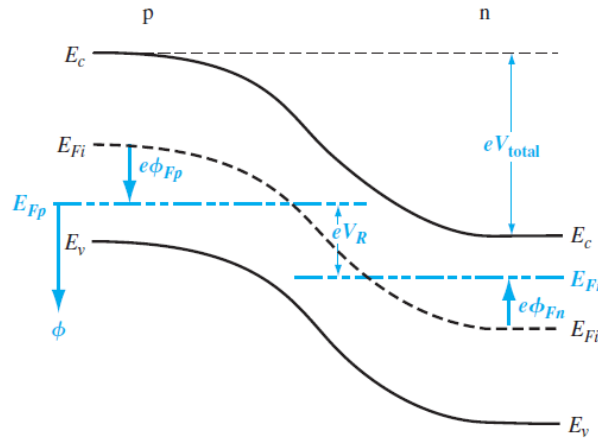
$$W = \left[ \frac{2\epsilon_s V_B}{e} \left( \frac{N_a N_d}{N_a + N_d} \right) \right]^{1/2} \quad (21)$$

If we apply a potential between the p and n-regions, then the pn-junction will no longer be in an equilibrium condition—the Fermi energy level will no longer be constant through the system. Figure 10 shows the energy-band diagram of the pn-junction for the case when a positive voltage is applied to the n-region with respect to the p-region.

As the positive potential is downward, the Fermi level on the n-side is below the Fermi level on the p-side. The difference between the two is equal to the applied voltage in units of energy. The total potential barrier, indicated by  $V_{total}$ , has increased. The applied potential is the reverse-biased condition. The total potential barrier is now given by

$$V_{total} = V_B + V_R \quad (22)$$

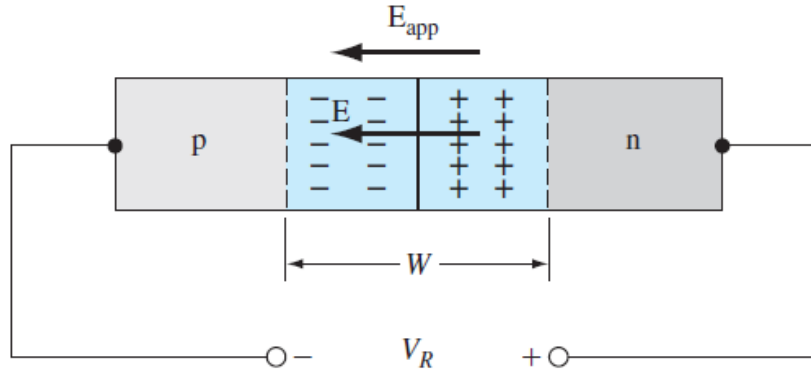
where  $V_R$  is the magnitude of the applied reverse-biased voltage and  $V_B$  is the same built-in potential barrier.



**Fig. 10: Energy-band diagram of a pn-junction under reverse bias.**

Figure 11 shows a pn-junction with an applied reverse-biased voltage  $V_R$ . Also indicated in the figure are the electric field in the space charge region and the electric field  $E_{app}$ , induced by the applied voltage. The electric fields in the neutral p and n regions are essentially zero, or at least very small, which means that the magnitude of the electric field in the space charge region must increase above the thermal-equilibrium value due to the applied voltage. The electric field originates on positive charge and terminates on negative charge; this means that the number of positive and negative charges must increase if the electric field increases. For given impurity doping concentrations, the number of positive and negative charges in the depletion region can be increased only if the space charge width  $W$  increases. The space charge width  $W$  increases, therefore, with an increasing reverse-biased voltage  $V_R$ . We are

assuming that the electric field in the bulk n and p regions is zero. This assumption will become clearer in the next chapter when we discuss the current–voltage characteristics.



**Fig. 11: A pn junction, with an applied reverse-biased voltage, showing the directions of the electric field induced by  $V_R$  and the space charge electric field.**

In all of the previous equations, the built-in potential barrier can be replaced by the total potential barrier. The total space charge width can be written from Equation (21) as,

$$W = \left[ \frac{2\epsilon_s(V_B + V_R)}{e} \left( \frac{N_a N_d}{N_a + N_d} \right) \right]^{1/2} \quad (23)$$

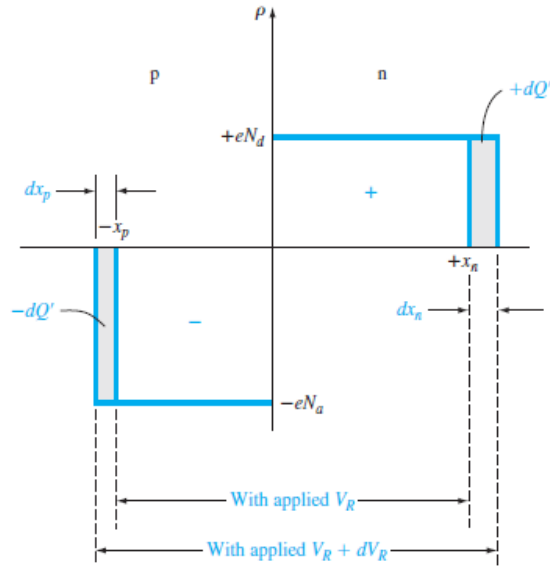
showing that the total space charge width increases as we apply a reverse-biased voltage. By substituting the total potential barrier  $V_{\text{total}}$  into Equations (7.28) and (7.29), the space charge widths in the n and p regions, respectively, can be found as a function of applied reverse-biased voltage.

## 2.28 Junction Capacitance:

Since there is a separation of positive and negative charges in the depletion region, a capacitance is associated with the pn junction. Figure 12 shows the charge densities in the depletion region for applied reverse-biased voltages of  $V_R$  and  $V_R + dV_R$ . An increase in the reverse-biased voltage  $dV_R$  will uncover additional positive charges in the n region

and additional negative charges in the p region. The junction capacitance is defined as

$$C = \frac{dQ'}{dV_R}, \text{ Where, } dQ' = eN_d dx_n = eN_a dx_p.$$



**Fig. 12: Differential change in the space charge width with a differential change in reverse-biased voltage for a uniformly doped pn-junction.**

The differential charge  $dQ'$  is in units of  $C/cm^2$  so that the capacitance  $C$  is in units of farads per square centimeter  $F/cm^2$ , or capacitance per unit area.

For the total potential barrier, Equation (18) may be written as,

$$x_n = \left[ \frac{2\epsilon_s(V_B + V_R)}{e} \left( \frac{N_a}{N_d} \right) \left( \frac{1}{N_a + N_d} \right) \right]^{1/2}$$

The junction capacitance can be written as,

$$C = \frac{dQ'}{dV_R} = eN_d \frac{dx_n}{dV_R}$$

$$C = \left[ \frac{e \epsilon_s N_a N_d}{2(V_B + V_R)(N_a + N_d)} \right]^{1/2}$$

Exactly the same capacitance expression is obtained by considering the space charge region extending into the p region  $x_p$ . The junction capacitance is also referred to as the depletion layer capacitance.

## 2.29 Metal-metal junction:

Formation of electronic devices requires putting together two or more dissimilar materials (semiconductors, metals, insulators). The interface between these materials becomes crucial because it affects the electrical properties (transport) of the devices. This interface is called the junction. An ideal junction is one where there are no defects formed at the interface. Forming ideal junctions is challenging and most real materials have defects at the interface which can affect the electronic properties.

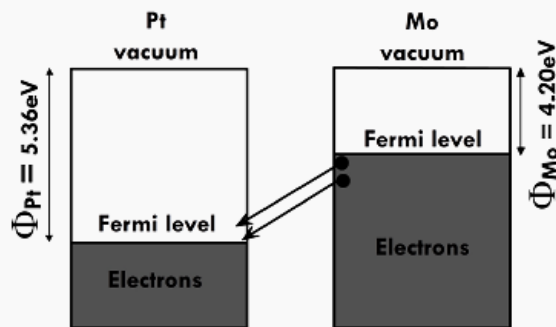
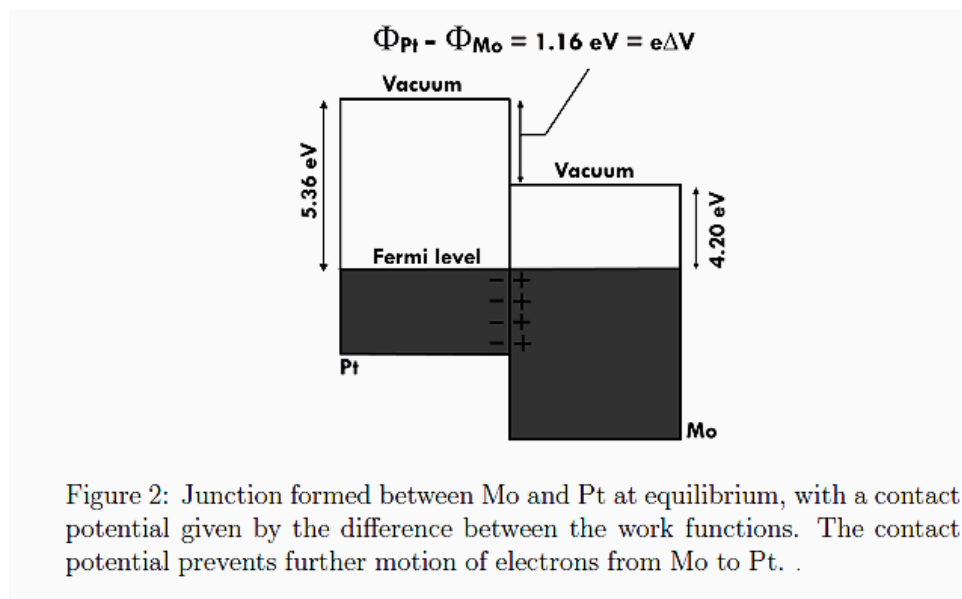


Figure 1: Junction between Mo and Pt. Mo has a smaller work function than Pt. When a junction is formed electrons from Mo move to Pt until the Fermi level lines up and the junction is in equilibrium.

Consider a junction formed between 2 metals with different work functions, as shown in figure 1. Metals are characterized by an incomplete energy band (valence and conduction band overlap) with the top of the energy band being the Fermi energy. The distance from the Fermi level to the vacuum level is called the work function ( $\Phi$ ). In figure 1 the two metals are

Pt and Mo with work functions of 5.36 and 4.20 eV respectively. Actual junctions are formed usually by vapour deposition of one metal on top of the other. Usually some form of post deposition annealing is also used to form the interface and reduce defect density. The most important rule when a junction is formed is that the Fermi levels must line up at equilibrium (no external bias). This can be understood by using figure 1, where there are electrons in Mo at a higher energy level than those in Pt. These electrons can occupy the empty energy levels located above the Pt Fermi level. Because electrons move from Mo to Pt a net positive charge develops on the Mo side and a net negative charge on the Pt side. Thus, a contact potential is developed at equilibrium between the two metals. This contact potential is related to the difference in the work functions, and is shown in figure 2. For Mo-Pt junction the contact potential is 1.16 V.



## 2.30 Metal-semiconductor junction:

A **Schottky contact** refers to a metal-semiconductor contact having a large barrier height (i.e.,  $q\phi_{BN}$  or  $q\phi_{BP} \gg kT$ ) and a low doping concentration that is less than the density of states in the conduction band or valence band. The current transport in a Schottky barrier is

due mainly to majority carrier, in contrast to a  $p$ - $n$  junction, where current transport is due mainly to minority carriers. For Schottky diodes operated at moderate temperature (e.g., 300 K), the dominate transport mechanism is thermionic emission of majority carriers from the semiconductor over the potential barrier into the metal. For a Schottky contact the metal work function should always be greater than the semiconductor work function  $\phi_m > \phi_{\text{semi}}$ .

**An ohmic contact** is defined as a metal-semiconductor contact that has a negligible contact resistance relative to the bulk or series resistance of the semiconductor. A satisfactory ohmic contact should not significantly degrade device performance and can pass the required current with a voltage drop that is small compared with the drop across the active region of the device. For a ohmic contact the metal work function should always be less than the semiconductor work function  $\phi_m < \phi_{\text{semi}}$

## 2.31 Schottky junction:

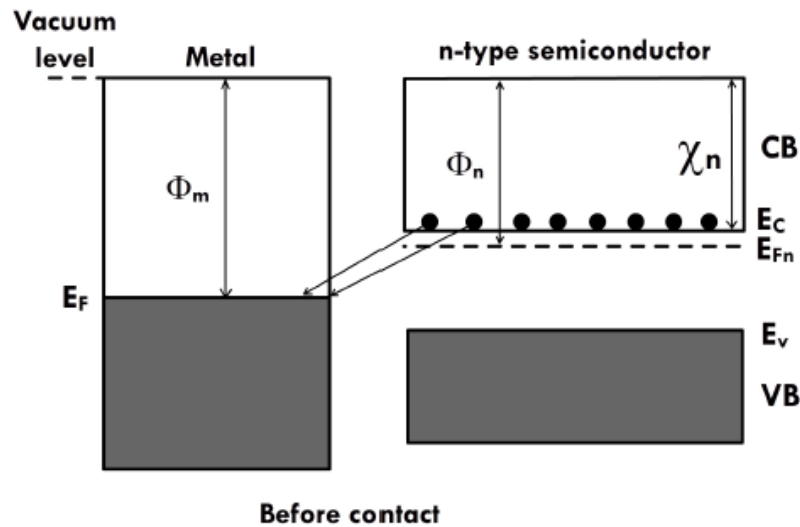
A **Schottky contact** refers to a metal-semiconductor contact having a large barrier height (i.e.  $q\phi_{\text{BN}}$  or  $q\phi_{\text{BP}} \gg kT$ ) and a low doping concentration that is less than the density of states in the conduction band or valence band. The current transport in a Schottky barrier is due to mainly because of majority carrier, in contrast to a  $p$ - $n$  junction, where current transport is due to mainly because of minority carriers. For Schottky diodes operated at moderate temperature (e.g., 300 K), the dominate transport mechanism is thermionic emission of majority carriers from the semiconductor over the potential barrier into the metal. For a Schottky contact the metal work function should always be greater than the semiconductor work function  $\phi_m > \phi_{\text{semi}}$ .

### Schottky contact in thermal equilibrium:

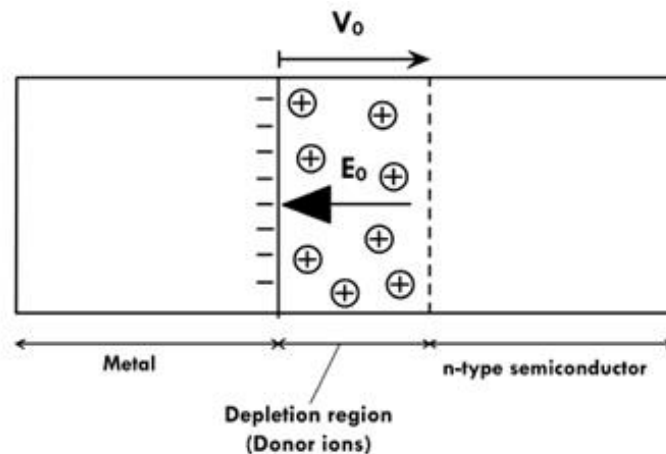
Consider a junction formed between a metal and  $n$ -type semiconductor, as shown in Figure 1. The Fermi level of the semiconductor is higher (since its work function is lower) than the metal. Similar to a metal-metal junction, when the metal semiconductor junction is formed the Fermi levels must line up at equilibrium. Another way to look at this is that there are electrons



in the conduction level of the semiconductor which can move to the empty energy states above the Fermi level of the metal. This leaves a positive charge on the semiconductor side and due to the excess electrons, a negative charge on the metal side, shown in figure 2, leading to a contact potential.



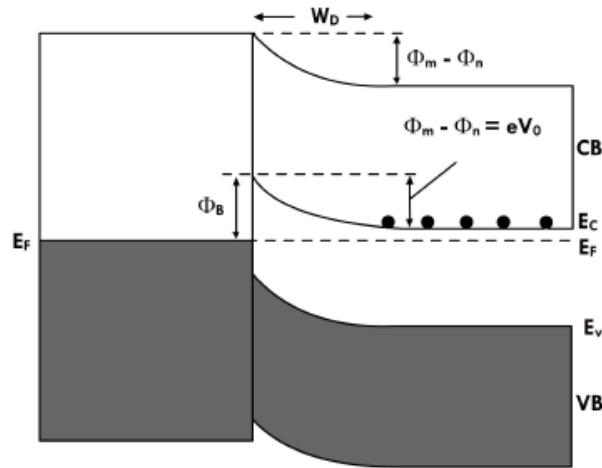
**Figure 1:** Schottky junction between metal and n-type semiconductor before contact. The work function of the semiconductor is smaller than the metal so that electrons can move from semiconductor to metal, forming a contact potential.



**Figure 2:** Schematic showing the metal, n-type semiconductor, and the Schottky junction between them. There is a depletion layer in the n-type semiconductor due to transfer of electrons to the metal. This leads to the formation of a contact potential.

When a contact is formed between two metals, the charges reside on the surface. This is due to the high electron density found in metals (typically  $10^{22} \text{ cm}^{-3}$ ). On the other hand, when a contact is formed between a metal and semiconductor, due to the low charge density on the semiconductor side (typically  $10^{17} \text{ cm}^{-3}$ ) the electrons are removed not only from the surface but also from a certain depth within the semiconductor. This leads to the formation of a depletion region within the semiconductor as shown in the figure 2.

Thus, when a Schottky junction is formed between the metal and semiconductor, the Fermi level lines up and also a positive potential is formed on the semiconductor side. Because the depletion region extends within a certain depth in the semiconductor there is bending of the energy bands on the semiconductor side. Bands bend up in the direction of the electric field (field goes from positive charge to negative charge, opposite of the potential direction). This means the energy bands bend up going from n-type semiconductor to metal, shown in figure 3. The Fermi levels line up and there is a certain region in the semiconductor (denoted by  $W$ ) where the bands bend (this is the depletion region). Another name for the depletion region is the space charge layer.



**Figure 3:** Schottky junction showing the band bending on the semiconductor side. Semiconductor bands bend up going from the semiconductor (positive) to metal (negative) since this is the same direction as the electric field.

There is a built in potential in the Schottky junction ( $V_0$ ), and from figure 3, this is given by the difference in work functions.

$$eV_0 = \phi_m - \phi_{semi} \longrightarrow (1)$$

### **Schottky barrier:**

The work function of the metal is a constant while the semiconductor work function depends on the dopant concentration (since this affects the Fermi level position). The contact potential then represents the barrier for the electrons to move from the n-type semiconductor to the metal. Initially, when the junction is formed electrons move to the metal to create the depletion region. The contact potential thus formed prevents further motion of the electrons to the metal. There is also a barrier for electrons to move from metal to semiconductor. This is called the Schottky barrier and denoted by  $\phi_B$  in figure 3. This is given by

$$\phi_B = (\phi_m - \phi_n) + (E_c - E_{Fn}) = \phi_m - \chi_n \longrightarrow (2)$$

Where,  $\chi_n$  is the electron affinity of the n-type semiconductor defined as the energy difference between the conduction band edge and the vacuum level in the semiconductor.

### **Biased Schottky Junction**

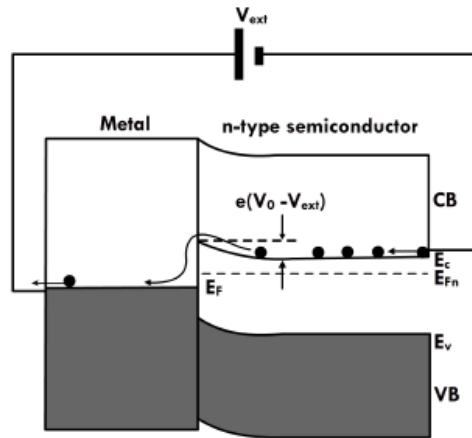
At equilibrium the motion of electrons from the semiconductor to metal is balanced by the contact potential so that there is no net current. The Schottky junction can be biased by application of an external potential. There are two types of bias:

- a) Forward bias - metal is connected to positive terminal and n-type semiconductor connected to negative terminal.
- b) Reverse bias - metal is connected to negative terminal and n-type semiconductor connected to positive terminal

The current now depends on the type of bias and the amount of applied external potential.

**a) Forward bias:**

In a forward biased Schottky junction the external potential is applied in such a way that it opposes the in-built potential. Since the region with the highest resistivity is the depletion region near the junction, the voltage drop is across the depletion region. Under external bias the Fermi levels no longer line up, but are shifted with respect to one another and the magnitude of the shift depends on the applied voltage. Energy band diagram of the Schottky junction under forward bias is shown in figure 4.



**Figure 4:** Schottky junction under forward bias.

Thus, electrons injected from the external circuit into the n-type semiconductor have a lower barrier to surmount before reaching the metal. This leads to a current in the circuit which increases with increasing external potential. The current in a Schottky diode under forward bias is given by

$$J = J_0 \left[ \exp\left(\frac{eV}{k_B T}\right) - 1 \right] \longrightarrow (3)$$

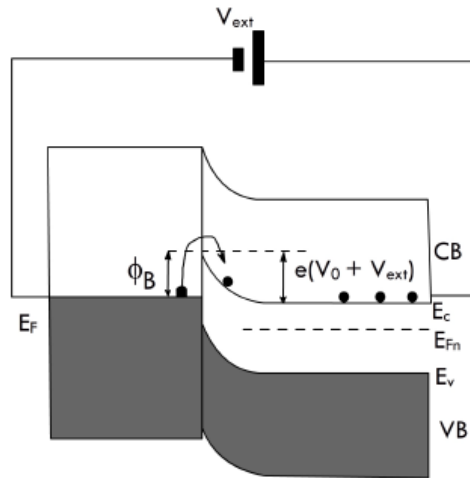
Here J is the current density for an applied potential of V.  $J_0$  is a constant and depends on the Schottky barrier ( $\phi_B$ ) for the system and the expression is,

$$J_0 = AT^2 \exp\left(-\frac{\phi_B}{k_B T}\right) \longrightarrow (4)$$

Where  $A$  is the Richardson constant for thermionic emission and is a material property. Equation (3) shows that in the forward bias the current exponentially increases with applied potential.

**b) Reverse bias:**

In the case of a reverse bias the external potential is applied in the same direction as the junction potential, as shown in figure 5.

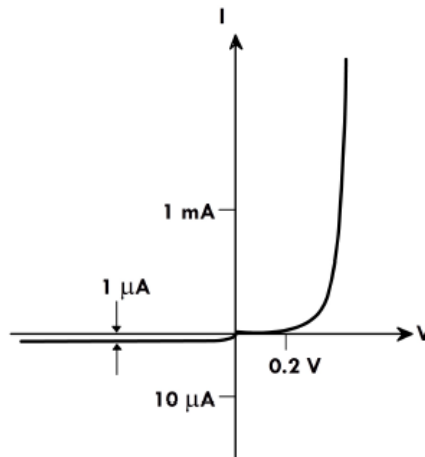


**Figure 5:** Schottky junction under reverse bias.

Once again the Fermi levels no longer line up but the barrier for electron motion from the n-type semiconductor to metal becomes higher. The electron flow is now from the metal to the semiconductor and the barrier for this is given by the Schottky barrier ( $\phi_B$ ). So there is a constant current in reverse bias, whose magnitude is equal to  $J_0$  (as given in equation (4)). From calculations it can be shown that the current in the forward bias is orders of magnitude higher than the current in reverse bias (this arises due to the exponential dependence on potential). So a Schottky junction acts as a rectifier i.e. it conducts in forward bias but not in reverse bias.

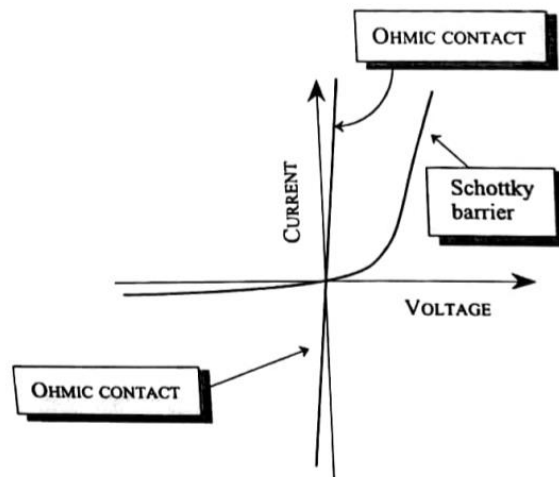
### I-V characteristics:

The I-V characteristics of the junction is shown in figure 6. There is an exponential increase in current in the forward bias (I quadrant) while there is a small current in reverse bias (IV quadrant).



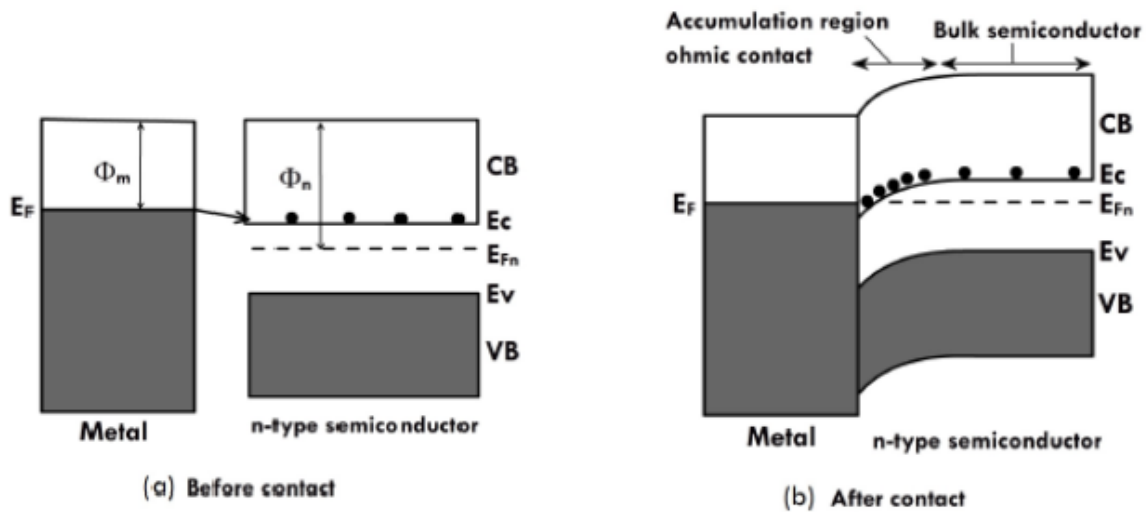
**Figure 6:** I-V characteristics of a Schottky junction showing rectifying properties.

### Comparison between schottky and ohmic contact I-V characteristics:



## 2.32 Ohmic junction:

When the semiconductor has a higher work function the junction formed is called the Ohmic junction. it is possible to draw the energy band diagram of the junction in equilibrium (Fermi levels line up). This is shown in figure 3.



**Figure 3. Ohmic contact demonstration**

At equilibrium, electrons move from the metal to the empty states in the conduction band so that there is an accumulation region near the interface (on the semiconductor side). The accumulation region has a higher conductivity than the bulk of the semiconductor due to this higher concentration of electrons. Thus, an Ohmic junction behaves as a resistor conducting in both forward and reverse bias. The resistivity is determined by the bulk resistivity of the semiconductor.

## 2.33 Differentiation between Schottky and PN-diode:

PN-junction diode		Schottky junction breakdown	
1	It is the junction between two different types of semiconductors. When a p-type semiconductor is brought into close contact with an n-type semiconductor, the assembly formed is called pn-junction.	1	<b>A Schottky contact</b> refers to a metal-semiconductor contact having a large barrier height (i.e., $q\phi_{BN}$ or $q\phi_{BP} \gg kT$ ) and a low doping concentration that is less than the density of states in the conduction band or valence band. For a Schottky contact the metal work function should always be greater than the semiconductor work function $\phi_m > \phi_{\text{semi}}$ .
2	Reverse current is due to minority carriers diffusing to the depletion layer and has strong temperature dependence.	2	Reverse current is due to majority carriers that overcome the barrier and has less temperature dependence.
3	Forward current is due to minority carrier's injection from n-region to p-region.	3	Forward current is due to the majority injection from the semiconductor to metal.
4	Forward bias needed to make the device conducting and therefore the cut-in voltage is large.	4	The cut-in voltage is quite small.
5	Switching speed controlled by recombination (or elimination) of minority injected carriers.	5	Switching speed controlled by thermalization of "hot" injected electrons across the barrier ~ few ps.