Module 3
Semiconductors and Dielectrics

3.1 INTRODUCTION

Most of the devices are made of semiconductors such as silicon and germanium. The applications of semiconductors and their devices are ever increasing in all fields of science and technology. The knowledge on the classification, structures and properties of semiconductors is more essential to identify the right materials for applications.

3.2 ENERGY BAND DIAGRAM

An energy band diagram is a graphic representation of the energy levels associated with top energy band and the next lower energy band in a solid. The energy band diagram shows two bands with a gap in-between (Fig. 3.1). The upper band is called the **conduction band** and the lower energy band is called the **valence band**. These two bands are separated by a forbidden gap. This energy gap is more popularly called band gap and is denoted by the symbol E_g. The conduction band corresponds to the energy values of free electrons that have broken their valence bonds, and hence have become free to move in the crystal. The bottom of the conduction band represents the smallest energy that the electron must possess to become free. Only the free electrons can move in the crystal under the influence of the externally applied electric field. Hence, these electrons are called **conduction electrons** and the energies of such electrons constitute the conduction band. The band showing the energy values of **valence electrons** that are engaged in covalent bonding is called the **valence band**.

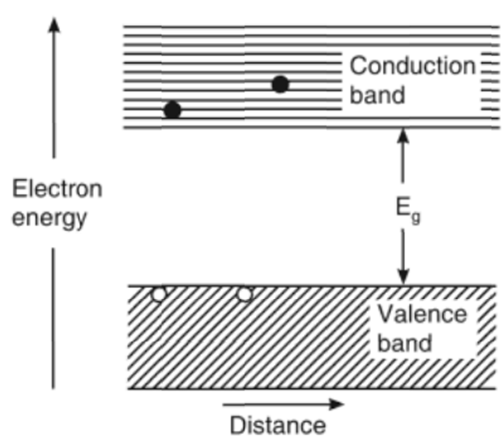


Fig 3.1 Energy band diagram

Classification of solids on the basis of band theory

Based on the energy band structure, the arrangement of electrons and forbidden bands, solid materials are classified into the following three categories:

(1) Conductors

(2) Insulators, and

(3) Semiconductors

Let us briefly discuss these materials briefly and elaborate more on semiconducting materials

1. Conductors

Materials which conduct electric current when a potential difference is applied across them are known as conductors. In case of a conductor, the valence band is completely filled, while the conduction band is half filled, as shown in Fig. 3.2.1.

Therefore, when a small potential difference is applied to a solid material, it provides sufficient energy to the electron in the valence band to shift to the conduction band. Thus, the shifting of electrons from the valence band to the unfilled conduction band results in the flow of current in the material. Examples for good conductors are copper, lithium, etc.

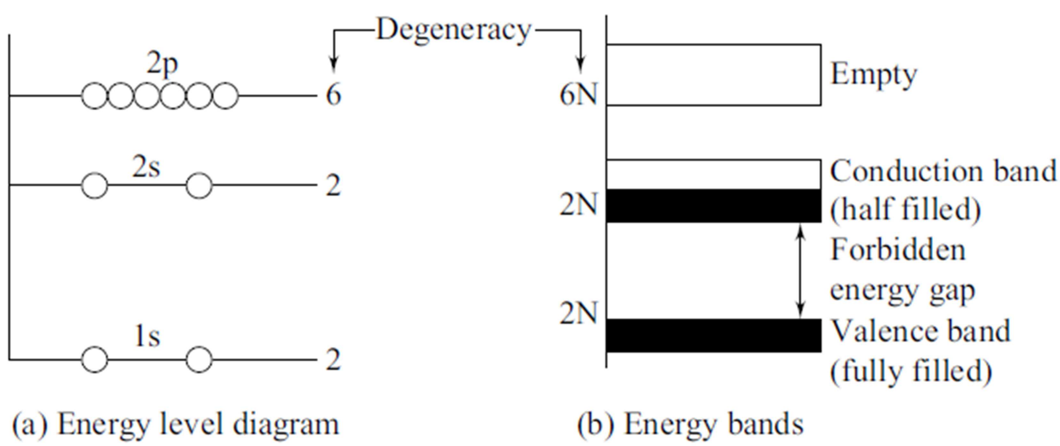


Fig. 3.2.1 Conductors–Lithium atom

2. Insulators

Solid materials which do not conduct electric current under normal conditions are known as insulators. In insulators, the valence band is completely filled and it has no electron in the conduction band. Further, the forbidden energy gap will be very high when compared with a conductor. The energy band diagram of an insulator (for example, ebonite) is shown in Fig. 3.2.2

Therefore, the energy required to shift an electron from the valence band, to the conduction band in order make electrical conduction possible, is very high. Hence, it is not possible to provide enough energy by an ordinary electric field. However, one can achieve electrical conduction in an insulator with very high voltage known as *breakdown voltage.*

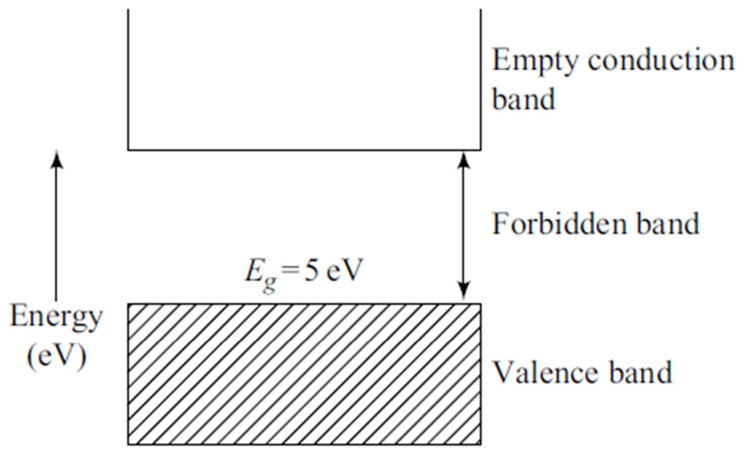


Fig 3.2.2 Energy bands in insulators

3. Semiconductors

Semiconductors (for example, silicon or germanium) are materials whose electrical conductivity lies between that of conductors and insulators. The conductivity of semiconductors is in the order of 10^4 to 10^{-4} mho m^{-1}. The magnitude of the forbidden energy gap of a semiconductor lies in between the forbidden energy gap of insulators and conductors, as shown in Fig. 3.2.3.

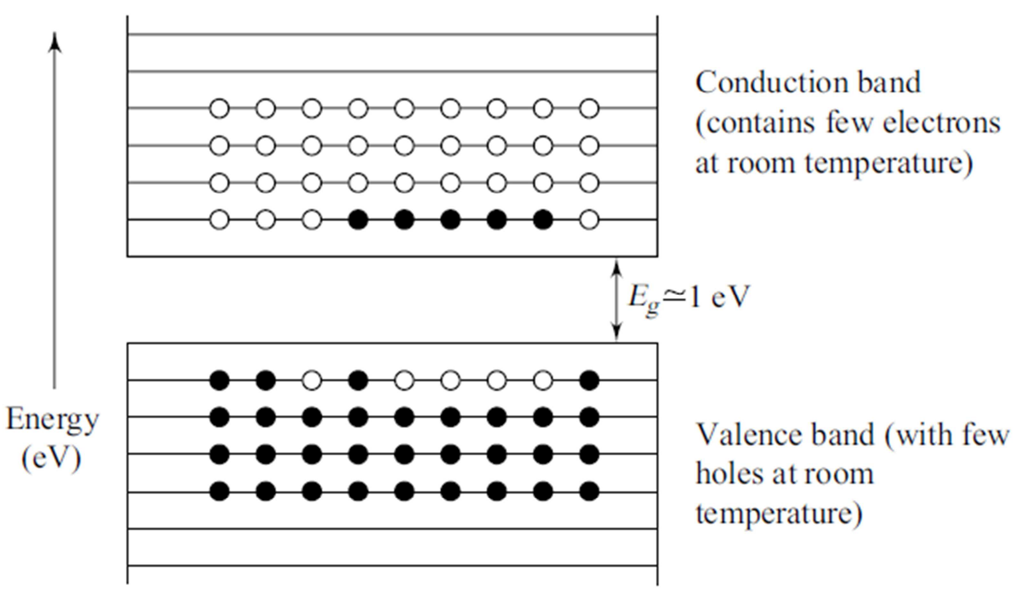


Fig. 3.2.3 Energy bands in semiconductors

Semiconducting materials, whether elemental, compound or oxide, are crystalline solids in nature. Well known semiconductors, such silicon and germanium, are elemental semiconductors, while gallium arsenide (GaAs), cadmium sulphide (CdS), etc., are known as compound semiconductors. Some of the oxide semiconductors are Bi_2O_3, Te_2O_3, ZnO_3, Cu_2O, etc.

Comparison Between Metal, Semi-Conductor and Insulator

Sr. No	*Conductors*	*Semiconductors*	*Insulators*
1.	The valance band is completely filled.	At room temperature, few electrons are filled.	The valence band is completely filled.
2.	The forbidden energy gap is zero.	The forbidden energy gap is ~1 or 2 eV which is very small when compared to an insulator.	The forbidden energy gap in the order of few MeV which is higher than semiconductor.
3.	Electrons are loosely bound to the nucleus.	Electrons are not tightly bound to the nucleus.	Electrons are tightly bound to the nucleus.
4.	It will conduct electricity at normal condition.	It will conduct electricity partially at normal condition.	It will not conduct the electricity at normal condition.
5.	The resistivity of conductor is very small and it is in the order of few mill ohm m.	The resistivity of semiconductor is very less and it is in the order of 0.5 to 10^3 ohm m.	The resistivity of an insulator is very high and it is in the order of 10^7 to 10^{12} ohm m.
6.	Example for conductor are copper lithium, gold, silver, etc.	Example for semiconductor are silicon, germanium, gallium Arsenide, cadmium sulphate, etc.	Example for insulator are ebonite, glass, rubber, glass fibre, porcelain, etc.

3.3 CLASSIFICATION OF SEMICONDUCTORS

Semiconductors are of two types and are classified on the basis of the concentration of electrons and holes in the materials:

(1) Pure or intrinsic semiconductors, and

(2) Doped or extrinsic semiconductors

Let us see the above two types of semiconductors in detail.

1 Pure or Intrinsic Semiconductors

Highly pure semiconductors are called intrinsic semiconductors, which means that the concentration of electrons must be equal to the concentration of holes.

In order to understand the electrical conducting property in a semiconducting material, let us first consider the arrangement of atoms in a material, say, silicon (or germanium). Both germanium and silicon have four valence electrons as shown in Fig. 3.3.1.

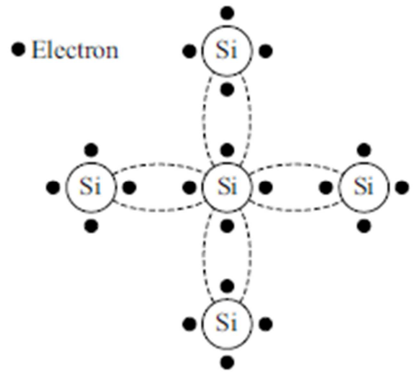


Fig 3.3.1 Two-dimensional arrangement of atoms in silicon at T = 0 K

Let us consider the crystal arrangement of silicon (or germanium). Silicon atom is represented by a circle, while the valence electron is marked by black dots. The central silicon atom is surrounded by four valence electrons constituting a covalent bond between two atoms, due to the sharing of valence electrons.

If we consider the crystal structure at 0 K, all the valence electrons are engaged in forming covalent bonds with the neighbours. Therefore, no free electron is available. At this stage, the material does not conduct current due to the lack of mobile charges, i.e., at 0 K, the material behaves as an insulator. Electrical conduction can be more lucidly explained by considering the energy level diagram of silicon at 0 K as shown in Fig. 3.3.2 At 0 K, the valence band is completely filled and there is no empty space in the valence band. Therefore, electrons can not shift from the valence band to the conduction band through the forbidden gap. The shifting electrons is not possible even for a large applied field strength at 0 K.

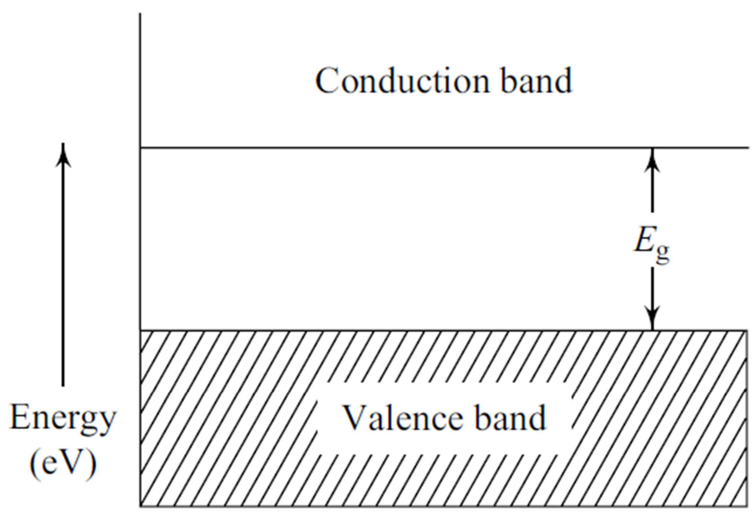


Fig 3.3.2 Energy band diagram of silicon at T = 0 K

However, the thermal energy is sufficient to liberate an electron from the valence band at room temperature. When the temperature of the intrinsic semiconductor is raised, the atoms in silicon vibrate in their mean position. This provides sufficient energy to the electrons due to which the backing of the covalent bonds take place as shown in Fig. 3.3.3. The broken electron is now said to be a free electron.

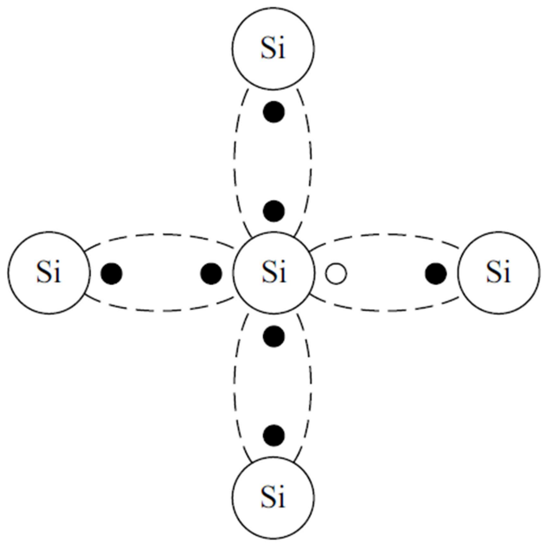


Fig. 3.3.3 Two dimensional arrangement of atoms in silicon at T > 0 K

When an electrical field is applied, the free electron acquires sufficient energy and shifts from the valence band to the conduction band. This results in the creation of a hole in the valence band as shown in Fig. 3.3.4. The free electrons move in the conduction band, while the holes move in the valence band. Thus, electrons and holes move opposite to each other as shown in Fig. 3.3.4.

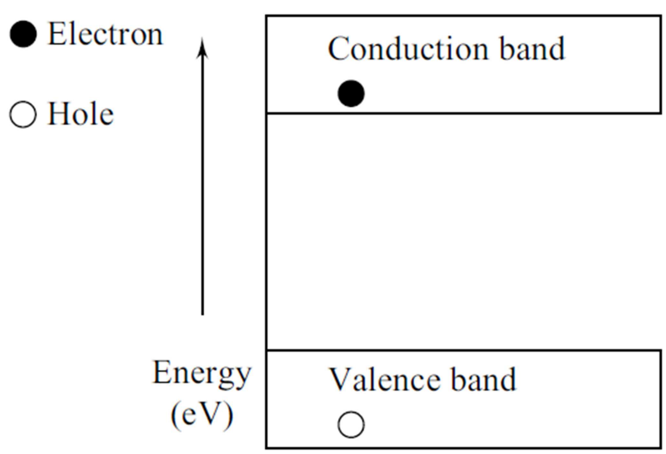


Fig. 3.3.4 Energy band diagram of silicon at T > 0 K

When a potential difference is applied across a silicon/germanium crystal, the electric force experienced by the electrons and holes are opposite. Thus, due to the opposite charges, the movement of electrons and holes give rise to an electric current in the same direction. The conductivity of germanium is higher than silicon due to its lower energy gap.

2. Extrinsic Semiconductors

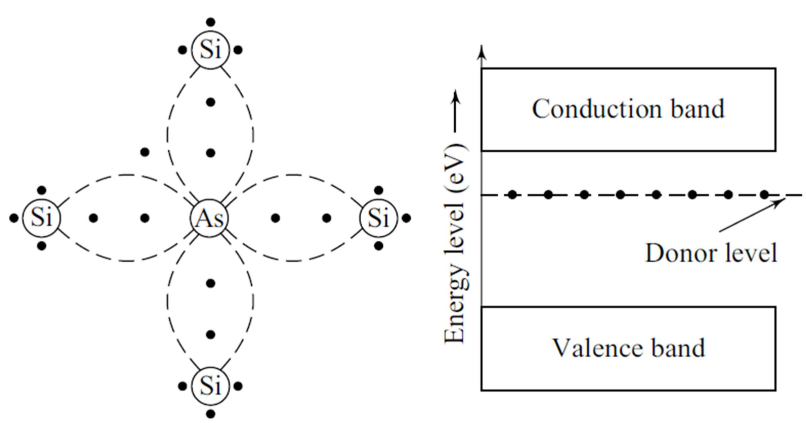
The application of intrinsic semiconductors is restricted due to its low conductivity. In electronic devices, high conducting semiconductors are more essential. The concentration of either electrons or holes in a semiconductor is essential. The concentration of either electrons or holes in a semiconductor is increased depending upon the requirements in the electronic devices.

This can be carried out simply by adding impurities (one atom in 107 host atoms) to the intrinsic semiconductors. The process of adding impurity to the intrinsic semiconductors is known as doping. The doped semiconductor is known as extrinsic semiconductor. The concentration of electrons and holes are not equal in an extrinsic semiconductor.

Extrinsic semiconductors are classified into two categories based on the concentration of the charge carriers namely,

(i) n-type semiconductors, and
(ii) p-type semiconductors

(i) n-type semiconductors When a pentavalent atom such as arsenic (antimony, bismuth, phosphorus) is added as a dopant to the tetravalent silicon atom, the arsenic atom will occupy one site of the silicon atom. Thus, out of five free electrons in arsenic, four electrons make covalent bonds with the four neighbouring silicons atoms and the fifth one is loosely bound to the silicon atom, as shown in Fig. 3.4



(a) With free electrons (b) Energy at donor level

Fig 3.4 Doping in n-type semiconductors

The energy required to ionise the fifth electrons is very less and hence, the thermal energy of the material shifts the free electrons to the conduction band. Each arsenic atom contributes one free electron to the crystal and hence, it is called a donor impurity. In this type of semiconductor, the concentration of charge carriers (i.e., electrons) is more than that of holes. Therefore, these semiconductors are called n-type semiconductors. In an n-type semiconductor, electrons are the majority current carriers while holes are the minority current carriers.

(ii) p-type semiconductors Instead of a pentavalent atom, the addition of a trivalent atom indium (In) to the tetravalent silicon atom, occupies the crystal site of the silicon atom as shown in Fig. 3.5.

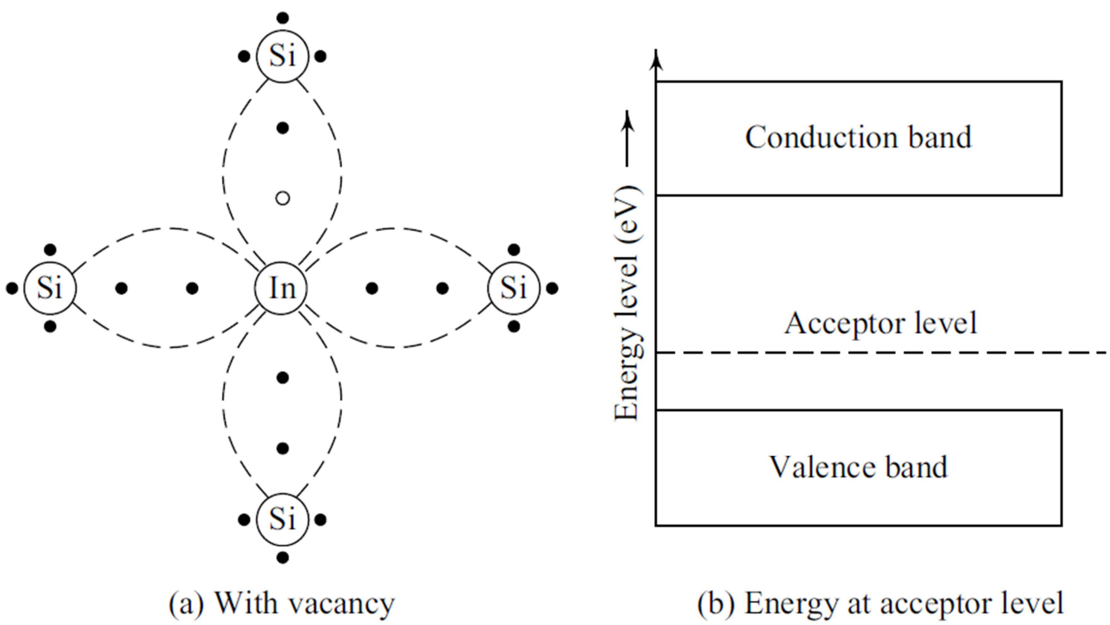


Fig. 3.5 Doping in p-type semiconductors

The three valence electrons in indium make covalent bonds with the three neighbouring silicon atoms, whereas the fourth bond has an empty space known as hole due to the deficiency of one electron. Therefore, when a trivalent atom is added to silicon, it creates a hole in the valence band. The dopant (indium) accepts an electron from the neighbouring silicon atom to form a covalent bond and hence, it is called an acceptor. The hole in the valence band moves freely and hence, the current flows through the material.

This type of electrical conduction will take place only when the dopant valency is less than that of the parent atom. Such semiconductors are called p-type semiconductors. In a p-type semiconductor, holes are the majority current carrier and electron are the minority current carriers.

3.4 Semiconductor materials

Semiconductors are those materials which have resistivity between a metal and an insulator. In semiconductors, the conduction band and the valence band is separated by an energy gap in the order of 1 eV. Some semiconductors such as Si and Ge are made up of only one type of atoms. These semiconductors are said to be elemental semiconductors. Some semiconductors like GaAs, InP, InSe, CdS, CdSe, etc., are made up of two different materials. They are binary semiconductors. A semiconductors made up of two or more than two different types of atoms is said to be a compound semiconductors. In a semiconductor, the electrical conductivity is due to the movement of the electrons and holes. The equation for the conductivity of a semiconductor is

$$\sigma = ne\mu_e + pe\mu_h$$

where n is the concentration of electron, p the concentration of hole and, μ_e and μ_h are the mobilities of electron and hole, respectively.

The energy gap of Si at room temperature is 1.1 eV and that of Ge is 0.72 eV. The properties of Si and Ge are listed in below Table 3.1.

Table 3.1 The advantages of Si over Ge

Property	Si	Ge
Energy gap	1.1 eV	0.66 eV
Upper temperature limit	150°C	100°C
Junction leakage current	less	more
Breakdown strength	higher	lower
Oxide quality	excellent	water soluble and unsuitable
Relative cost of electronic grade	1	10

GaAs is a semiconductor of group III and group V in the periodic table. It is one of the mostly studied semiconducting material. It has a direct band gap of 1.42 eV at room temperature. The gallium arsenide devices are nearly 2.5 times faster than Si devices and the noise of the GaAs devices is low. Since the cost of a GaAs device is nearly 10 times greater than that of a Si devices and the density of GaAs is high, it is not possible to materialise GaAs for device applications.

Table 3.2 lists the energy gap of some groups IV, group II–VI and groups III–V semiconductors. Consider the semiconductors with one element common such as ZnS, ZnSe and ZnTe. From Table 3.2, is found that the energy gap decreases with the increase in atomic weight. Some compound semiconductors, like GaP, InP, InAs and InSb, belonging to groups III–V of the periodic table are important. Generally, compound semiconductors are slightly ionic in character. This ionic character increases as we go from group III–V to II–VI elements.

Table 3.2 Energy Gap of Some Semiconductors

Semiconductor	E_g *(eV)*
Group IV	
Diamond	5.3
Si	1.11
Ge	0.72
Group II–VI	
ZnS	3.5
ZnSe	2.8
ZnTe	0.85
Group III–V	
InP	1.27
InAs	0.33
InSb	0.18

ZnO, ZnS, ZnSe, CdS, CdTe, CdSe and HgS are some of the group II–VI compound semiconductors, while PbS, PbSe and PbTe are some of the group II–VI compound semiconductors. Some ternary alloys such as $Al_xGa_{1-x}As$ or quaternary alloys such as $Al_xGa_{1-x}As_ySb_{1-y}$ are mostly used in device applications. GaAs, $GaAs_{1-x}P_x$ are mostly used for opto electronic devices such as LEDs. $Al_xGa_{1-x}As$ is also used in Modulation Doped Filled Effect (MODFET) transistor.

3.5 FABRICATION OF IC

Integrated Circuits (IC)

The integral circuit (or) IC IS a miniature, low cost electronic circuit consisting of active and passive components that are joined together in a single chip of silicon. It consists of components such as transistors, diodes, resistors and capacitors interconnected in a single chip. Thickness of an IC is 0.25 mm. It has so many advantages such as extremely small size, low power consumption, low cost, high processing speed, easy replacement, and small weight. IC can work as amplifier, timer, oscillator, counter, computer memory etc.

Production of Metallurgical Grade Silicon

The starting material is pure sand, which is available in plenty on earth's crust. Sand (SiO_2) is heated with carbon in an electric furnace to reduce it according to the reaction:

$$\mathbf{SiO_2 + 2C \longrightarrow Si + 2CO}$$

The silicon thus obtained is of 99% purity and is called the metallurgical grade silicon. This has to be purified further to a very low level of impurity content to make it suitable for use in devices.

Semiconductor Grade Silicon

The semiconductor grade silicon has only a few parts per billion of impurities. Starting from the metallurgical grade, it can be produced by the zone refining process. The other common method is a chemical process.

The metallurgical grade silicon is dissolved in HCl

$$\mathbf{Si + 3HCl \longrightarrow SiHCl_3 + H_2}$$

The product trichlorosilane ($SiHCl_3$) is a liquid at room temperature. The fractional distillation of this liquid removes chlorides of dopants and of other impurities such as iron and copper and also $SiCl_4$. A mixture of the purified $SiHCl_3$ and H_2 is then evaporated and passed through a reactor, which contains "slim rods" of high purity silicon. The gaseous mixture now undergoes the reverse reaction:

$$\mathbf{SiHCl_3 + H_2 \longrightarrow Si + 3HCl}$$

Solid silicon is deposited on the heated 'slim rods', which grow radially. This process can produce rods of semiconductor-grade, polycrystalline silicon up to 200 mm in diameter and 2–3 m long.

Basic Planer processes

The basic processes used to fabricate ICs are

I) Si-wafer preparation II) oxidation III) ion implantation IV) photolithography V) Isolation technique VI) Metallization VII) Assembly processing & packaging VII) Diffusion

Wafer production:

The first step is wafer production. The wafer is a round slice of semiconductor material such as silicon. It is the base or substrate for entire chip.

- Purification of polycrystalline silicon from the sand.
- Heating it to produce molten liquid.
- A small piece of solid silicon is dipped on the molten liquid and solid silicon is slowly pulled from the melt.
- The liquid is cooled down to form single crystal.
- A thin round wafer of silicon is cut using wafer slicer having thickness about 0.01-0.025".
- Damaged surface is smoothened by polishing.
- The wafers are cleaned using high purity low particle chemicals.
- The silicon wafers are now exposed to ultra pure oxygen.

Oxidation:

It consists of growing a thin film of SiO_2 on the surface of Si wafer. It servers two important purposes.

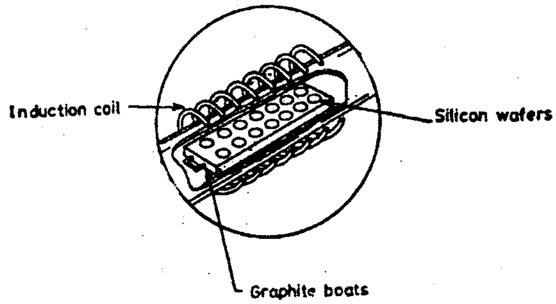
➔ SiO_2 is an extremely hard protective coating and is unaffected by almost all reagents except hydrofluoric acid.

➔ By selective etching of SiO_2 diffusion of impurities can be accomplished to fabricate various components.

Steps: 1) The silicon wafers are stacked up in a quartz boat and then inserted into quartz furnace tube.

2) The Si-wafers are raised to a Temperature in the range of 950 – 1150 ^{O}C and at the same time, exposed to a gas containing O_2 or H_2O or both. The chemical reaction is

$$\mathbf{Si + 2H_2O \longrightarrow SiO_2 + 2H_2}$$



A diagrammatic representation of a system for growing silicon epitaxial films

Oxidation provides surface passivation and isolates one device from another. This oxidation process is known as thermal oxidation because high temperature is used to grow the oxide layer. The thickness is usually in the order of 0.02 to 2μm.

Photolithography:

It is a technique used to produce microscopically small circuit and device patterns on silicon wafers.

Photolithographic process uses UV light exposure and device dimension or line width as small as 25 μm can be obtained.

It involves two processes.

Making two processes.

I) Making of a photographic mask.

II) Photo etching

i) Making of a photographic mask:

It involves following sequence of operations:

➔ Preparation of initial art work and

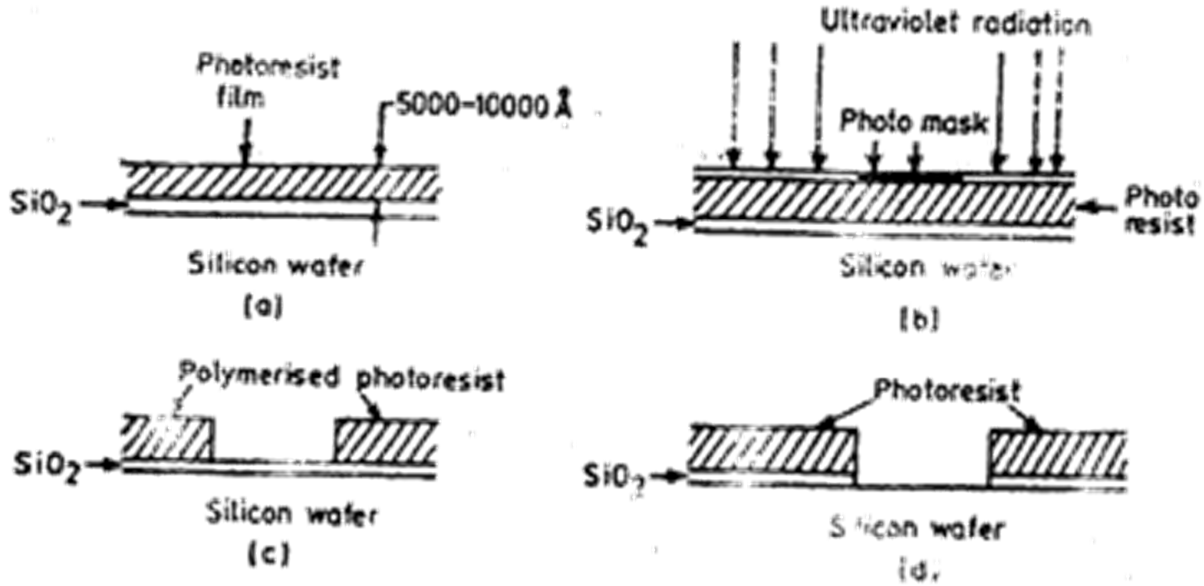
➔ Its reduction

Initial artwork is done at a scale hundred times larger than financial dimensions of finished circuit. This is because larger the artwork, more accurate is the final mask. For e.g. it is required to making an opening of width of 1 mil (25 μm). Clearly, this cannot be managed by any draftsman even with his thinnest of sketch pens. So the drawings are magnified and often by a factor of 500. So, now 1 mil is modified to width of 500 mils.

For photographic purpose, artwork should not contain any line drawings, but must be of alternate clear & opaque regions.

ii) Photo etching:

Used for removal of SiO_2 from desired regions so that desired impurities can be diffused.



Various steps for photo etching

Steps:

- Silicon wafer is coated with a film of photosensitive emulsion (Kodak photo resist KPR)
- The thickness of the film is in the range of 5000 – 10000A^0.
- The mask as prepared by described earlier is placed over the photo resist coated wafer, and exposed to UV light that KPR becomes polymerized beneath the transparent regions of the mask.
- The mask is then removed and wafer is developed using a chemical trichloroethylene, which dissolves unexposed regions on the photo resist and leaves the pattern as shown in figure 'c'
- Next, the photo resist is fixed (or) cured, so that it becomes immune to certain chemicals called etchants used in subsequent processing steps.

Diffusion:

Important process in fabrications of IC is diffusion of impurities in the silicon chip.

Steps:

- This uses a high temperature furnace having a flat temperature profile over a 20" length.
- A quartz boat containing about 20 cleaned wafers is pushed into the hot zone with temperature maintained at about 1000^0C.
- Impurities diffused are B_2O_3 (boron oxide), BCL_3 (boron chloride) for Boron and P_2O_5 (Phosphorous pentoxide) and $POCL_3$ (Phosphorous oxy chloride) for phosphorous.
- Next a carrier gas such as N_2 or O_2 is used for carrying impurities to the high temperature zone.
- The depth of diffusion depends upon the time of diffusion (which normally extends to 2 hrs.)
- The diffusion if an impurity normally takes place both laterally as well as vertically.

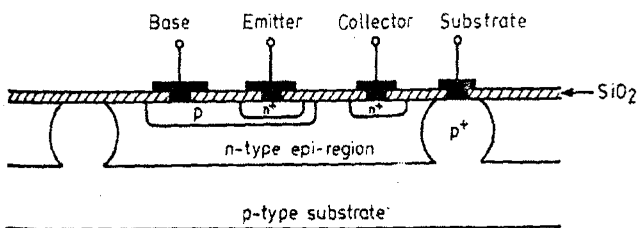


Fig The cross-section of an n-p-n transistor showing curved junction profiles as a result of lateral diffusion.

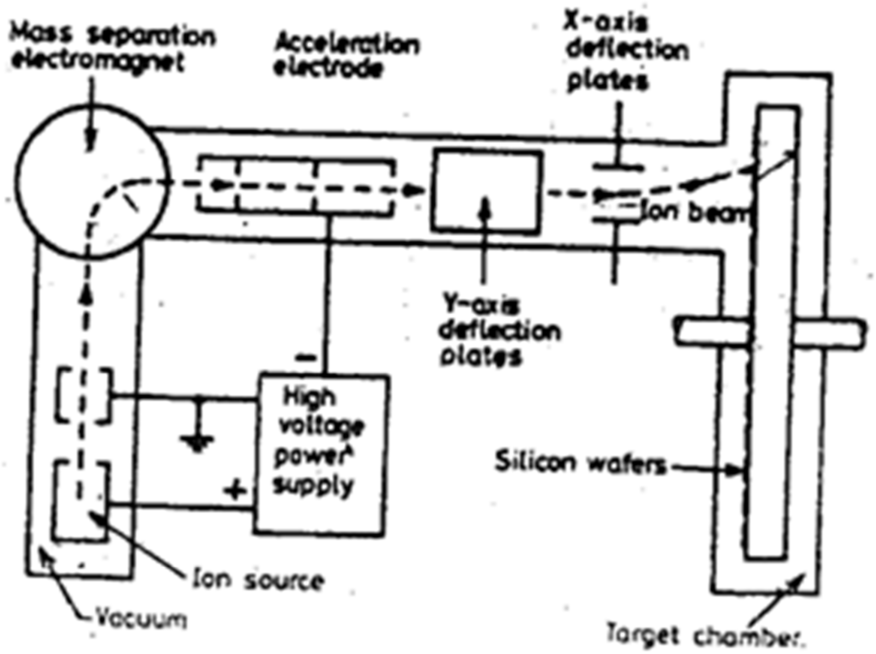
Ion implantation:

Ion implantation is a technique used to introduce impurities into a silicon wafer.

Steps:

- Silicon wafers are placed in a vacuum chamber and are scanned by a beam of high energy dopant ions (borons for p-type and phosphorous for n-type).
- Those ions are accelerated by energies between 20KV to 250 KV.
- As the ions strike the silicon wafers, they penetrate into the wafer.
- Depth of penetration increases with increasing acceleration voltage.
- Then the chip is immersed in the etching solution of hydrofluoric acid which removes silica from the areas which are not protected by KPR as shown in figure.

- After diffusion of impurities, photo resist is removed by hot sulfuric acid and mechanical abrasion.
- The etching process described above is a wet etching process, since the chemical reagents are in the liquid form.
- A new process used these days is a dry etching process called plasma etching.
- A major advantage of dry process is that it is possible to achieve smaller openings (1μm) compared to wet process.



Ion implantation system

Metallization:

Process for creating contact silicon and its interconnections on chip.

- Deposition of thin layer of aluminum over the whole wafer.
- Making successive layers.
- The process such as etching, masking and doping are repeated for each successive layers until all IC's are completed.
- Silicon dioxide is used as insulator between the components.
- Aluminum is deposited to make contact pads.
- The fabrication includes about three layers which are separated by dielectric layers. For electrical and physical isolation, solid layer of dielectric is surrounded in each component for purpose of isolation.

- Final dielectric layer is deposited to avoid damage and contamination of circuit.
- The individual IC is tested again for electrical function.
- Then by checking the functionality of each chip on wafer, those chips are not passed in the test will be rejected.

Assembly and packaging:

Each of the wafers contains lots of chips. These chips are separated and packaged by method known as cleaving and scribing.

- The wafer is similar to a piece of glass and diamond saw is used for cutting the wafer into single chips.
- For separation of the individual chips through the rectangular grid, diamond tipped tool is used.
- Those chips are discarded which are failed in electrical test.
- Observation under microscope before packaging.
- The good chip is then sent for packaging.
- For protection, thin wire is connected using ultrasonic bonding
- The chip is tested again before delivered to customer.
- There are three configurations available for packaging.

1) Metal can package
2) Dual in line package.
3) Ceramic flat package

- The chip is assembled in ceramic packages for military applications.

This complete IC's are sealed in anti-static plastic bags.

3.6 Some Semiconductor Devices

A number of semiconductor devices such as junction rectifiers, transistors, photocells, solar batteries and thermistors are known. They all use intrinsic or extrinsic semiconductor crystals. The same crystal consists of regions with different dopants giving them the *n*-type or the *p*-type characteristics. The boundary between two regions of opposite characteristics is called a *p-n* junction. It has special electrical properties such as the rectifying action. A three region crystal (*p-n-p* or *n-p-n*) exhibits transistor action. Here, the superposition of the properties of the two junctions (*p-n* and *n-p*) close to each other acts as an amplifier of electric signals. The small size, reliability, low cost and low power consumption of the solid state devices have brought about a major revolution in the electronics industry in the last two decades.

3.6.1 PN JUNCTION DIODE

PN JUNCTION WITH NO APPLIED VOLTAGE OR OPEN CIRCUIT CONDITION:

In a piece of sc, if one half is doped by p type impurity and the other half is doped by n type impurity, a PN junction is formed. The plane dividing the two halves or zones is called PN junction. As shown in the fig the n type material has high concentration of free electrons, while p type material has high concentration of holes. Therefore at the junction there is a tendency of free electrons to diffuse over to the P side and the holes to the N side. This process is called diffusion. As the free electrons move across the junction from N type to P type, the donor atoms become positively charged. Hence a positive charge is built on the N-side of the junction. The free electrons that cross the junction uncover the negative acceptor ions by filing the holes. Therefore a negative charge is developed on the p –side of the junction..This net negative charge on the p side prevents further diffusion of electrons into the p side. Similarly the net positive charge on the N side repels the hole crossing from p side to N side. Thus a barrier sis set up near the junction which prevents the further movement of charge carriers i.e. electrons and holes. As a consequence of induced electric field across the depletion layer, an electrostatic potential difference is established between P and N regions, which are called the potential barrier, junction barrier, diffusion potential or contact potential, Vo. The magnitude of the contact potential Vo varies with doping levels and temperature. V_o is 0.3V for Ge and 0.72 V for Si.

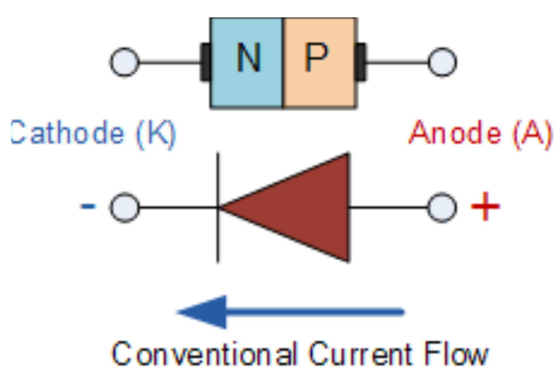


Fig 3.6.1 Symbol of PN Junction Diode

The electrostatic field across the junction caused by the positively charged N-Type region tends to drive the holes away from the junction and negatively charged p type regions tend to drive the electrons away from the junction. The majority holes diffusing out of the P region leave behind negatively charged acceptor atoms bound to the lattice, thus exposing a negatives pace charge in a previously neutral region. Similarly electrons diffusing from the N region expose positively ionized donor atoms and a double space charge builds up at the junction as shown in the fig. 3.6.2

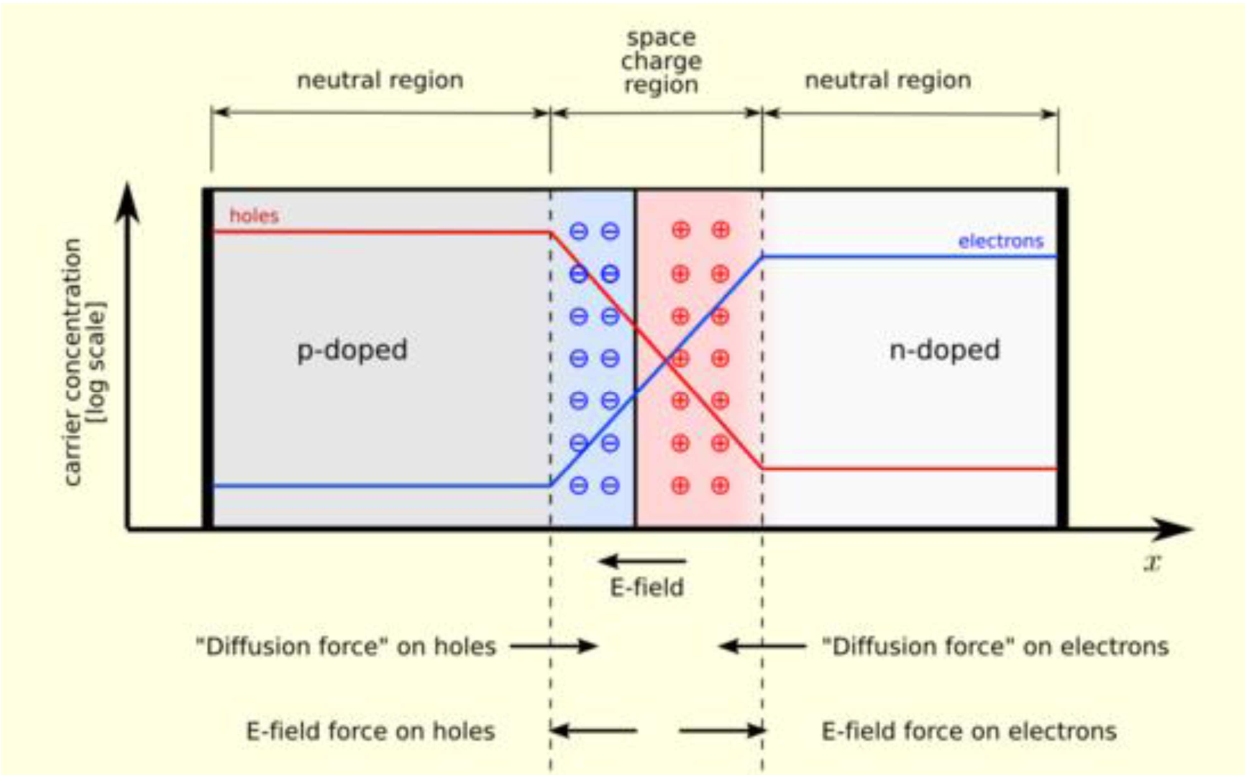


Fig 3.6.2

It is noticed that the space charge layers are of opposite sign to the majority carriers diffusing into them, which tends to reduce the diffusion rate. Thus the double space of the layer causes an electric field to be set up across the junction directed from N to P regions, which is in such a direction to inhibit the diffusion of majority electrons and holes as illustrated in fig 3.6.3. The shape of the charge density, ρ, depends upon how diode id doped. Thus the junction region is depleted of mobile charge carriers. Hence it is called depletion layer, space region, and transition region. The depletion region is of the order of 0.5 µm thick. There are no mobile carriers in this narrow depletion region. Hence no current flows across the junction and the system is in equilibrium. To the left of this depletion layer, the carrier concentration is $p = N_A$ and to its right it is $n = N_D$.

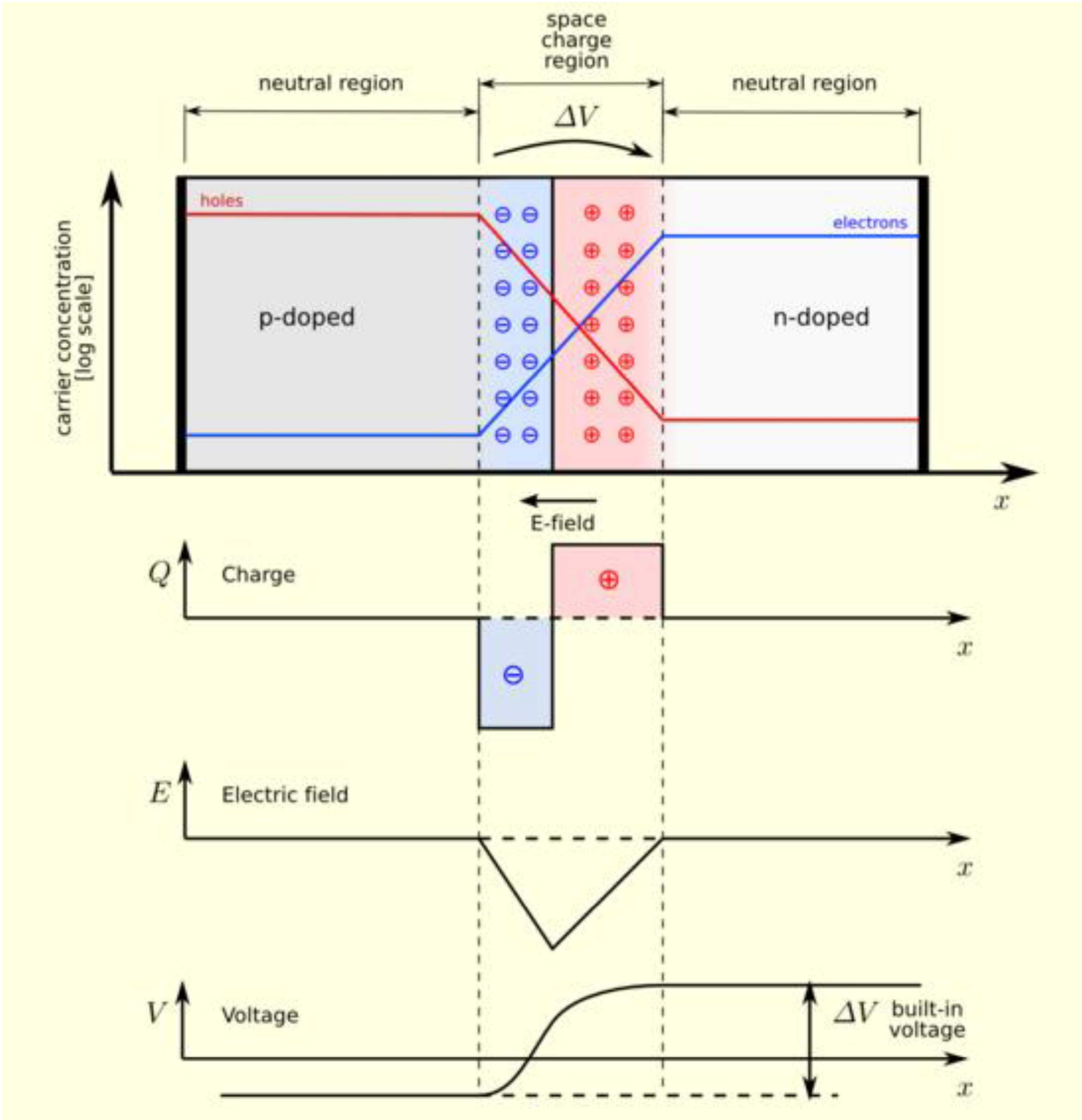


Fig 3.6.3

FORWARD BIASED JUNCTION DIODE

When a diode is connected in a **Forward Bias** condition, a negative voltage is applied to the N-type material and a positive voltage is applied to the P-type material. If this external voltage becomes greater than the value of the potential barrier, approx. 0.7 volts for silicon and 0.3 volts for germanium, the potential barriers opposition will be overcome and current will start to flow. This is because the negative voltage pushes or repels electrons towards the junction giving them the energy to cross over and combine with the holes being pushed in the opposite direction towards the junction by the positive voltage. This results in a characteristics curve of zero current flowing up to this voltage point, called the "knee" on the static curves and then a high current flow through the diode with little increase in the external voltage as shown below.

Forward Characteristics Curve for a Junction Diode

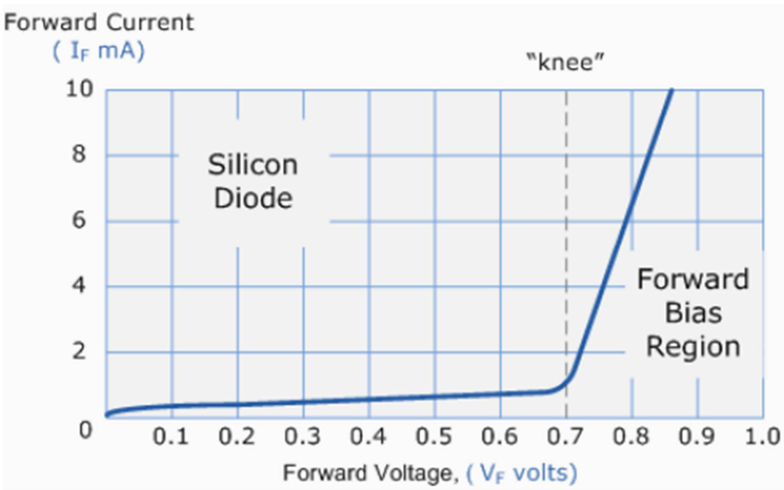


Fig 3.6.5: Diode Forward Characteristics

The application of a forward biasing voltage on the junction diode results in the depletion layer becoming very thin and narrow which represents a low impedance path through the junction thereby allowing high currents to flow. The point at which this sudden increase in current takes place is represented on the static I-V characteristics curve above as the "knee" point.

Forward Biased Junction Diode showing a Reduction in the Depletion Layer

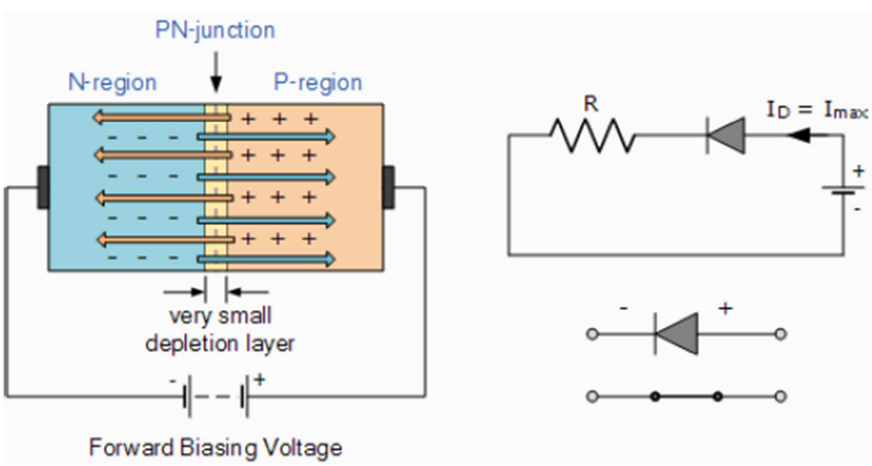


Fig 3.6.6: Diode Forward Bias

This condition represents the low resistance path through the PN junction allowing very large currents to flow through the diode with only a small increase in bias voltage. The actual potential difference across the junction or diode is kept constant by the action of the depletion layer at approximately 0.3v for germanium and approximately 0.7v for silicon junction diodes. Since the diode can conduct "infinite" current above this knee point as it effectively becomes a short circuit, therefore resistors are used in series with the diode to limit its current flow. Exceeding its maximum forward current specification causes the device to dissipate more power in the form of heat than it was designed for resulting in a very quick failure of the device.

PN JUNCTION UNDER REVERSE BIAS CONDITION:

Reverse Biased Junction Diode

When a diode is connected in a **Reverse Bias** condition, a positive voltage is applied to the N-type material and a negative voltage is applied to the P-type material. The positive voltage applied to the N-type material attracts electrons towards the positive electrode and away from the junction, while the holes in the P-type end are also attracted away from the junction towards the negative electrode. The net result is that the depletion layer grows wider due to a lack of electrons and holes and presents a high impedance path, almost an insulator. The result is that a high potential barrier is created thus preventing current from flowing through the semiconductor material.

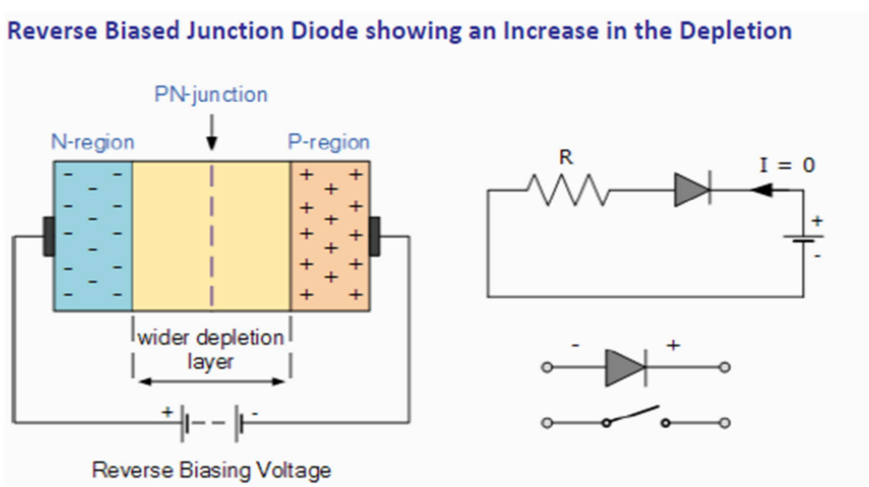


Fig 3.6.7: Diode Reverse Bias

This condition represents a high resistance value to the PN junction and practically zero current flows through the junction diode with an increase in bias voltage. However, a very small **leakage current** does flow through the junction which can be measured in microamperes, (μA). One final point, if the reverse bias voltage Vr applied to the diode is increased to a sufficiently high enough value, it will cause the PN junction to overheat and fail due to the avalanche effect around the junction. This may cause the diode to become shorted and will result in the flow of maximum circuit current, and this shown as a step downward slope in the reverse static characteristics curve below.

Reverse Characteristics Curve for a Junction Diode

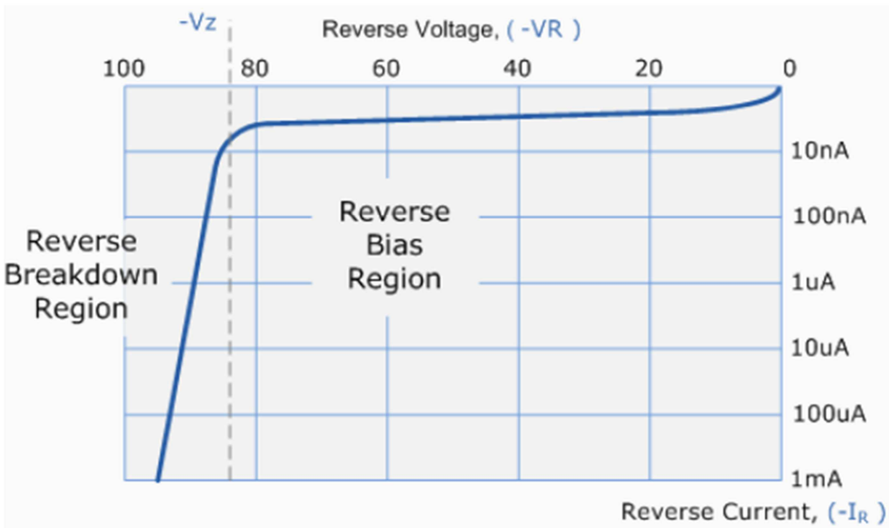


Fig 3.6.8: Diode Reverse Characteristics

Sometimes this avalanche effect has practical applications in voltage stabilizing circuits where a series limiting resistor is used with the diode to limit this reverse breakdown current to a preset maximum value thereby producing a fixed voltage output across the diode. These types of diodes are commonly known as **Zener Diodes**

VI CHARACTERISTICS AND THEIR TEMPERATURE DEPENDENCE

Diode terminal characteristics equation for diode junction current:

$$I_D = I_o(e^{\frac{v}{nV_T}-1})$$

Where V_T = KT/q;

V_D_ diode terminal voltage, Volts

I_o _ temperature-dependent saturation current, µA

T _ absolute temperature of p-n junction, K

K _ Boltzmann's constant 1.38x 10^{-23}J/K

q _ electron charge 1.6x10^{-19} C

η = empirical constant, 1 for Ge and 2 for Si

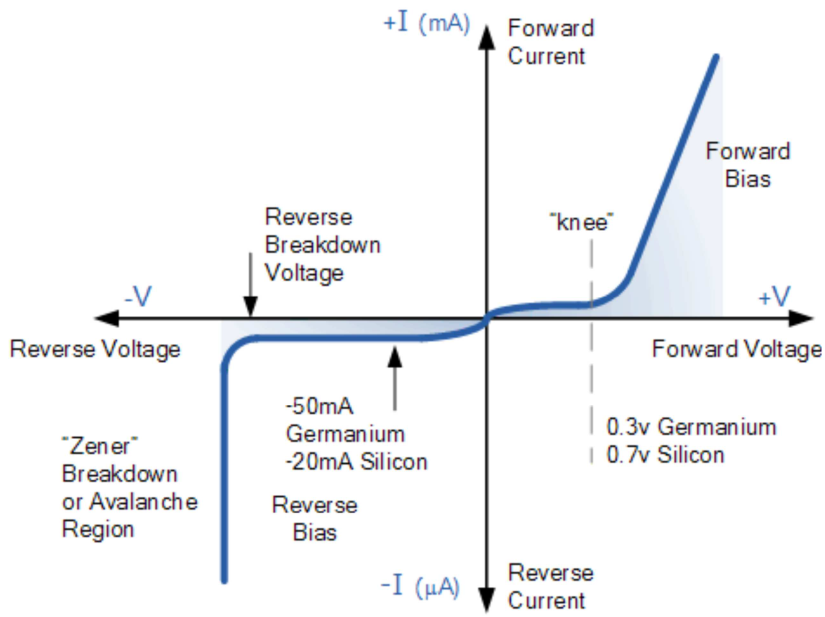


Fig 3.6.9: Diode Characteristics

APPLICATIONS

A p-n junction diode is used as a rectifier, a switch, a clamper etc. We shall study here the application of diode as a rectifier.

Rectifiers

Government supplies electrical energy in the form of ac voltage because it is more economic and efficient to distribute ac power. However, dc voltage is required for operating most of the electronic equipments and gadgets. It is therefore necessary to convert the incoming mains supply to dc voltage. The process of converting ac voltage into dc voltage is called **rectification** and the device or circuit that converts ac voltage into dc voltage is called **rectifier**. The unidirectional current conduction property of diode is used in rectifiers. A single diode or more diodes can be connected into a circuit to form different types of rectifier circuits.

Half-Wave Rectifier

A circuit employing a single diode whereby the diode conducts during only one half-cycle of the input ac cycle is called a half-wave rectifier. The basic half wave rectifier circuit is shown in Fig. 3.7 (a). In the circuit, an alternating voltage is applied to a single diode connected in series with resistor R_L. Resistor R_L represents a device to which the circuit delivers power. As the device draws power and performs work, it is called load. Therefore, R_L is called load resistor. The input voltage u i is a sinusoidal voltage, which changes in polarity 100 times in a second. During the positive half cycle of the input voltage the diode is forward biased and offers a very low resistance. As a result, a large current I_L flows through the load resistor. A voltage u0 develops across R_L. The voltage drop across the diode is usually negligible. Therefore, the entire power appears across R_L. During the negative half cycle of the input waveform, the diode is reverse-biased and offers a very high resistance. Therefore, only a negligible reverse saturation current flows through the circuit. As such $I_L = 0$ and hence $V_0 = 0$. It means that the negative half-cycle of input voltage waveform is suppressed and is not utilized for delivering power to the load. In this condition all the input voltage appears across the diode itself. The input and output waveforms are shown in Fig. 3.7 (b). It is seen that the output voltage is no longer an ac voltage. It is a unidirectional fluctuating voltage. It has an average dc value over which a number of ac components are superimposed. The undesired ac components are called the ripple. In case of the half-wave rectifier, the lowest ripple frequency is the same as the frequency of the input voltage.

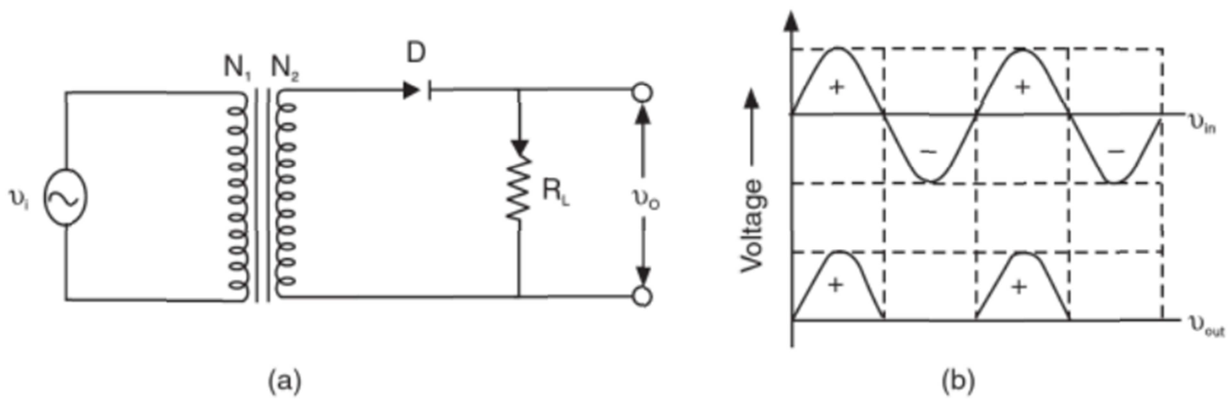


Fig 3.7 Half-wave rectifier (a) Circuit diagram (b) input and output voltage wave forms

Full Wave Rectifier

A full wave rectifier uses two diodes and rectifies both the half cycles of the input voltage. The circuit is shown in Fig. 3.8 (a). The input is supplied to the full wave rectifier from a transformer with a center-tapped secondary winding. During the positive half-cycles of the secondary voltage, the diode D1 is forward biased and the diode D2 is reverse biased. The current iL flows through the diode D1, the load resistor R_L and the upper half-winding of the secondary. The output voltage V_0 develops across RL. During the negative half cycles, the diode D1 is reverse biased and the diode D2 is forward biased. The current now flows through the diode D2, the load resistor R_L and the lower half winding of the secondary. The direction of the current flow through the resistor R_L is the same in both the cases. Hence during both the alternations, current i_L passes through R_L and produces an output voltage Vo. Thus, in the full wave rectifier, both the half cycles are utilized to produce the output. The output consists of a continuous series of positive half cycles of alternating voltage. The input and output waveforms are shown in Fig. 3.8 (b). The pulsating dc voltage obtained from a rectifier is smoothened out with the help of filter circuits and a stable dc voltage is obtained.

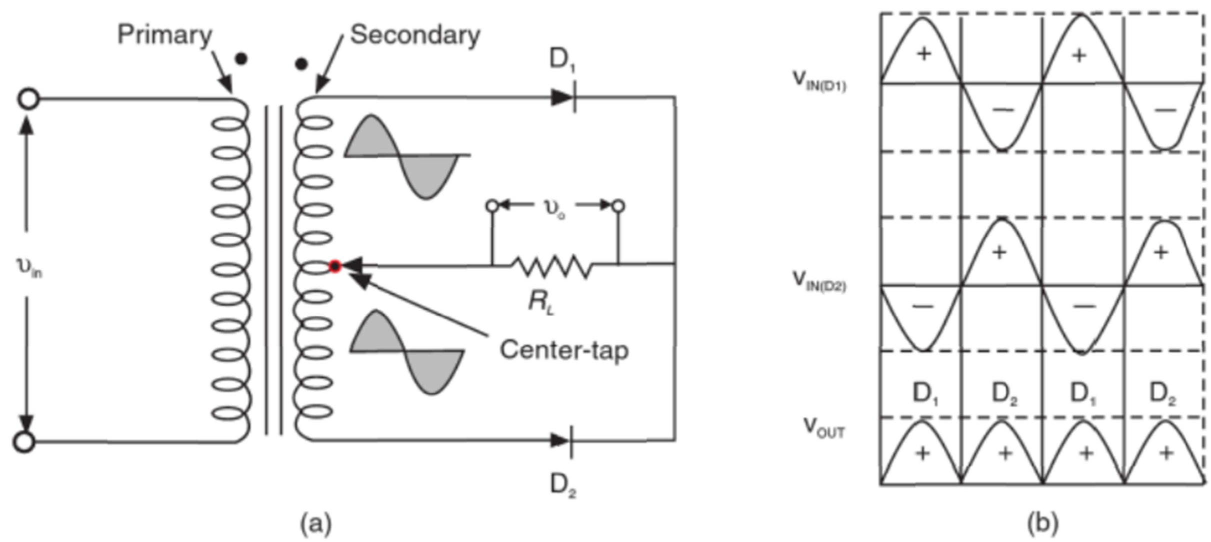


Fig 3.8 Full-wave rectifier (a) Circuit diagram (b) input and output voltage wave forms

3.6.2 LIGHT EMITTING DIODE (LED)

A light emitting diode (LED) is a semiconductor diode that gives off light when it is forward biased. LEDs are generally fabricated using III-IV compound semiconductors, such as GaAs, which have a direct band gap. Principle: When a p-n junction is forward biased, minority carriers flow in large numbers into regions where they can recombine with majority carriers producing light in the visible or infra red region. The wavelength of light is given by

$$\lambda = \frac{hc}{E_g} = \frac{1.24}{E_g\,(eV)}\mu\text{ m}$$

This effect is known as injection electroluminescence. A significant light output is obtained only when there is large number of electro-hole recombinations occurring per second. To ensure this, the p and n regions are heavily doped.

Theory: The energy band diagram of a heavily doped p-n junction is shown in Fig. 3.9 a. There is a large concentration of electrons in the conduction band of n-region and a large concentration of holes in the valence band of p-region. When forward bias is applied the electrons push into the depletion region and occupy energy levels in the conduction band. Similarly, holes push forward into the depletion region and occupy energy levels in the valence band. The electrons in the conduction band are directly above the holes at the edge of the valence band (Fig. 3.9 b). The situation is highly conducive for direct recombination of electrons and holes. When an electron from the conduction band jumps into the hole in the valence band, recombination occurs and the excess energy is emitted in the form of a light photon.

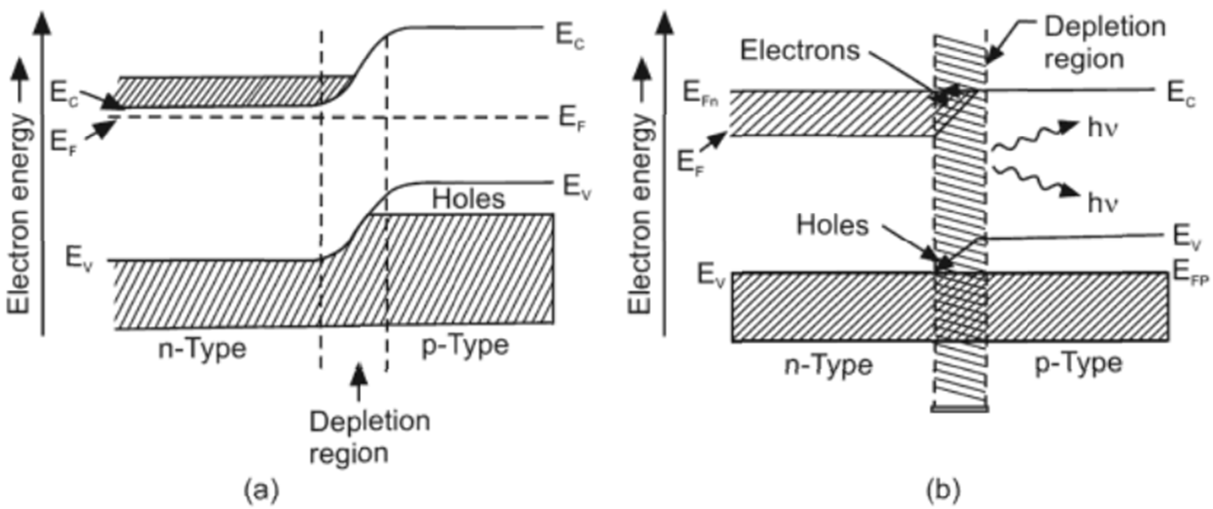


Fig. 3.9 Energy band diagram of an LED- (a) without bias (b) under forward bias

<u>Construction</u>

We describe here the structure of a surface emitting LEDs. These LEDs emit light in a direction perpendicular to the p-n junction plane. The construction of a surface emitting LED is shown in Fig. 3.10 (a). An n-type layer is grown on a substrate and a p-type layer is grown on it by the process of diffusion. The p-layer is made very thin to prevent loss of photons due to absorption in the layer. Metal connections are made at the edges of the p-layer in order to allow more central surface for the light to escape. A metal film is deposited at the bottom of the substrate for reflecting as much light as possible towards the surface of the device and also to provide electrode connection. The light generated at the junction may not emerge from the surface of the device as it is likely to suffer total internal reflection at the semiconductor-air boundary. Therefore, the device is encapsulated in a clear epoxy resin of suitable refractive index (Fig. 3.10 b).

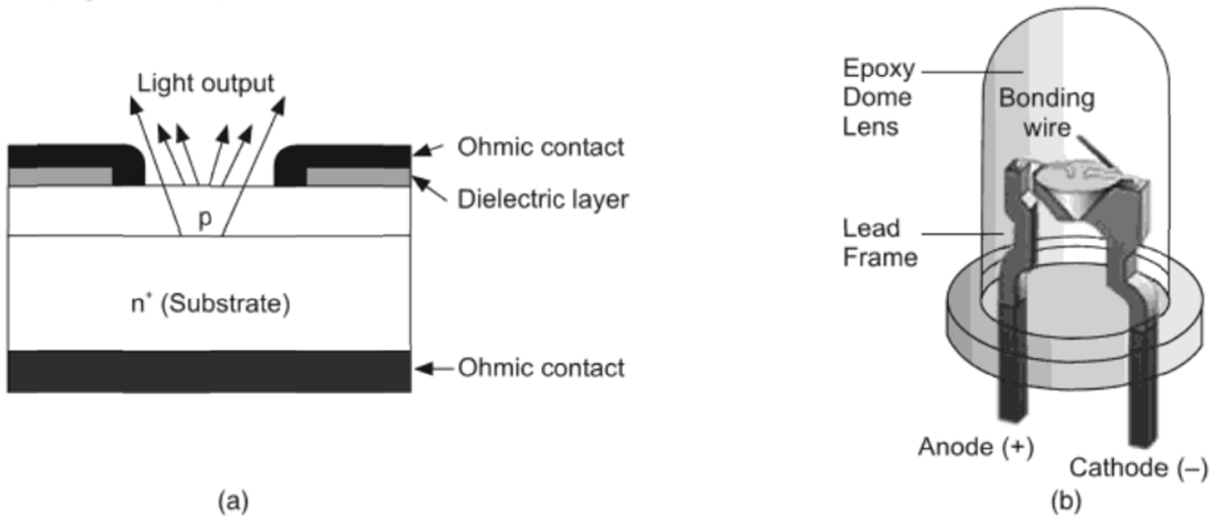


Fig 3.10

Working

The circuit symbol of LED and a simple circuit to illustrate the working of an LED are shown in Fig. 3.11 LED is always forward biased. The forward voltage across an LED is considerably greater than an ordinary diode. Typically the maximum forward voltage for LED is between 1.2 V and 3.2 V depending on the device. The LED emits light in response to a sufficient forward current. The amount of light emitted is directly proportional to the forward current, as shown in Fig. 3.11 (c). The reverse breakdown voltage of LED is of the order of 3V and an LED is never reverse biased.

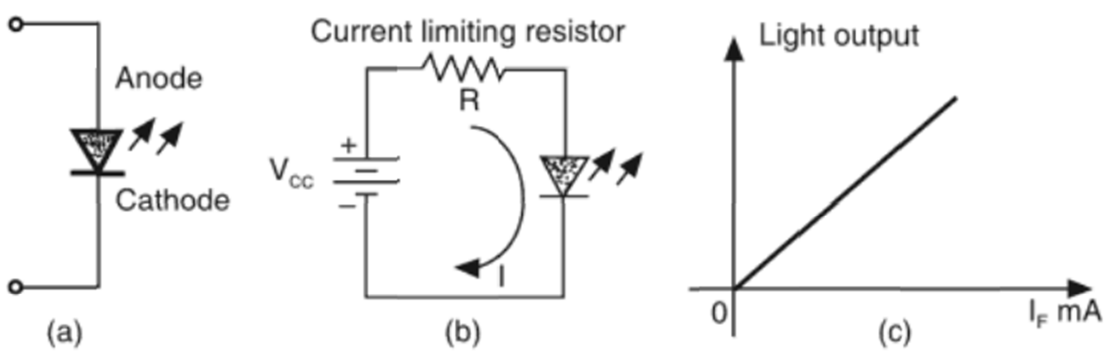


Fig. 3.11

Applications LEDs are used in many applications. Discrete LEDs are used as indicators and as light sources in fibre-optic communications. A number of LEDs may be grouped to form a display. The LEDs may be arranged in the form of a seven-segment display where by energizing a proper combination of segments the decimal numbers 0 to 9 may be displayed. Or they may be arranged in the form of a 5 × 7 matrix which may be used to generate a decimal number or alphabetical character.

3.6.3 Photon Detectors

The electrons in the valence band can be excited into the conduction band by visible light, when the energy gap of a semiconductor is in the right range 1.60–3.20 eV. The additional holes and electrons created by the incident light can lead to an increase in the current in an appropriate external circuit. The current is a direct measure of the incident light intensity. *Photoconductors* are devices used for detecting and measuring light energy.

The sensitivity of a photoconductor is a maximum, if it remains essentially an insulator in the dark, that is, with a minimum of thermal excitation of charge carriers. Cadmium sulphide (E_g = 2.42 eV) is a good photoconductor and responds to light in the green region. CdSe has a smaller gap of 1.74 eV, corresponding to the red end of the visible spectrum. CdTe with a gap of 1.45 eV is further out in the infrared region. PbS, PbSe and PbTe photoconductors have energy gaps in the range of 0.3 eV. Photoconductors are used as burglar alarms, for automatic door opening and for switching on street lights, as the sun goes down.

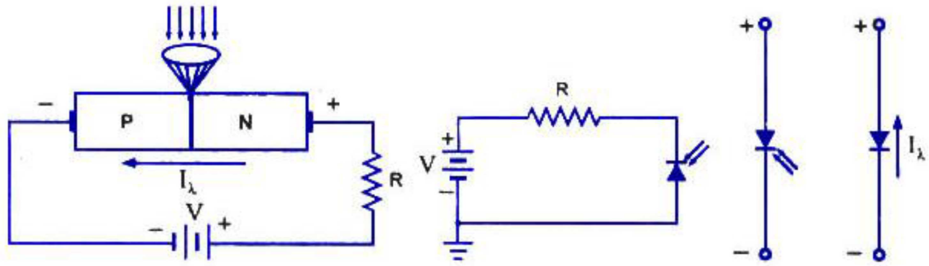
PHOTO DIODE

The photo diode is a semiconductor p-n junction device whose region of operation is limited to the reverse biased region. The figure below shows the symbol of photodiode

Fig 3.12:Symbol of photodiode.

Principle of operation:

A photodiode is a type of photo detector capable of converting light into either current or voltage, depending upon the mode of operation. The common, traditional solar cell used to generate electric solar power is a large area photodiode. A photodiode is designed to operate in reverse bias. The deletion region width is large. Under normal conditions it carries small reverse current due to minority charge carriers. When light is incident through glass window on the p-n junction, photons in the light bombard the p-n junction and some energy s imparted to the valence electrons. So valence electrons break covalent bonds and become free electrons. Thus more electron-hole pairs are generated. Thus total number of minority charge carriers increases and hence reverse current increases. This is the basic principle of operation of photo diode.



3.13 Basic Biasing Arrangement and construction of photodiode and symbols

Characteristics of photodiode:

When the P-N junction is reverse-biased, a reverse saturation current flows due to thermally generated holes and electrons being swept across the junction as the minority carriers. With the increase in temperature of the junction more and more hole-electron pairs are created and so the reverse saturation current I_0 increases. The same effect can be had by illuminating the junction. When light energy bombards a P-N junction, it dislodges valence electrons. The more light striking the junction the larger the reverse current in a diode. It is due to generation of more and more charge carriers with the increase in level of illumination. This is clearly shown in figure for different intensity levels. The dark current is the current that exists when no light is incident. It is to be noted here that current becomes zero only with a positive applied bias equals to V_Q. The almost equal spacing between the curves for the same increment in luminous flux reveals that the reverse saturation current I_0 increases linearly with the luminous flux as shown in figure. Increase in reverse voltage does not increase the reverse current significantly, because all available charge carriers are already being swept across the junction. For reducing the reverse saturation current I_0 to zero, it is necessary to forward bias the junction by an amount equal to barrier potential. Thus the photodiode can be used as a photoconductive device.

On removal of reverse bias applied across the photodiode, minority charge carriers continue to be swept across the junction while the diode is illuminated. This has the effect of increasing the concentration of holes in the P-side and that of electrons in the N-side But the barrier potential is negative on the P-side and positive on the N-side, and was created by holes flowing from P to N-side and electrons from N to P-side during fabrication of junction. Thus the flow of minority carriers tends to reduce the barrier potential.

When an external circuit is connected across the diode terminals, the minority carrier; return to the original side via the external circuit. The electrons which crossed the junction from P to N-side now flow out through the N-terminal and into the P-terminal This means that the device is behaving as a voltage cell with the N-side being the negative terminal and the P-side the positive terminal. Thus, the photodiode is & photovoltaic device as well as photoconductive device.

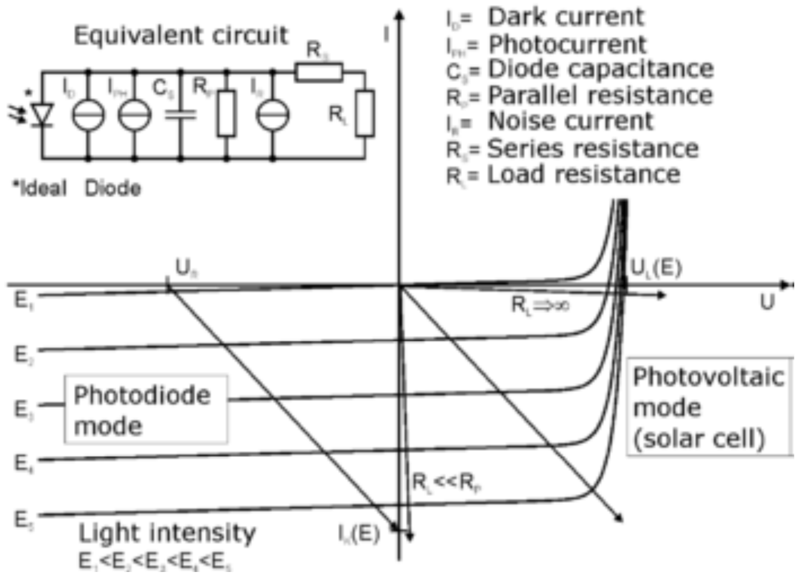


Fig 3.14 characteristics of photodiode

Advantages:

The advantages of photodiode are:

1. It can be used as variable resistance device.
2. Highly sensitive to the light.
3. The speed of operation is very high.

Disadvantages:

1. Temperature dependent dark current.
2. poor temperature stability.
3. Current needs amplification for driving other circuits.

Applications:

1. Alarm system.
2. counting system.
3. Fiber optical communication systems.

3.6.4 Bipolar Junction Transistor

TRANSISTOR STRUCTURE

A **transistor** is a semiconductor device consisting of three regions separated by two distinct p-n junctions. The central region is called the **base**. It may be a p-type or n-type semiconductor. The two outer regions are called emitter and collector (Fig. 3.15). They are of the same type extrinsic semiconductor but different from that of base. Thus, if the base is p-type the emitter and collector are n-type and if the base is n-type the **emitter** and **collector** are p-type. Thus, two types of transistors are available. They are called npn and pnp transistors. The npn transistor is constructed using n type material as the emitter and collector while the base is made of p type material. The pnp transistor is constructed using p type material as the emitter and collector while the base is made of n type. The n-region contains free electrons which are negative charge carriers and p-region contains mobile holes which are positive carriers. Thus, two types of charge carriers, namely electrons and holes, are involved in current flow through an npn or pnp transistor. Therefore, these transistors are known as bipolar junction transistors.

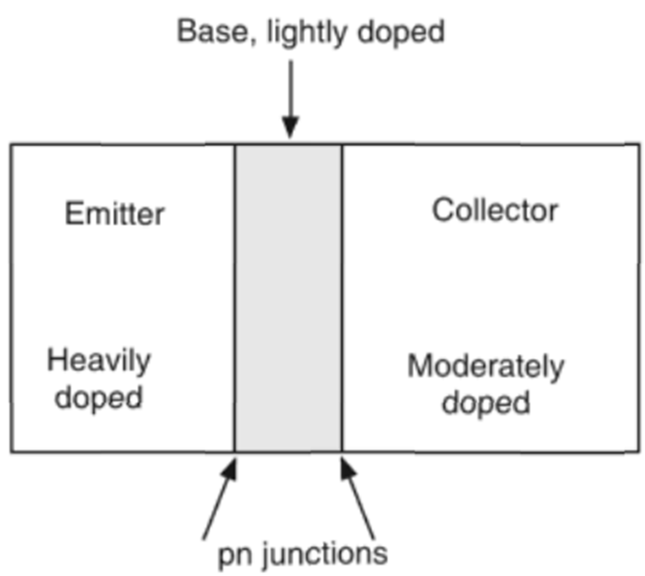


Fig 3.15

The function of each element is as follows: (i) The emitter provides the majority carriers necessary to support current flow. (ii) The base controls the flow of the majority carriers within all elements of the transistor. (iii) The collector supports the majority of the current flow in the transistor. In most cases the current that flows through the collector accomplishes the work done by a transistor.

SCHEMATIC REPRESENTATION

The schematic symbols of npn and pnp transistors are shown in Fig. 3.16 In the symbols, the emitter is always indicated by an arrowhead. This arrowhead serves to tell us three things: (i) the location of the emitter; (ii) the type of the transistor that is being represented; and (iii) the direction in which the conventional current flows.

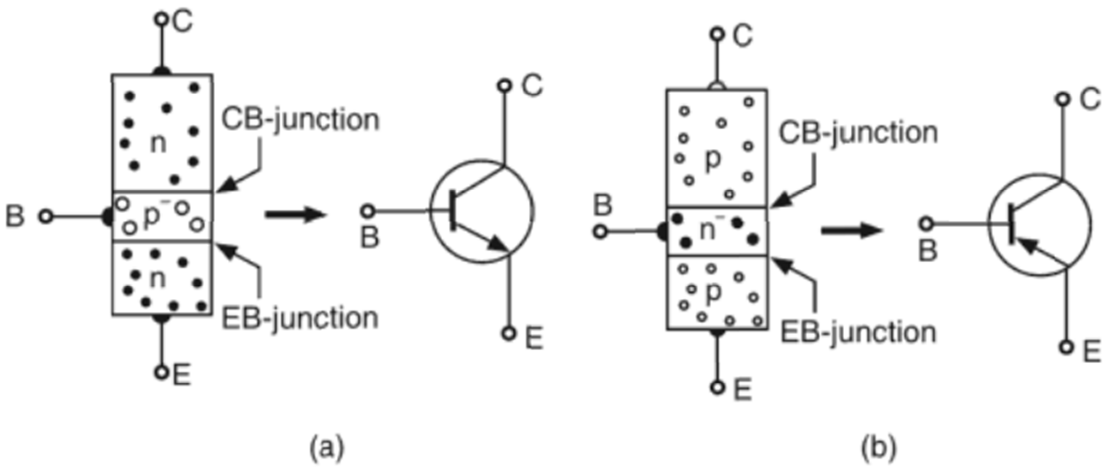


Fig.3.16: Circuit symbol of transistor (a) n-p-n transistor (n) p-n-p transistor

In the npn transistor, the arrow points away from the base. In this device electrons flow from the emitter into the base and hence the current flows from the base to the emitter. In the pnp transistor, the arrow points toward the base. The holes flow from the emitter into the base and the current flows from the emitter into the base. A simple way of remembering the direction of arrow is as follows. In an npn transistor, n-regions are outside p-region and the arrow points outward and in a pnp-transistor, n-region is in between p-regions and the arrow points inward.

Thus, npn ….. n outside, arrow outward

pnp ….. n inside, arrow inward

BIASING THE TRANSISTOR

The two junctions of a transistor can be biased in four different ways.

(i) Both the junctions may be forward biased. It causes large currents to flow across the junctions. The currents join together in the base and flow down the common lead. Then the transistor is said to be operating in **saturation region**.

(ii) Both the junctions may be reverse biased. Very small currents flow through the junctions. The transistor is said to be in **cut-off region**.

(iii) EB-junction may be reverse biased and CB-junction forward biased. The transistor is said to operate in an **inverted mode**.

(iv) EB junction may be forward biased and the CB junction reverse biased. Such biasing arrangement causes a large current to flow across the EB-junction as well as CB-junction. Further, the collector current is controlled by the emitter current or base current. With such biasing, the transistor is said to operate in **active region** or in **normal mode**.

We are interested in the particular biasing where the transistor operates in normal mode. In the normal mode, when the current flows from emitter to base, it meets with a low resistance of about 15-20 Ω. Hence, the current is usually of the order of a few milliamperes for a voltage of a fraction of volt applied suitably between the emitter and the base. On the contrary the current encounters a large resistance of the order of 100 kΩ when it flows from the base to collector. It will be seen later that this is very advantageous for setting up amplifier circuits using transistors.

CIRCUIT CONFIGURATIONS

A transistor is a three-terminal device. There are three possible ways in which it may be connected into a circuit. They are known as circuit configurations. When the transistor is connected with its base terminal common to both the EB-junction and CB-junction, the configuration is known as common-base (CB) configuration. The other configurations are known as common emitter and common collector configurations. The circuit configurations for an npn transistor are shown in Fig. 3.17.

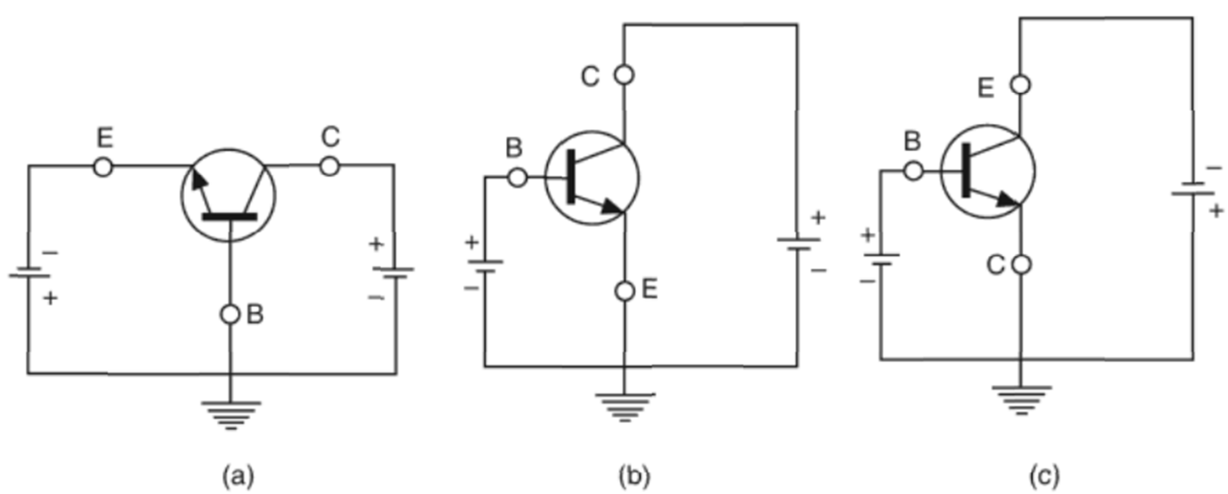


Fig.3.17: The three different ways of connecting a transistor (a) Common-base configuration. (b) Common-emitter configuration. (c) Common-collector configuration.

TRANSISTOR AS AN AMPLIFIER

An amplifier is an electronic circuit that causes an increase in the voltage or power level of a given signal. Fig. 3.18 shows an npn transistor connected in common base configuration. The transistor is biased to operate in the active region. The battery V_{EE} forward biases the EB junction and the battery V_{CC} reverse biases the CB junction. A signal source υ_i is connected in the input circuit. A load resistance R_L is connected in the output circuit. An output voltage υ_o is developed across R_L. The dc voltage V_{EE} is a fixed voltage and causes a dc current I_E to flow through E_B junction. When the ac voltage υ_i is superimposed on V_{EE}, the emitter-base voltage varies with time. For example, if V_{EE} is10 volts and the peak voltage of the signal υ_i is 1 V, the emitter-base voltage swings from 9 volts to 11 volts. This variation causes corresponding variations in emitter current IE and the collector current I_C. The varying current flows through the load resistor R_L and develops a varying voltage across it. It is the output voltage υ_o.

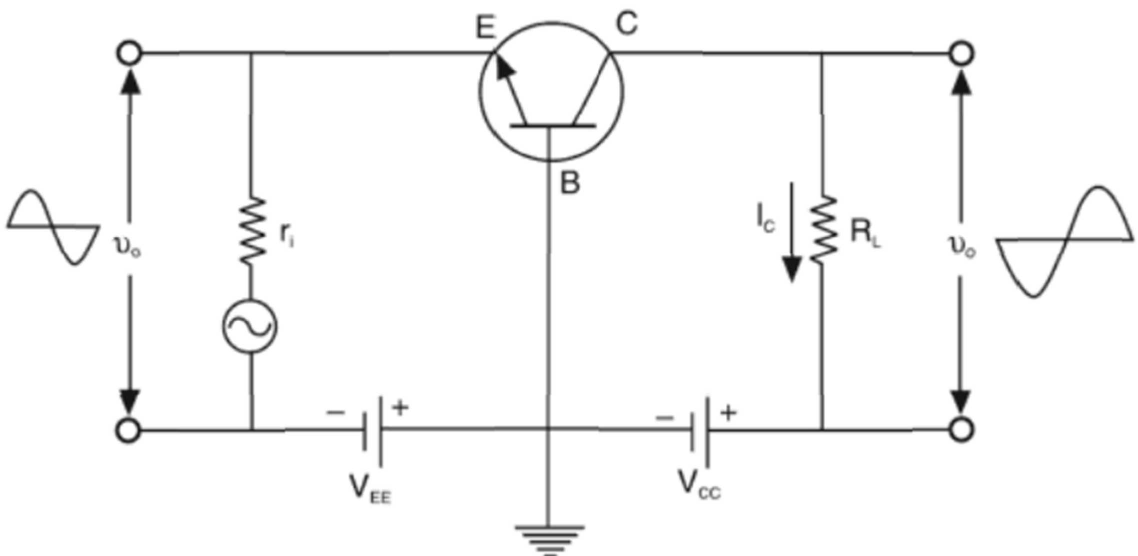


Fig. 3.18: Transistor amplifier

Since the EB junction is forward biased, it has a low resistance ri of the order of 100 ohms. The emitter current variation ΔI_E due to emitter-base voltage variation can be expressed as

$$\mathbf{\Delta I_E = \upsilon_i / r_i}$$

The collector current I_C changes by ΔI_C due to the variation ΔI_E caused in I_E.

$$\mathbf{\Delta I_C = \alpha_{dc} \Delta I_E}$$

This current ΔIC flows through RL causing a voltage drop (ΔI_C) R_L. Therefore,

$$\mathbf{\upsilon o = (\Delta I_C) R_L}$$

$$= \alpha\ (\Delta I_E)\ R_L$$

$$= \alpha\ (\upsilon_i/r_i)\ R_L$$

The voltage gain of an amplifier is defined as the ratio of output signal voltage υ_o to the input signal voltage υ_i. Thus,

$$\text{Gain} = \text{Output Voltage, } \upsilon_o \,/\, \text{Input Voltage, } \upsilon_i$$

$$= \alpha\ R_L/r_i$$

$$\cong R_L/r_i\ (\alpha \approx 1)$$

R_L is of the order of kilo ohms and is far larger than r_i. . Consequently, υ_o is larger than υ_i and the gain of the circuit is larger than unity. It means that the transistor amplifies a small input voltage to give a larger output voltage.