SEMICONDUCTOR DIODE

CLASSIFICATION OF METALS:

On the basis of the relative values of electrical conductivity (σ) or resistivity ($\rho = \frac{1}{\sigma}$) , the solids are broadly classified as:

(i) Metal (Conductors): These are those solids which have high conductivity and low resistivity.

$\rho \approx 10^{-2} to 10^{-8} \Omega m$

$\sigma \approx 10^{2} to 10^{8} \Omega^{-1} m^{-1}$

Examples: Al, Cu, Ag, etc.

(ii) *Insulators:* These are those solids which have very low conductivity and high resistivity.

$\rho \approx 10^{11} to 10^{19} \Omega m$

$\sigma \approx 10^{-11} to 10^{-19} \Omega^{-1} m^{-1}$

Examples: Rubber, Plastic, Putty etc.

(iii) *Semiconductors:* These are those solids which have conductivity and resistivity in between metal conductors and insulators.

$\rho \approx 10^{-5} to 10^{-6} \Omega m$

$\sigma \approx 10^{5} to 10^{-6} \Omega^{-1} m^{-1}$

Examples:

1. Elemental semiconductors: Silicon (Si) and Germanium (Ge), Selenium and Carbon.
2. Compound semiconductors:
   Inorganic semiconductors: CdS, GaAs, CdSe, InP etc.
   Organic semiconductors: anthracene, doped pthalocyanines, etc.
   Organic polymers semiconductors: polypyrole, polyaniline, polythiophene, etc.

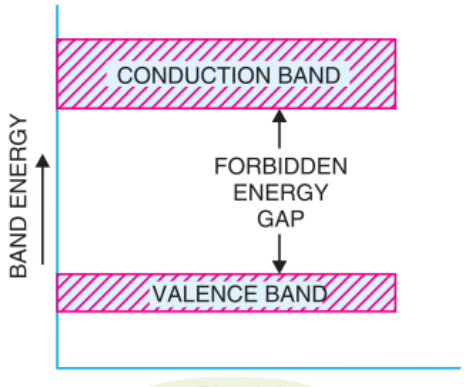
A subgroup of 3rd, 4th, 5th, 6th, and 7th columns of periodic table are shown in figure. The elements enclosed by the line are semiconductors.

III A	IV A	V A	VI A	VII A
B	C	N	O	F
Al	Si	P	S	Cl
Ga	Ge	As	Se	Br
In	Sn	Sb	Te	I
Tl	Pb	Bi	Po	At

Most of the currently available semiconductor devices are based on elemental semiconductors Si or Ge and compound inorganic semiconductors.

Energy Bands:

The range of energies possessed by an electron in a solid is known as **energy band.** The figure shows the energy level diagram of an atom.



Valence band:

The range of energies (i.e. band) possessed by valence electrons is known as **valence band.**

The electrons in the outermost orbit of an atom are known as valence electrons. In a normal atom, valence band has the electrons of highest energy. This band may be completely or partially filled. For instance, in case of inert gases, the valence band is full whereas for other materials, it is only partially filled. The partially filled band can accommodate more electrons.

Conduction band:

The range of energies (i.e. band) possessed by conduction band electrons is known as **conduction band.**

In certain materials (*e.g.* metals), the valence electrons are loosely attached to the nucleus. Even at ordinary temperature, some of the valence electrons may get detached to become free electrons. In fact, it is these free electrons which are responsible for the conduction of current in a conductor. For this reason, they are called *conduction electrons*.

All electrons in the conduction band are free electrons. If a substance has empty conduction band, it means current conduction is not possible in that substance. Generally, insulators have empty conduction band. On the other hand, it is partially filled for conductors.

Forbidden energy gap:

The separation between conduction band and valence band on the energy level diagram is known as **forbidden energy gap.** (The energy difference between the bottom of the Conduction and the top of the Valence bands is called the Band Gap.)

No electron of a solid can stay in a forbidden energy gap as there is no allowed energy state in this region. The greater the energy gap, more tightly the valence electrons are bound to the nucleus. In order to push an electron from valence band to the conduction band (i.e. to make the valence electron free), external energy equal to the forbidden energy gap must be supplied.

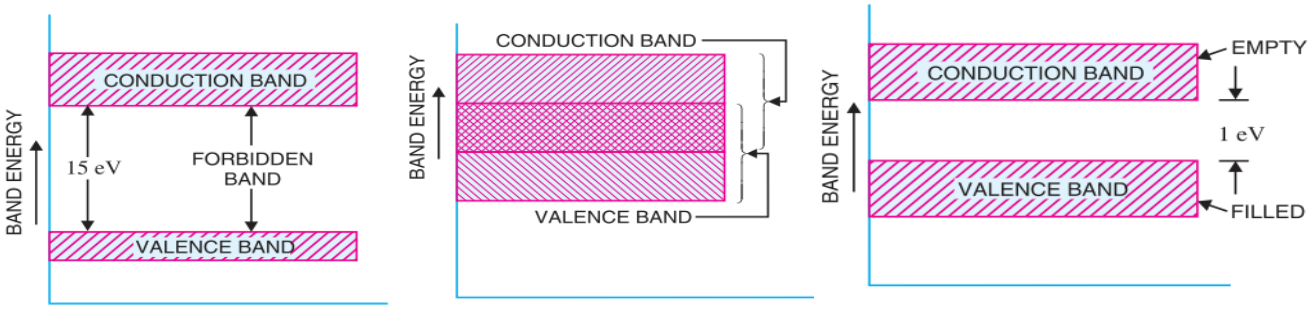
Fermi energy: Fermi energy is the maximum possible energy possessed by-free electrons of a material at absolute zero temperature (i.e.0^0K). The value of different materials is different.

Fermi level in semiconductor: It is that energy level in energy band diagram of semiconductor for which the probability of occupancy (i.e., the presence of main current carrier's electrons or -holes) becomes half.

Classification of Solids and Energy Bands:

Insulators:

Insulators (*e.g.* wood, glass etc.) are those substances which do not allow the passage of electric current through them. In terms of energy band, the valence band is full while the conduction band is empty. Further, the energy gap between valence and conduction bands is very large ($\approx$15 eV) as shown in figure. Therefore, a very high electric field is required to push the valence electrons to the conduction band.



Insulators **Conductors** **Semiconductors**

Conductors:

Conductors (*e.g.* copper, aluminium) are those substances which easily allow the passage of electric current through them. It is because there are a large number of free electrons available in a conductor. In terms of energy band, the valence and conduction bands overlap each other as shown in figure. Due to this overlapping, a slight potential difference across a conductor causes the free electrons to constitute electric current.

Semiconductors:

Semiconductors (*e.g.* germanium, silicon etc.) are those substances whose electrical conductivity lies in between conductors and insulators. In terms of energy band, the valence band is almost filled and conduction band is almost empty. Further, the energy gap between valence and conduction bands is very small as shown in figure. Therefore, comparatively smaller electric field (smaller than insulators but much greater than conductors) is required to push the electrons from the valence band to the conduction band. In short, a semiconductor has:

(a) Almost full valence band
(b) Almost empty conduction band
(c) Small energy gap ($\approx$1 eV) between valence and conduction bands.

At low temperature, the valence band is completely full and conduction band is completely empty. Therefore, a semiconductor virtually behaves as an insulator at low temperatures. However, even at room temperature, some electrons (about one electron for 10^{10} atoms) cross over to the conduction band, imparting little conductivity to the semiconductor. As the temperature is increased, more valence electrons cross over to the conduction band and the conductivity increases. This shows that electrical

conductivity of a semiconductor increases with the rise in temperature *i.e.* a semiconductor has negative temperature co-efficient of resistance.

Properties of Semiconductors:

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

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

(*iii*) When a suitable metallic impurity (*e.g.* arsenic, gallium etc.) is added to a semiconductor, its current conducting properties change appreciably. This property is most important and is discussed later in detail.

Commonly Used Semiconductors

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

Energy Band Description of Semiconductors

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

Fig. 1(a) and fig. 1(b) shows the energy band diagrams of germanium and silicon respectively. It may be seen that forbidden energy gap is very small; being 1.1 eV for silicon and 0.7 eV for germanium. Therefore, relatively small energy is needed by their valence electrons to cross over to the conduction band. Even at room temperature, some of the valence electrons may acquire sufficient energy to enter into the conduction band and thus become free electrons. However, at this temperature, the number of free electrons available is very small. Therefore, at room temperature, a piece of germanium or silicon is neither a good conductor nor an insulator. For this reason, such substances are called semiconductors.

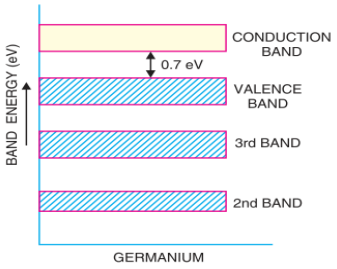


Fig.1 (a)

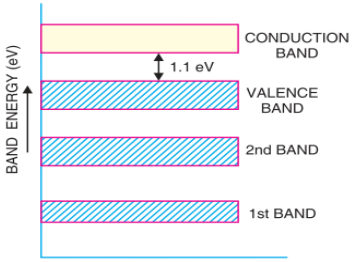


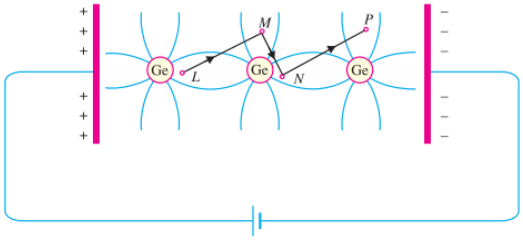
Fig.1 (b)

Hole Current:

At room temperature, some of the co-valent bonds in pure semiconductor break, setting up free electrons. Under the influence of electric field, these free electrons constitute electric current. At the same time, another current – the hole current – also flows in the semiconductor. *When a covalent bond is broken due to thermal energy, the removal of one electron leaves a vacancy i.e. a missing electron in the covalent bond. This missing electron is called a hole which acts as a positive charge.*

The current conduction by holes can be explained as follows:

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



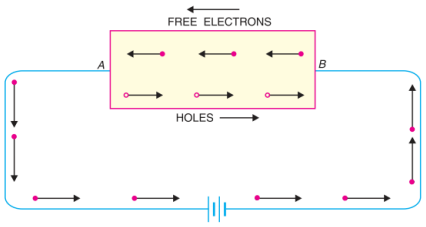
(It may be noted that hole current is due to the movement of valence electrons from one covalent bond to another bond. The reader may wonder why to call it a hole current when the conduction is again by electrons (of course *valence electrons*!). The answer is that the basic reason for current flow is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

Intrinsic Semiconductor

A semiconductor in an extremely pure form is known as an **intrinsic semiconductor**. E.g. Ge, Si, etc

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

$$I = I_e + I_h$$



Extrinsic Semiconductor

A doped semiconductor or a semiconductor with suitable impurity atom added to it is called doped or extrinsic semiconductor. The process of adding impurities to a semiconductor is known as *doping*.

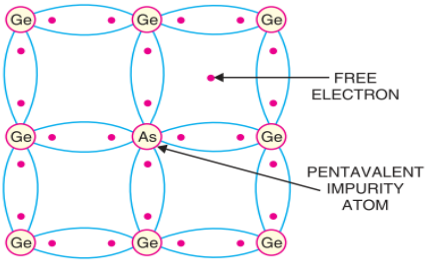
The purpose of adding impurity is to increase either the number of free electrons or holes in the semiconductor crystal. Depending upon the type of impurity added, extrinsic semiconductors are classified into:

(*i*) *n*-type semiconductor. **(*ii*)** *p*-type semiconductor.

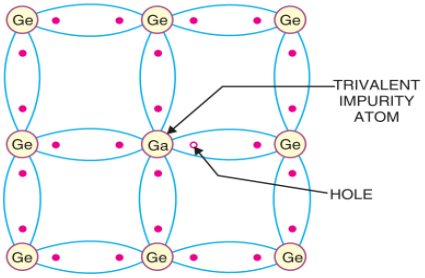
N-type Semiconductor

When a small amount of pentavalent impurity is added to a pure semiconductor, it is known as **n-type semiconductor.**

The addition of pentavalent impurity provides a large number of free electrons in the semiconductor crystal. Typical examples of pentavalent impurities are arsenic (At. No.33) and antimony (At. No.51). The impurities which produce n-type semiconductor are known as donor impurities because they donate or provide free electrons to the semiconductor crystal.



N-Type Semiconductor



P-Type Semiconductor

P-type Semiconductor

When a small amount of trivalent impurity is added to a pure semiconductor, it is called **p-type semiconductor.**

The addition of trivalent impurity provides a large number of holes in the semiconductor. Typical examples of trivalent impurities are gallium (At. No.31) and indium (At. No. 49). The impurities which produce p-type semiconductor are known as acceptor impurities because the holes created can accept the electrons.

Majority and Minority Charge Carriers:

In N-type semiconductor, electrons are majority carriers and holes are minority carrier's i.e. $n_e >> n_h$

In P-type semiconductor, holes are majority carriers and electrons are minority carrier's i.e. $n_h >> n_e$

Under thermal equilibrium, the product of the free negative and positive concentration is a constant quantity i.e.

$$n_e . n_h = n_i^2$$

This relation is known as mass-action law. Where n_e and n_h are the number density of electrons and holes respectively and n_i is number density of intrinsic carrier (i.e. electrons or holes) in a pure semiconductor.

Drift velocity and Mobility of semiconductor:

When an electric field is applied to a semiconductor, the electrons drifted towards the positive electrode and form the current. When an electron leaves the bond, an empty space is created. This empty space is called hole and constitutes a net positive charge. This hole moves towards the negative electrode which constitutes the current. *(The average velocity with which the free electrons get drifted towards the positive end of the conductor under the influence of the applied electric field is called the drift velocity of electrons.)*

Thus, when a potential difference is applied across a crystal, the charge carriers attain a drift velocity v_d , which is directly proportional to the applied electric field intensity E . Thus,

$$v_d \propto E \Rightarrow v_d = \mu E$$

Where, μ is called mobility of charge carriers. So, the mobility of charge carrier is the drift velocity of charge carrier produced by unit field strength i.e.

$$\mu = \frac{v_d}{E}$$

Mobility of an electron and hole are represented by μ_e and μ_h respectively i.e.

For electrons, $\mu_e = -\frac{v_{de}}{E}$ and for holes, $\mu_h = \frac{v_{dh}}{E}$

The SI unit of mobility is $\mu = \frac{v_d}{E} = \frac{meter/second}{volt/meter} = \frac{m^2}{volt.Second}$

The mobility of an electron or hole generally decreases with increasing temperature because of increased thermal vibrations of lattice (phonons) which scatter the moving electrons and holes.

Mobility of an electron:

Mobility of electrons is responsible for electric current and it is defined as the magnitude of drift velocity acquired per unit strength of the electric field applied across the conductor. It is denoted by μ .

If v_d is the drift velocity attained by free electrons on applying electric field 'E', then electron mobility is given by

$$\mu = \frac{|\vec{v_d}|}{|\vec{E}|}$$

$\because \quad |\vec{v_d}| = \frac{eE}{m}\tau$ $\qquad \left[\because \quad \vec{V_d} = \frac{-e\vec{E}}{m}\tau = \frac{-eV}{ml}\tau\right]$

$$\therefore \quad \mu = \frac{\frac{eE\tau}{m}}{E} = \frac{e\tau}{m}$$

The S.I. unit of electron mobility is $ms^{-1}N^{-1}C$ or $m^2s^{-1}volt^{-1}$.

Relation between drift velocity of electron and electric current:

Let us consider a conductor of length 'l' and uniform area of cross-section 'A'. If 'n' is the number of free electrons per unit volume of the conductor, then the total number of free electrons in the conductor is

$$N = nAl$$

If 'e' is the charge on each electron, then total charge on all the free electrons in the conductor

$$q = neAl$$

When a battery of 'V' volt is connected across the two ends of the conductor then the electric field setup across the conductor is given by,

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

Due to the electric field, if the electrons drifted towards the positive end of the conductor with drift velocity $\vec{v_d}$, then time taken by the free electrons to cross the conductor is

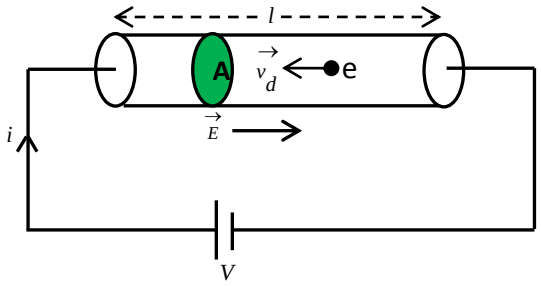
$$t = \frac{l}{v_d}$$

Hence, the electric current flowing through the conductor is

$$i = \frac{q}{t} = \frac{neAl}{\frac{l}{v_d}}$$

$$\Rightarrow i = neAv_d$$

Since, $n, e \text{ and } A$ are constant $\therefore i \propto v_d$



Current density:

Current density is defined as the current per unit cross sectional area, i.e., the charge crossing per unit area per second. It is represented by J and its unit is Amp/m^2 .

If I is the current flowing through cross sectional area A, then current density is given by

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

Conductivity of semiconductor:

Electrical conductivity of a semiconductor is different from that of a conductor. The charge carriers in a conductor are only electrons. However in a semiconductor, both electrons and holes are charge carriers.

Conductivity is defined as the current density per unit applied electric field. If J is the current density due to an applied electric field E, then the conductivity σ is given by,

$$\sigma = \frac{J}{E}$$

The SI unit of σ is

$$\sigma = \frac{J}{E} = \frac{amp/m^2}{volt/meter} = \frac{amp}{volt.meter} = \frac{1}{ohm.\ meter} = moh/meter = Siemens/meter\ .$$

Expression for electrical conductivity of a semiconductor:

Let us consider a semiconductor of area of cross-section A in which current I is flowing. Let V_d be the velocity electrons whose flow constitute the electric current. The electrons move through distance V_d in one second. As per assumption of free electron theory a large number of free electrons flow freely through the semiconductor whose area of cross section is A .

The volume swept by the electrons per second = AV_d

If N_e be the number of electrons per unit volume and e is the magnitude of electric charge on the electron, then the charge flow per second $= N_e e AV_d$

Since charge flow per second is the current I, therefore,

$$I = N_e e AV_d$$

Therefore, current density

$$J = \frac{I}{A} = \frac{N_e e AV_d}{A} = N_e e\ V_d \qquad \text{------ (1)}$$

The electron mobility μ_e is given by

$$\mu_e = \frac{V_d}{E} \Rightarrow V_d = \mu_e E \qquad \text{----- (2), where E is the electric field.}$$

Now, equation (1) becomes,

$$J = N_e e\, \mu_e E \qquad \text{------ (3)}$$

If σ_e is the conductivity due to electrons in the semiconductor materials then the ohm's law in terms of current density is given by the equation

$$J = \sigma_e\ E \qquad \text{----- (4)}$$

Comparing equation (3) and (4), we have

$$\sigma_e = N_e e\, \mu_e \qquad \text{----- (5)}$$

For an intrinsic semiconductor, the number of electrons will be equal to the number of holes, hence σ_h will be

$$\sigma_h = N_h e\, \mu_h$$

Thus the conductivity for an intrinsic semiconductor can be calculated as

$$\sigma = \sigma_e + \sigma_h = N_e e \mu_e + N_h e \mu_h$$

Since, $N_e = N_h = N_i$ therefore, conductivity for an intrinsic semiconductor is given by

$$\sigma = N_i e (\mu_e + \mu_h)$$

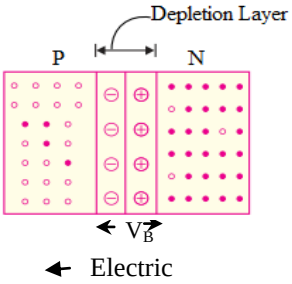
Where, N_i is the intrinsic carrier concentration.

P-N Junction

When a p-type semiconductor is suitably joined to n-type semiconductor, the contact surface is called **p-n junction.**

Properties of P-N Junction

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



Once *pn* junction is formed and depletion layer created, the diffusion of free electrons stops. In other words, the depletion region acts as a barrier to the further movement of free electrons across the junction. The positive and negative charges set up an electric field. This is shown by a black arrow in figure. The electric field is a barrier to the free electrons in the *n*-region. There exists a potential difference across the depletion layer and is called **barrier potential** (V_B). The typical barrier potential is approximately:

For silicon, V_B= 0.7 V; for germanium, V_B = 0.3 V

Barrier formation in PN junction diode:

In PN junction diode, one region is a p-type semiconductor and the adjacent region is a n-type semiconductor. The interface separating the n and p regions is called metallurgical junction (fig.1). Now, as we know in p-region, holes are in majority while electrons are in minority. Similarly, in n-region, electrons are in majority and holes are in minority. Due to concentration gradient, (i.e. large difference in number of electrons and holes on two sides of p-n junction) diffusion of charge carrier begins.

(i) Electrons start diffusing from n-region to p-region.

(ii) Holes start diffusing from p-region to n-region (fig.1)

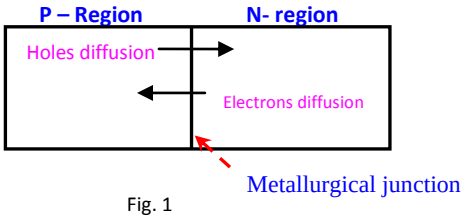


Fig. 1

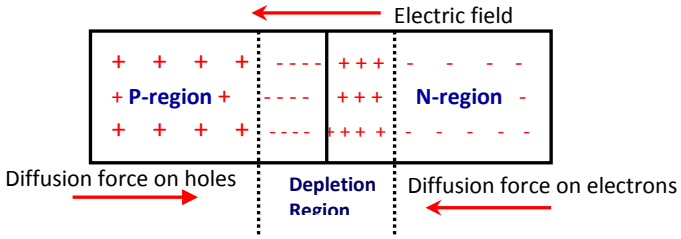


Fig. 2

As there is no external connection, diffusion of electrons and holes cannot continue indefinitely.

(i) As electrons diffuse from n-region to p-region, positively charged atoms are left behind.

(ii) As holes diffuse from p-region, negatively charge atoms are left behind.

These net negative and positive charges in p and n regions induce an electric field near the metallurgical junction, in the direction from positive to the negative charge and they produces a narrow region at the junction called depletion layer. The electric field produces another force i.e., drifts forces on electrons (from p-region to n-region) and drift forces on holes (from n-region to p-region). The above negatively and positively charges region opposes the diffusion of electrons and holes. So, it is name as barrier. As, this region is depleted of electron and hole it is also called depletion region.

PN junction diode fabrication (Simple Idea):

Most semiconductor devices contain at least one junction between P type and N type material. This P-N junction is the fundamental to the performance of function such as rectification and switching in electronic circuits. Here the discussion is about how the P-N junction diodes are made. The technology of junction fabrication is a broad subject which includes the accumulated knowledge and experience of many research and manufacturing groups forming junctions and making contacts to them in mountings suitable for device.

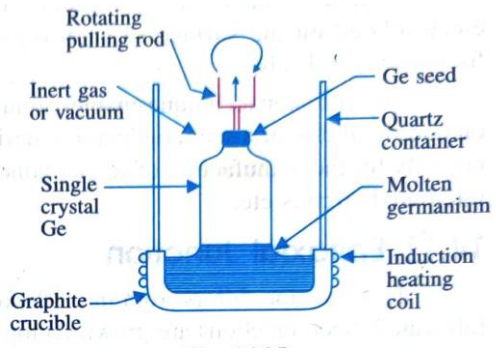
The first and foremost requirement is to obtain an extremely pure germanium or silicon. Impurity of less than one part in ten billion (10^{10}) is required for most semiconductor device fabrication to-day. For obtaining pure semiconductor material it is first purified chemically. For reducing the impurities further, and to ensure the formation of a mono crystalline structure, a technique known as *floating zone* is quite often employed. The mono crystalline structure is formed through the use of a small seed of semiconductor *(*e.g., silicon or germanium). The seed itself is a monocrystal that has been very carefully cut along the face of its cubic lattice called wafers.

There are basically 5(Five) methods for the fabrication on PN junction diode

(i) Grown junction method.
(ii) Alloy junction method.
(iii) Diffusion junction method.
(iv) Epitaxial growth junction method.
(v) Point Contact junction method.

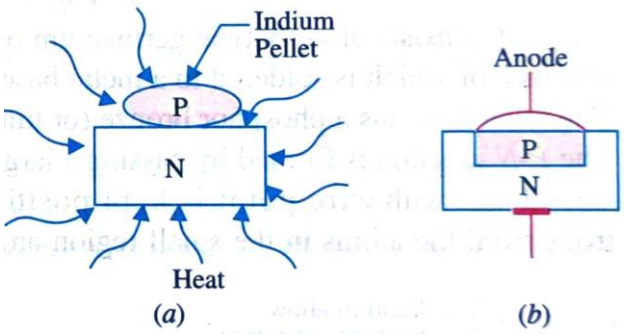
Among these techniques we have used only diffusion technology.

(i). Grown Junction method: In this technique a single crystal seed is immerse in the molten semiconductor material. The molten semiconductor material is contained in a graphite crucible. There is an induction heating coil to melt the Ge. All the material is kept in a quartz container in inert gas or vacuum atmosphere. Now the single crystal is slowly withdrawn using rotating pulling road. When the single crystal is pulled out impurities of P & N are alternately added to produce a PN junction. Then the large specimen is cut into a large number (say in thousands) of smaller-area semiconductor diodes. Though such diodes, because of larger area, are capable of handling large currents but larger area also introduces more capacitive effects, which are undesirable. Such diodes are used for low frequencies. Grown junction technique is an old technique.

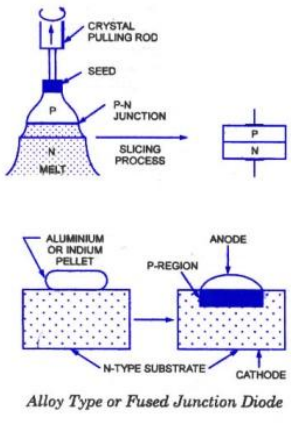


(ii). Alloy junction method: In this process a tiny dot or pellet of indium or aluminums or any P-type impurity is placed on the surface of an N-type silicon wafer. Then heat is given to the system ($\approx 150\ ^0C$) until liquefaction occurs where the two materials meet i.e. As a result of which indium melts and get dissolved in silicon. Then the temperature is lowered. So we get a PN junction.

The alloy junction method produces junction diodes that have (i) high pick inverse voltage (PIV), (ii) high current ratings. This is possible due to their large junction area.



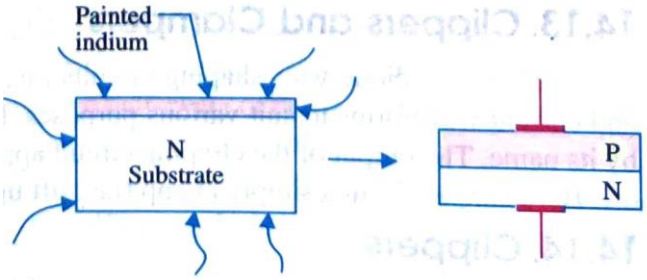
OR



Alloy Type or Fused Junction Diode

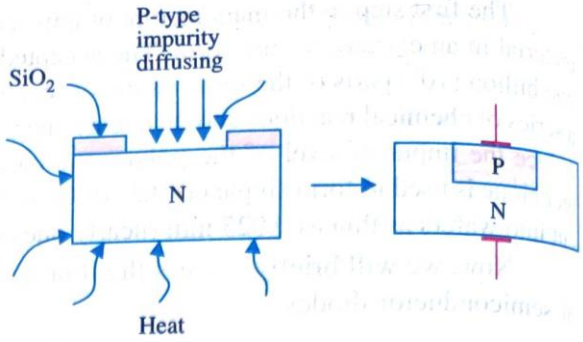
(iii). Diffused Junction method: In this process particle of an area of high concentration drift (move) to surrounding region of lesser concentration. There are two types of diffusion junction:

(i) **Solid diffusion technique:** In solid diffusion technique, p-type impurities say indium is painted on N-type substrate. Then it is heated from all sides until the impurity diffuses a short distance in to the substrate to form the PN junction.



(ii) **Gaseous diffusion technique:** In the gaseous diffusion process an N-type material is heated in a chamber containing a high concentration of P-type impurity in vapour form. Some of the P-type atoms are diffused or absorbed in to the N-type substrate to form the P-type layer thus creating a PN junction. The size of the P-region can be controlled by covering a thin coating of SiO_2. Metal contacts are electroplated on the surface of P & N region for connecting leads.

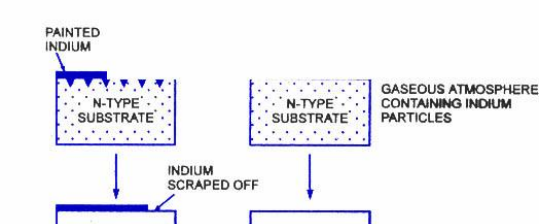
The diffusion process takes more times than alloy process but is relatively cheaper and more accurately controllable. The diffusion technique enables simultaneous fabrication of many hundreds of diodes on one small disc of semiconductor material. That is why it is the most frequently used technique not only for the manufacture of semiconductor diodes but also for production of transistors and integrated circuits.



OR

Diffused Junction method: Diffusion is a process by which a heavy concentration of particles diffuses into a surrounding region of lower concentration. The main difference between the diffusion and alloy process is the fact that liquefaction is not reached in the diffusion process. In the diffusion process heat is applied only to increase the activity of elements involved. For formation of such diodes, either solid or gaseous diffusion process can be employed. The process of solid diffusion starts with formation of layer of an acceptor impurity on an N- type substrate and heating the two until the impurity diffuses into the substrate to form the P-type layer, as illustrated in figure. A large P-N junction is divided into parts by cutting process. Metallic contacts are made for connecting anode and cathode leads.

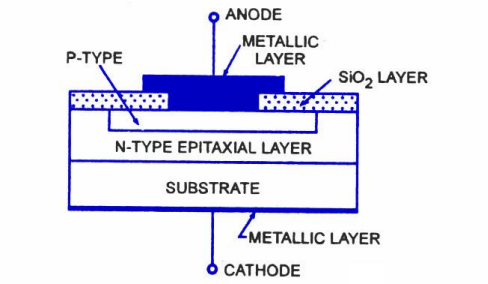
In the process of gaseous diffusion instead of layer formation of an acceptor impurity, an N- type substrate is placed in a gaseous atmosphere of acceptor impurities and then heated. The impurity diffuses into the substrate to form P- type layer on the N- type substrate. Though, the diffusion process requires more time than the alloy process but it is relatively inexpensive, and can be very accurately controlled. The diffusion technique leads itself to the simultaneous fabrication of many hundreds of diodes on one small disc of semiconductor material and is most commonly used in the manufacture of semiconductor diodes. This technique is also used in the production of transistors and ICs (integrated circuits).



(iv). Epitaxial Junction technique: Epitaxial junction is similar to diffusion junction. Here growth proceeds atom by atom and hence are exactly similar to the crystal lattice of the wafer on which it is grown. Such junctions have the advantage of low resistance.

OR

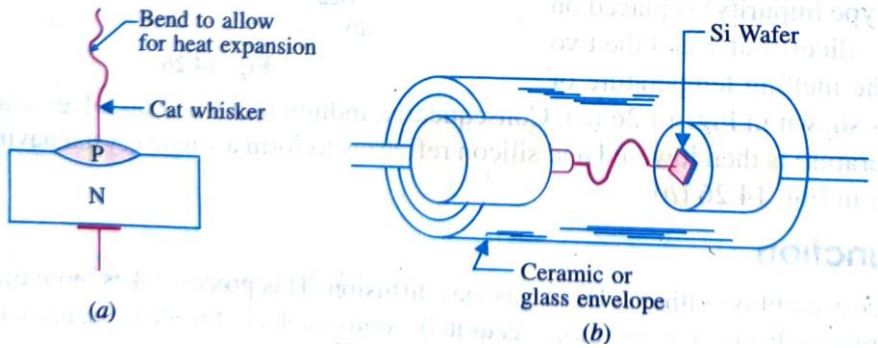
Epitaxial Growth or Planar Diffused Diode: The term "epitaxial" is derived from the Latin terms *epi* meaning 'upon' and *taxis* meaning "arrangement". To construct an epitaxially grown diode, a very thin (single crystal) high impurity layer of semiconductor material (silicon or germanium) is grown on a heavily doped substrate (base) of the same material. This complete structure then forms the N- region on which P- region is diffused. $Si0_2$ layer is thermally grown on the top surface, photo-etched and then aluminium contact is made to the P- region. A metallic layer at the bottom of the substrate forms the cathode to which lead is attached.This process is usually employed in the fabrication of IC chips.



Epitaxially Grown or Planar Diffused Diode

(v). Point Contact Junction technique:

In this technique an N-type Ge or Si wafer is soldered to a metal base by radio frequency heating. The other face of the wafer is pressed by a spring made of Phosphor bronze or tungsten using a large current for one or two seconds. The heat so produced drives away some of the electrons from the atoms at the point of contact leaving holes behind. This small region of N-type material is consequently converted into P-type material. In this way a PN junction is formed.



OR

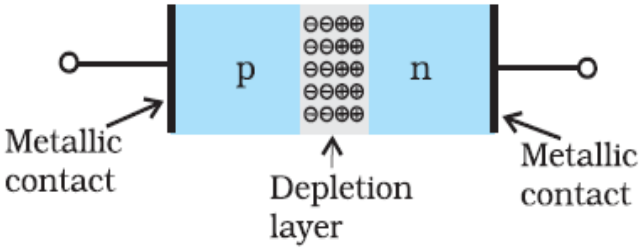
Point Contact Diode: It consists of an N-type germanium or silicon wafer about 12.5 mm square by 0.5 mm thick, one face of which is soldered to a metal base by radio-frequency heating and the other face has a phosphor bronze or tungsten spring pressed against it. A barrier layer is formed round the point contact by a pulsating current forming process. This causes a P-region to be formed round the wire and since pure germanium is N-type, a very small P-N junction in the shape of a hemisphere is formed round the point contact. The forming process cannot be controlled with precision. Because of small area of the junction, point contact

diode can be used to rectify only very small currents (of the order of m A). On the other hand, the shunting capacitance of point contact diodes are very valuable in equipment operating at super high frequencies (as high as 25,000 MHz).

SEMICONDUCTOR DIODE

A semiconductor diode is basically a p-n junction with metallic contacts provided at the ends for the application of an external voltage. It is a two terminal device.

A p-n junction diode is symbolically represented as shown in fig. (b). The direction of arrow indicates the conventional direction of current.



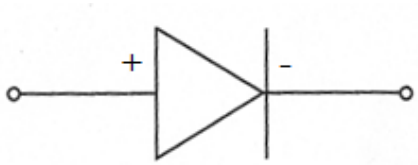


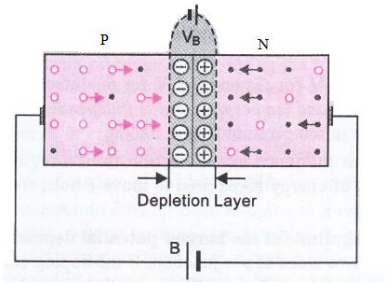
Fig. (b)

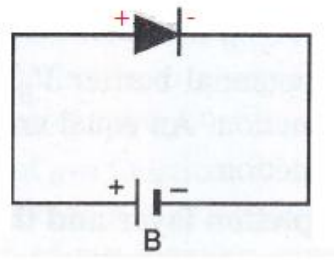
P-N Junction diode under forward bias:

Forward biasing:

When an external voltage *V* is applied across a semiconductor diode such that p-side is connected to the positive terminal of the battery and n-side to the negative terminal, it is said to be *forward biased*. (When external d.c. voltage applied to the junction is in such a direction that it cancels the potential barrier, thus permitting current flow, it is called **forward biasing.)**

The circuit for forward biasing of p-n junction is shown in figure below. In forward biasing, the applied forward potential establishes an electric field which acts against the field due to potential barrier. Therefore, the resultant field is weakened and the barrier height is reduced at the junction as shown in figure. As potential barrier voltage is very small (0.1 to 0.3 V), therefore, a small forward voltage is sufficient to completely eliminate the barrier. Once the potential barrier is eliminated by the forward voltage, junction resistance becomes almost zero and a low resistance path is established for the entire circuit. Therefore, current flows in the circuit. This is called *forward current*.

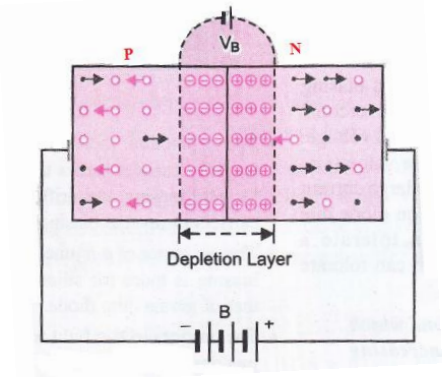


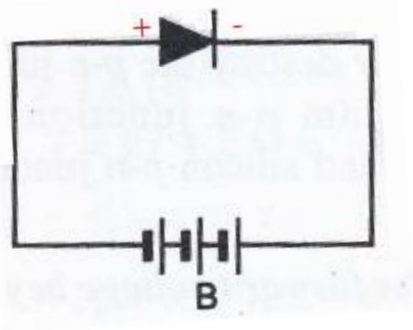


P-N junction diode under reverse bias

When an external voltage (*V*) is applied across the diode such that n-side is positive and p-side is negative, it is said to be *reverse biased. (When the external d.c. voltage applied to the junction is in such a direction that potential barrier is increased, it is called* **reverse biasing.)**

The circuit for reverse biasing of p-n junction is shown in figure below. In reverse biasing, the applied reverse voltage establishes an electric field which acts in the same direction as the field due to potential barrier. Therefore, the resultant field at the junction is strengthened and the barrier height is increased as shown in figure. The increased potential barrier prevents the flow of charge carriers across the junction. Thus, a high resistance path is established for the entire circuit and hence the current does not flow.





Volt-Ampere Characteristics of p-n Junction diode:

Volt-ampere or *V-I* characteristic of a *pn* junction diode is the curve between voltage across the junction and the circuit current. Usually, voltage is taken along *x*-axis and current along *y*-axis. The circuit arrangement for determining the *V-I* characteristics of a *pn* junction are shown in fig.1(a) and fig.1(b). The characteristics can be studied under three heads, namely; zero external voltage, forward bias and reverse bias.

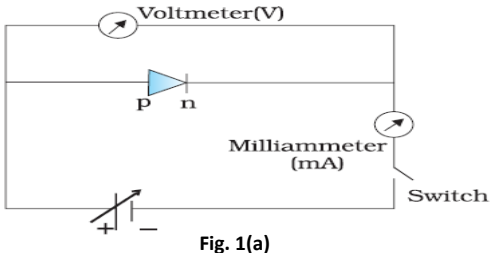


Fig. 1(a)

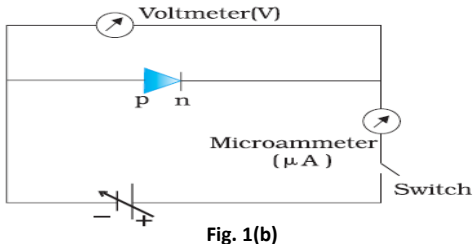
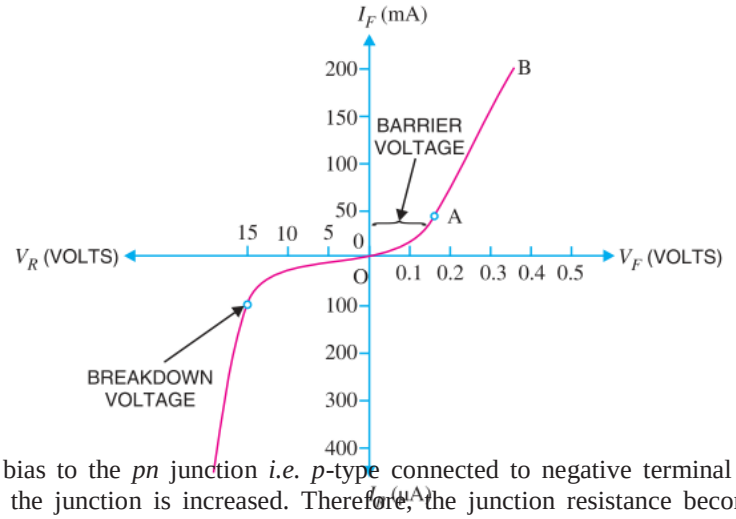


Fig. 1(b)

Zero external voltage: When the external voltage is zero, *i.e.* circuit is open at switch; the potential barrier at the junction does not permit current flow. Therefore, the circuit current is zero as indicated by point *O* in fig.1(c).

Forward bias: With forward bias to the *pn* junction *i.e. p*-type connected to positive terminal and *n*-type connected to negative terminal, the potential barrier is reduced. At some forward voltage, the potential barrier is altogether eliminated and current starts flowing in the circuit. From now onwards, the current increases with the increase in forward voltage. Thus, a rising curve *OB* is obtained with forward bias as shown in fig. 1(c). From the forward characteristic, it is seen that at first (*region OA*), the current increases very slowly and the curve is non-linear. It is because the external applied voltage is used up in overcoming the potential barrier. However, once the external voltage exceeds the potential barrier voltage, the *pn* junction behaves like an ordinary conductor. Therefore, the current rises very sharply with increase in external voltage (*region AB on the curve*). The curve is almost linear.



Reverse bias. With reverse bias to the *pn* junction *i.e. p*-type connected to negative terminal and *n*-type connected to positive terminal, potential barrier at the junction is increased. Therefore, the junction resistance becomes very high and practically no current flows through the circuit. However, in practice, a very small current (of the order of μ *A*) flows in the circuit with reverse bias as shown in the reverse characteristic. This is called reverse saturation current (Is) or leakage current and is due to the minority carriers. It may be recalled that there are a few free electrons in *p*-type material and a few holes in *n*-type material. These undesirable free electrons in *p*-type and holes in *n*-type are called *minority carriers*.

If reverse voltage is increased continuously, the kinetic energy of electrons (minority carriers) may become high enough to knock out electrons from the semiconductor atoms. At this stage breakdown of the junction occurs, characterised by a sudden rise of reverse current and a sudden fall of the resistance of barrier region. This may destroy the junction permanently.

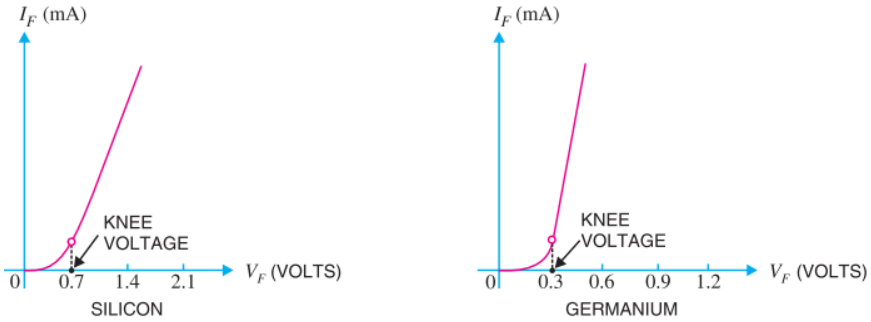
Note. The forward current through a *pn* junction is due to the *majority carriers* produced by the impurity. However, reverse current is due to the *minority carriers* produced due to breaking of some co-valent bonds at room temperature.

Important Terms:

Two important terms often used with pn junction diode are breakdown voltage and knee voltage.

(i) Breakdown voltage: It is the minimum reverse voltage at which pn junction breaks down with sudden rise in reverse current.

(ii) Knee voltage: It is the forward voltage (0.7 V for Si and 0.3 V for Ge) beyond which the current through the junction starts increasing rapidly with voltage, showing the linear variation. But below the knee voltage the variation is non-linear.



(iii) Peak Inverse Voltage (PIV): It is the maximum reverse voltage that can be applied to the diode without destruction. Other names and abbreviation used for it are: (a) Peak reverse voltage (PRV), (b) Reverse breakdown voltage(V_R or V_{RB}), (c) Maximum reverse voltage (V_{RM} or $V_{R.\max}$).

(iv) Avalance breakdown: If the reverse bias is made very high, the covalent bonds near the junction break down and a large number of electron – hole pairs are liberated. The reverse current then increases abrouply to a relatively large value. This is known as "Avalance breakdown" and may damage the junction by excessive heat generated unless the current is limitted by external circuit. This phenomenon is used in making Zener diode.

The maximum voltage that a junction diode can bear without breakdown is called zener voltage or reverse breakdown voltage.

(v) hh

Resistance of Diode:

It has already been discussed that a forward biased diode conducts easily whereas a reverse biased diode practically conducts no current. It means that *forward resistance* of a diode is quite small as compared with its *reverse resistance*. An ideal diode must offer zero resistance in forward bias and infinitely large resistance in reverse bias.

1. Forward resistance. The resistance offered by the diode to forward bias is known as *forward resistance*. This resistance is not the same for the flow of direct current as for the changing current.

Accordingly; this resistance is of two types, namely; *d.c. forward resistance or static resistance* and *a.c. forward resistance or dynamic resistance*.

(i) *Static or d.c. forward resistance.* It is the opposition offered by the diode to the direct current. It is measured by the ratio of d.c. applied voltage across the diode to the resulting d.c. current through it. Thus, referring to the forward characteristic in Fig. 6.5, it is clear that when forward voltage is *OA*, the forward current is *OB*.

$\therefore$ d.c. forward resistance, $R_f = \dfrac{OA}{OB}$

It is important to mention here that static resistance is not constant but depends on the operating point on V-I characteristic of diode.

(ii) *Dynamic or a.c. forward resistance.* It is the opposition offered by the diode to the changing forward current. It is measured by the ratio of change in applied voltage across diode to the resulting change in current through it *i.e.*

a.c. forward resistance, $R_f = \dfrac{\text{Change in voltage across diode}}{\text{Corresponding change in current through diode}}$

The a.c. forward resistance is more significant as the diodes are generally used with alternating voltages. The a.c. forward resistance can be determined from the forward characteristic as shown in Fig. 6.6. If *P* is the operating point at any instant, then forward voltage is *ob* and forward current is *oe*.

To find the a.c. forward resistance, vary the forward voltage on both sides of the operating point equally as shown in Fig.6.6, where *ab* = *bc*. It is clear from this figure that:

For forward voltage *oa*, circuit current is *od*.

For forward voltage *oc*, circuit current is *of*.

$\therefore$ a.c. forward resistance, $r_f = \dfrac{\text{Change in forward voltage}}{\text{Change in forward current}} = \dfrac{oc - oa}{of - od} = \dfrac{ac}{df}$

It may be mentioned here that forward resistance of a crystal diode is very small, ranging from 1 to 25 Ω.

Reverse Resistance: The resistance offered by the diode to the reverse bias is known as *reverse resistance*. It can be d.c. reverse resistance or a.c. reverse resistance depending upon whether the reverse bias is direct or changing voltage. Ideally, the reverse resistance of a diode is infinite. However, in practice, the reverse resistance is not infinite because for any value of reverse bias, there does exist a small leakage current. It may be emphasised here that reverse resistance is very large compared to the forward resistance. In germanium diodes, the ratio of reverse to forward resistance is 40000 : 1 while for silicon this ratio is 1000000 : 1

Derivation of expression for Height of Potential barrier or Junction Voltage and Width of Depletion Layer (i.e. Barrier width):

When P-N junction is formed, the diffusion of holes from P region to N region and electrons from N region to P region takes place for a short time and a force, called a potential barrier is automatically developed across the junction. The potential barrier stops the further diffusion of holes and electrons from one side to the other. The difference of potential from one side of the barrier to the other side is called the height of the barrier. Figure 33.9 shows the variation of potential with distance along the junction.

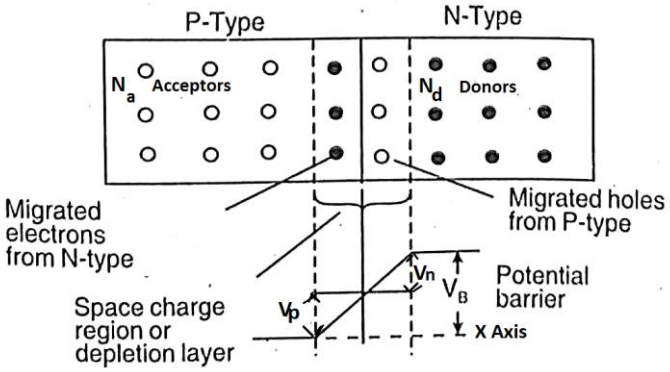
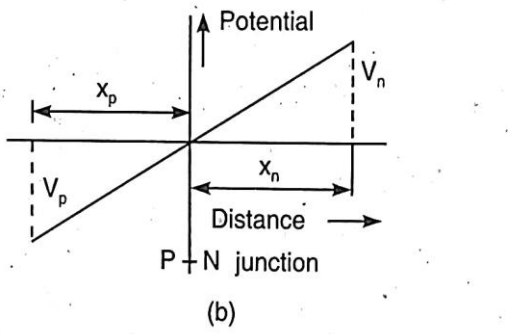


Fig. (a)



(b)

Let V_p be the magnitude of the potential fall in P region and V_n be the magnitude of the potential rise in N region then barrier potential V_B is given by

$$V_B = |V_p| + |V_n| \qquad \text{---------- (1)}$$

If x_p and x_n are the widths of depletion layer in P and N regions respectively then the width of the depletion region is given by

$$x = x_p + x_n \qquad \text{---------- (2)}$$

To find the distribution of barrier potential in the space charge region, we use Poisson's equation

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\varepsilon} \qquad \text{---------- (3)}$$

Here ρ is the volume density of charge and ε is the permittivity of the medium.

The charge density in P-side of depletion layer is given by

$$\rho = -eN_a$$

Here, N_a is the density of completely ionized acceptor atoms.

Substituting the value of ρ in equation (3), we get

$$\frac{d^2V}{dx^2} = \frac{eN_a}{\varepsilon}$$

Integrating it, we get

$$\frac{dV}{dx} = \frac{eN_a}{\varepsilon}x + C_1 \qquad \text{-------- (4)}$$

Here, C_1 is a constant of integration. Applying boundary condition,

At $x = -x_p$, $\frac{dV}{dx} = 0$, we get, $C_1 = \frac{eN_a}{\varepsilon}x_p$

Substituting the value of C_1 in equation (4), we get

$$\frac{dV}{dx} = \frac{eN_a}{\varepsilon}x + \frac{eN_a}{\varepsilon}x_p$$

$$\Rightarrow \frac{dV}{dx} = \frac{eN_a}{\varepsilon}\left(x + x_p\right) \qquad \text{--------- (5)}$$

Integrating equation (5), we get

$$V = \frac{eN_a}{\varepsilon}\left(\frac{x^2}{2} + x_p x\right) + C_2 \qquad \text{----- (6)}$$

Here, C_2 is another constant of integration. Its value is obtained by applying the condition that at $x = 0$, $V = 0$ and $C_2 = 0$.

$$\therefore \quad V = \frac{eN_a}{\varepsilon}\left(\frac{x^2}{2} + x_p x\right) \qquad \text{------ (7)}$$

For P region, at $x = -x_p$, $V = V_p$

$$\therefore \quad V_p = \frac{eN_a}{\varepsilon}\left(\frac{{x_p}^2}{2} - {x_p}^2\right) = -\frac{eN_a}{\varepsilon}\frac{{x_p}^2}{2}$$

$$\Rightarrow |V_p| = \frac{eN_a}{2\varepsilon}{x_p}^2 \qquad \text{---------- (8)}$$

Similarly, the Poisson's equation for N-side of depletion layer is

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\varepsilon} = \frac{eN_d}{\varepsilon}$$

Here, N_d is the density of completely ionized donor atoms.

Proceeding as above and applying boundary conditions:

at $x = 0$, $V = 0$,

and $x = x_n$, $\frac{dV}{dx} = 0$ and $V = V_n$

We get $V_n = -\frac{eN_d}{2\varepsilon}{x_n}^2$

$$\Rightarrow |V_n| = \frac{eN_d}{2\varepsilon}{x_n}^2 \qquad \text{--------- (9)}$$

Therefore, the height of potential barrier is given by

$$V_B = |V_p| + |V_n|$$

$$\Rightarrow V_B = \frac{eN_a}{2\varepsilon}{x_p}^2 + \frac{eN_d}{2\varepsilon}{x_n}^2$$

$$\therefore \quad V_B = \frac{e}{2\varepsilon}\left(N_a {x_p}^2 + N_d\ {x_n}^2\right) \qquad \text{----- (10)}$$

Since the crystal as a whole is electrically neutral, the number of charge carriers on both sides must be equal, i.e. ,

$$N_a x_p = N_d\ x_n$$

$$\Rightarrow x_n = \frac{N_a}{N_d}\, x_p$$

Substituting the value of x_n in equation (10), we get

$$V_B = \frac{e}{2\varepsilon}\left(N_a {x_p}^2 + N_d\ \frac{{N_a}^2}{{N_d}^2}\, {x_p}^2 \right)$$

$$\Rightarrow V_B = \frac{e}{2\varepsilon} N_a {x_p}^2 \left(1 + \frac{N_a}{N_d}\right) \qquad \text{--------- (11)}$$

$$\Rightarrow {x_p}^2 = \frac{2\varepsilon V_B}{eN_a\left(1+\frac{N_a}{N_d}\right)}$$

$$\Rightarrow x_p = \left[\frac{2\varepsilon V_B}{eN_a\left(1+\frac{N_a}{N_d}\right)}\right]^{1/2}$$

$$\Rightarrow x_p = \left[\frac{2\varepsilon V_B}{e}\,\frac{1}{N_a\left(\frac{N_d+N_a}{N_d}\right)}\right]^{1/2}$$

$$\Rightarrow x_p = \left[\frac{2\varepsilon V_B}{e}\,\frac{\frac{N_d}{N_a}}{N_a+N_d}\right]^{1/2} \qquad \text{----- (12)}$$

Similarly,

$$x_n = \left[\frac{2\varepsilon V_B}{e}\,\frac{\frac{N_a}{N_d}}{N_a+N_d}\right]^{1/2} \qquad \text{------ (13)}$$

Therefore, total width of depletion layer

$$x = x_p + x_n$$

$$x = \left\{\frac{2\varepsilon V_B}{e\left(N_a+N_d\right)}\right\}^{1/2}\left[\left(\frac{N_d}{N_a}\right)^{1/2} + \left(\frac{N_a}{N_d}\right)^{1/2}\right] \qquad \text{------ (14)}$$

If $N_d >> N_a$, then $\left(N_a + N_d\right) \approx N_d$ and $\frac{N_a}{N_d} \approx 0$

Equation (14) becomes

$$x = \left\{\frac{2\varepsilon V_B}{eN_d}\right\}^{1/2}\left(\frac{N_d}{N_a}\right)^{1/2}$$

$$\Rightarrow x = \left\{\frac{2\varepsilon V_B}{eN_a}\right\}^{1/2} \qquad \text{----- (15)}$$

Equation (15) gives the width of depletion layer. The width of depletion layer decreases with increase in impurity concentration.