SOLIDS AND SEMICONDUCTORS DEVICES

Electronic instruments are being utilized in various fields like telecommunication, entertainment, computers, nuclear physics and many more. Although the history started with the advent of vacuum tubes, however the rapid advancement in electronics which we see today is due to the valuable contributions of semiconductor devices.

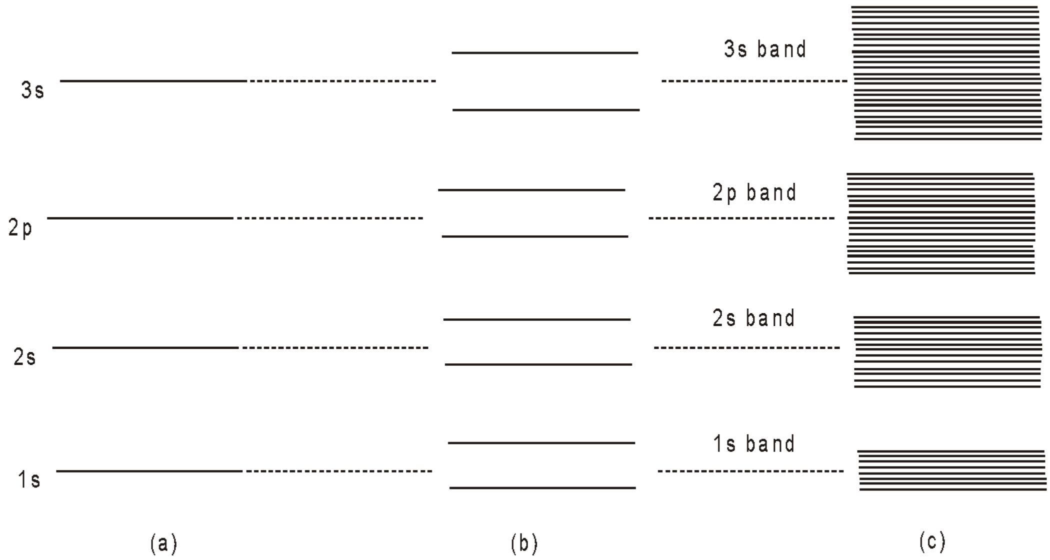
Semiconductor devices are not only small in size, consumes less power, have long life times and are more efficient than vacuum tubes but also are of low cost. That is why these have replaced vacuum tubes nearly in all applications. As an example we can consider the case of a computer. In early days, the vacuum tube based computers were as big as the size of a room and were capable of performing simple calculations only. At present the personal computer (PC) that you see in laboratory or at your home is much smaller in size and capable of performing many operations. Needless to say this is possible because of the advances in semiconductor technology.

We will learn the basic concept of semiconductors. This will enable us to understand the operation of many semiconductor devices and then we will be discussing few semiconductor devices like diode, transistor along with their applications.

1. ENERGY LEVELS AND ENERGY BANDS IN SOLIDS

The electrons of an isolated atom are restricted to well defined energy levels. The maximum number of electrons which can be accommodated in any level is determined by the Pauli exclusion principle. The electrons belonging to the outermost energy level are called valence electrons. For example, the electronic configuration of sodium (atomic number 11) is –$1s_2\ 2s_2\ 2p_6\ 3s_1$, here the electron belonging to the 3s level is the valence electron. Most of the solids including metals with which we are familiar occur in crystalline form.As we know a crystal is a regular periodic arrangement of atoms separated from each other by very small distance called lattice costant. The value of lattice constant is different for different crystalline solids, however it is of the order of linear dimension of atoms {~Å}. Obviously at such a short separation between various neighbouring atoms, electrons in an atom cannot only be subjected to the Coulombic force of the nucleus of this atom but also oby Coulombic forces due to nuclei and electrons of the neighboring atoms. In fact it is this interaction which results in the bonding between various atoms which leads to the formation of crystals.

When atoms are interacting (such as in crystal) then the energy level scheme for the individual atoms as shown in figure(a) does not quite hold. The interaction between atoms markedly affect the electron energy levels, as a result there occurs a splitting of energy levels belonging to various atoms. To understand this phenomenon in more clear terms, let us first consider the simplest case of two interacting identical atoms. Let us assume that initially they are far apart i.e. the forces of interaction between them can be neglected. [If the distance between two atoms is much larger (~50Å) compared to their linear dimensions (~ 10Å) this assumption is reasonably correct] .In such a case we may treat them as isolated with energy levels like that for the case of an isolated atom as shown in figure(a).

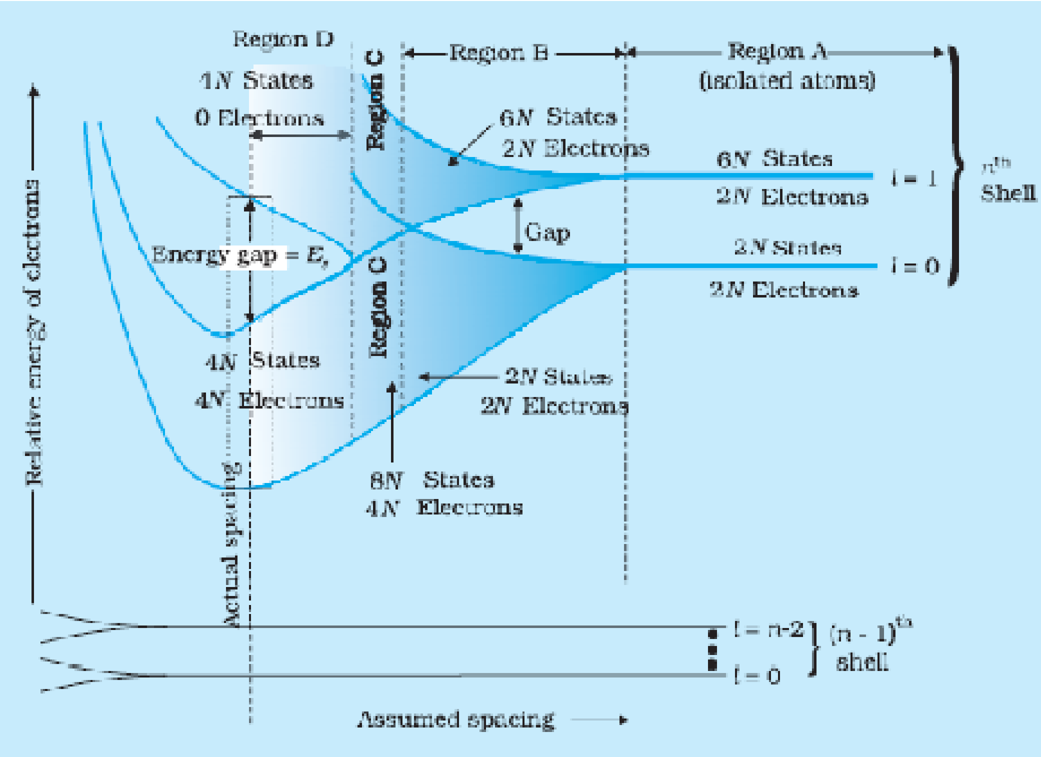


In crystals the number of atoms, N is very large of the order of 10^{22} to 10^{23} per cubic centimetre, so each energy band contains as many levels as the number of atoms. The spacing between various levels within a band is therefore very small. If for example we assume the total width of a band of energies as 1 eV and 10^{22} levels are to be accommodated with in this band, then the average spacing between the adjacent levels is about 10^{-22} eV. For all practical purposes, therefore, energy within a band can be assumed to vary continuously. The formation of bands in a solid is shown schematically in figure(c).

Energy Bands :

This theory is based on the Pauli exclusion principle.

In isolated atom the valence electrons can exist only in one of the allowed orbitals each of a sharply defined energy called energy levels. But when two atoms are brought nearer to each other, there are alterations in energy levels and they spread in the form of bands.



Energy bands are of following types

(1) **Valence band :** The energy band formed by a series of energy levels containing valence electrons is known as valence band. At 0 K, the electrons fils the energy levels in valence band starting from lowest one.

(i) This band is always filed with electrons.

(ii) This is the band of maximum energy.

(iii) Electrons are not capable of gaining energy from external electric field.

(iv) No flow of current due to electron present in this band.

(v) The highest energy level which can be occupied by an electron in valence band at 0 K is called fermi level.

(2) **Conduction band :** The higher energy level band is called the conduction band.

(i) It is also called empty band of minimum energy.

(ii) This band is partially filled by the electrons.

(iii) In this band the electron can gain energy from external electric field.

(iv) The electrons in the conduction band are called the free electrons. They are able to move any where within the volume of the solid.

(v) Current flows due to such electrons.

(3) **Forbidden energy gap (ΔE_g) :** Energy gap between conduction band and valence band

$$\Delta E_g = (C.B.)_{min} - (V.B.)_{max}$$

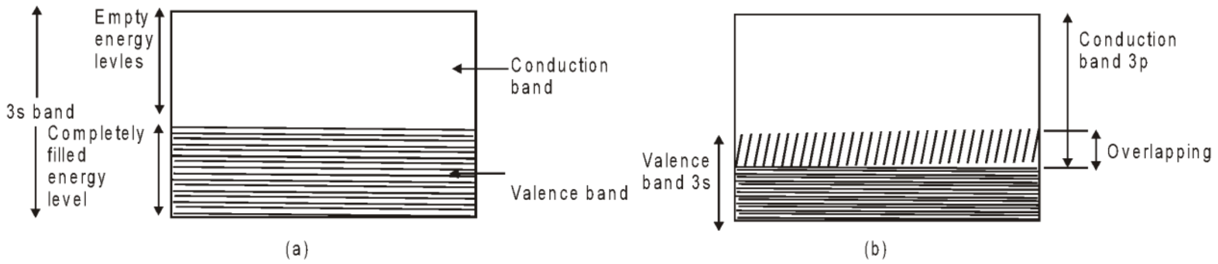
2. CONDUCTOR, INSULATOR AND SEMICONDUCTOR :

The electrical conductivity of materials is a physical quantity which varies over a large span. On one hand we know about metals having very large values of electrical conductivity and on the other hand we have insulators like quartz and mica having negligible conductivity. Beside these there are materials having conductivity (at room temperature) much smaller, than that of metals but much larger than that of insulators these materials are called semiconductors e.g. Silicon and Germanium. Not only that the conductivity of a semiconductor is intermediate, to that of metals and insulators the conductivity of semiconductor varies substantially with temperature. For very low temperature (around 0K) semiconductor behaves like insulator, however, its conductivity increases with increase in temperature.

(a) Conductors :

These are solids in which either the energy band containing valence band is partially filled or the energy band containing valence electrons overlaps with next higher band to give a new band which is partially filled too. For both these situations there are enough free levels available for electrons to which they can be excited by receiving energy from an applied electric field.

Let us consider an example of sodium which is a monovalent metal. Its band structure is such that 1s, 2s and 2p bands are filled with electrons to their capacity however, the 3s band is only half filled. The reason for such a band structure is that for an isolated sodium atom in its electronic structure $1s_2$, $2s_2$, $2p_6$, $3s_1$ the energy levels 1s, 2s and 2p are filled while 3s contains only one electron against its capacity of accommodating two electrons. The completely filled 1s, 2s and 2p bands do not contribute to electrical conduction because an applied electric field cannot bring about intra band transitions in them. Electrons can also not make band to band transitions from Is to 2s or from 2s to 2p band as for both these situations unfilled energy levels are not available. However, electrons belonging to 3s band can take part in intra band transitions as half of the energy levels present in this band are available. An applied electric field can impart them an amount of energy sufficient for the transition to free energy levels, and take part in the process of conduction. Thus the conduction properties of sodium are due to this partially filled band which is shown in figure(a). The lower half portion of this band is called valence band and upper half portion is called conduction band as it is in this part when electron reach after receiving energy from electric field the process of conduction starts. All monovalent metals have a half filled conduction band like sodium.



The bivalent elements belonging to the second group of the periodic table e.g magnesium, zinc etc are also metallic. In the solid state of these materials there is an overlapping between the highest filled band and next higher unfilled band. For example magnesium atom (atomic number = 12) has electronic structure - $1s_2$ $2s_2$ $2p_6$ $3s_2$ and in atomic state there is some energy gap between completely filled 3s level and next higher but unfilled 3p level. However, during the process of crystal formation, the splitting of energy levels take place in such a manner that the 3p band overlaps with 3s band. In the 'hybrid' band' so formed now electrons have sufficient number of unfilled levels for transition. In such situation if 3s band is called valence band then 3p band is conduction band and the two bands overlap as shown in figure (B).

We can conclude that for both the above metals there is no energy gap between maximum energy of valence band and the minimum energy of the conduction band.

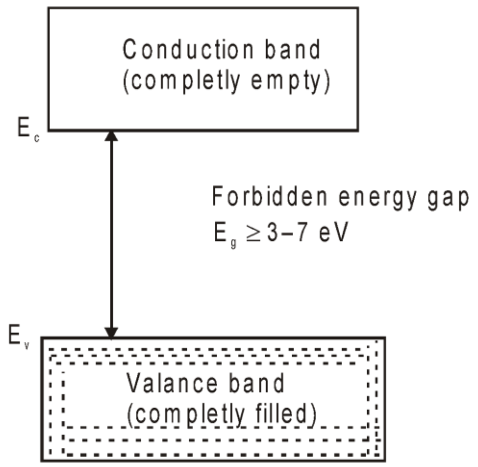
The energy that an electron gains from an ordinary current source usually is 10_{-4} to 10_{-8} eV which is sufficient to cause transition between levels inside a partially filled band. As the difference between the adjacent levels is infinitesimal, for such bands the electron can absorb infinitesimal energy in a manner like free electron. Such electrons when reach unfilled higher levels contribute to the process of electric conduction. In metals both the number of free electrons and the vacant energy levels for transitions are very large that is why metals are good conductors of electricity and heat. For metals at ordinary temperature the electrical conductivities are in range 10_2 mho/metre to 10_8 mho/metre indicating this fact.

(b) Insulator :

It is a solid in which the energy band formation takes place in such a manner, that the valence band is completely filled while the conduction band is completely empty. In addition to this these two bands are separated by a large energy gap called forbidden energy gap or band gap. If E_c and E_v respectively denotes the minimum energy in conduction band and the maximum energy in valence band then band gap E_g is defined as

$E_g = E_c - E_v$

For insulators E_g ~ 3 to 7 eV. As in an empty band no electron is there to take part in the process of electric conduction, such a band does not contribute in conduction. In a completely filled band very large number of electrons are present but no vacant levels to which these electrons make transition are available and hence again there will not be any conduction non such a band. As explained earlier ordinary current sources provide only a very small energy to an electron in a solid and so electrons cannot be excited from valence band to conduction band. Also not only at ordinary temperatures but at elevated temperatures too, the thermal energy is much smaller than the band gap energy E_g so electrons cannot be excited from valence band to conduction band by thermal means. Consequently solids with such large band gaps are insulators.



For diamond, $E_g \approx 6$ eV hence it is insulator.

In general electrical conductivities· of insulators are in the range 10^{-12} mho/metre to 10^{-18} mho/metre (resitivity in the range 10^{11} ohm-metre to 10^{19} ohm metre.]

(c) Semiconductors :

In case of semiconductors, the band structure is essentially of the same type as that for insulators with the only difference that of a relatively smaller forbidden gap. In case of a semiconductor this is typically of the order of 1eV. At absolute zero temperature, the valence band is completely filled and the conduction band is completely empty and consequently no electrical conduction can result. This is the same behaviour as observed in insulators. i.e at absolute zero a semiconductor behaves like an insulator.

At finite temperatures (room temperature and above) some of the electrons from near the top of valence band acquire enough thermal energy to move into the otherwise empty conduction band. These electrons contribute to the conduction of electricity in a semiconductor.

Also the above said transitions create some unfilled levels in the valence band and the electrons of this band can move into these levels again resulting in conduction. Thus the electrical conductivity of a semiconductor is larger than that of an insulator at room temperature. However since the number of electrons made available to conduction band via this process of thermal excitation is very small as compared to what available for conduction in metals, the conductivity of semiconductors is much smaller than that of metals at a given temperature. Thus the conductivity of semiconductor lies between that of metals and insulators, that is why these are named so. The conductivity of semiconductor increases with temperature.

Note : Free electron and Hole in semiconductors.

(1) When an electron is removed from a covalent bond, it leaves a vacancy behind. An electron from a neighbouring atom can move into this vacancy, leaving the neighbour with a vacancy. In this way the

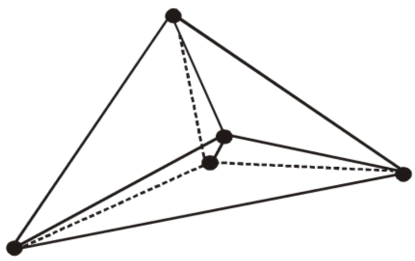
vacancy formed is called hole (or cotter), and can travel through the material and serve as an additional current carriers.
(2) A hole is considered as a seat of positive charge, having magnitude of charge equal to that of an electron.
(3) Holes acts as virtual charge, although there is no physical charge on it.
(4) Effective mass of hole is more than electron.
(5) Mobility of hole is less than electron.
(6) Free electron move in CB, while hole in VB in opposite direction.
(7) Imobile charge is at rest.

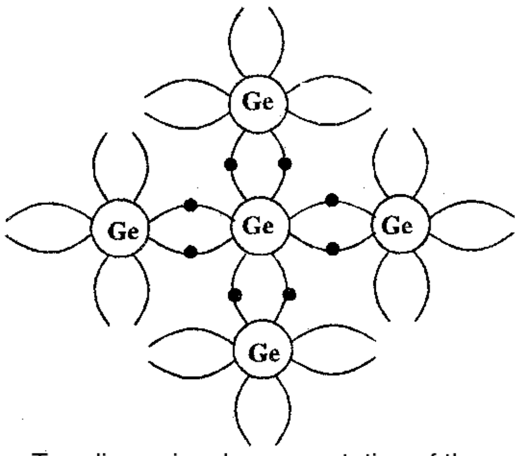
3. INTRINSIC SEMICONDUCTORS :

A semiconductor free from impurities is called an intrinsic semiconductor. Ideally an intrinsic semiconductor crystal should contain atoms of this semiconductor only but it is not possible in practice to obtain crystals with such purities. However if the impurity is less than 1 in 10^8 part of semiconductor it can be treated as intrinsic. For describing the properties of intrinsic semiconductor we are taking examples of silicon and germanium Both silicon and germanium are members of the group IV of periodic table of elements and are tetravalent. Their electronic configuration is as follows:

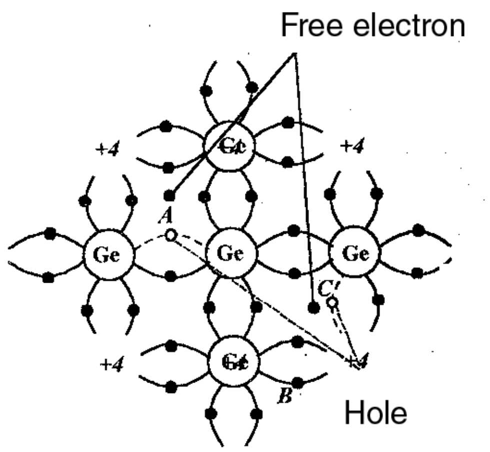
Si(14)=$1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^2$
Ge(32)= $Is^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 3d^{10}\ 4s^2\ 4p^2$

Both elements crystallize in such a way that each atom in the crystal is inside a tetrahedron formed by the four atoms which are closest to it. Figure shows one of these tetrahedral units. Each atom shares its four valence electrons with its immediate neighbours on a one to one basis, so that each atom is involved in four covalent bonds. For convenience, a two dimensional representation of the crystal structure for germanium is shown in figure, which can also be used for silicon (as only covalent bands are being shown).





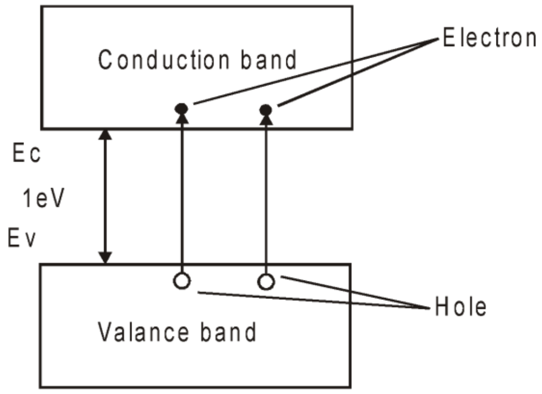
Two dimensional representation of the crystal structure of germanium at OK



Formation of electron hole pair

At 0K, all the valence electrons are involved in the bonding and so the crystal is a perfect insulator as there are no free electrons available for conduction. At higher temperatures, however, some of the valence electrons have sufficient energy to break away from the bond and move in the crystal in random manner. Under an applied electric field these electrons drift and conduct electricity.

When an electron escapes from a band it leaves behind a vacancy in the lattice. This vacancy is termed as a "hole". The absence of electron amounts to the presence of a positive charge of same magnitude. As explained later, holes also take part in conduction in semiconductors. When a covalent bond is broken, all electron- hole pair is contributed. At room temperature (300K) many electron - hole pairs are present in the crystal. The process of electron - hole generation is explained in figure. Let due to thermal energy an electron is set free from the covalent bond at site A whereby a hole is created at this site. An electron from the covalent bond of a neighbouring atom site B may jump to vacant site A then bond is completed at A but a hole is created at B. In this process a very small energy is involved compared to what is required for an electron - hole pair generation. It is because the electron is jumping from one bond to the other and all electrons in bonding are on an average of same energy. As shown in the figure when an electron jumps from C to B a hole is created at C and so on. In effect then such a vacancy or hole can be considered as mobile. Thus in a semiconductor both electrons and holes act as· charge carriers and contribute in electric conduction.



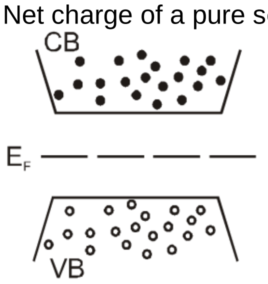
The number of electrons and holes generated by thermal means is equal for an intrinsic semiconductor. If n_e and n_h represents the electron and hole concentrations respectively then

$$n_i = n_e = n_h$$

$$n_e n_h = n_i^2$$

Here n_i is intrinsic concentration.

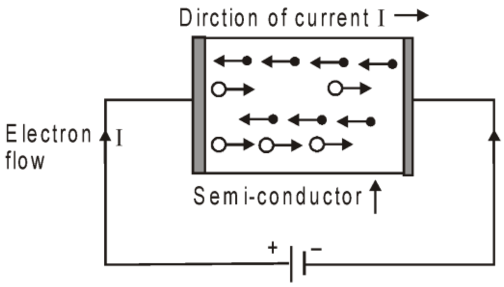
Note : (1) A pure semiconductor is called intrinsic semiconductor. It has thermally generated current carries.
(2) They have four electrons in the outermost orbit of atom and atoms are held together by covalent bond.
(3) Free electrons and holes both are charge carriers and n_e (in C.B.) = n_h (in V.B.)
(4) The drift velocity of electrons (v_e) is greater than that of holes (v_0).
(5) For them fermi energy level lies at the centre of the C.B. and V.B.
(6) In pure semiconductor, impurity must be less than 1 in 10^8 parts of semiconductor.
(7) In intrinsic semiconductor
$n_{e(0)} = n_{h(0)} = n_i$; where $n_{e(0)}$ = Electron density in conduction band, $n_{h(0)}$ = Hole density in V.B., n_i = Density of intrinsic carriers.
(8) The fraction of electron of valance band present in conduction band is given by $f \propto e^{-E_g/kT}$; where E_g = Fermi energy or k = Boltzmann's constant and T = Absolute temperature.
(9) Because of less number of charge carriers at room temperature, intrinsic semiconductors have low conductivity so they have no partical use.
(10) Number of electrons reaching from valence band to conduction band $n = AT^{3/2}e^{-E_g/2kT}$ where A is positive constant.
(11) Net charge of a pure semiconductor is zero.

(12) 

Fermi level is at the middle of ΔE_g

(a) Electrical conductivity of intrinsic semiconductor:

A semiconductor, at room temperature, contains electrons in the conduction band and holes in the valence band. When an external electric field is applied, the electrons move opposite to the field and the holes move in the direction of the field, thus constituting current in the same direction. The total current is the sum of the electron and hole currents.



Let us consider a semiconductor block of length ℓ, area of cross-section A and having electron concentration n_e and hole concentration is n_h, across the ends of the semiconductor creates an electric field E given by

$$E = V/\ell. \quad \text{.........(i)}$$

Under the field E, the electrons and the holes both drift in opposite directions and constiute currents i_e and i_h respectively in the direction of the field. The total current flowing through the semiconductor is

$$i = i_e + i_h$$

If v_e be the drift velocity of the electrons in the conduction band and v_h the drift velocity of the holes in the valence band, then we have

$$i_e = n_e e A v_e \quad \text{and} \quad i_h = n_h e A v_h$$

Where e is the magnitude of electron charge

$$\therefore i = i_e + i_h = eA (n_e v_e + n_h v_h)$$

or $$\frac{i}{A} = e (n_e v_e + n_h v_h). \quad \text{.......(ii)}$$

Let R be the resistance of the semiconductor block and ρ the resistivity of the block material. Then

$$\rho = RA/\ell. \quad \text{.......(iii)}$$

Dividing equation (i) by equation (iii), we have

$$\frac{E}{\rho} = \frac{V}{RA} = \frac{i}{A}, \quad \text{(Since V = iR by Ohm's law).}$$

Substituting in it the value of i/A from equation (ii), we get

$$\frac{E}{\rho} = e (n_e v_e + n_h v_h)$$

or $$\frac{1}{\rho} = e\left(n_e \frac{v_e}{E} + n_h \frac{v_h}{E}\right). \quad \text{........(iv)}$$

Let us now introduce a quantity μ, called mobility which is defined as the drift velocity per unit field and is expressed in meter^2/(volt-second). Thus, the mobilities of electron and hole are given by

$$\mu_e = \frac{v_e}{E} \quad \text{and} \quad \mu_h = \frac{v_h}{E}$$

Introducing μ_e and μ_h in eqaution (iv), we get

$$\frac{1}{\rho} = e (n_e \mu_e + n_h \mu_h)$$

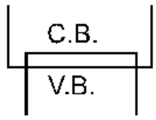
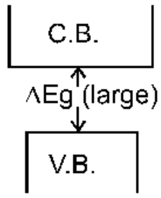
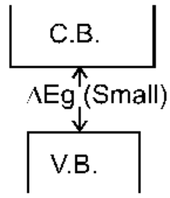
The electrical conductivity σ is the reciprocal of the resistively ρ . Thus, the electrical conductivity of the semiconductor is given by

$$\sigma = e (n_e \mu_e + n_h \mu_h). \quad \because n_e = n_h = n_i$$

So $$\sigma = e n_i (\mu_e + \mu_h)$$

This is the required expression. It shows that the electrical conductivity of a semiconductor depends upon the electron and hole concentrations (number densities) and their mobilities. The electron mobility is higher than the hole mobility.

As temperature rises, both the concentrations n_e and n_h increase due to breakage of more covalent bonds. The mobilities μ_e and μ_h, however, slightly decrease with rise in temperature but this decrease is offset by the much greater increase in n_e and n_h. Hence, the conductivity of a semiconductor increases (or the resistivity decreases) with rise in temperature.

Properties	Conductors	Insulators	Semiconductors
Electrical conductivity	10^{7} to 10^{8} mho/m	10^{-19} to 10^{-11} mho/m	10^{5} to 10^{-6} mho/m
Resistivity	10^{-2} to 10^{-8} Ω-m (negligible)	10^{11} to 10^{19} Ω-m	10^{5} to 10^{6} Ω-m
Band Structure			
Energy gap (E_g)	Zero or very small	Very large : for diamond it is 6 eV	Ge → 0.7 eV Si → 1.1 eV GaAs → 1.3 eV GaF_2 → 2.8 eV
Current carriers	Free electrons	—	Free electrons and holes
Condition of V.B. and C.B. at ordinary temperature	V.B. and C.B. are completely filled or C.B is some what empty	V.B – Completely .field C.B.–Completely unfilled	V.B– some what empty C.B.- some what filled
Temperature co-efficient of resistance	Positive	Zero	Negative
Effect of temperature on conductivity	Decreases	—	Increases
Effect of temperature on resistance	Increases	—	Decreases
Examples	Cu, Ag, Au, Na, Pt, Hg etc.	Wood, plastic, mica diamond, glass etc.	Ge, Si GaAs etc,
Electron density	$10^{29}/m^3$	—	Ge ~ $10^{19}/m^3$ Si ~ 10^{16} / m^3

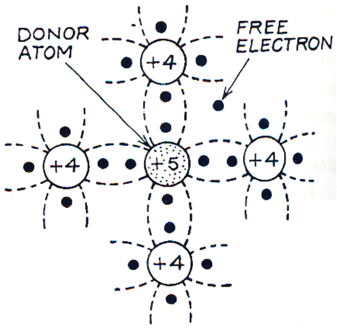
4. EXTRINSIC SEMICONDUCTORS :

The electrical conductivity of intrinsic (pure) semiconductor is too small to be of any practical use. If, however, a small quantity of some pentavalent or trivalent impurity is added to a pure semiconductor, the conductivity of the semiconductor is significantly increased. Such impure semiconductors are called 'extrinsic' or 'impurity' or 'doped semiconductors.

Doping : The process of adding impurity to an intrinsic semiconductor in a controlled manner is called 'doping'. It increases significantly the electrical conductivity of the semiconductor. The impurity atoms added are called 'dopants'.

Extrinsic semiconductor are of two types : n–type and p–type.

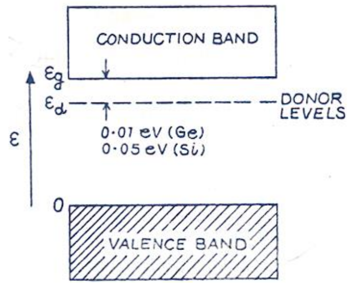
(a) n–type semiconductor : When a pentavalent impurity atom (antimony, phosphorus or arsenic) is added to a Ge(or Si) crystal, it replaces a Ge (or Si) atom in the crystal lattice. Four of the five valence electrons of the impurity atom form covalent bonds with one with each valence electron of four Ge (or Si) atoms surrounding.



Thus, by adding pentavalent impurity to pure Ge(or Si), the number of free electrons increases, that is, the conductivity of the crystal increases. The impure Ge (or Si) crystal is called an 'n–type' semiconductor because it has an excess of 'negative' charge–carrier (electrons). The impurity atoms are called 'donor' atoms because they donate the conducting electrons to the crystal.

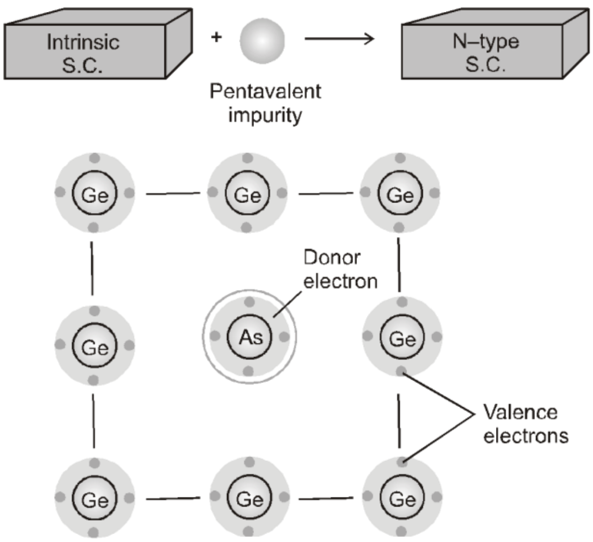
The fifth valence electrons of the impurity atoms occupy some discrete energy levels just below the condition band fig. These are called 'donor levels' and are only 0.01 eV below the conduction band in case of Ge, and 0.05 eV below in case of Si. Therefore, at room temperature, the "fifth" electrons of almost all the donor atoms are thermally excited from the donor levels into the conduction band where they move as charge–carriers when an external electric field is applied.

At ordinary temperature, almost all the electrons in the conduction band come from the donor levels, only a few come from the valence band. Therefore, the main charge–carriers responsible for conduction are the electrons contributed by the donors. Since the excitation from the valence band is small, there are very few holes in this band. The current contribution of the holes is therefore small. Thus, in an n–type semiconductor the electrons are the 'majority carriers' and the holes are the 'minority carriers.'



Note : N-Type Semiconductor

These are obtained by adding a small amount of pentavalent impurity to a pure sample of semiconductor (Ge).



(1) Majority charge carriers – electrons
Minority charge carriers – hole

(2) $n_e >> n_h$; $i_e >> i_h$

(3) Conductivity $\sigma = n_e \mu_e e$

(4) Donor energy level lies just below the conduction band.

(5) **Electrons and hole concentration :** In a doped semiconductor, the electron concentration n_e and the hole concentration n_h are not equal (as they are in an instrinsic semiconductor). It can be shown that

$n_e \approx n_h = n_{i2}$

where n_i is the intrinsic concentration.

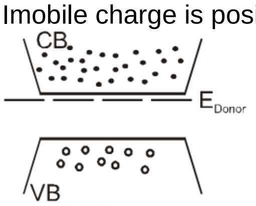
In an n–type semiconductor, the concentration of electrons in conduction band is nearly equal to the concentration of donor atoms (N_d) and very large compared to the concentration of holes in valence band. That is :

$n_e \;\; N_d >> n_h$

(6) Impurity atom called donar atom which is elements of V group of periodic table.

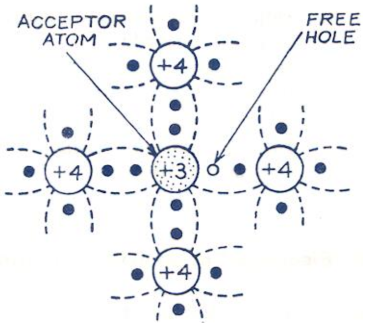
(7) Net charge on N type crystal is zero.

(8) Imobile charge is positive charge



(9) N-Type

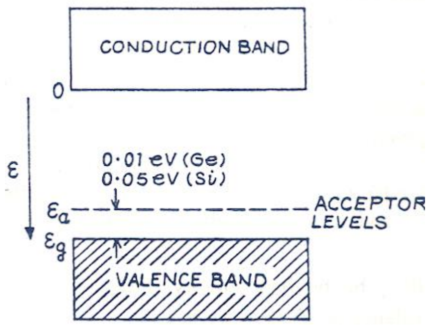
(b) p–type semiconductor : When a trivalent impurity atom (boron, aluminium, gallium or indium) is added to a Ge (or Si) crystal, it also replaces one of the Ge (or Si) atoms in the crystal lattice . Its three valence electrons form covalent bonds with one each valence electron of these Ge (or Si) atoms surrounding it. Thus, there remains an empty space, called a 'p–type' semiconductor because it has an excess of positive 'acceptor' atoms because they create holes which accept electrons.



The impurity atoms inductance vacant discrete levels just above the top of the valence band. These are called 'acceptor levels'. At room temperature, electrons are easily excited from the valence band into the acceptor levels. The corresponding holes created in the valence band are the main charge–carries in the crystal when an electric field is applied.

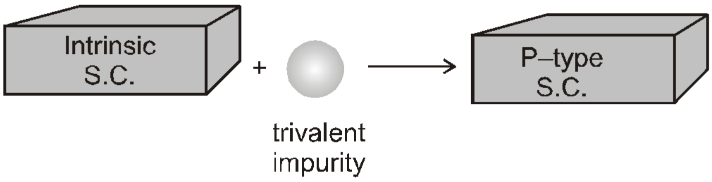
Thus, in a p–type semiconductor the holes are the 'majority carriers' and the few electrons, thermally excited from the valence band into the conduction band, are 'minority carriers'.

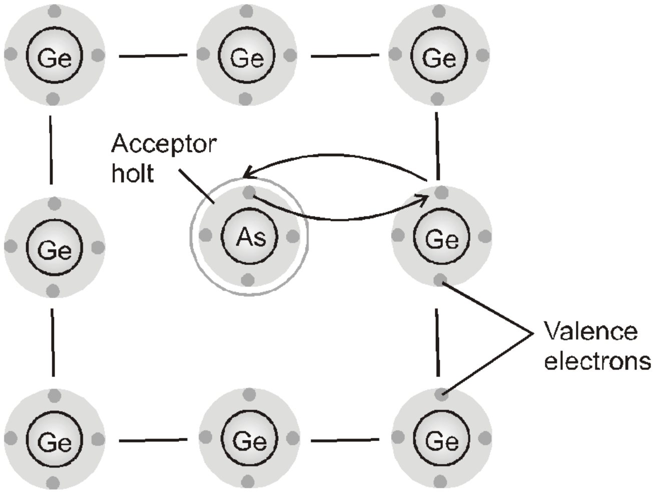
Electron and hole concentration :



Note: P-Type Semiconductor

There are obtained by adding a small amount of trivalent impurity to a pure sample of semiconductor (Ge).





(1) Majority charge carries – holes
Minority charge carries – electrons

(2) $n_h >> n_e$; $i_h >> i_e$

(3) Conductivity $\sigma \approx n_h \mu_h e$

(4) P–type semiconductor is also electrically neutral (not positively charged)

(5) Impurity is called Acceptor impurity which is element of III group of the periodic table.

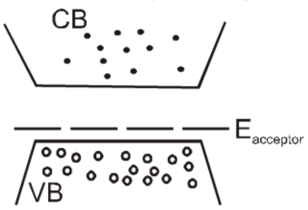
(6) Acceptor energy level lies just above the valency band.

(7) **Electron and hole concentration :** In a p–type semiconductor, the concentration of holes in valence band is nearly equal to the concentration of acceptor atoms (N_a) and very large compared to the concentration of electron in conduction band. That is

$n_h = N_a >> n_e$

(8) Net charge on p-type crystal is zero.

(9) Imobile charge is negative charge.



(10) P-Type

Distinction between intrinsic and extrinsic semiconductors :

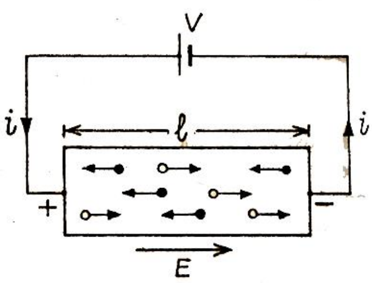
	Instrinsic Semiconductor		Extrinsic Semiconductor
1	It is a pure, natural semiconductor, such as pure Ge and pure Si.	1	It is prepared by adding a small quantity of impurity to a pure semiconductor, such as n- and p-type semiconductors.
2	In it the concentration of electrons and holes are equal.	2	In it the two concentrations are unequal. There is an excess of electrons in n-type semiconductors and an excess of holes in p-type semiconductors.
3	Its electrical conducitivity is very low.	3	Its electrical conductivity is siginificantly high.
4	Its conductivity cannot be controlled.	4	Its conductivity can be controlled by adjusting the quantity of the impurity added.
5	Its conductivity increases exponentially with temperature.	5	Its conductivity also increases with temperature, but not exponentially.

Distribution between n–type and p–type semiconductor :

	n-type semiconductor		p-type semiconductor
1	It is an extrinsic semiconductor obtained by adding a pentavalent impurity to a pure intrinsic semiconductor.	1	It is also an extrinsic semiconductor obtained by adding a trivalent impurity to a pure intrinsic semiconductor.
2	The impurity atoms added provides extra free electrons to the crystal lattice and are called donor atoms.	2	The impurity atoms added create holes in the crystal lattice and are called acceptor atoms because the created holes accept electrons.
3	The electrons are majority carriers and the holes are minority carriers.	3	The holes are majority carriers and the electrons are minority carriers.
4	The electrons concentration is much more than the hole concentration ($n_e >> n_h$).	4	The hole concentration is much more than the electron concentration ($n_h >> n_e$).

(c) Electrical conductivity of extrinsic semiconductors :

A semiconductor, at room temperature, contains electrons in the conduction band and holes in the valence band. When an external electric field is applied, the electrons move opposite to the field and the hole move in the direction of the field, thus constituting current in the same direction. The total current is the sum of the electron and hole currents.



Let us consider semiconductor block of length l, area of cross–section A and having electrons concentration n_e and hole concentration n_h . A potential difference V applied across the ends of the semiconductor creates an electric field E given by :

$$E = V/l \qquad \text{...............(i)}$$

Under the field E, the electrons and the holes both drift in opposite directions and constitute currents i_e and i_h respectively in the direction of the field. The total current flowing through the semiconductor is,

$$i = i_e + i_h$$

If v_e , be the drift velocity of the electrons in the conduction band and v_h the drift velocity of the holes in the valence band, then we have

$$i_e = n_e eA\, v_e \qquad \text{and} \qquad i_h = n_h e\, A\, v_h$$

where e is the magnitude of electron charge

$$\therefore \quad i = i_e + i_h = eA(n_e v_e + n_h v_h)$$

or $$\frac{i}{A} = e(n_e\, v_e + n_h\, v_h) \qquad \text{...............(ii)}$$

Let R be the resistance of the semiconductor block and ρ the resistivity of the block material. Then

$$\rho = R\,A / l \qquad \text{...............(iii)}$$

Dividing eq.(i) by eq.(ii) we have

$$\frac{E}{\rho} = \frac{V}{RA} = \frac{i}{A},$$

Because, V = iR (Ohm's law). Substituting in it the value of i/A from eq.(ii), we get

$$\frac{E}{\rho} = e(n_e v_e + n_h\, v_h)$$

or $$\frac{1}{\rho} = e\left(n_e \frac{v_e}{E} + n_h \frac{v_h}{E}\right) \qquad \text{...............(iv)}$$

Let us introduce a quantity μ, called mobility which is defined as the drift velocity per unit field and is expressed in $metre^2$ / (volt/second). Thus, the mobilities of electrons and hole are given by :

$$\mu_e = \frac{v_e}{E} \qquad \text{and } \mu_h = \frac{v_h}{E}$$

Introducing μ_e and μ_h in eq. (iv), we get

$$\frac{1}{\rho} = e(n_e\mu_e + n_h\mu_h)$$

The electrical conductivity ρ is the reciprocal of the resitivity ρ. Thus, the electrical conductivity of the semiconductor is given by

$$\rho = e(n_e\, \mu_e + n_h\, \mu_h)$$

This is the required expression. It shows that the electrical conductivity of a semiconductor depends upon the electron and hole concentrations (number densities) and their mobilities. The mobility of electrons is higher than the hole mobility.

As temperature rises, both the concentration n_e and n_h increases due to breakage of more convalent bonds. The mobilities μ_e and μ_h , however, slightly decrease with rise in temperature but this decrease is offset by the much greater increase in n_e and n_h. Hence, the conductivity of a semiconductor increases (or the resistivity decreases) with rise in temperature.

Solved Example

Example 1. The majority charge carriers in P-type semiconductor are

(1) Electrons (2) Protons (3) Holes (4) Neutrons

Solution : (3) In P-type semiconductors, holes are the majority charge carriers

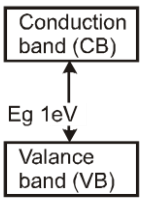
Example 2. When a semiconductor is heated, its resistance

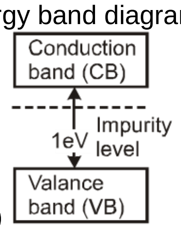
(1) Decreases (2) Increases

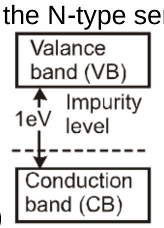
(3) Remains unchanged (4) Nothing is definite

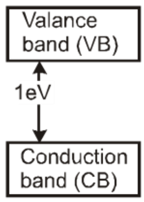
Ans. (1)

Example 3. Which of the following energy band diagram shows the N-type semiconductor

(1) 

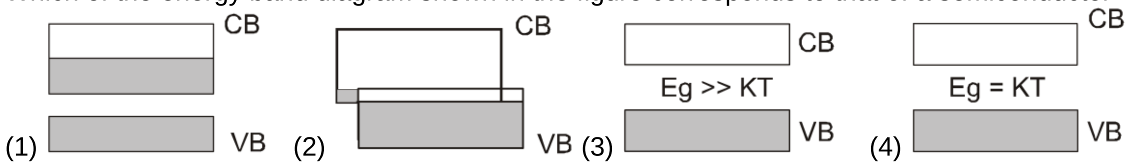
(2) 

(3) 

(4) 

Solution : (2) In N-type semiconductor impurity energy level lies just below the conduction band.

Example 4. Which of the energy band diagram shown in the figure corresponds to that of a semiconductor



Solution : (4) In semiconductors, the forbidden energy gap between the valence band and conduction band is very small, almost equal to kT. Moreover, valence band is completely filled where as conduction band is empty.

Example 5. The P-N junction is-

(1) an ohmic resistance (2) an non ohmic resistance

(3) a positive resistance (4) a negative resistance

Ans. (2)

Example 6. The mean free path of conduction electrons in copper is about 4×10^{-8} m. For a copper block, find the electric field which can give, on an average, 1eV energy to a conduction electron.

Solution : Let the electric field be E. The force on an electron is eE. As the electron moves through a distance d, the work done on it is eEd. This is equal to the energy transferred to the electron. As the electron travels an average distance of 4×10^{-8} m before a collision, the energy transferred is $eE(4 \times 10^{-8}$ m). To get 1 eV energy from the electric field,

$eE(4 \times 10^{-8}\text{ m}) = 1\text{ eV}$ or $E = 2.5 \times 10^{7}$ V/m.

Example 7. The band gap in germanium is $\Delta E = 0.68$ eV. Assuming that the number of hole–electron pairs is proportional to $e^{-\Delta E/2kT}$, find the percentage increase in the number of charge carries in pure germanium as the temperature is increased from 300 K to 320 K.

Solution : The number of charge carries in an intrinsic semiconductor is double the number of hole–electron pairs. If N_1 be the number of charge carries at temperature T_1 and N_2 at T_2, we have

$$N_1 = N_0 e^{-\Delta E/2kT_1}$$

and $$N_2 = N_0 e^{-\Delta E/2kT_2}$$

The percentage increase as the temperature is raised from T_1 to T_2 is

$$f = \frac{N_2 - N_1}{N_1} \times 100 = \left(\frac{N_2}{N_1} - 1\right) \times 100 = 100\left[e^{\frac{\Delta E}{2k}\left(\frac{1}{T_1} - \frac{1}{T_2}\right)} - 1\right]$$

Now $$\frac{\Delta E}{2k}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) = \frac{0.68\text{eV}}{2 \times 8.62 \times 10^{-5}\text{ eV/K}}\left(\frac{1}{300\text{K}} - \frac{1}{320\text{K}}\right) = 0.82$$

Thus $f = 100 \times [e^{0.82} - 1] \approx 127$.

Thus, the number of charge carries increase by about 127%.

Example 8. A silicon specimen is made into a p-type semiconductor by doping on an average one indium atom per 5×10^7 silicon atoms. If the number density of atoms in the silicon specimen is 5×10^{28} atoms/m^3; find the number of acceptor atoms in silicon per cubic centimeter.

Solution : The doping of one indium atoms in silicon semiconductor will produce one acceptor atom in p-type semiconductor. Since one indium atom has been dopped per 5×10^7 silicon atoms, so

number density of acceptor atoms in silicon = $\frac{5 \times 10^{28}}{5 \times 10^7} = 10^{21}$ atom/m^3 = 10^{15} atoms/cm^3.

Example 9. Pure Si at 300K has equal electron (n_e) and hole (n_h) concentrations of 1.5×10^{16} m^{-3}. Dopping by indium increases n_h to 3×10^{22} m^{-3}. Calculate n_e in the doped Si.

Solution : For a doped semi-conductor in thermal equilibrium $n_e n_h = n_i^2$ (Law of mass action)

$$n_e = \frac{n_i^2}{h_h} = \frac{(1.5 \times 10^{16})^2}{3 \times 10^{22}} = 7.5 \times 10^{9}\text{ m}^{-3}$$

Example 10. Pure Si at 300 K has equal electron (n_e) and hole (n_h) concentrations of 1.5×10^{16} m^{-3}. Doping by indium increases n_h to 4.5×10^{22} m^{-3}. Calculate n_e in the doped Si-

Solution : $n_e n_h = n_i^2$

$n_h = 4.5 \times 10^{22}$ m^{-3}

so, $n_e = 5.0 \times 10^9$ m^{-3}

Example 11. The energy of a photon of sodium light (λ = 589 nm) equals the band gap of a semiconducting material. (a) Find the minimum energy E required to create a hole-electron pair. (B) Find the value of E/kT at a temperature of 300 K.

Solution : (a) The energy of the photon is $E = \frac{hc}{\lambda}$

$= \frac{1242 \text{ eV} - \text{nm}}{589 \text{ nm}} = 2.1$ eV.

Thus the band gap is 2.1 eV. This is also the minimum energy E required to push an electron from the valence band into the conduction band. Hence, the minimum energy required to create a hole–electron pair is 2.1 eV.

(B) At T = 300 K,

$kT = (8.62 \times 10^{-5}$ eV/K$)$ (300 K)

$= 25.86 \times 10^{-3}$ eV.

Thus, $\frac{E}{kT} = \frac{2.1 \text{ eV}}{25.86 \times 10^{-3} \text{ eV}} = 81.$

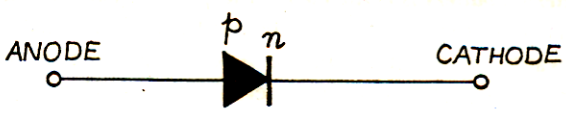
So it is difficult for the thermal energy to create the hole–electron pair but a photon of light can do it easily.

5. Junction Diode

A junction diode is a basic semiconductor device. It is a semiconductor crystal having acceptor impurities in one region (P – type crystal) and donor impurities in the other region (n–type crystal). The boundary between the two regions is called 'p–n junction'.

Circuit Symbol for a p-n Junction Diode:

In electronic circuits, the semiconductor devices are represented by their symbols. The symbol for the basic device, the p-n junction diode, is shown in Fig. The arrow-head represents the p -region and the bar represents the n -region of the diode. The direction of the arrow is from p to nand indicates the direction of conventional current flow under forward bias. The p -side is called 'anode' and the n -side is called 'cathode'.

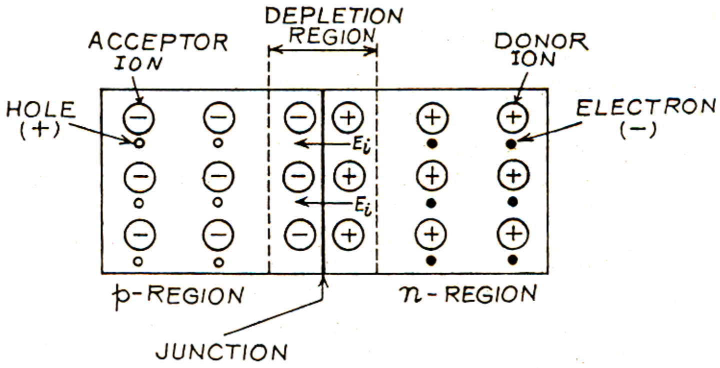


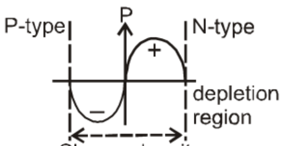
(a) Formation of p-n Junction :

A p-n junction is not the interface between p -type and n – type semiconductor crystals pressed together. It is a single piece of semiconductor crystal having an excess of acceptor impurities into one side and of donor impurities into the other. P-type semiconductor is grown on one side of metallic film while n-type is grown on other side.

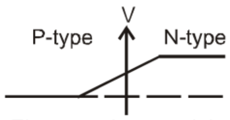
(b) Potential Barrier at the Junction: Formation of Depletion Region:

A p-n junction is shown in Fig. The p -type region has (positive) holes as majority charge-carriers, and an equal number of fixed negatively-charged acceptor ions. (The material as a whole is thus neutral). Similarly, the n -type region has (negative) electrons as majority charge-carriers, and an equal number of fixed positively-charged donor ions.

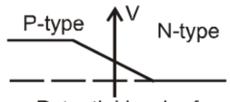




Charge density



Electrostatic potential energy barrier for holes (potential barrier for holes)



Potential barrier for electrons

The region on either of the junction which becomes depleted (free) of the mobile charge-carriers is called the 'depletion region'. The width of the depletion region is of the order of 10^{-6} m. The potential difference developed across the depletion region is called the 'potential barrier'. It is about 0.3 volt for Ge, p-n junction and about 0·7 volt for silicon p-n junction. It, however, depends upon the dopant concentration in the semiconductor.
The magnitude of the barrier electric field for a silicon junction is

$$E_i \approx \frac{V}{d} \approx \frac{0.7}{10^{-6}} = 7 \times 10^5 \text{ Vm}^{-1}$$

Diffusion & Drift Current : Due to concentration difference hole try to diffuse from p side to n side but due to depletion layer only those hole are able to diffuse from p to n side which have high kinetic energy. Similarly electron of high kinetic energy also diffuse from n to p so diffusion current flow from p to n side. Due to thermal collision or increase in temperature some valence electron comes in conduction band. If these occurs in depletion reign then hole move to p side & electron move to n side so a current produce from n to p side it is called drift current in steady state both are equal & opposite.

Solved Example

Example 12. In a p-n junction with open ends,
(1) there is no systematic motion of charge carriers
(2) holes and conductor electrons systematically go from the p-side and from the n-side to the p-side respectively
(3) there is no net charge transfer between the two sides
(4) there is a constant electric field near the junction

Ans. (2,3,4)

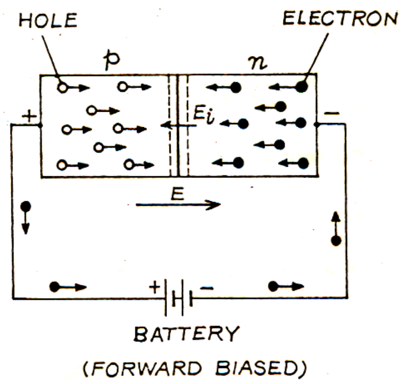
Example 13. A potential barrier of 0.50 V exists across a P-N junction. If the depletion region is 5.0×10^{-7} m wide, the intensity of the electric field in this region is
(1) 1.0×10^6 V/m (2) 1.0×10^5 V/m (3) 2.0×10^5 V/m (4) 2.0×10^6 V/m

Solution : (1) $E = \frac{V}{d} = \frac{0.5}{5 \times 10^{-7}} = 10^6$ V/m

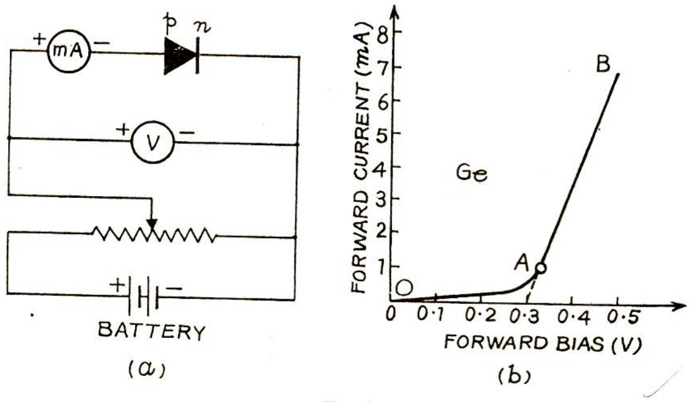
(c) Forward and Reverse Biasing of Junction Diode

The junction diode can be connected to an external battery in two ways, called 'forward biasing' and 'reverse biasing' of the diode. It means the way of connecting emf source to P-N junction diode. It is of following two types

(i) Forward Biasing :
A junction diode is said to be forward-biased when the positive terminal of the external battery is conncted to the p -region and the negative tenninal to the n -region of the diode.



Forward-Biased Characteristics : The circuit connections are shown in Fig. The positive terminal of the battery is connected to the p -region and the negative terminal to the n -region of the



junction diode through a potential-divider arrangement which enables to change the applied voltage. The voltage is read by a voltmeter V and the current by a milliammeter mA. Starting with a low value, the forward bias voltage is increased step by step and the corresponding forward current is noted. A graph is then plotted between voltage and current. The resulting curve OAB (Fig. b) is the forward characteristic of the diode.

In the beginning, when the applied voltage is low, the current through the junction diode is almost zero. It is because of the potential barrier (about 0·3 V for Ge p-n junction and about 0·7 V for Si junction) which opposes the applied voltage. With increase in applied voltage, the current increases very slowly and non-linearly until the applied voltage exceeds the potential barrier. This is represented by the portion OA of the characteristic curve. With further increase in applied voltage, the current increases very rapidly and almost linearly Now the diode behaves as an ordinary conductor. This is represented by the straight-line part AB of the characteristic. If this straight line is projected back. it intersects the voltage-axis at the barrier potential voltage.

Note :

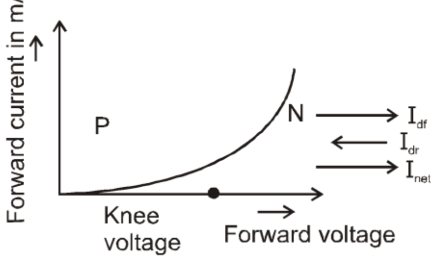
(i) In forward biasing width of depletion layer decreases

(ii) In forward biasing resistance offered $R_{Forward} \approx 10\Omega - 25\Omega$

(iii) Forward bias opposes the potential barrier and for $V > V_B$ a forward current is set up across the junction.

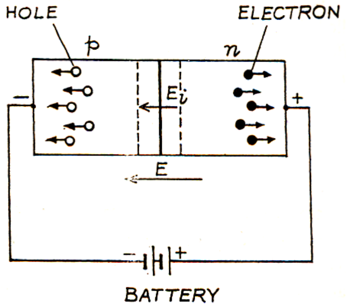
(iv) Cut-in (Knee) voltage : The voltage at which the current starts to increase rapidily. For Ge it is 0.3 V and for Si it is 0.7V.

(v) df–diffusion dr–drift



(ii) Reverse Biasing:

A junction diode is said to be reverse-biased when the positive terminal of the external battery is connected to the n -region and the negative terminal to the p -region of the diode (Fig.)

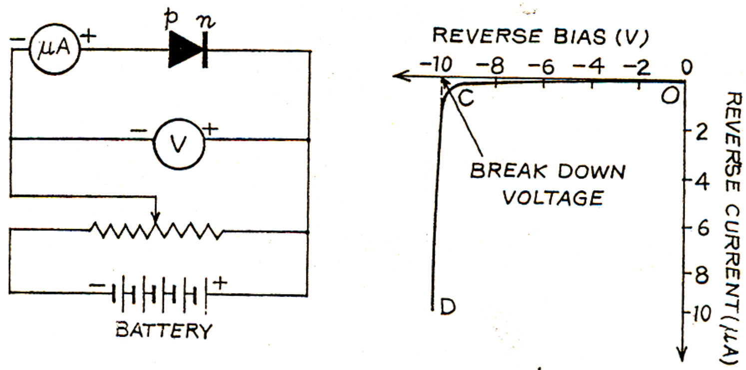


(REVERSE BIASED)

In this condition, the external field E is directed from n toward p and thus aids the internal barrier field E_i Hence holes in the p -region and electrons in the n -region are both pushed away from the junction, that is, they cannot combine at the junction. Thus, there is almost no current due to flow of majority carriers.

Reverse-Biased Characteristic:

The circuit connections are shown in Fig. (a) in which the positive terminal of the battery is connected to the n -region and the negative terminal to the p -region of the junction diode. In reverse-biased diode, a very small current (of the order of micro Ampere) flows across the junction due to the motion of the few thermally-generated minority-carriers (electrons in p -region and holes in n -region) whose motion is aided by the applied voltage. The small reverse current remains almost constant over a sufficiently long range of reverse bias (applied voltage)*. increasing very little with increasing bias. This is represented by the part OC of the reverse characteristic curve (Fig. b).



(a) (b)

Note:

(i) In reverse biasing width of depletion layer increases

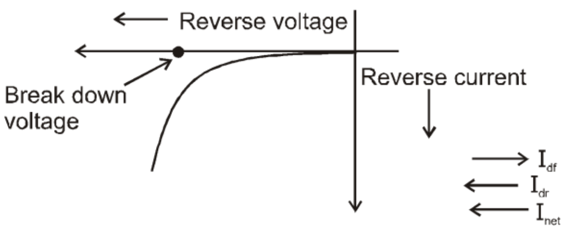
(ii) In reverse biasing resistance offered $R_{Reverse} \approx 10^5 \Omega$

(iii) Reverse bias supports the potential barrier and no current flows across the juction due to the diffusion of the majority carriers.

(A very small reverse currents may exist in the circuit due to the drifting of minority carriers across the juction)

(iv) Break down voltage : Reverse voltage at which break down of semiconductor occurs. For Ge it is 25V and for Si it is 35 V.

(v)



(vi) Reverse saturation current is temperature sensitive and nearely doubles for every 10ºC rise.

(d) Avalanche Breakdown :

If the reverse bias is made very high, the minority-carriers acquire kinetic energy enough to break the covalent bonds near the junction, thus liberating electron-hole pairs. These charge-carriers are accelerated and produce, in the same way, other electron-hole pairs. The process is cumulative and an avalanche of electron-hole pairs is produced. The reverse current then increases abruptly to a relatively large value (part CD of the characteristic). This is known as 'avalanche breakdown' and may damage the junction by the excessive heat generated. The reverse bias voltage at which the reverse current increase abruptly is called the 'breakdown voltage' or 'Zener voltage'. The numerical value of the breakdown voltage varies from tens of volts to several hundred volts depending on the number density of the impurity atoms doped into the diode.

(e) Dynamic Resistance of a Junction Diode

The current-voltage curve of junction diode shows that the current does not vary linearly with the voltage, that is, Ohm's law is not obeyed. In such situation, a quantity known as 'dynamic resistance' (or a.c. resistance) is defined.

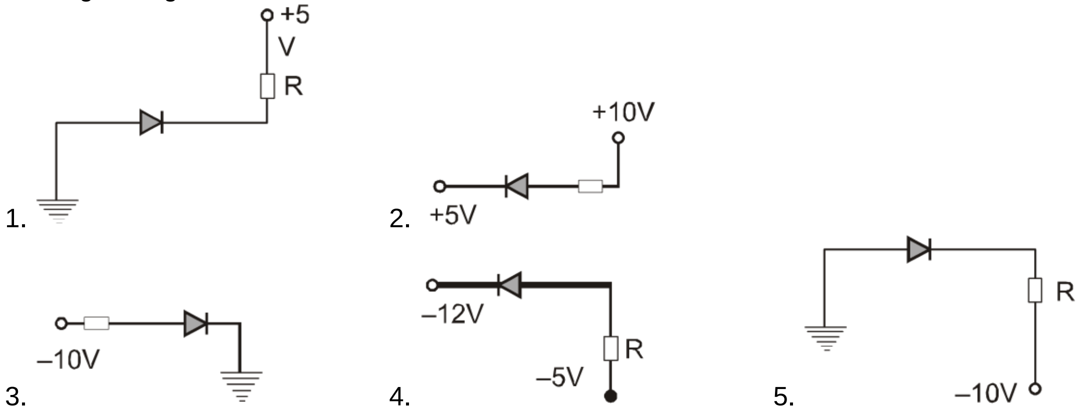
The dynamic resistance of ajunction diode is defined as the ratio of a small change in applied voltage (ΔV) to the corresponding small change in current (Δi) , that is

$$R_d = \frac{\Delta V}{\Delta i}$$

In the forward characteristic of p-n junction diode, beyond the turning point (knee), however, the current varies almost linearly with voltage. In this region, R_d is almost independent of V and Ohm's law is obeyed.

Solved Example

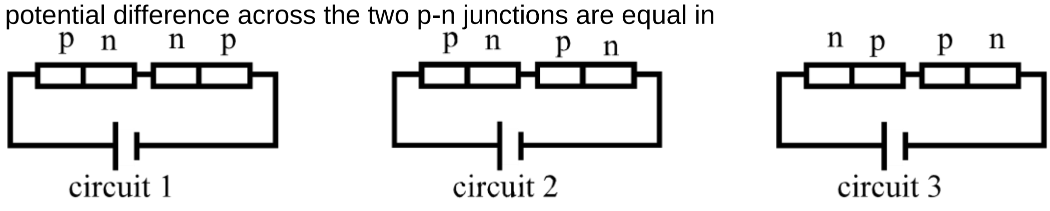
Example 14. In the given figure, which of the diodes are forward biased ?



(1) 1, 2, 3 (2) 2, 4, 5 (3) 1, 3, 4 (4) 2, 3, 4

Solution : (2) In figure 2,4 and 5. P-crystals are more positive as compared to N-crystals.

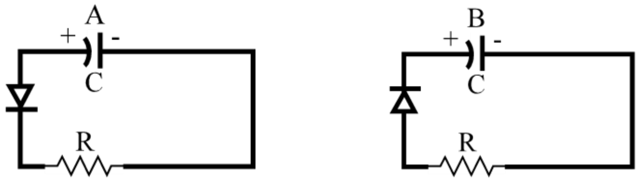
Example 15. Two identical p-n junction may be connected in serices with a battery in three ways fig. The potential difference across the two p-n junctions are equal in



(1) circuit 1 and circuit 2 (2) circuit 2 and circuit 3
(3) circuit 3 and circuit 1 (4) circuit 1 only

Ans. **(2)**

Example 16. Two identical capacitors A and B are charged to the same potential V and are connected in two circuits at t = 0 as shown in fig. The charges on the capacitor at a time t = CR are, respectively,

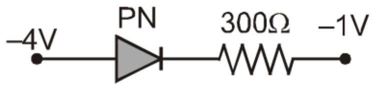


(A) (B)

(1) VC, VC (2) VC/e, VC (3) VC, VC/e (4) VC/e, VC/e

Ans. 2

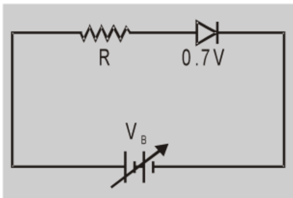
Example 17. What is the current in the circuit shown below



(1) 0 amp (2) 10^{-2} amp (3) 1 amp (4) 0.10 amp

Solution : (1) The potential fof P-side is more negative that of N-side, hence diode is in reverse biasing. In reverse biasing it acts as open circuit, hence no current flows.

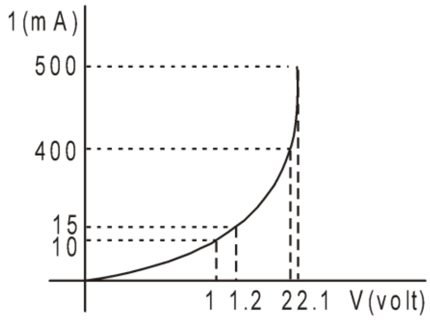
Example 18. Assume that the junction diode in the following circuit requires a minimum current of 1 mA to be above the knee point (0.7V) of its I-V characteristic curve. Also assume that the voltage across the diode is independent of current above the knee point. If V_B = 5V, what should be the maximum value of R so that the voltage is above the knee point-



(1) 4.3 kΩ (2) 860 kΩ (3) 4.3 Ω (4) 860 Ω

Ans. (1)

Example 19. The i-V characteristic of a p-n junction diode is shown in figure. Find the approximate dynamic resistance of the p-n junction when (a) a forward bias of 1 volt is applied, (b) a forward bias of 2 volt is applied



(a) The current at 1 volt is 10 mA and at 1.2 volt it is 15 mA. The dynamic resistance in this region is

$$R = \frac{\Delta V}{\Delta i} = \frac{0.2\,\text{volt}}{5\,\text{mA}} = 40\ \Omega$$

(b) The current at 2 volt is 400 mA and at 2.1 volt it is 800 mA. The dynamic resistance in the region is

$$R = \frac{\Delta V}{\Delta i} = \frac{0.1\,\text{volt}}{400\,\text{mA}} = 0.25\ \Omega .$$

6. p-n Junction Diode as a Rectifier

An electronic device which converts alternating current / voltage into direct current/voltage is called 'rectifier'.

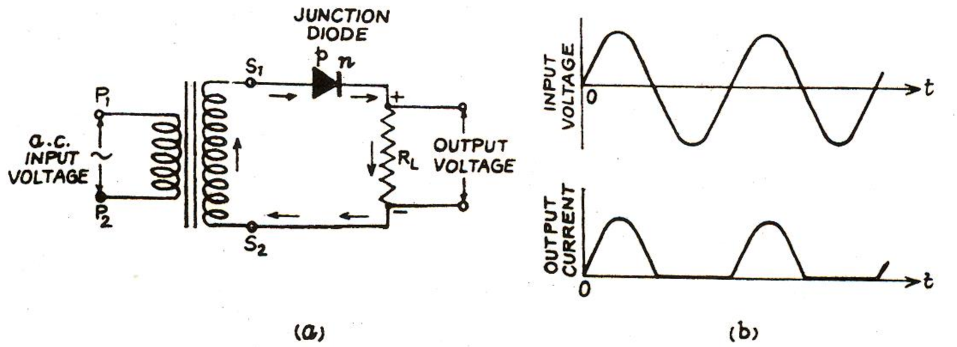
A p-n junction diode offers a low resistance for the current to flow, when forward-biased, but a very high resistance, when reverse-biased. It thus passes current only in one direction and acts as a rectifier.

The junction diode can be used either as an half-wave rectifier, when it allows current only during the positive half-cycles of the input a.c. supply; or as a full-wave rectifier when it allows current in the same direction for both half-cycles of the input a.c.

(a) p-n Junction Diode as Half·wave Rectifier:

The half-wave rectifier circuit is shown in Fig. (a) and the input and output wave forms in Fig. (b). The a.c. input voltage is applied across the primary P_1P_2 of a transformer. S_1S_2 is the secondary coil of the same transformer. S_1 is connected to the p -type crystal of the junction diode and S_2 is connected to the n -type crystal through a load resistance R_L.

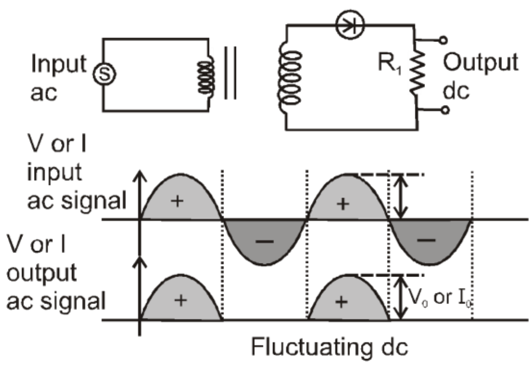
During the first half-cycle of the a.c. input, when the terminal S_1 of the secondary is suppose positive and S_2 is negative, the junction diode is forward-biased. Hence it conducts and current flows through the load R_L in the direction shown by arrows. The current produces across the load an output voltage of the same shape as the half-cycle of the input voltage. During the second half-cycle of the a.c. input, the terminal S_1 is negative and S_2 is positive. The diode is now reverse-biased. Hence there is almost zero current and zero output voltage across R_L. The process is repeated. Thus, the output current is unidirectional, but intermittantand pulsating, as shown in lower part of Fig. (b).



Since the output- current corresponds to one half of the input voltage wave, the other half being missing, the process is called half-wave rectification.

The purpose of the transformer is to supply the necessary voltage to the rectifier. If direct current at high voltage is to be obtained from the rectifier, as is necessary for power supply, then a step-up transformer is used, as shown in Fig. (a). In many solid-state equipments, however, direct current of low voltage is required. In that case, a step-down transformer is used in the rectifier.

Note:



(i) During positive half cycle
Diode –→ forward biased
Output signal –→ obtained

(ii) During negative half cycle
Diode –→ reverse biased
Output signal –→ not obtained

(iii) Output voltage is obtained across the load resistance R_L. It is not constant but pulsating (mixture of ac and dc) in nature.

(iv) Average output in one cyle

$I_{dc} = \frac{I_0}{\pi}$ and $V_{dc} = \frac{V_0}{\pi}$; $I_0 = \frac{V_0}{r_f + R_L}$

(r_f = forward biased resistance)

(v) r.m.s. output : $I_{rms} = \frac{I_0}{2}$, $V_{rms} = \frac{V_0}{2}$

(vi) The ratio of the effective alternating component of the output voltage or current o the dc component is known as ripple factor.

$r = \frac{I_{ac}}{I_{dc}} = \left[\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1\right]^{1/2} = 1.21$

(vii) Peak inverse voltage (PIV) : The maximum reverse biased voltage that can be applied before commoncement of Zener region is called the PIV. When diode is not conducting PIV across it = V_0

(viii) Efficiency : It is given by % $\eta = \frac{P_{out}}{P_{in}} \times 100 = \frac{40.6}{1 + \frac{r_f}{R_L}}$

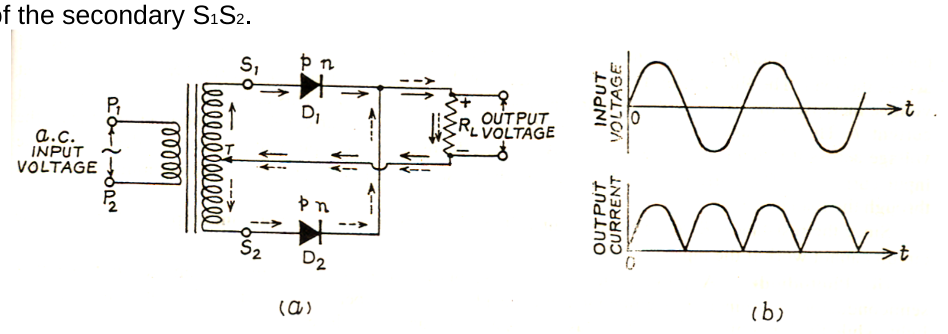
If $R_L >> r_f$ then η = 40.6 % If $R_L = r_f$ then η = 20.3 %

(ix) Form factor = $\frac{I_{rms}}{I_{dc}} = \frac{\pi}{2} = 1.57$

(x) The ripple frequency (ω) for half wave rectifier is same as that of ac.

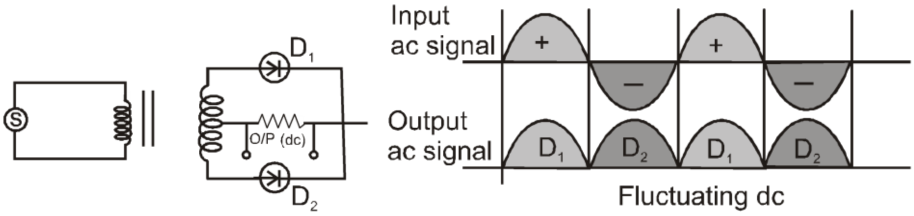
(b) p-n Junction Diode as Full-wave Rectifier: In a full-wave rectifier, a unidirectional, pulsating output current is obtained for both halves of the a.c. input voltage. Essentially, it requires two junction diodes so connected that one diode rectifies one half and the second diode rectifies the second half of the input.

The circuit for a full-wave rectifier is shown in Fig. 8 (a) and the input and output wave forms in Fig. (b). The a.c. input voltage is applied across the primary P_1P_2 of a transformer. The terminals S_1 and S_2 of the secondary are connected to the p -type crystals of the junction diodes D_1 and D_2 whose n -type crystals are connected to each other. A load resistance R_L is connected across the n -type crystals and the central-tap T of the secondary S_1S_2.



During the first half-cycle of the a.c. input voltage, the terminal S_1 is suppose positive relative to T and S_2 is negative. In this situation, the junction diode D_1 is forward-biased and D_2 is reverse-biased. Therefore, D_1 conducts while D_2 does not. The conventional current flows through diode D_1 , load R_L and the upper half of the secondary winding, as shown by soiid arrows. During the second half-cycle of the input voltage, S_1 is negative relative to T and S_2 is positive. Now, D_1 is reverse-biased and does not conduct while D_2 is forward-biased and conducts. The current now flows through D_2, load R_L and the lower half of the secondary, as shown by dotted arrows. It may be seen that the current in the load R_L flows in the same direction for both half-cycles of the a.c. input voltage. Thus, the output current is a continuous series of unidirectional pulses. However, it can be made fairly steady by means of smoothing filters.

Note :



(i) During positive half cycle
Diode : D_1 –→ forward biased
D_2 –→ reverse biased
Output signal –→ obtained due to D_1 only

(ii) During negative half cycle
Diode : D_1 –→ reversed biased
D_2 –→ forward biased
Output signal –→ obtained due to D_2 only

(iii) Fluctuating dc –→ Filter –→ constant dc.

(iv) Output voltage is obtained across the load resistance R_L. It is not constant but pulsating in nature.

(v) Average output : $V_{av} = \frac{2V_0}{\pi}$, $I_{av} = \frac{2I_0}{\pi}$

(vi) r.m.s. output = $V_{rms} = \frac{V_0}{\sqrt{2}}$, $I_{rms} = \frac{I_0}{\sqrt{2}}$

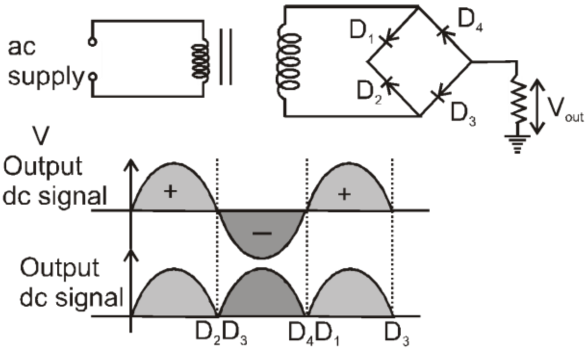
(vii) Ripple factor : r = 0.48 = 48%

(viii) Ripple frequency : The ripple frequency of full wave rectifier = 2 × (Frequency of input ac)

(ix) Peak inverse voltage (PIV) : It's value is $2V_0$.

(x) Efficiency : $\eta_{\%} = \frac{81.2}{1+\frac{r_1}{R_L}}$ for $r_f << R_L$, η = 81.2%

(3) Full wave bridge rectifier : Four diodes D_1, D_2, D_3 and D_4 are used in the circuit.
During positve half cycle D_1 and D_3 are forward biased and D_2 and D_4 are reverse biased. During negative half cycle D_2 and D_4 are forward bieased and D_1 and D_3 are reverse biased



(c) Different Types of Junction Diode

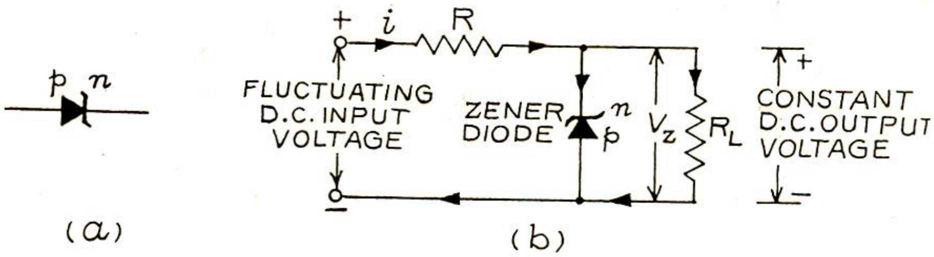
The junction diodes are of many types. The important types are Zener diode, photodiode, light-emitting diode (LED) and solar cell.

(i) Zener Diode: It is a voltage-regulating device based upon the phenomenon of avalanche breakdown in a junction diode.

When the reverse-bias applied to a junction diode is increased, there is an abrupt rise in the (reverse) current when the bias reaches a certain value, known as 'breakdown voltage' or 'Zener voltage' (Fig. 6b).

Thus, in this region of the reverse characteristic curve, the voltage across the diode remains almost constant for a large range of currents. Hence the diode may be used to stabilize voltage at a pre-determined value. It is then known as 'Zener diode'. It can be designed, by properly controlled doping of the diode, to stabilize voltage at any desired value between 4 –100 volt. .

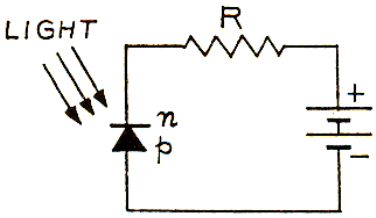
Fig. (a) shows the symbol of a Zener diode and Fig. (b) shows a simple circuit for stabilizing voltage across a load R_L. The circuit consists of a series voltage-dropping resistance R and a Zener diode in



parallel with the load R_L. The Zener diode is selected with a Zener voltage V_Z equal to the voltage desired across the load. The fluctuating d.c. input voltage may be the d.c. output of a rectifier. Whenever the input voltage increases, the excess voltage is dropped across the resistance R . This causes an increase in the input current i. This increase is conducted by the Zener diode, while the current through the load and hence the voltage across it remains constant at V_Z. Likewise, a decrease in the input voltage causes a decrease in the input current i. The current through the diode decreases correspondingly, again maintaining the current through the load constant.

Since the resistance R absorbs the input voltage fluctuations to give a constant output voltage V_Z, the circuit cannot work if the input voltage falls below V_Z. .

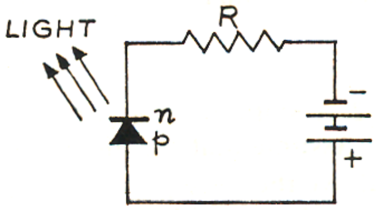
(ii) Photodiode : A photodiode is a reverse-biased p-n junction made from a photosensitive semiconductor. The junction is embedded in clear plastic. The upper surface across the junction is open to light, while the remaining sides of the plastic are painted black or enclosed in a metallic case. The entire unit is extremely small, of the order of a 0·1 inch size.



The circuit is shown in Fig. . When no light is falling on the junction and the reverse-bias is of the order of a few tenths of a volt, an almost constant small current (≈ μA) is obtained. This "dark" current is the reverse saturation current due to the thermally-generated minority-carriers (electrons in p -region and holes in n -region). When light of appropriate frequency is made incident on the junction, additional electron-hole pairs are created near the junction (due to breaking of covalent bonds). These light-generated minority-carriers cross the· (reverse-biased) junction and contribute to the (reverse) current due to thermally-generated carriers. Therefore, the current in the circuit increases (a fraction of a mA). This, so-called 'photoconductive' current varies almost linearly with the incident light flux.

The p--n photodiodes can operate at frequencies of the order of 1 MHz. Hence they are used in high-speed reading of computer punched cards, light-detection systems, light-operated switches, electronic counters etc.

(iii) Light-Emitting Diode (LED) : When a p--n junction diode is forward-biased. both the electron and the holes move towards the junction. As they cross the junction, the electrons fall into the holes (recombine). Hence, energy is released at the junction (because the electrons fall from a higher to a lower energy level). In case of Ge and Si diodes, the energy released is infra-red radiation. If, however, the diode is made of gallium arsenide or indium phosphide, the energy released is visible light. The diode is then called a 'light-emitting diode' (LED).



LEDs have replaced incandescent lamps in many applications because of their low input power, long life and fast on-off switching.

LEDs that can emit red, yellow, orange, green and blue light are commercially available. The semiconductor used for fabrication of visible LEDs must at least have a band gap of 1.8 eV (spectral range of visible light is from about 0.4 μm to 0.7 μm, i.e., from about 3 eV to 1.8 eV). The compound semiconductor Gallium Arsenide – Phosphide ($GaAs_{1-x}P_x$) is used for making LEDs of different colours. $GaAs_{0.6}P_{0.4}$ (E_g ~ 1.9 eV) is

used for red LED. GaAs (E_g ~ 1.4 eV) is used for making infrared LED. These LEDs find extensive use in remote controls, burglar alarm systems, optical communication, etc. Extensive research is being done for

developing white LEDs which can replace incandescent lamps. They are extensively used in fancy electronic devices like calculators, etc.

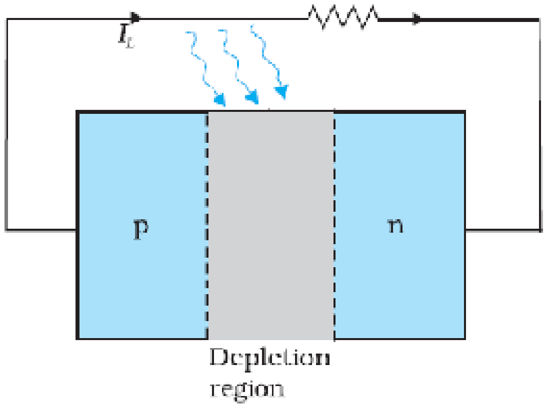
(iv) Solar Cell :

A solar cell is basically a p-n junction which generates emf when solar radiation falls on the p-n junction. It works on the same principle (photovoltaic effect) as the photodiode, except that no external bias is applied and the junction area is kept much larger for solar radiation to be

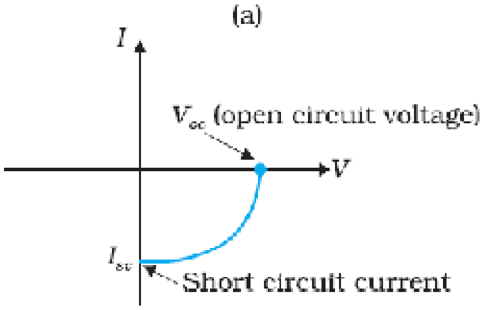
incident because we are interested in more power.

A p-Si wafer of about 300 μm is taken over which a thin layer (~0.3 μm) of n-Si is grown on one-side by diffusion process. The other side of p-Si is coated with a metal (back contact). On the top of n-Si layer, metal finger electrode (or metallic grid) is deposited. This acts as a front contact. The metallic grid occupies only a very small fraction of the cell area (<15%) so that light can be incident on the cell from the top. The generation of emf by a solar cell, when light falls on, it is due to the following three basic processes: generation, separation and collection—(i) generation of e-h pairs due to light (with $h\nu > E_g$) close to the junction ; (ii) separation of electrons and holes due to electric field of the depletion region. Electrons are swept to n-side and holes to p-side ;

(iii) the electrons reaching the n-side are collected by the front contact and holes reaching p-side are collected by the back contact. Thus p-side becomes positive and n-side becomes negative giving rise to photovoltage. When an external load is connected (a) a photocurrent I_L flows through the load. A typical I-V characteristics of a solar cell is shown. Note that the I – V characteristics of solar cell is drawn in the fourth quadrant of the coordinate axes. This is because a solar cell does not draw current but supplies the same to the load. Semiconductors with band gap close to 1.5 eV are ideal materials for solar cell fabrica tion. Solar cells are made with semiconductors like Si (E_g = 1.1 eV), GaAs (E_g = 1.43 eV), CdTe (E_g = 1.45 eV), CuInSe2 (E_g = 1.04 eV), etc. The important criteria for the selection of a material for solar cell fabrication are (i) band gap (~1.0 to 1.8 eV), (ii) high optical absorption (~10^4 cm^{-1}), (iii) electrical conductivity, (iv) availability of the raw material, and (v) cost. Note that sunlight is not always required for a solar cell. Any light with photon energies greater than the bandgap will do.



(a)



(b)

Solar cells are used to power electronic devices in satellites and space vehicles and also as power supply to some calculators. Production of low-cost photovoltaic cells for large-scale solar energy is a topic for research.

Solved Examples

Example 20. A zener diode of voltage V_Z (= 6) is used to maintain a constant voltage across a load resistance R_L (=1000Ω) by using a series resistance R_S (=100Ω). If the e.m.f. or source is E (= 9V), calculate the value of current through series resistance, Zener diode and load resistance. What is the power being dissipated in Zener diode.

Solution : Here, E = 9V; V_Z = 6; R_L = 1000Ω and R_s = 100Ω,

Potential drop across series resistor $V = E - V_Z = 9 - 6 = 3$ V

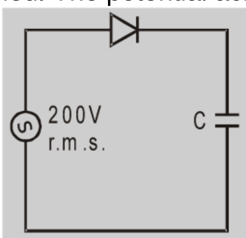
Current through series resistance R_S is $I = \frac{V}{R} = \frac{3}{100} = 0.03A$

Current through load resistance R_L is $I_L = \frac{V_Z}{R_L} = \frac{6}{1000} = 0.006A$

Current through Zener diode is $I_Z = I - I_L = 0.03 - 0.006 = 0.024$ A

Power dissipated in Zener diode is $P_Z = V_Z I_Z = 6 \times 0.024 = 0.144$ Watt

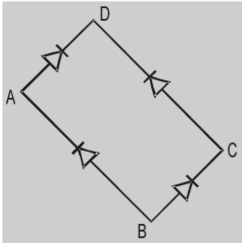
Example 21. In the figure, an A.C. of 200 rms voltage is applied to the circuit containing diode and the capacitor and it is being rectified. The potential across the capacitor C in volt will be-



(1) 500 (2) 200 (3) 283 (4) 141

Ans. (3)

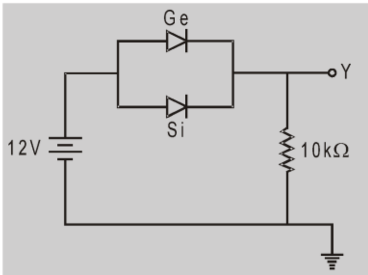
Example 22. In the figure, input is applied across A and C and output is taken across B and D, then the output is-



(1) Zero (2) Same as input (3) Full wave rectified (4) Half wave rectified

Ans. (3)

Example 23. Two junction diodes one of germanium (Ge) and other of silicon (Si) are connected as shown in figure to a battery of emf 12 V and a load resistance 10 kΩ. The germanium diode conducts at 0.3 V and silicon diode at 0.7 V. When a current flows in the circuit, the potential of terminal Y will be-



(1) 12 V (2) 11 V (3) 11.3 V (4) 11.7 V

Ans. (4)

Example 24. Potential barrier developed in a junction diode opposes-
(1) Minority carriers in both regions only (2) Majority carriers
(3) Electrons in N-region (4) Holes in P-region

Ans. (2)

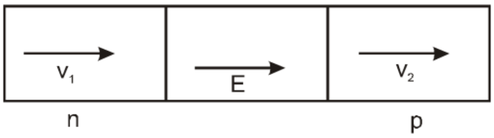
Example 25. Avalanche breakdown in a semiconductor diode occurs when-
(1) Forward current exceeds a certain value (2) Reverse bias exceeds a certain value
(3) Forward bias exceeds a certain value (4) The potential barrier is reduced to zero

Ans. (2)

Example 26. A potential barrier of 0.50 V exists across a p-n junction.
(a) If the depletion region is 5.0×10^{-7} m wide, what is the intensity of the electric field in this region ?
(b) An electron with speed 5.0×10^{5} m/s approaches the p-n junction from the n-side. With what speed will it enter the p-side ?

Solution : (a) The electric field is E = V/d $= \dfrac{0.50\,V}{5.0 \times 10^{-7}\,m} = 1.0 \times 10^{6}\,V/m.$

(b)



Suppose the electron has a speed υ_1 when it enters the depletion layer and υ_2 when it comes out of it (figure). As the potential energy increases by e × 0.50 V, from the principle of conservation of energy,

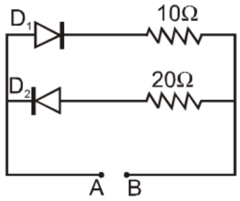
$$\frac{1}{2}m\upsilon_1^2 = e \times 0.50V + \frac{1}{2}m\upsilon_2^2$$

or, $\frac{1}{2} \times (9.1 \times 10^{-31}\,kg) \times (5.0 \times 10^{5}\,m/s)^2 = 1.6 \times 10^{-19} \times 0.5\,J + (9.1 \times \frac{1}{2}\,10^{-31}\,kg)\,\upsilon_2^2$

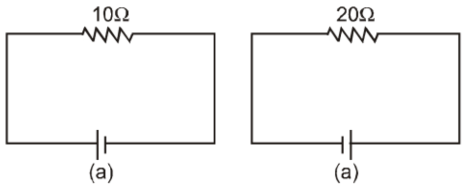
or, $1.13 \times 10^{-19}\,J = 0.8 \times 10^{-19}\,J + (4.55 \times 10^{-31} kg)\,\upsilon_2^2$.

Solving this, $\upsilon_2 = 2.7 \times 10^{5}$ m/s

Example 27. A 2 V battery may be connected across the points A and B as shown in figure. Assume that the resistance of each diode is zero in forward bias and infinity in reverse bias. Find the current supplied by the battery if the positive terminal of the battery is connected to (a) the point A (b) the point B.

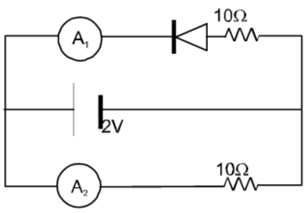


Solution : (a) When the positive terminal of the battery is connected to the point A, the diode D_1 is forward-biased and D_2 is reverse-biased. The resistance of the diode D_1 is zero, and it can be replaced by a resistance less wire. Similarly, the resistance of the diode D_2 is infinity, and it can be replaced by a broken wire. The equivalent circuit is shown in figure. The current supplied by the battery is 2 V/10 Ω = 0.2 A.



(b) When the positive terminal of the battery is connected to the point B, the diode D_2 is forward-biased and D_1 is reverse biased. The equivalent circuit is shown in figure. The through the battery is 2 V/20 Ω = 0.1 A.

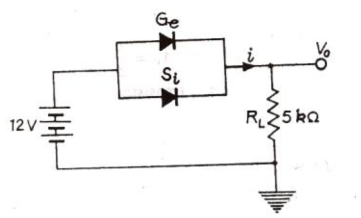
Example 28. What are the reading of the ammeters A_1 and A_2 shown in figure. Neglect the resistance of the meters.



Ans. zero 0.2 A

Example 29. Calculate the value of V_0 and if the Si diode and the Ge diode start conducting at 0.7 V and 0.3 V respectively, in the given circuit.
If the Ge diode connection be reversed, what will be the new values of V_0 and I ?

Solution : The effective forward voltage across Ge diode is 12 V – 0.3 = 11.7.
This will appear as the output voltage across the lad, that is, V_0 = 11.7 V
The current in the load is



$$i = \frac{V_0}{R_L} = \frac{11.7}{5K\Omega} = 2.34 \text{ mA.}$$

On reversing the connections of Ge diode, it will be reverse-biased and conduct no current. Only Si diode will conduct. The effective forward voltage across Si diode is 12 V – 0.7V = 11.3 V. This will appear as output, that is V_0 = 11.3 V

The current in the load is $i = \frac{V_0}{R_L} = \frac{11.3}{5k\Omega} = 2.26$ mA.

7. JUNCTION TRANSISTOR :

Transistor structure and action :

A transistor has three doped regions forming two p–n junctions between them. There are two types of transistors, as shown in figure.

(i) n–p–n transistor : Here two segments of n–type semiconductor (emitter and collector) are separated by a segment of p–type semiconductor (base).

(ii) p–n–p transistor : Here two segments of p–type semiconductor (termed as emitter and collector) are separated by a segment of n–type semiconductor (termed as base).

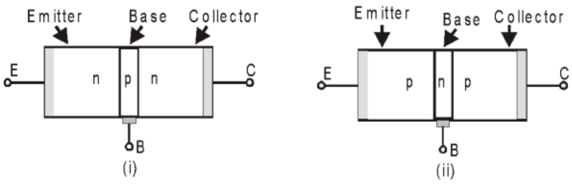
The schematic representations of an n–p–n and a p–n–p configuration are shown in figure. All the three segments of a transistor have different thickness and their doping levels are also different. In the schematic symbols used for representing p–n–p and n–p–n transistors (figure b) the arrowhead shows the direction of conventional current in the transistor. A brief description of the three segments of a transistors is given below:

Emitter : This is the segment on one side of the transistor shown in fig.(a). It is of moderate size and heavily doped. It supplies a large number of majority carriers for the current flow through the transistor.

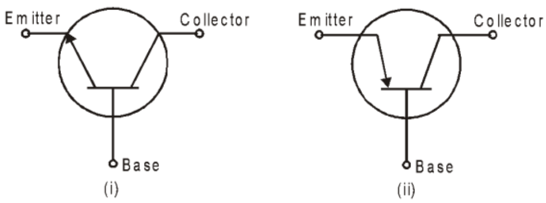
Base : This is the central segment. It is very thin and lightly doped.

Collector : This segment collects a major portion of the majority carries supplied by the emitter. The collector side is moderately doped and larger in size as compared to the emitter.

In case of a p–n junction, there is a formation of depletion region across the junction. In case of a transistor, depletion regions are formed at the emitter base–junction and the base collector junction.



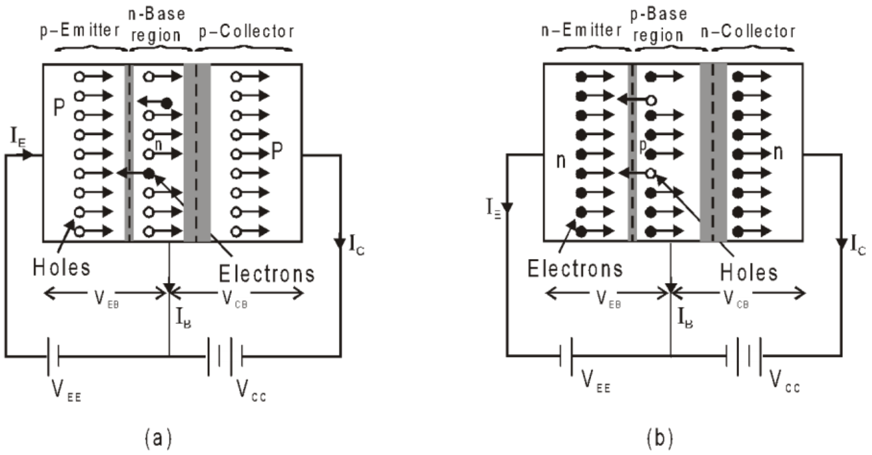
(a) Schematic representations of a n–p–n transistor and p–n–p transistor



(b) Symbols for n–p–n and p–n–p transistors.

The transistor works as an amplifier, with its emitter–base junction forward biased and the base–collector junction reverse biased. This situation is shown in figure, where V_{CC} and V_{EE} are used for creating the respective biasing. When the transistor is biased in this way it is said to be in active state. We represent the voltage between emitter and base as V_{EB} and that between the collector and base as V_{CB} . In figure, base is a common terminal for the two power supplies whose other terminals are connected to emitter and collector, respectively. So, the two power supplies are represented as $V_{EE'}$ and $V_{CC'}$ respectively. In circuits, where emitter is the common terminal, the power supply between the base and emitter is represented as V_{BB} and that between collector and emitter as $V_{CC'}$.

The heavily doped emitter has a high concentration of majority carriers, which will be holes in a p–n–p transistor and electrons in an n–p–n transistor. These majority carriers enter the base region in large numbers. The base is thin and lightly doped. So, the majority carriers there would be few. In a p–n–p transistor the majority carries in the base are electrons since base is of n–type semiconductor. The large number of holes entering the base from the emitter swamps the small number of electrons there. As the base collector–junction is reverse biased, these holes, which appear as minority carriers at the junction, can easily cross the junction and enter the collector. The holes in the base could move either towards the base terminal to combine with the electrons entering from outside or cross the junction to enter into the collector and reach the collector terminal. The base is made thin so that most of the holes find themselves near the reverse-biased base-collector junction and so cross the junction instead of moving to the base terminal.



Bias Voltage applied on : (a) p–n–p transistor and (b) n–p–n transistor

Note : Due to forward bias a large current enters the emitter–base junction, but most of it is diverted to adjacent reverse–biased base–collector junction and the current coming out of the base becomes a very small fraction of the current that entered the junction. If we represent the hole current and the electron current crossing the forward biased junction by the sum $I_h + I_e$. We see that the emitter current $I_E = I_h + I_e$ but the base current $I_B << I_h + I_e$, because a major part of I_E goes to collector instead of coming out of the base terminal. The base current is thus a small fraction of the emitter current.

It is obvious from the above description and also from a straight forward application of Kirchoff's law to figure(a) that the emitter current is the sum of collector current and base current :

$I_E = I_C + I_B$

We also see that $I_C \approx I_E$.

Our description of the direction of motion of the holes is identical with the direction of the conventional current. But the direction of motion of electrons is just opposite to that of the current. Thus in a p–n–p transistor the current enters from emitter into base whereas in a n–p–n transistor it enters from the base into the emitter. The arrowhead in the emitter shows the direction of the conventional current.

We can conclude that in the active state of the transistor the emitter–base junction acts as a low resistance while the base collector acts as a high resistance.

In a transistor, only three terminals are available viz emitter (E), base (B) and collector (C). Therefore in a circuit the input/output connections have to be such that one of these (E,B or C) is common to both the input and the output. Accordingly, the transistor can be connected in either of the following three configurations :

Common Emitter (CE), Common Base (CB), Common Collector (CC).

Working of Transistor

(1) There are four possible ways of biasing the two P-N junctions (emitter junction and collector junction) of transistor.

(i) Active mode : Also known as linear mode operation.

(ii) Saturation mode : Maximum collector current flows and transistor acts as a closed switch from collector to emitter terminals.

(iii) Cut-off mode : Denotes operation like an open switch where only leakage current flows.

(iv) Inverse mode : The emitter and collector are inter changed.

Different modes of operation of a transistor

Operating mode	**Emitter base bias**	**Collector base bias**
Active	Forward	Reverse
Saturation	Forward	Forward
Cut off	Reverse	Reverse
Inverse	Reverse	Forward

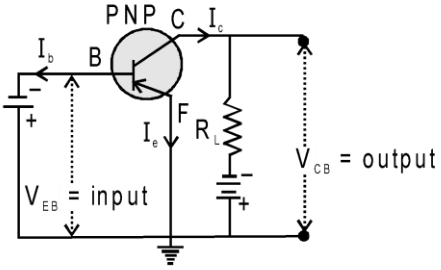
(2) A transistor is mostly used in the active region of operation i.e., emitter base junction is forward biased and collector base junction is reverse biased.

(3) From the operation of junction transistor it is found that when the current in emitter circuit changes. There is corresponding change in collector current.

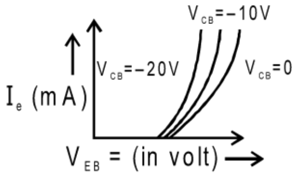
(4) In each state of the transistor there is an input port and an output port. In general each electrical quantity (V or I) obtained at the output is controlled by the input.

Transistor Configurations

A transistor can be connected in a circuit in the following three different configurations.
Common base (CB), Common emitter (CE) and Common collector (CC) configuration.
(1) CB configurations : Base is common to both emitter and collector.



(i) Input current = I_e (ii) Input voltage = V_{EB} (iii) Output voltage = V_{CB} (iv) Output current = I_C
With small increase in emitter-base voltage V_{EB}, the emitter current I_e increases rapidly due to small input resistance.
(v) **Input characteristics :** If V_{CB} = constant, curve between I_e and V_{EB} is known as input charactersistics. It is also known as emitter charactersistics :

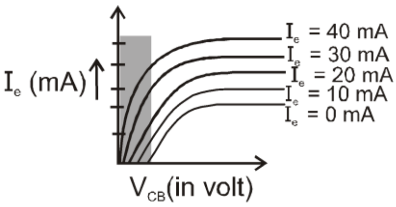


Input characteristics of NPN transistor are also similar to the above figure but I_e and V_{EB} both are negative and V_{CB} is positive. Dynamic input resistance of a transistor is given by

$R_i = \left(\frac{\Delta V_{EB}}{\Delta I_e}\right)_{V_{CB=constant}}$ {R_i is of the order of 100Ω}

(vi) **Output characteristics :** Taking the emitter current i_e constant, the curve drawn between I_C and V_{CB} are known as output characteristics of CB configuration.

Dynamics output resistance $R_0 = \left(\frac{\Delta V_{CB}}{\Delta i_C}\right)_{i_e=constant}$



Note: Transistor as CB amplifier

(i) ac current gain $\alpha_c = \frac{\text{Small change in collector current } (\Delta i_c)}{\text{Small changeincollectorcurrent} (\Delta i_e)}$

(ii) dc current gain α_{dc} (or α) = $\frac{\text{Collector current} (i_c)}{\text{Emitter current} (i_e)}$

value of α_{dc} lies between 0.95 to 0.99

(iii) Voltage gain $A_v = \frac{\text{Change in output voltage } (\Delta V_0)}{\text{Change in input voltage } (\Delta V_i)}$

⇒ $A_v = \alpha_{ac}$ × Resistance gain

(iv) Power gain = $\frac{\text{Change in output power} (\Delta P_0)}{\text{Change in input power} (\Delta P_C)}$

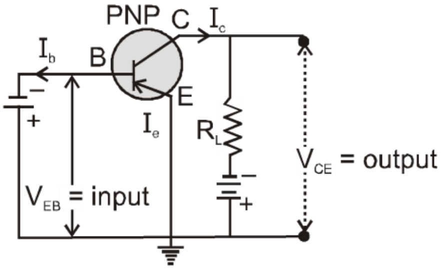
⇒ Power gain = α_{2ac} × Resistance gain

Common Emitter(CE) : The transistor is most widely used in the CE configuration.

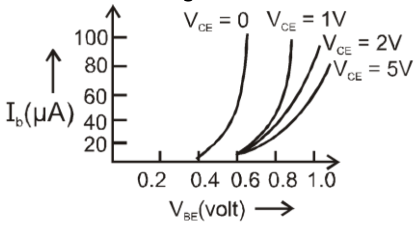
When a transistor is used in CE configuration, the input is between the base and the emitter and the output is between the collector and the emitter. The variation of the base current I_B with the base–emitter voltage V_{BE} is called the input characteristic. The output characteristics are controlled by the input characteristics. This implies that the collector current changes with the base current.

CE configurations : Emitter is common to both base and collector.

The graphs between voltages and currents when emitter of a transistor is common to input and output circuits are known as CE charactersistics of a transistor.

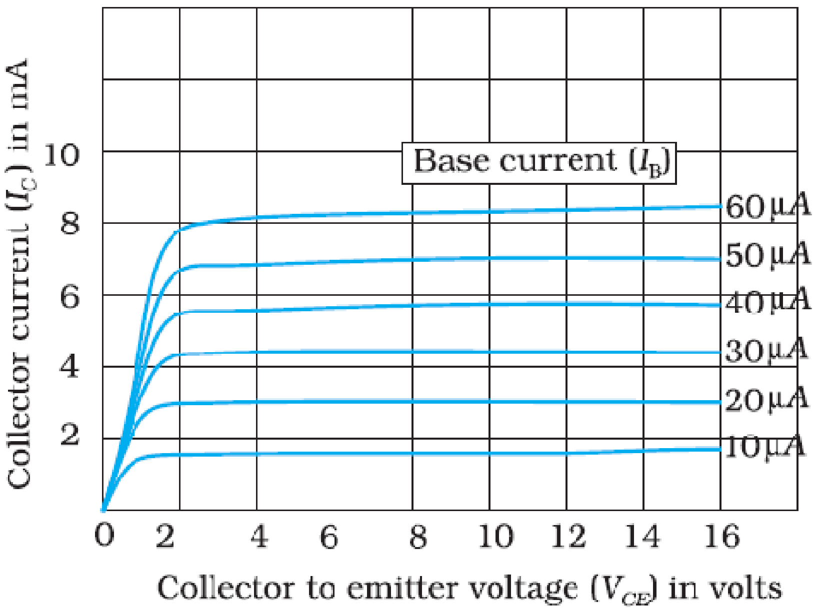


Input charactersistics : Input charactersistics curve is drawn between base current I_b and emitter base voltage V_{EB}, at constant collector emitter voltage V_{CE}.



Dynamic input resistance $R_i = \left(\frac{\Delta V_{BE}}{\Delta I_B}\right)_{V_{CE}\to constant}$

Output characteristics : Variation of collector current I_C with V_{CE} can be noticed for V_{CE} between 0 to 1 V only. The value of V_{CE} up to which the I_C changes with V_{CE} is called knee voltage. The transistor are operated in the region above knee voltage.



Dynamic output resistance $R_0 = \left(\frac{\Delta V_{CE}}{\Delta I_C}\right)_{I_B\to constant}$

(b) Transistor as a device :

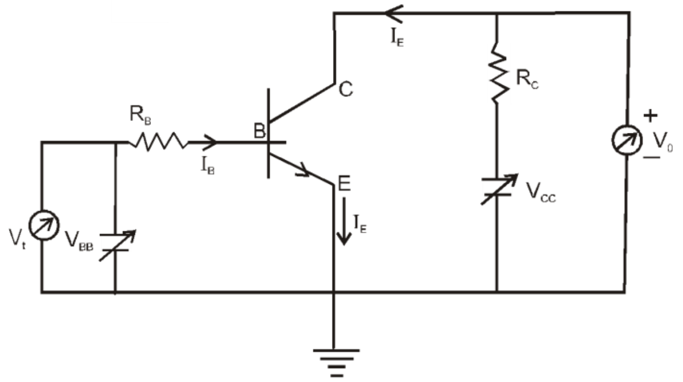
The transistor can be used as a device application depending on the configuration used (namely CB, CC and CE), the biasing of the E-B and B-C junction and the operation region namely cutoff, active region and saturation.

When the transistor is used in the cutoff or saturation state it acts as a switch. On the other hand for using the transistor as an amplifier, it has to operate in the active region.

(i) Transistor as a switch:

We shall try to understand the operation of the transistor as a switch by analysing the behaviour of the base-biased transistor in CE configuration as shown in fig. (a). Applying Kirchhoff's voltage rule to the input and output sides of this circuit, we get

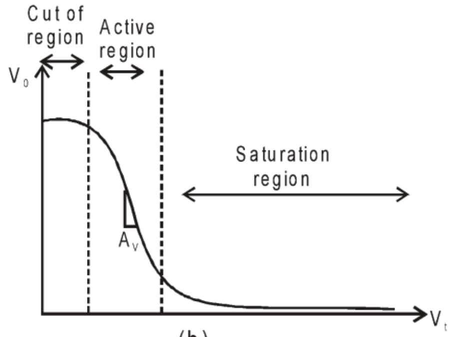
$V_{BB} = I_B R_B + V_{BE}$ and $V_{CE} = V_{CC} - I_C R_C$.



(a)

We shall treat V_{BB} as the dc input voltage V_i and V_{CE} as the dc output voltage V_o.So, we have

$V_i = I_B R_B + V_{BE}$ and $V_0 = V_{CC} - I_C R_C$.



(b)

Let us see how V_0 changes as V_i increases from zero onwards. In the case of Si transistor, as long as input V_i is less than 0.6 V, the transistor will be in cut off state and current I_C will be zero. Hence $V_O = V_{CC.}$

When V_i becomes greater than 0.6 V the transistor is in active state with some current I_C in the output path and the output V_0 decrease as the term $I_C R_C$ increases. With increase of V_i , I_C increases almost linearly and so V_0 decreases linearly till its value becomes less than about 1.0 V.

Beyond this, the change becomes non linear and transistor goes into saturation state. With further increase in V_i the output voltage is found to decrease further towards zero though it may never become zero. If we plot the V_0 vs V_i curve, [also called the transfer characteristics of the base-biased transistor (figure b], we see that between cut off state and active state and also between active state and saturation state there are regions of non-linearity showing that the transition from cutoff state to active state and from active state to saturation state are not sharply defined.

As long as V_i is low and unable to forward-bias the transistor, V_0 is high (at V_{CC}). If V_i is high enough to drive the transistor into saturation very near to zero. When the transistor is not conducting it is said to be switched off and when it is driven into saturation it is said to be switched on. This shows that if we define low and high states as below and above certain voltage levels corresponding to cutoff and saturation of the transistor, then we can say that a low input switches the transistor off and a high input switches it on.

(ii) Transistor as an Amplifier (CE-Configuration) : To perate the transistor as an amplifier it is necessary to fix its operating point somewhere in the middle of its active region. If we fix the value of V_{BB} corresponding to a point in the middle of the linear part of the transfer curve then the dc base current I_B would be constant and corresponding collector current I_C will also be constant. The dc voltage $V_{CE} = V_{CC} - I_C R_C$ would also remain constant. The operating values of V_{CE} and I_B determine the operating point, of the amplifier. If a small sinusoidal voltage with amplitude v_s is superposed on the dc base bias by connecting the source of that signal in series with the V_{BB} supply, then the base current will have sinusoidal variations superimposed on the value of I_B. As a consequence the collector current also will have sinusoidal variations superimposed on the value of I_C producing in turn corresponding change in the value of V_0. We can measure the ac variations across the input and output terminals by blocking the dc voltages by larger capacitors.

In the discription of the amplifier given above we have not considered any ac signal. In general, amplifiers are used to amplify alternating signals. Now let us superimpose an ac input signal v_i (to be amplified) on the bias V_{BB} (dc) as shown in Figure. The output is taken between the collector and the ground.

The working of an amplifier can be easily understood, if we first assume that $v_i = 0$. Then applying Kirchhoff's law to the output loop, we get

$$V_{CC} = V_{CE} + I_C R_L$$

Likewise, the input loop gives

$$V_{BB} = V_{BE} + I_B R_B$$

when v_i is not zero, we get

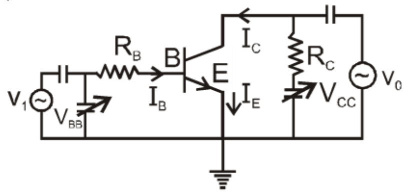
$$V_{BE} + v_i = V_{BE} + I_B R_B + \Delta I_B (R_B + r_i)$$

The change in V_{BE} can be related to the input resistance r_i and the change in I_B. Hence

$$v_i = \Delta I_B (R_B + r_i)$$
$$= r\Delta I_B$$

The change in I_B causes a change in I_C. We define a parameter β_{ac}, which is similar to the β_{dc} defined in equation as

$$\beta_{ac} = \frac{\Delta I_C}{\Delta I_B} = \frac{i_c}{i_b}$$



which is also known as the ac current gain A_i. Usually β_{ac} is close to β_{dc} in the linear region of the output characteristics.

The change in I_C due to a change in I_B causes a change in V_{CE} and the voltage drop across the resistor R_L because V_{CC} is fixed.

These changes can be given by Eq. as

$\Delta V_{CC} = \Delta V_{CE} + R_L \Delta I_C = 0$ or $\Delta V_{CE} = -R_L \Delta I_C$

The change in V_{CE} is the output voltage v_0. From equation we get $v_0 = \Delta V_{CE} = -\beta_{ac} R_L \Delta I_B$

The voltage gain of the amplifier is $A_v = \frac{v_0}{v_i} = \frac{\Delta V_{CE}}{r\Delta I_B} = -\frac{\beta_{ac} R_L}{r}$

The negative sign represents that output voltage is opposite with phase with the input voltage.

From the dissussion of the transistor characteristics you have seen that there is a current gain β_{ac} in the CE configuration. Here we have also seen the voltage gain A_v. Therefore the power gain A_p can be expressed as the product of the current gain and voltage gain. Mathematically

$$A_p = \beta_{ac} \times A_v$$

Since β_{ac} and A_v are greater than 1, we get ac power gain. However it should be realised that transistor is not a power generating device. The energy for the higher ac power at the output is supplied by the battery.

Note: Transistor as CE amplifier

(i) ac current gain $\beta_{ac} = \left(\frac{\Delta i_C}{\Delta i_b}\right) V_{CE}$ = constant

(ii) dc current gain $\beta_{dc} = \frac{i_c}{i_b}$

(iii) Votage gain : $A_v = \frac{\Delta V_0}{\Delta V_i}$ = β_{ac} × Resistance gain

(iv) Power gain = $\frac{\Delta P_0}{\Delta P_i}$ = β_{2ac} × Resistance

(v) Trans conductance (g_m) : The ratio of the change in collector current to the change in emitter base voltage is called trans conductance. i.e. $g_m = \frac{\Delta i_c}{\Delta V_{EB}}$. Also $g_m = \frac{A_v}{R_L}$; R_L = Load resistance.

(3) Relation between α and β : $\beta = \frac{\alpha}{1-\alpha}$ or $\alpha = \frac{\beta}{1+\beta}$

Solved Examples

Example 30. Let i_E, i_C and i_B represent the emitter current, the collector current and the base current respectively in a transistor. Then

(1) i_C is slightly smaller than i_E.
(2) i_C is slightly greater than i_E.
(3) i_B is much smaller than i_E.
(4) i_B is much greater than i_E.

Ans. (1,3)

Example 31. In a common base transistor amplifier, the input and the output resistance are 500 Ω and 40kΩ, and the emitter current is 1.0mA. Find the input and the output voltages. Given α = 0.95.

Solution : The input voltage is emitter current multiplied by input resistance, that is,

$V_{in} = i_E \times R_{in} = (1.0 \times 10^{-3}\ A) \times 500\Omega = 0.5\ V$

Similarly, the output voltage is

$V_{out} = i_C \times R_{out} = \alpha\ i_E \times R_{out} = 0.95\ (1.0 \times 10^{-3}\ A) \times (40 \times 10^3\ \Omega) = 38\ V$.

Example 32. A P–N–P transistor is used in common–emitter mode in an amplifier circuit. A change of 40μA in the base current brings a change of 2mA in collector current and 0.04 V in base–emitter voltage. Find the:

(i) input resistance ($R_{inp.}$), and

(ii) the base current amplification factor (β).

If a load of 6kΩ is used, then also find the voltage gain of the amplifier.

Solution : Given $\Delta I_B = 40\mu A = 40 \times 10^{-6}\ A$

$\Delta I_C = 2mA = 2 \times 10^{-3}\ A$

$\Delta V_{BE} = 0.04$ volt, $R_L = 6k\Omega = 6 \times 10^3\ \Omega$

(i) Input Resistance,

$R_{inp.} = \frac{\Delta V_{BE}}{\Delta I_B} = \frac{0.04}{40 \times 10^{-6}} = 10^3\ \Omega = 1\ k\Omega$

(ii) Current amplification factor,

$\beta = \frac{\Delta I_C}{\Delta I_B} = \frac{2 \times 10^{-3}}{40 \times 10^{-6}} = 50$

(iii) Voltage gain in common–emitter configuration,

$A_v = \beta \frac{R_L}{R_{inp.}} = 50 \times \frac{6 \times 10^3}{1 \times 10^3} = 300.$

Example 33. In an N–P–N transistor 10^{10} electrons enter the emitter in 10^{-6} s. 2% of the electrons are lost in the base. Calculate the current transfer ratio and current amplification factor.

Solution : We know that current = charge/time

The emitter current (I_E) is given by $I_E = \frac{Ne}{t} = \frac{10^{10} \times (1.6 \times 10^{-19})}{10^{-6}} = 1.6\ mA$

The base current (I_B) is given by $I_B = \frac{2}{100} \times 1.6 = 0.032\ mA$

In a transistor, $I_E = I_B + I_C$ $I_C = I_E - I_B = 1.6 - 0.032 = 1.568\ mA$

Current transfer ratio = $\frac{I_C}{I_E} = \frac{1.568}{1.6} = 0.98$

Current amplification factor = $\frac{I_C}{I_B} = \frac{1.568}{0.032} = 49.$

Example 34. When the voltage between emitter and the base V_{EB} of a transistor is changed by 5mV while keeping the collector voltage V_{CE} fixed when then its emitter current changes by 0.15 mA. Calculate the input resistance of the transistor.

Ans. 33.33 ohm

Example 35. A transistor is used in common-emitter mode in an amplifier circuit. When a signal of 20 mV is added to the base–emitter voltage, the base current changes by 20μA and the collector current changes by 2 mA. The load resistance is 5 kΩ. Calculate (a) the factor β, (B) the input resistance $R_{BE'}$ (C) the transconductance and (D) the voltage gain.

Solution : (A) $\beta = \frac{\Delta I_C}{\Delta I_B} = \frac{2mA}{20\mu A} = 100$

(B) The input resistance $R_{BE} = \frac{\Delta V_{BE}}{\Delta I_B} = \frac{20\,mV}{20\,\mu A} = 1k\Omega$

(C) Transconductance = $\frac{\Delta I_c}{\Delta V_{BE}} = \frac{2mA}{20\,mV} = 0.1$ mho.

(D) The change in output voltage is $R_L \Delta I_C$ = (5 kW) (2mA) = 10V.

The applied signal voltage = 20 mV.

Thus, the voltage gain is, = $\frac{10\,V}{20\,mV} = 500.$

Example 36. The a-c current gain of a transistor is β= 19. In its common-emitter configuration, what will be the change in the collector-current for a change of 0.4 mA in the base-current ? What will be the change in the emitter current ?

Solution : By definition, the a-c current gain β is given by

β (a-c) = $\frac{\Delta_{iC}}{\Delta_{iB}}$ ∴ $\Delta_{iC} = \beta \times \Delta_{iB}$ = 19 × 0.4 mA = 7.6 mA.

The emitter - current is the sum of the base- current and the collector-current ($i_E = i_B + i_C$)

∴ $\Delta_{iE} = \Delta_{iB} + \Delta_{iC}$ = 0.4 mA + 7.6 mA = 80 mA.

Example 37. A transistor is connected in common-emitter (C-E) configuration. The collector-supply is 8 V and the voltage drop across a resistor of 800 Ω in the collector circuit is 0.5 V. If the current-gain factor (α) is 0.96, find the base-current.

Solution : The alternating-current gain is $\beta = \frac{\alpha}{1-\alpha} = \frac{0.96}{1-0.96} = 24$

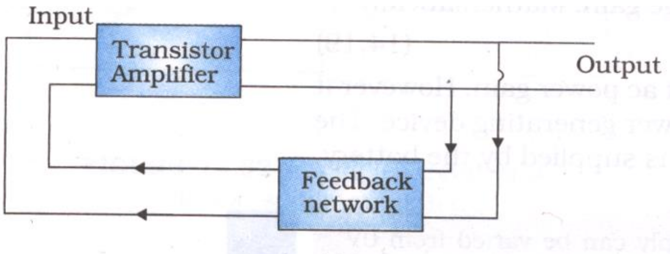
The collector - current is

$i_C = \frac{\text{voltage} - \text{drop across collector resistor}}{\text{resistance}} = \frac{0.5\,V}{800\,\Omega} \times 10^{-3}$ A.

But $\beta = \frac{i_C}{i_B}$, where i_B is base - current. ∴ $i_B = \frac{i_C}{\beta} = \frac{0.625 \times 10^{-3}\,A}{24}$ = 26 × 10^{-6} A - 26 μA.

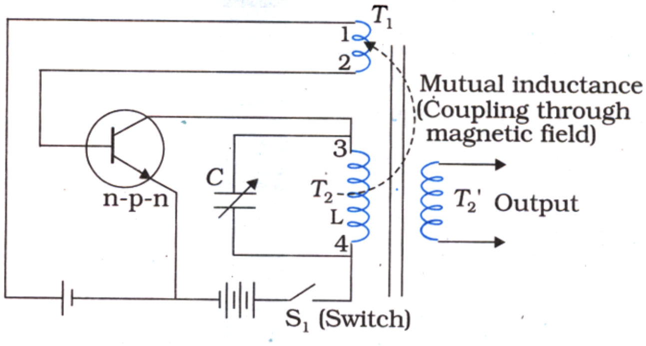
8. Feedback amplifier and transistor oscillator:

In an oscillator, we get ac output without any external input singnal. A portion of the output power is returned back (feedback) to the input in phase with the starting power (this process is termed positive feedback) as shown in figure(a). The feedback can be achieved by inductive coupling (through mutual inductance) or LC or RC networks.



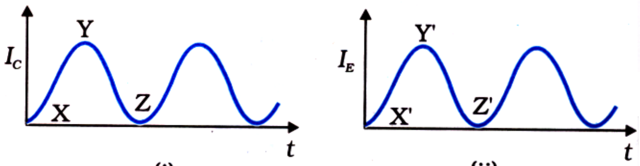
(a)

Suppose switch S_1 is put on to apply proper bias for the first time. Obviously, a surge of collector current flows in the transistor. This current flows through the coil T_2 where terminals are numbered 3 and 4 (Fig. b).



(b)

This current does not reach full amplitude instantaneously but increases from X To Y, as shown in figure(C). The inductive coupling between coil T_2 and coil T_1 now causes a current to flow in the emitter circuit (note that this actually is the 'feedback' from input to output). As a result of this positive feedback, this current (in T_1 emitter current) also increases from X' to Y' Fig. (C) (ii).



(i) (ii)

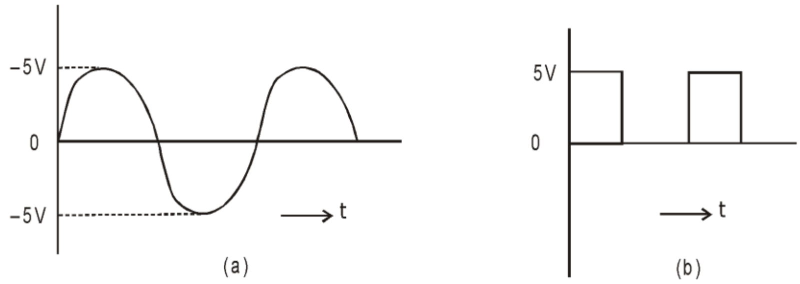
(c)

The current in T_2 (collector current), connected in the collector circuit acquires the value Y when the transistor becomes saturated. This means that maximum collector current is flowing and can increase no further. Since there is no further change in collector current, the magnetic field around T_2 ceases to grow. As soon as the field becomes static, there will be no further feedback from T_2 to T_1. Without continued feedback, the emitter current begins to fall. Consequently, collector current decreases causes the magnetic field to decay around the coil T_2. Thus, T_1 is now seeing a decaying field in T_2 (opposite from what it saw when the field was growing at the initial start operation). This causes a further decrease in means that both I_E and I_C cease to flow. Therefore, the transistor has reverted back to its original state (when the power was first switched on). The whole process now repeat itself. The transistor is driven to saturation, then to cut-off, and then back to saturation. The time for change from saturation to cut-off and back is determined by the constant of the tank circuit or tuned circuit (inductance L of Coil T_2 and C connected in parallel to it). The resonance frequency (v) of this tuned circuit determines the frequency at which the oscillator will oscillate. $\nu = \frac{1}{2\pi\sqrt{LC}}$

9. Analogue Circuits and Digital Circuits and signal :

There are two types of electronic circuits : analogue circuits and digital circuits :

In analogue circuits, the voltage (or current) varies continuously with time (figure a). Such a voltage (or current) signal is called an 'analogue signal'. Figure shows a typical voltage analogue signal varying sinusoidally between 0 and 5V.



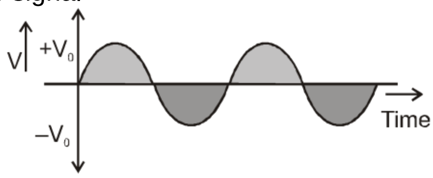
On the other hand, in digital circuits, the voltage (or current) has only two levels, either zero or some constant value of voltage (figure b). A signal having only two levels of voltage (or current) is called a ‘digital signal’. Figure shows a typical digital signal in which the voltage at any time is either 0 or 5V.

In digital circuits, the binary number system is used, according to which the two levels of the (digital) signal are represented by the digits 0 and 1 only.

The digital circuits are the basis of calculators, computers, etc.

Note: Voltage Signal :

(a) Analogue voltage singal : The signal which represents the continuous variation of voltage with time is known as analogue voltage signal



(b) Digital voltage signal : The signal which has only two values. i.e., either a constant high value of voltage or zero value is called digital voltage signal.

Decimal and Binary Number System :

(1) Decimal number system : In a decimal number system, we have ten digits i.e. 0,1,2,3,4,5,6,7,8,9

A decimal number system has a bse of ten (10)

$1971 = 1000 + 900 + 70 + 1$

$= 1 \times 10^3 + 9 \times 10^2 + 7 \times 10^1 + 1 \times 10^0$

e.g. MSD LSD

LSD = Least significant digit

MSD = Most significant digit

(2) Binary number system : A number system which has only two digits i.e. 0 (Low) and 1 (High) is known as binary system. The base of binary number system is 2.

(i) Each digit in binary system is known as a bit and a group of bits is known as a byte.

(ii) The electrical circuit which operatres only in these two state i.e. (On or High) and 0 (i.e. Off or Low) are known as digital circuits.

State Code	**1**	**0**
	On	Off
	Up	Down
	Close	Open
Name for the State	Excited	Unexcited
	True	False
	Pluse	No pulse
	High	Low
	Yes	No

(3) Decimal to binary conversion

(i) Divide the given decimal number by 2 and the successive quotients by 2 till the quotient becomes zero.

(ii) The sequence of remainders obtained during divisions gives the binary equivalent of decimal number.

(iii) The most significant digit (or bit) of the binary number so obtained is the last remainder and the least significant digit (or bit) is the first remainder obtained during the division.

For Example : Binary equivalence of 61

2	61	Remainder
2	30	1 LSD
2	15	0
2	7	1
2	3	1
2	1	1
	0	**1 MSD**

$\Rightarrow (61)_{10} = (111101)_2$

(4) Binary to decimal conversion : The least significant digit in the binary number is the coefficient of 2 with power zero. As we move towards the left side of LSD, the power of 2 goes on increasing.

For Example : $(11111100101)_2 = 1 \times 2^{10} + 1 \times 2^9 + 1 \times 2^8 + 1 \times 2^7 + 1 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 2021$

10. Logic Gates :

A logic gate is a digital circuit which works according to some logical relationship between input and output voltages. It either allows a signal to pass through or stops it.

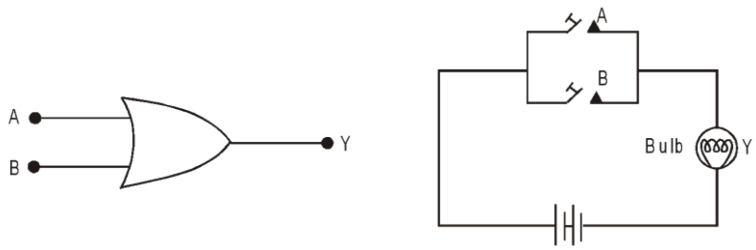
The logic gates are the building blocks of digital circuits. There are three basic logic gates.

(a) OR gate (b) AND gate (c) NOT gate

(a)The OR Gate :

The OR gate is a device that has two input variables A and B and one output variable Y, and follows the Boolean expression, $A + B = Y$,

read as' A OR B equal Y'. Its logic symbol is shown in figure.



(a) (b)

A	B	Y(= A+ B)
0	0	0
0	1	1
1	0	1
1	1	1

(c)

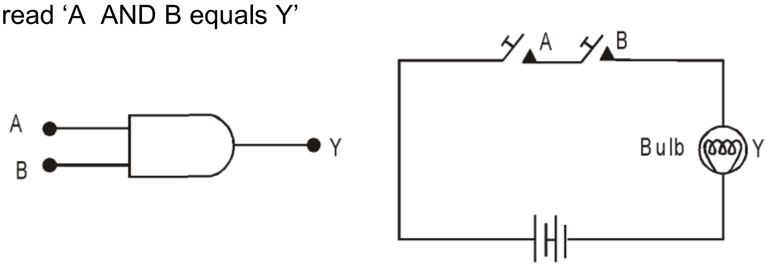
The possible combinations of the inputs A and B and the output Y of the OR gate can be known with the help of an electrical circuit, shown in figure. In this circuit, two switches A and B (inputs) are connected in parallel with a battery and a bulb Y (output).

(b) The AND Gate :

The AND gate is also a two-input and one-output logic gate. It combines the inputs A and B to give the output Y, according to the Boolean expression

$$A.B = Y$$

read 'A AND B equals Y'



(a) (b)

A	B	Y(= A. B)
0	0	0
0	1	0
1	0	0
1	1	1

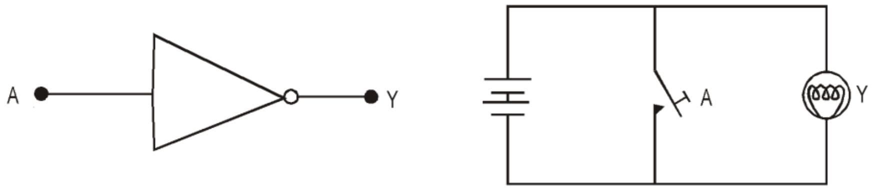
(c)

(c) The NOT Gate :

The NOT gate has only one input and one output. It combines the input A with the output Y, according to the Boolean expression

$$\overline{A} = Y \quad ,$$

read as 'NOT A equals Y'. It means that Y is negation (or inversion) of A. Since there are only two digits 0 and 1 in the binary system, we have, Y = 0, if A = 1and Y = 1 if A = 0 . The logic symbol of the NOT gate is shown in figure.

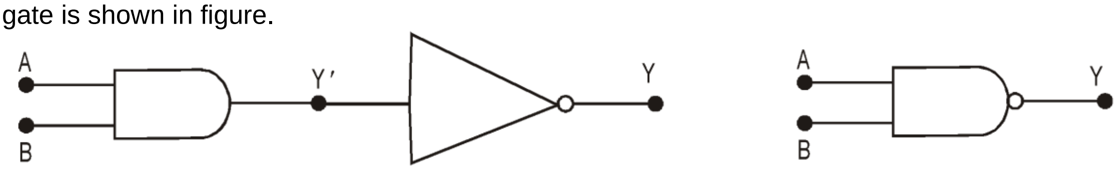


A	Y(= $\bar{A}$)
0	1
1	0

(a) (b) (c)

The possible combinations of the input A and the output Y of the NOT gate can be known with the help of electric circuit, shown in figure. In this circuit, a swith A (input) is connected in parallel to a battery and a bulb Y(output). The working of the circuit is as follows :
If switch A is open (A = 0), the bulb wil glow (Y = 1).
If switch A is closed (A = 1), the bulb will not glow (Y = 0).
These two possible combinations of input A and output Y are tabulated in figure, which is the truth table of the NOT gate.

Combinations of gates :
Various combinations of the three basic gates, namely, OR, AND and NOT, produce complicated digital circuits, which are also called 'gates'. The commonly used combinations of basic gates are NAND gate, NOR, gate. These are also called universal gates.
(i) The NAND gate :
This gate is a combination of AND and NOT gates. If the output Y' of AND gate is connected to the input of NOT gate, as shown in figure, the gate so obtained is called NAND gate. The logic symbol of NAND gate is shown in figure.



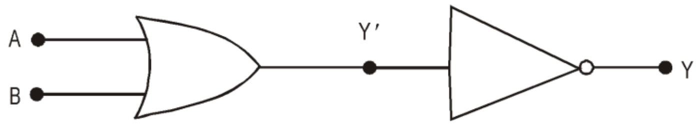
(a) (b)

The Boolean expression for the NAND gate is $\overline{A \cdot B} = Y$
read as 'A AND B negated equals Y'.
The truth table of the NAND gate can be obtained by logically combining the truth tables of AND and NOT gates. In figure, the output Y' of the truth table of AND gate have been negated (NOT operation) to obtain the corresponding outputs Y for the NAND gate. The resulting table is the truth table of the NAND gate.

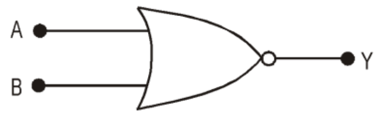
A	B	Y' (= A . B)	Y (= $\overline{A.B}$) = $\overline{Y'}$
0	0	0	1
0	1	0	1
1	0	0	1
1	1	1	0

A	B	Y
0	0	1
0	1	1
1	0	1
1	1	0

(ii) The NOR Gate :
The NOR gate is a combination of OR and NOT gates. If the output Y′ of OR gate is connected to the input of NOT gate, as shown in figure, the gate so obtained is NOR gate.



(a)



(b)

The Boolean expression for the NOR gate is

$$\overline{A + B} = Y$$

read as 'A OR B negated equals Y' :

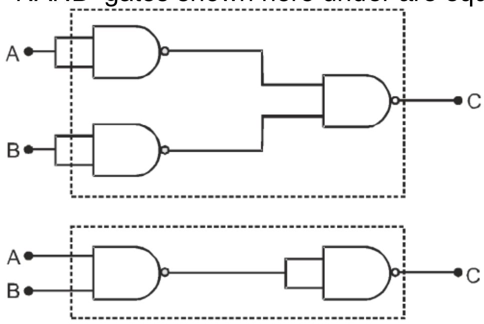
A	B	Y′ (= A+B)	Y (= $\overline{A+B}$) = $\overline{Y'}$
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0

A	B	Y
0	0	1
0	1	0
1	0	0
1	1	0

The truth table of the NOR gate can be obtained by logically combining the truth tables of OR and NOT gates. In figure(a), the outputs Y′ of the truth table of OR gate have been negated to obtain the corresponding outputs Y for the NOR gate.

Self Practice Problems

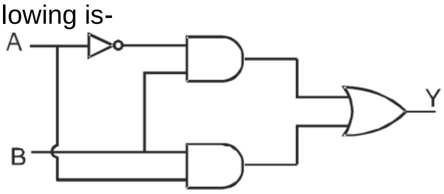
Example 1. The combination of 'NAND' gates shown here under are equivalent to-



(1) An OR gate and an AND gate respectively (2) An AND gate and a NOT gate respectively
(3) An AND gate and an OR gate respectively (4) An OR gate and a NOT gate respectively

Ans. **(1)**

Example 2. Truth table for the following is-



(1)

A	B	Y
0	0	0
0	1	1
1	0	1
1	1	0

(2)

A	B	Y
0	0	0
0	1	1
1	0	0
1	1	1

(3)

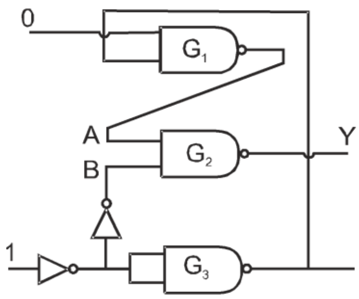
A	B	Y
0	0	1
0	1	0
1	0	0
1	1	1

(4)

A	B	Y
0	0	0
0	1	1
1	0	1
1	1	1

Ans. (2)

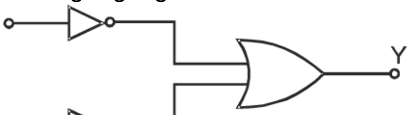
Example 3. In circuit in following figure the value of Y is-



(1) zero (2) 1
(3) fluctuates between 0 and 1 (4) indeterminate as the circuit cannot be realized

Ans. **(1)**

Example 4. The output Y for the following logic gate circuit will be-



(1) AB (2) $\bar{A} \cdot \bar{B}$ (3) $\overline{A + B}$ (4) $\overline{A \cdot B}$

Ans. (4)

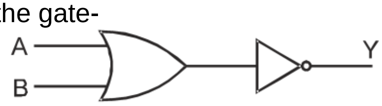
Example 5. The following truth-table belongs to which one of the four gates-

A	B	X
1	1	0
0	1	0
1	0	0
0	0	1

(1) OR (2) NAND (3) XOR (4) NOR

Ans. (4)

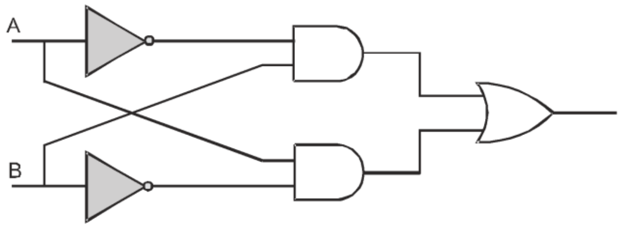
Example 6. The given circuit is for the gate-



(1) NOR (2) NAND (3) NOT (4) XOR

Ans. (1)

Example 7. The truth table of the logic circuit shown-



(1)

A	B	Y
0	0	0
1	0	1
1	0	1
1	1	1

(2)

A	B	Y
0	0	0
0	1	1
1	0	1
1	1	0

(3)

A	B	Y
0	0	1
0	1	0
1	0	0
1	1	0

(4)

A	B	Y
0	0	0
1	0	0
0	1	0
1	1	1

Ans. (2)