Lectures in electronics

Ass. Prof. Khaled Salah

2023

الكلية: العلوم

المستوي: الثاني

القسم:الفيزياء ، الفيزياء الجوية والفلك (لائحة جديدة)

تاريخ النشر:2023

عدد الصفحات:109

اعداد:د. خالد صلاح الدين

1

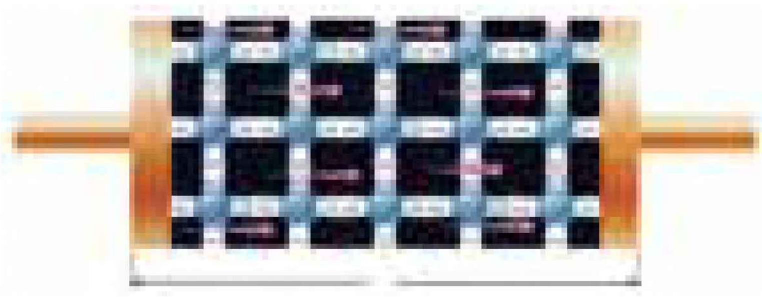
Introduction

GENERAL

In this fast developing society, *electronics* has come to stay as the most important branch of engineering. Electronic devices are being used in almost all the industries for quality control and automation and they are fast replacing the present vast army of workers engaged in processing and assembling in the factories. Great strides taken in the industrial applications of electronics during the recent years have demonstrated that this versatile tool can be of great importance in increasing production, efficiency and control.

The rapid growth of electronic technology offers a formidable challenge to the beginner, who may be almost paralysed by the mass of details. However, the mastery of fundamentals can simplify the learning process to a great extent. The purpose of this chapter is to present the elementary knowledge in order to enable the readers to follow the subsequent chapters.

1.1 Electronics


Current conduction through semiconductor

The branch of engineering which deals with current conduction through a vacuum or gas or semiconductor is known as ***electronics.**

Electronics essentially deals with electronic devices and their utilisation. An *electronic device* is that in which current flows through a vacuum or gas or semiconductor. Such devices have valuable properties which enable them to function and behave as the friend of man today.

Importance. Electronics has gained much importance due to its numerous applications in industry. The electronic devices are capable of performing the following functions :

(*i*) Rectification. The conversion of a.c. into d.c. is called *rectification*. Electronic devices can convert a.c. power into d.c. power (See Fig. 1.1) with very high efficiency. This d.c. supply can be used for charging storage batteries, field supply of d.c. generators, electroplating etc.

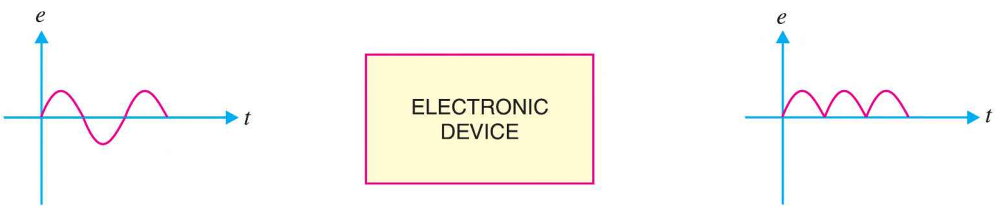


Fig. 1.1

(*ii*) Amplification. The process of raising the strength of a weak signal is known as *amplification*. Electronic devices can accomplish the job of amplification and thus act as amplifiers (See Fig. 1.2). The amplifiers are used in a wide variety of ways. For example, an amplifier is used in a radio-set where the weak signal is amplified so that it can be heard loudly. Similarly, amplifiers are used in public address system, television etc.

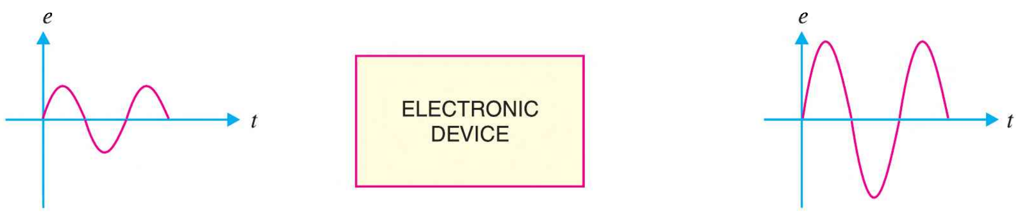


Fig. 1.2

(*iii*) Control. Electronic devices find wide applications in automatic control. For example, speed of a motor, voltage across a refrigerator etc. can be automatically controlled with the help of such devices.

(*iv*) Generation. Electronic devices can convert d.c. power into a.c. power of any frequency (See Fig. 1.3). When performing this function, they are known as *oscillators*. The oscillators are used in a wide variety of ways. For example, electronic high frequency heating is used for annealing and hardening.

* The word *electronics* derives its name from electron present in all materials.

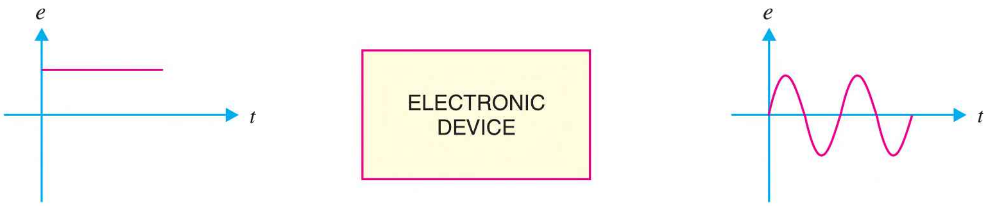


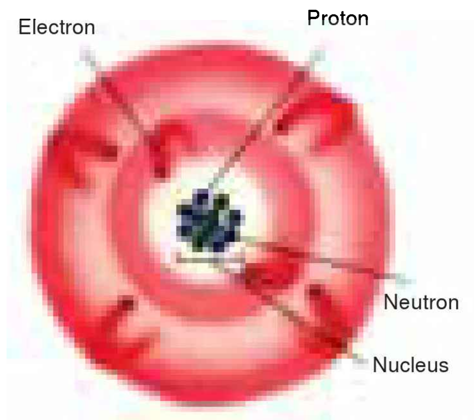
Fig. 1.3

(*v*) Conversion of light into electricity. Electronic devices can convert light into electricity. This conversion of light into electricity is known as *photo-electricity*. Photo-electric devices are used in Burglar alarms, sound recording on motion pictures *etc.*

(*vi*) Conversion of electricity into light. Electronic devices can convert electricity into light. This valuable property is utilised in television and radar.

1.2 Atomic Structure

According to the modern theory, matter is electrical in nature. All the materials are composed of very small particles called *atoms*. The atoms are the building bricks of all matter. An atom consists of a central *nucleus* of positive charge around which small negatively charged particles, called *electrons* revolve in different paths or orbits.



Carbon Atom

(1) Nucleus. It is the central part of an atom and *contains *protons* and *neutrons*. A proton is a positively charged particle, while the neutron has the same mass as the proton, but has no charge. Therefore, the nucleus of an atom is positively charged. The sum of protons and neutrons constitutes the entire weight of an atom and is called atomic weight. It is because the particles in the extra nucleus (*i.e.* electrons) have negligible weight as compared to protons or neutrons.

$\therefore$ atomic weight = no. of protons + no. of neutrons

(2) Extra nucleus. It is the outer part of an atom and contains *electrons* only. An electron is a negatively charged particle having negligible mass. The charge on an electron is equal but opposite to that on a proton. Also, the number of electrons is equal to the number of protons in an atom under ordinary conditions. Therefore, an atom is neutral as a whole. The number of electrons or protons in an atom is called *atomic number i.e.*

atomic number = no. of protons or electrons in an atom

The electrons in an atom revolve around the nucleus in different orbits or paths. The number and arrangement of electrons in any orbit is determined by the following rules :

(*i*) The number of electrons in any orbit is given by $2n^2$ where n is the number of the orbit. For example,

First orbit contains 2×1^2 = 2 electrons

Second orbit contains 2×2^2 = 8 electrons

Third orbit contains 2×3^2 = 18 electrons

* Although the nucleus of an atom is of complex structure, yet for the purpose of understanding electronics, this simplified picture of the nucleus is adequate.

and so on.

(*ii*) The last orbit cannot have more than 8 electrons.

(*iii*) The last but one orbit cannot have more than 18 electrons.

1.3 Structure of Elements

We have seen that all atoms are made up of protons, neutrons and electrons. The difference between various types of elements is due to the different number and arrangement of these particles within their atoms. For example, the structure* of copper atom is different from that of carbon atom and hence the two elements have different properties.

The atomic structure can be easily built up if we know the atomic weight and atomic number of the element. Thus taking the case of copper atom,

Atomic weight = 64

Atomic number = 29

$\therefore$ No. of protons = No. of electrons = 29

and No. of neutrons = 64 − 29 = 35

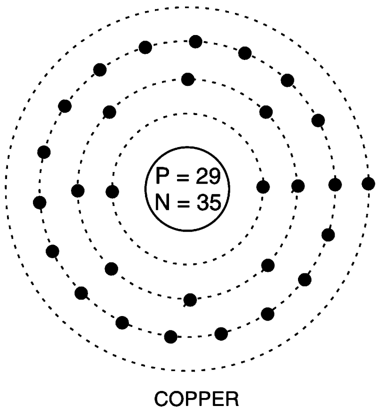


Fig. 1.4

Fig. 1.4 shows the structure of copper atom. It has 29 electrons which are arranged in different orbits as follows. The first orbit will have 2 electrons, the second 8 electrons, the third 18 electrons and the fourth orbit will have 1 electron. The atomic structure of all known elements can be shown in this way and the reader is advised to try for a few commonly used elements.

1.4 The Electron

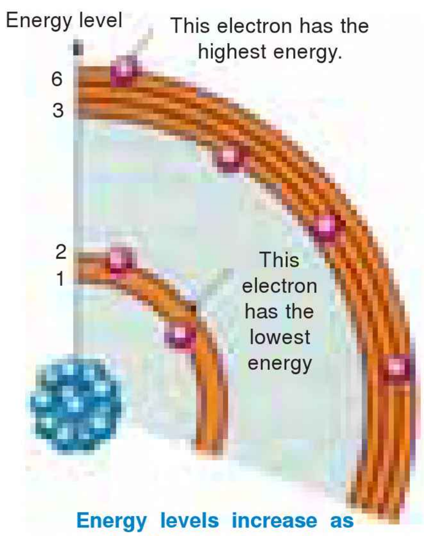
Since electronics deals with tiny particles called electrons, these small particles require detailed study. As discussed before, an electron is a negatively charged particle having negligible mass. Some of the important properties of an electron are :

(*i*) Charge on an electron, e = 1.602×10^{-19} coulomb

(*ii*) Mass of an electron, m = 9.0×10^{-31} kg

(*iii*) Radius of an electron, r = 1.9×10^{-15} metre

The ratio *e/m* of an electron is 1.77×10^{11} coulombs/kg. This means that mass of an electron is very small as compared to its charge. It is due to this property of an electron that it is very mobile and is greatly influenced by electric or magnetic fields.



Energy levels increase as the distance from the nucleus increases

1.5 Energy of an Electron

An electron moving around the nucleus possesses two types of energies viz. kinetic energy due to its motion and potential energy due to the charge on the nucleus. The total energy of the electron is the sum of these two energies. The energy of an electron increases as its distance from the nucleus increases. Thus, an electron in the second orbit possesses more energy than the electron in the first orbit; electron in the third

* The number and arrangement of protons, neutrons and electrons.

orbit has higher energy than in the second orbit.It is clear that electrons in the last orbit possess very high energy as compared to the electrons in the inner orbits. These last orbit electrons play an important role in determining the physical, chemical and electrical properties of a material.

1.6 Valence Electrons

The electrons in the outermost orbit of an atom are known as **valence electrons.**

The outermost orbit can have a maximum of 8 electrons *i.e.* the maximum number of valence electrons can be 8. The valence electrons determine the physical and chemical properties of a material. These electrons determine whether or not the material is chemically active; metal or non-metal or, a gas or solid. These electrons also determine the electrical properties of a material.

On the basis of electrical conductivity, materials are generally classified into *conductors, insulators* and *semi-conductors.* As a rough rule, one can determine the electrical behaviour of a material from the number of valence electrons as under :

(i) When the number of valence electrons of an atom is less than 4 (*i.e.* half of the maximum eight electrons), the material is usually *a metal and a conductor.* Examples are sodium, magnesium and aluminium which have 1, 2 and 3 valence electrons respectively (See Fig. 1.5).

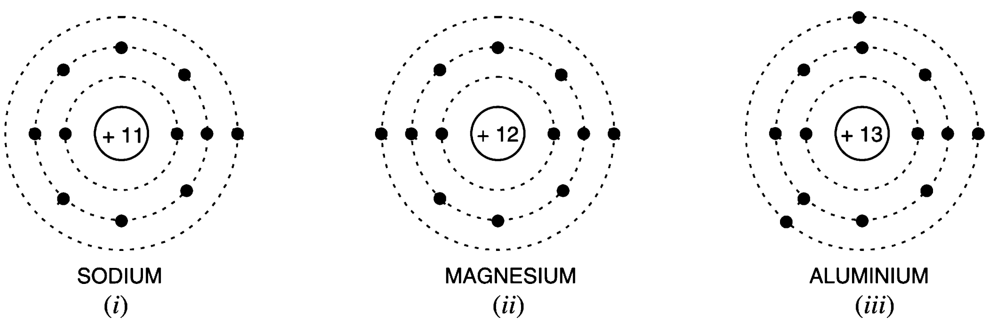


Fig. 1.5

(ii) When the number of valence electrons of an atom is more than 4, the material is usually *a non-metal and an insulator.* Examples are nitrogen, sulphur and neon which have 5, 6 and 8 valence electrons respectively (See Fig. 1.6).

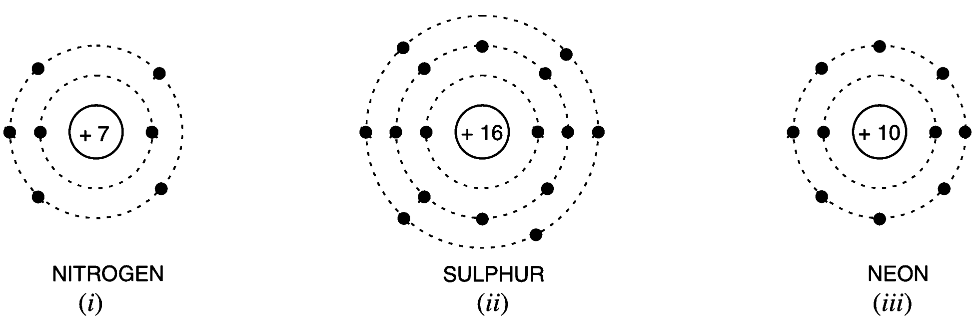
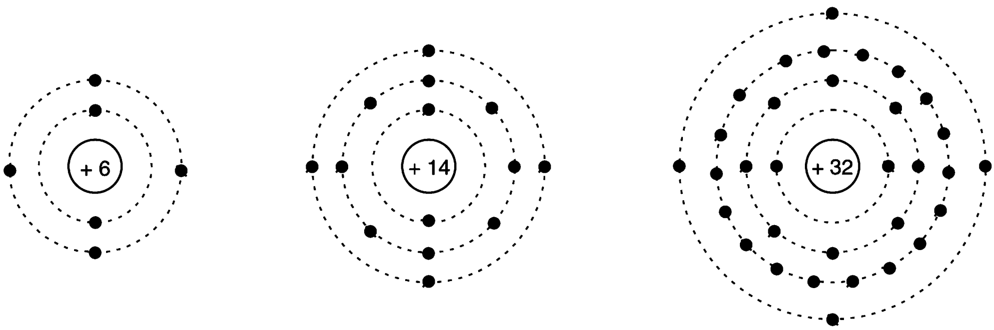


Fig. 1.6

(iii) When the number of valence electrons of an atom is 4 (*i.e.* exactly one-half of the maximum 8 electrons), the material has both metal and non-metal properties and is usually a *semi-conductor.* Examples are carbon, silicon and germanium (See Fig. 1.7).



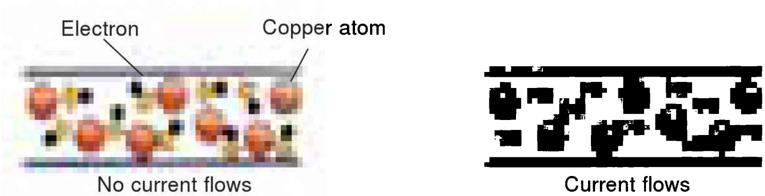
CARBON (i) | SILICON (ii) | GERMANIUM (iii)

Fig. 1.7

1.7 Free Electrons

The valence electrons of different materials possess different energies. The greater the energy of a valence electron, the lesser it is bound to the nucleus. In certain substances, particularly metals, the valence electrons possess so much energy that they are very loosely attached to the nucleus. These loosely attached valence electrons move at random within the material and are called *free electrons*.

The valence electrons which are very loosely attached to the nucleus are known as **free electrons.**



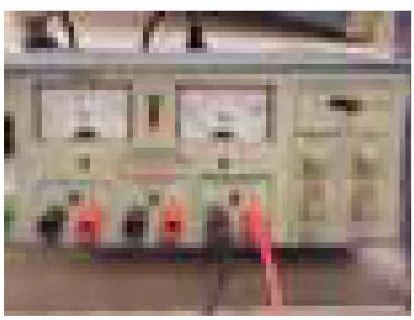
Current moves through materials that conduct electricity.

The free electrons can be easily removed or detached by applying a small amount of external energy. As a matter of fact, these are the free electrons which determine the electrical conductivity of a material. On this basis, conductors, insulators and semiconductors can be defined as under :

(*i*) A *conductor* is a substance which has a large number of free electrons. When potential difference is applied across a conductor, the free electrons move towards the positive terminal of supply, constituting electric current.

(*ii*) An *insulator* is a substance which has practically no free electrons at ordinary temperature. Therefore, an insulator does not conduct current under the influence of potential difference.

(*iii*) A *semiconductor* is a substance which has very few free electrons at room temperature. Consequently, under the influence of potential difference, a semiconductor *practically* conducts no current.



Voltage source

1.8 Voltage Source

Any device that produces voltage output continuously is known as a *voltage source*. There are two types of voltage sources, namely ; direct voltage source and alternating voltage source.

(*i*) **Direct voltage source.** A device which produces direct voltage output continuously is called *a direct voltage source*. Common examples are cells and d.c. generators. An important characteristic of a direct voltage source is that it maintains the same polarity of the output voltage *i.e.* positive and negative terminals remain the same. When load resistance R_L is connected across such a source,*current flows from positive terminal to negative terminal *via* the load [See Fig. 1.8 (*i*)]. This is called *direct current* because it has just one direction. The current has one direction as the source maintains the same polarity of output voltage. The opposition to load current inside the d.c. source is known as *internal resistance* R_i. The equivalent circuit of a d.c. source is the generated *e.m.f.* E_g in series with internal resistance R_i of the source as shown in Fig. 1.8 (*ii*). Referring to Fig. 1.8 (*i*), it is clear that:

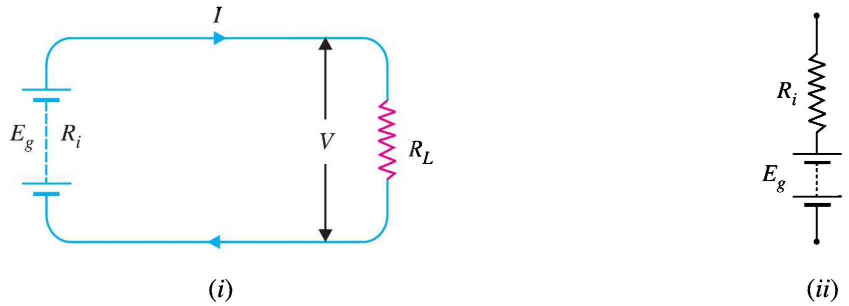


Fig. 1.8

$$\text{Load current, } I = \frac{E_g}{R_L + R_i}$$

$$\text{Terminal voltage, } V = (E_g - I R_i) \quad \text{or} \quad I R_L$$

***(ii)* Alternating voltage source.** A device which produces alternating voltage output continuously is known as *alternating voltage source e.g.* a.c. generator. An important characteristic of alternating voltage source is that it periodically reverses the polarity of the output voltage. When load impedance Z_L is connected across such a source, current flows through the circuit that periodically reverses in direction. This is called *alternating current.*

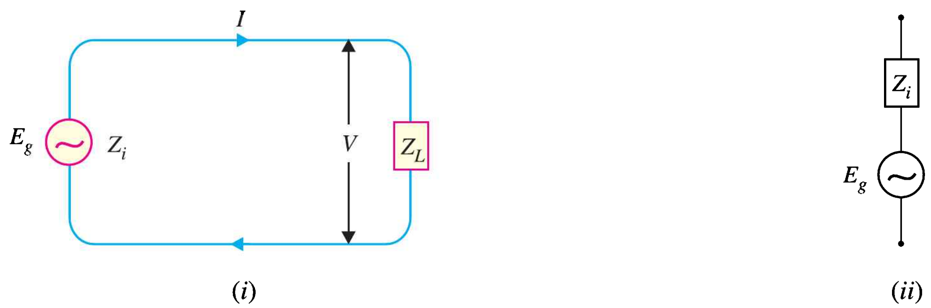


Fig. 1.9

The opposition to load current inside the a.c. source is called its *internal impedance* Z_i. The equivalent circuit of an a.c. source is the generated *e.m.f.* E_g (*r.m.s.*) in series with internal impedance Z_i of the source as shown in Fig. 1.9 (*ii*). Referring to Fig. 1.9 (*i*), it is clear that :

$$\text{Load current, } I\,(r.m.s.) = \frac{E_g}{Z_L + Z_i}$$

$$\text{Terminal voltage, } V = (E_g - I Z_i)^{**} \quad \text{or} \quad I Z_L$$

* This is the conventional current. However, the flow of electrons will be in the opposite direction.

** Vector difference since a.c. quantities are vector quantities.

1.9 Constant Voltage Source

*A voltage source which has very low internal *impedance as compared with external load impedance is known as a* **constant voltage source.**

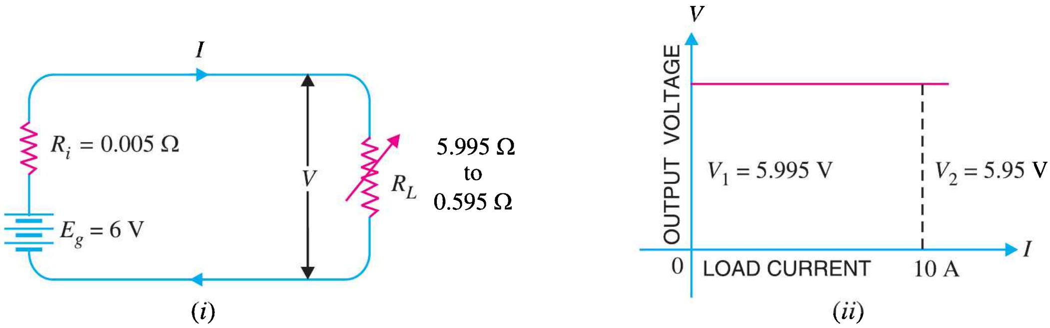


Fig. 1.10

In such a case, the output voltage nearly remains the same when load current changes. Fig. 1.10 (*i*) illustrates a constant voltage source. It is a d.c. source of 6 V with internal resistance R_i = 0.005 Ω. If the load current varies over a wide range of 1 to 10 A, for any of these values, the internal drop across R_i (= 0.005 Ω) is less than 0.05 volt. Therefore, the voltage output of the source is between 5.995 to 5.95 volts. This can be considered constant voltage compared with the wide variations in load current.

Fig. 1.10 (*ii*) shows the graph for a constant voltage source. It may be seen that the output voltage remains constant inspite of the changes in load current. Thus as the load current changes from 0 to 10 A, the output voltage essentially remains the same (*i.e.* $V_1 = V_2$). A constant voltage source is represented as shown in Fig. 1.11.

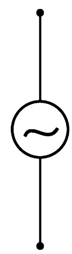


Fig. 1.11

Example 1.1. *A lead acid battery fitted in a truck develops 24V and has an internal resistance of 0.01 Ω. It is used to supply current to head lights etc. If the total load is equal to 100 watts, find :*

(i) voltage drop in internal resistance

(ii) terminal voltage

Solution.

Generated voltage, E_g = 24 V

Internal resistance, R_i = 0.01 Ω

Power supplied, P = 100 watts

(*i*) Let *I* be the load current.

Now $P = E_g \times I$ (∵ For an ideal source, $V \simeq E_g$)

∴ $I = \dfrac{P}{E_g} = \dfrac{100}{24} = 4.17$ A

∴ Voltage drop in R_i = $I R_i = 4.17 \times 0.01 =$ **0.0417 V**

(*ii*) Terminal Voltage, $V = E_g - I R_i$

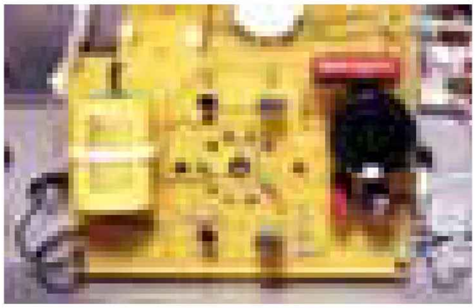
$= 24 - 0.0417 =$ **23.96 V**

---

* resistance in case of a d.c. source.

Comments : It is clear from the above example that when internal resistance of the source is quite small, the voltage drop in internal resistance is very low. Therefore, the terminal voltage substantially remains constant and the source behaves as a constant voltage source irrespective of load current variations.

1.10 Constant Current Source

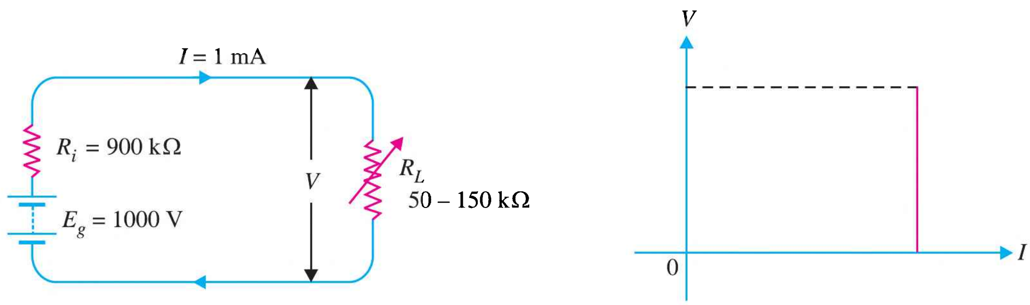


Constant Current Source

*A voltage source that has a very high internal *impedance as compared with external load impedance is considered as a* **constant current source.**

In such a case, the load current nearly remains the same when the output voltage changes. Fig. 1.12 (*i*) illustrates a constant current source. It is a d.c. source of 1000 V with internal resistance R_i = 900 kΩ. Here, load R_L varies over 3 : 1 range from 50 k Ω to 150 k Ω. Over this variation of load R_L, the circuit current I is essentially constant at 1.05 to 0.95 mA or approximately 1 mA. It may be noted that output voltage V varies approximately in the same 3 : 1 range as R_L, although load current essentially remains **constant at 1mA. The beautiful example of a constant current source is found in vacuum tube circuits where the tube acts as a generator having internal resistance as high as 1 MΩ.

Fig. 1.12 (*ii*) shows the graph of a constant current source. It is clear that current remains constant even when the output voltage changes substantially. The following points may be noted regarding the constant current source :



(*i*) (*ii*)

Fig. 1.12

(*i*) Due to high internal resistance of the source, the load current remains essentially constant as the load R_L is varied.

(*ii*) The output voltage varies approximately in the same range as R_L, although current remains constant.

(*iii*) The output voltage V is much less than the generated voltage E_g because of high $I R_i$ drop.

Fig. 1.13 shows the symbol of a constant current source.

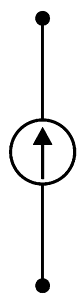


Fig. 1.13

* Resistance in case of a d.c. source

** Now $I = \dfrac{E_g}{R_L + R_i}$. Since $R_i >> R_L$, $I = \dfrac{E_g}{R_i}$

As both E_g and R_i are constants, I is constant.

Example 1.2. *A d.c. source generating 500 V has an internal resistance of 1000 Ω. Find the load current if load resistance is (i) 10 Ω (ii) 50 Ω and (iii) 100 Ω.*

Solution.

Generated voltage, E_g = 500 V

Internal resistance, R_i = 1000 Ω

(*i*) When R_L = 10 Ω

$$\text{Load current, } I = \frac{E_g}{R_L + R_i} = \frac{500}{10 + 1000} = \mathbf{0.495\ A}$$

(*ii*) When R_L = 50 Ω

$$\text{Load current, } I = \frac{500}{50 + 1000} = \mathbf{0.476\ A}$$

(*iii*) When R_L = 100 Ω

$$\text{Load current, } I = \frac{500}{100 + 1000} = \mathbf{0.454\ A}$$

It is clear from the above example that load current is essentially constant since $R_i >> R_L$.

1.11 Conversion of Voltage Source into Current Source

Fig. 1.14 shows a constant voltage source with voltage *V* and internal resistance R_i. Fig. 1.15 shows its equivalent current source. It can be easily shown that the two circuits behave electrically the same way under all conditions.

(*i*) If in Fig. 1.14, the load is open-circuited (*i.e.* $R_L \to \infty$), then voltage across terminals *A* and *B* is *V*. If in Fig. 1.15, the load is open-circuited (*i.e.* $R_L \to \infty$), then all current *I* (= V/R_i) flows through R_i, yielding voltage across terminals $AB = I R_i = V$. Note that open-circuited voltage across *AB* is *V* for both the circuits and hence they are electrically equivalent.

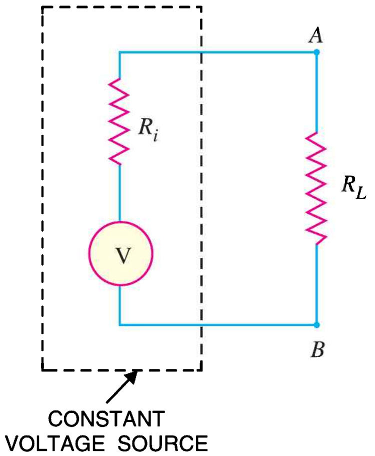


Fig. 1.14

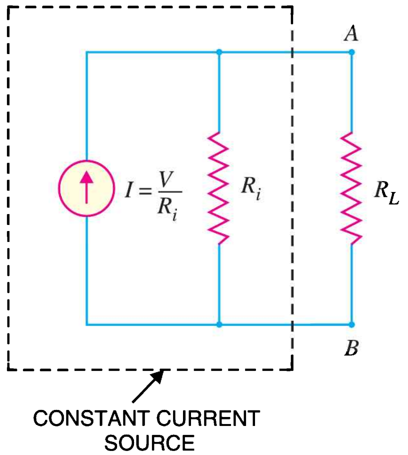


Fig. 1.15

(*ii*) If in Fig. 1.14, the load is short-circuited (*i.e.* $R_L = 0$), the short circuit current is given by:

$$I_{short} = \frac{V}{R_i}$$

If in Fig. 1.15, the load is short-circuited (*i.e.* $R_L = 0$), the current *I* (= V/R_i) bypasses R_i in favour of short-circuit. It is clear that current (= V/R_i) is the same for the two circuits and hence they are electrically equivalent.

Thus to convert a constant voltage source into a constant current source, the following procedure may be adopted :

(*a*) Place a short-circuit across the two terminals in question (terminals *AB* in the present case) and find the short-circuit current. Let it be *I*. Then *I* is the current supplied by the equivalent current source.

(*b*) Measure the resistance at the terminals with load removed and sources of *e.m.f.s* replaced by their internal resistances if any. Let this resistance be *R*.

(*c*) Then equivalent current source can be represented by a single current source of magnitude *I* in parallel with resistance *R*.

Note. To convert a current source of magnitude *I* in parallel with resistance *R* into voltage source,

Voltage of voltage source, $V = I R$

Resistance of voltage source, $R = R$

Thus voltage source will be represented as voltage *V* in series with resistance *R*.

Example 1.3. *Convert the constant voltage source shown in Fig. 1.16 into constant current source.*

Solution. The solution involves the following steps :

(*i*) Place a short across *AB* in Fig. 1.16 and find the short-circuit current *I*.

Clearly, $I = 10/10 = 1$ A

Therefore, the equivalent current source has a magnitude of 1 A.

(*ii*) Measure the resistance at terminals *AB* with load *removed and 10 V source replaced by its internal resistance. The 10 V source has negligible resistance so that resistance at terminals *AB* is $R = 10\ \Omega$.

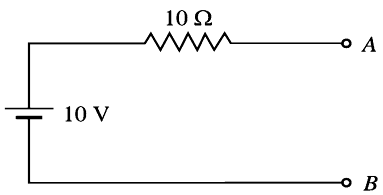


Fig. 1.16

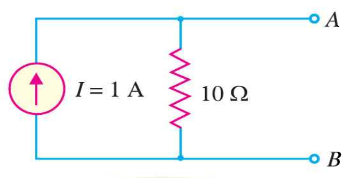


Fig. 1.17

(*iii*) The equivalent current source is a source of 1 A in parallel with a resistance of 10 Ω as shown in Fig. 1.17.

Example 1.4. *Convert the constant current source in Fig. 1.18 into equivalent voltage source.*

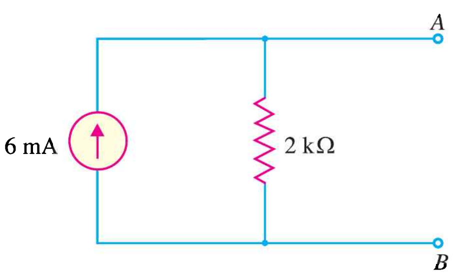


Fig. 1.18

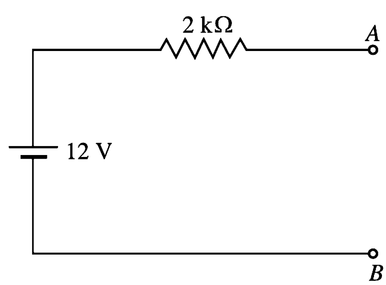


Fig. 1.19

Solution. The solution involves the following steps :

* Fortunately, no load is connected across *AB*. Had there been load across *AB*, it would have been removed.

(*i*) To get the voltage of the voltage source, multiply the current of the current source by the internal resistance *i.e.*

Voltage of voltage source $= IR = 6\text{ mA} \times 2\text{ k}\Omega = 12\text{V}$

(*ii*) The internal resistance of voltage source is 2 k Ω.

The equivalent voltage source is a source of 12 V in series with a resistance of 2 k Ω as shown in Fig. 1.19.

Note. The voltage source should be placed with +ve terminal in the direction of current flow.

1.12 Maximum Power Transfer Theorem

When load is connected across a voltage source, power is transferred from the source to the load. The amount of power transferred will depend upon the load resistance. If load resistance R_L is made equal to the internal resistance R_i of the source, then maximum power is transferred to the load R_L. This is known as *maximum power transfer theorem* and can be stated as follows :

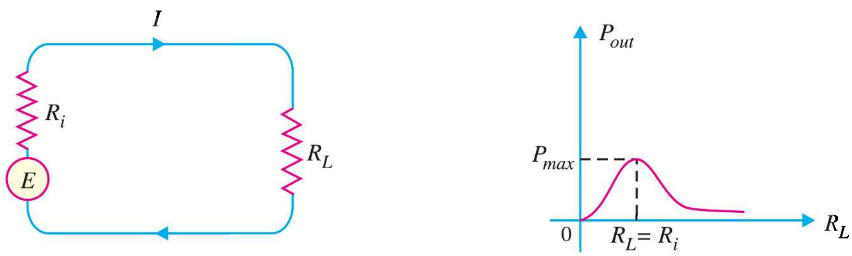
Maximum power is transferred from a source to a load when the load resistance is made equal to the internal resistance of the source.

This applies to d.c. as well as a.c. power.*

To prove this theorem mathematically, consider a voltage source of generated voltage E and internal resistance R_i and delivering power to a load resistance R_L [See Fig. 1.20 (*i*)]. The current I flowing through the circuit is given by :

$$I = \frac{E}{R_L + R_i}$$

Power delivered to the load, $$P = I^2 R_L = \left(\frac{E}{R_L + R_i}\right)^2 R_L \quad ...(i)$$



(*i*) (*ii*)

Fig. 1.20

For a given source, generated voltage E and internal resistance R_i are constant. Therefore, power delivered to the load depends upon R_L. In order to find the value of R_L for which the value of P is maximum, it is necessary to differentiate eq. (*i*) w.r.t. R_L and set the result equal to zero.

Thus, $$\frac{dP}{dR_L} = E^2\left[\frac{(R_L + R_i)^2 - 2R_L(R_L + R_i)}{(R_L + R_i)^4}\right] = 0$$

or $$(R_L + R_i)^2 - 2R_L(R_L + R_i) = 0$$

or $$(R_L + R_i)(R_L + R_i - 2R_L) = 0$$

or $$(R_L + R_i)(R_i - R_L) = 0$$

* As power is concerned with resistance only, therefore, this is true for both a.c. and d.c. power.

Since $(R_L + R_i)$ cannot be zero,

$$\therefore \quad R_i - R_L = 0$$

or $$R_L = R_i$$

i.e. *Load resistance* = *Internal resistance*

Thus, for maximum power transfer, load resistance R_L must be equal to the internal resistance R_i of the source.

Under such conditions, the load is said to be *matched* to the source. Fig. 1.20 (*ii*) shows a graph of power delivered to R_L as a function of R_L. It may be mentioned that efficiency of maximum power transfer is *50% as one-half of the total generated power is dissipated in the internal resistance R_i of the source.

Applications. Electric power systems never operate for maximum power transfer because of low efficiency and high voltage drops between generated voltage and load. However, in the electronic circuits, maximum power transfer is usually desirable. For instance, in a public address system, it is desirable to have load (*i.e.* speaker) "matched" to the amplifier so that there is maximum transference of power from the amplifier to the speaker. In such situations, efficiency is **sacrificed at the cost of high power transfer.

Example 1.5. *A generator develops 200 V and has an internal resistance of 100 Ω. Find the power delivered to a load of (i) 100 Ω (ii) 300 Ω. Comment on the result.*

Solution.

Generated voltage, $E = 200$ V

Internal resistance, $R_i = 100\ \Omega$

(*i*) When load $R_L = 100\ \Omega$

$$\text{Load current, } I = \frac{E}{R_L + R_i} = \frac{200}{100 + 100} = 1 \text{ A}$$

$$\therefore \quad \text{Power delivered to load} = I^2 R_L = (1)^2 \times 100 = \mathbf{100\ watts}$$

$$\text{Total power generated} = I^2 (R_L + R_i) = 1^2 (100 + 100) = 200 \text{ watts}$$

Thus, out of 200 W power developed by the generator, only 100W has reached the load *i.e.* efficiency is 50% only.

(*ii*) When load $R_L = 300\ \Omega$

$$\text{Load current, } I = \frac{E}{R_L + R_i} = \frac{200}{300 + 100} = 0.5 \text{ A}$$

$$\text{Power delivered to load} = I^2 R_L = (0.5)^2 \times 300 = \mathbf{75\ watts}$$

$$\text{Total power generated} = I^2 (R_L + R_i) = (0.5)^2 (300 + 100) = 100 \text{ watts}$$

Thus, out of 100 watts of power produced by the generator, 75 watts is transferred to the load *i.e.* efficiency is 75%.

Comments. Although in case of $R_L = R_i$, a large power (100 W) is transferred to the load, but there is a big wastage of power in the generator. On the other hand, when R_L is ***not*** equal to R_i, the

* $$\text{Efficiency} = \frac{\text{output power}}{\text{input power}} = \frac{I^2 R_L}{I^2 (R_L + R_i)}$$

$$= R_L / 2R_L = 1/2 = 50\% \qquad (\because R_L = R_i)$$

** Electronic devices develop small power. Therefore, if too much efficiency is sought, a large number of such devices will have to be connected in series to get the desired output. This will distort the output as well as increase the cost and size of equipment.

power transfer is less (75 W) but smaller part is wasted in the generator *i.e.* efficiency is high. Thus, it depends upon a particular situation as to what the load should be. If we want to transfer maximum power (*e.g.* in amplifiers) irrespective of efficiency, we should make $R_L = R_i$. However, if efficiency is more important (*e.g.* in power systems), then internal resistance of the source should be considerably smaller than the load resistance.

Example 1.6. *An audio amplifier produces an alternating output of 12 V before the connection to a load. The amplifier has an equivalent resistance of 15 Ω at the output. What resistance the load need to have to produce maximum power ? Also calculate the power output under this condition.*

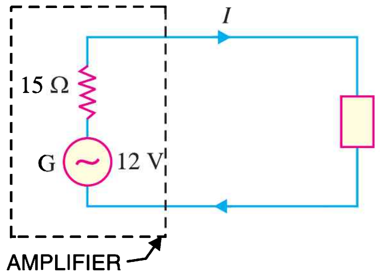


Fig. 1.21

Solution. In order to produce maximum power, the load (*e.g.* a speaker) should have a resistance of 15 Ω to match the amplifier. The equivalent circuit is shown in Fig. 1.21.

$$\therefore \quad \text{Load required, } R_L = \mathbf{15\ \Omega}$$

$$\text{Circuit current, } I = \frac{V}{R_T} = \frac{12}{15 + 15} = 0.4 \text{ A}$$

$$\text{Power delivered to load, } P = I^2 R_L = (0.4)^2 \times 15 = \mathbf{2.4\ W}$$

Example 1.7. *For the a.c. generator shown in Fig. 1.22 (i), find (i) the value of load so that maximum power is transferred to the load (ii) the value of maximum power.*

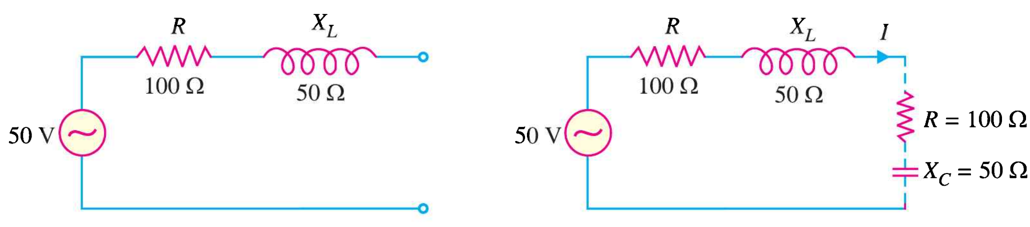


Fig. 1.22

Solution.

(*i*) In a.c. system, maximum power is delivered to the load impedance (Z_L) when load impedance is conjugate of the internal impedance (Z_i) of the source. Now in the problem, $Z_i = (100 + j50)\Omega$. For maximum power transfer, the load impedance should be conjugate of internal impedance *i.e.* Z_L should be $(100 - j50)\ \Omega$. This is shown in dotted line in Fig. 1.22 (*ii*).

$$\therefore \quad Z_L = \mathbf{(100 - j50)\ \Omega}$$

(*ii*) Total impedance, $Z_T = Z_i + Z_L = (100 + j50) + (100 - j50) = 200\ \Omega$*

$$\text{Circuit current, } I = \frac{V}{Z_T} = \frac{50}{200} = 0.25 \text{ A}$$

Maximum power transferred to the load $= I^2 R_L = (0.25)^2 \times 100 = \mathbf{6.25\ W}$

* Note that by making internal impedance and load impedance conjugate, the reactive terms cancel. The circuit then consists of internal and external resistances only. This is quite logical because power is only consumed in resistances as reactances (X_L or X_C) consume no power.

1.13 Thevenin's Theorem

Sometimes it is desirable to find a particular branch current in a circuit as the resistance of that branch is varied while all other resistances and voltage sources remain constant. For instance, in the circuit shown in Fig. 1.23, it may be desired to find the current through R_L for five values of R_L, assuming that R_1, R_2, R_3 and E remain constant. In such situations, the *solution can be obtained readily by applying *Thevenin's theorem* stated below :

Any two-terminal network containing a number of e.m.f. sources and resistances can be replaced by an equivalent series circuit having a voltage source E_0 in series with a resistance R_0 where,

E_0 = open circuited voltage between the two terminals.

R_0 = the resistance between two terminals of the circuit obtained by looking "in" at the terminals with load removed and voltage sources replaced by their internal resistances, if any.

To understand the use of this theorem, consider the two-terminal circuit shown in Fig. 1.23. The circuit enclosed in the dotted box can be replaced by one voltage E_0 in series with resistance R_0 as shown in Fig. 1.24. The behaviour at the terminals AB and $A'B'$ is the same for the two circuits, independent of the values of R_L connected across the terminals.

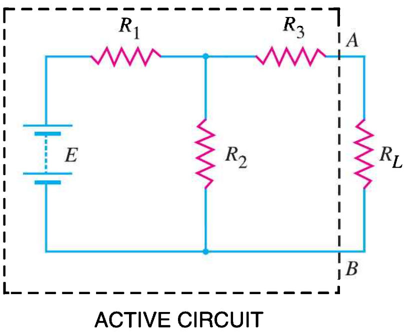


Fig. 1.23

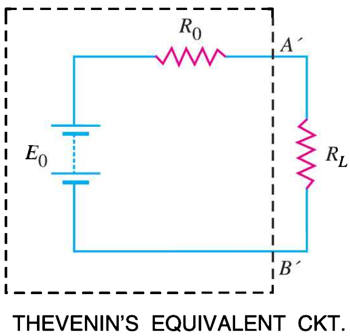


Fig. 1.24

(*i*) Finding $\mathbf{E_0}$. This is the voltage between terminals A and B of the circuit when load R_L is removed. Fig. 1.25 shows the circuit with load removed. The voltage drop across R_2 is the desired voltage E_0.

$$\text{Current through } R_2 = \frac{E}{R_1 + R_2}$$

$$\therefore \quad \text{Voltage across } R_2,\ E_0 = \left(\frac{E}{R_1 + R_2}\right) R_2$$

Thus, voltage E_0 is determined.

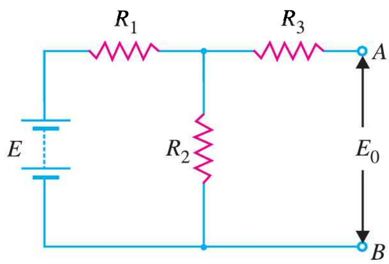


Fig. 1.25

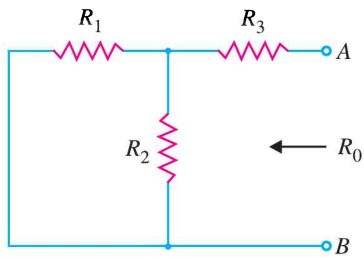


Fig. 1.26

* Solution can also be obtained by applying Kirchhoff's laws but it requires a lot of labour.

(ii) **Finding R_0.** This is the resistance between terminals *A* and *B* with load removed and e.m.f. reduced to zero (See Fig. 1.26).

∴ Resistance between terminals *A* and *B* is

$$R_0 = \text{parallel combination of } R_1 \text{ and } R_2 \text{ in series with } R_3$$

$$= \frac{R_1 R_2}{R_1 + R_2} + R_3$$

Thus, the value of R_0 is determined. Once the values of E_0 and R_0 are determined, then the current through the load resistance R_L can be found out easily (Refer to Fig. 1.24).

1.14 Procedure for Finding Thevenin Equivalent Circuit

(i) Open the two terminals (*i.e.* remove any load) between which you want to find Thevenin equivalent circuit.

(ii) Find the open-circuit voltage between the two open terminals. It is called Thevenin voltage E_0.

(iii) Determine the resistance between the two open terminals with all ideal voltage sources shorted and all ideal current sources opened (a non-ideal source is replaced by its internal resistance). It is called Thevenin resistance R_0.

(iv) Connect E_0 and R_0 in series to produce Thevenin equivalent circuit between the two terminals under consideration.

(v) Place the load resistor removed in step (*i*) across the terminals of the Thevenin equivalent circuit. The load current can now be calculated using only Ohm's law and it has the same value as the load current in the original circuit.

Example 1.8. *Using Thevenin's theorem, find the current through 100 Ω resistance connected across terminals A and B in the circuit of Fig. 1.27.*

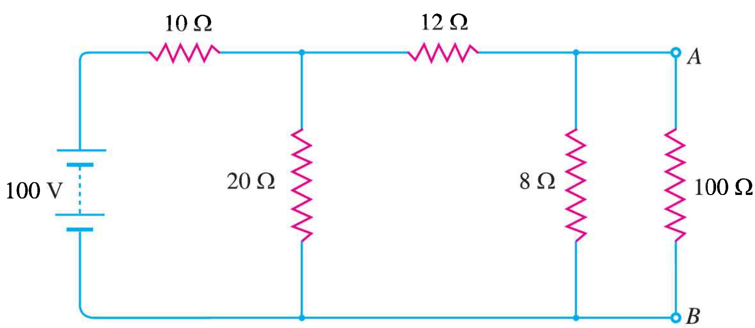


Fig. 1.27

Solution.

(i) **Finding E_0.** It is the voltage across terminals *A* and *B* with 100 Ω resistance removed as shown in Fig. 1.28.

$$E_0 = (\text{Current through } 8\ \Omega) \times 8\ \Omega = 2.5^* \times 8 = 20 \text{ V}$$

* By solving this series-parallel circuit.

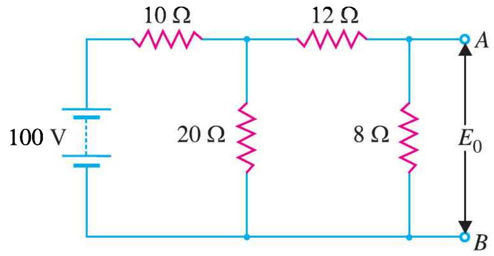


Fig. 1.28

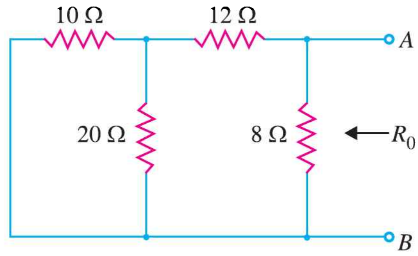


Fig. 1.29

(*ii*) **Finding R_0.** It is the resistance between terminals *A* and *B* with 100 Ω removed and voltage source short circuited as shown in Fig. 1.29.

R_0 = Resistance looking in at terminals *A* and *B* in Fig. 1.29

$$= \frac{\left[\frac{10 \times 20}{10 + 20} + 12\right] 8}{\left[\frac{10 \times 20}{10 + 20} + 12\right] + 8}$$

$= 5.6\ \Omega$

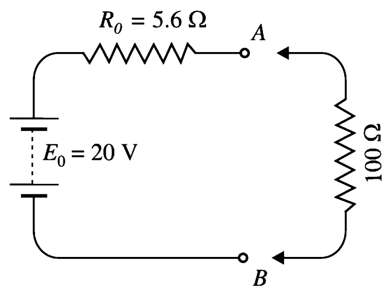


Fig. 1.30

Therefore, Thevenin's equivalent circuit will be as shown in Fig. 1.30. Now, current through 100 Ω resistance connected across terminals *A* and *B* can be found by applying Ohm's law.

$$\text{Current through 100 } \Omega \text{ resistor} = \frac{E_0}{R_0 + R_L} = \frac{20}{5.6 + 100} = \mathbf{0.19\ A}$$

Example 1.9. *Find the Thevenin's equivalent circuit for Fig. 1.31.*

Solution. The Thevenin's voltage E_0 is the voltage across terminals *A* and *B*. This voltage is equal to the voltage across R_3. It is because terminals *A* and *B* are open circuited and there is no current flowing through R_2 and hence no voltage drop across it.

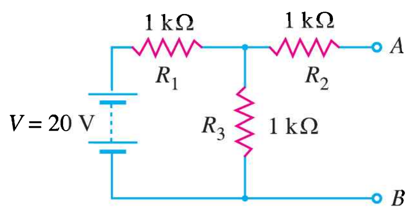


Fig. 1.31

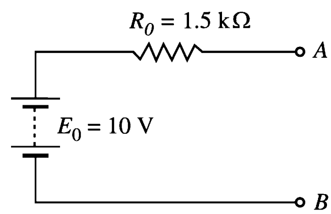


Fig. 1.32

$$\therefore \quad E_0 = \text{Voltage across } R_3$$

$$= \frac{R_3}{R_1 + R_3} \times V = \frac{1}{1+1} \times 20 = 10\ \text{V}$$

The Thevenin's resistance R_0 is the resistance measured between terminals *A* and *B* with no load (*i.e.* open at terminals *A* and *B*) and voltage source replaced by a short circuit.

$$\therefore \quad R_0 = R_2 + \frac{R_1 R_3}{R_1 + R_3} = 1 + \frac{1 \times 1}{1 + 1} = 1.5\ \text{k}\Omega$$

Therefore, Thevenin's equivalent circuit will be as shown in Fig. 1.32.

Example 1.10. *Calculate the value of load resistance R_L to which maximum power may be transferred from the circuit shown in Fig. 1.33 (i). Also find the maximum power.*

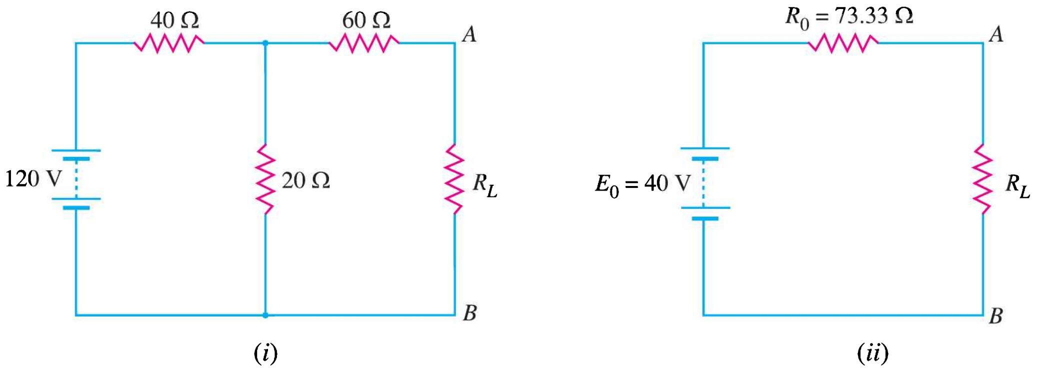


Fig. 1.33

Solution. We shall first find Thevenin's equivalent circuit to the left of terminals *AB* in Fig. 1.33 (*i*).

E_0 = Voltage across terminals *AB* with R_L removed

$$= \frac{120}{40 + 20} \times 20 = 40 \text{ V}$$

R_0 = Resistance between terminals *A* and *B* with R_L removed and 120 V source replaced by a short

$$= 60 + (40\ \Omega \parallel 20\ \Omega) = 60 + (40 \times 20)/60 = 73.33\ \Omega$$

The Thevenin's equivalent circuit to the left of terminals *AB* in Fig. 1.33 (*i*) is E_0 (= 40 V) in series with R_0 (= 73.33 Ω). When R_L is connected between terminals *A* and *B*, the circuit becomes as shown in Fig. 1.33 (*ii*). It is clear that maximum power will be transferred when

$$R_L = R_0 = \mathbf{73.33\ \Omega}$$

$$\text{Maximum power to load} = \frac{E_0^2}{4 R_L} = \frac{(40)^2}{4 \times 73.33} = \mathbf{5.45\ W}$$

Comments. This shows another advantage of Thevenin's equivalent circuit of a network. Once Thevenin's equivalent resistance R_0 is calculated, it shows at a glance the condition for maximum power transfer. Yet Thevenin's equivalent circuit conveys another information. Thus referring to Fig. 1.33 (*ii*), the maximum voltage that can appear across terminals *A* and *B* is 40 V. This is not so obvious from the original circuit shown in Fig. 1.33 (*i*).

Example 1.11. *Calculate the current in the 50 Ω resistor in the network shown in Fig. 1.34.*

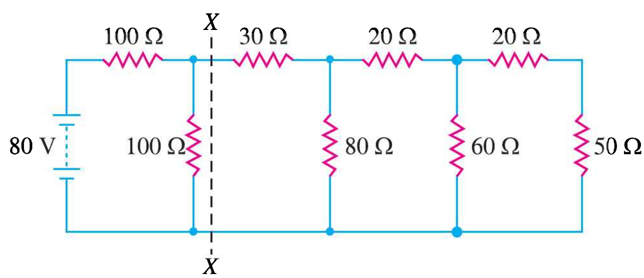


Fig. 1.34

Solution. We shall simplify the circuit shown in Fig. 1.34 by the repeated use of Thevenin's theorem. We first find Thevenin's equivalent circuit to the left of *XX.

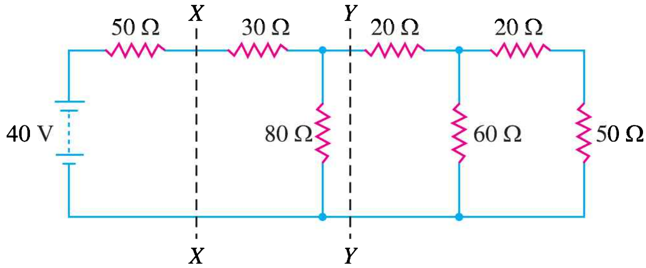


Fig. 1.35

$$E_0 = \frac{80}{100 + 100} \times 100 = 40\text{V}$$

$$R_0 = 100 \parallel 100 = \frac{100 \times 100}{100 + 100} = 50\ \Omega$$

Therefore, we can replace the circuit to the left of XX in Fig. 1.34 by its Thevenin's equivalent circuit *viz.* E_0 (= 40V) in series with R_0 (= 50 Ω). The original circuit of Fig. 1.34 then reduces to the one shown in Fig. 1.35.

We shall now find Thevenin's equivalent circuit to left of YY in Fig. 1.35.

$$E_0' = \frac{40}{50 + 30 + 80} \times 80 = 20\text{ V}$$

$$R_0' = (50 + 30) \parallel 80 = \frac{80 \times 80}{80 + 80} = 40\ \Omega$$

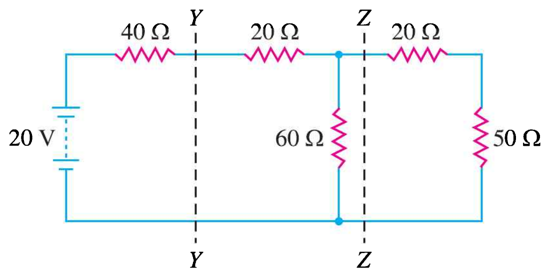
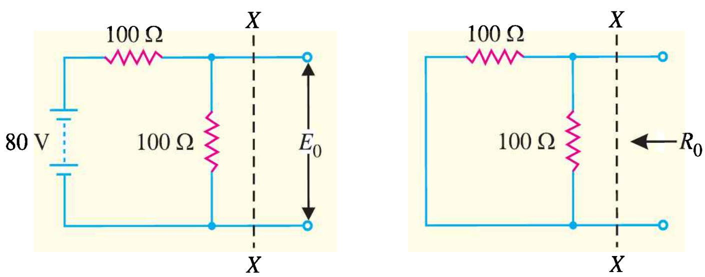


Fig. 1.36

We can again replace the circuit to the left of YY in Fig. 1.35 by its Thevenin's equivalent circuit. Therefore, the original circuit reduces to that shown in Fig. 1.36.

---

*



(*a*) (*b*)

$$E_0 = \text{Current in } 100\ \Omega \times 100\ \Omega = \frac{80}{100 + 100} \times 100 = 40\text{V} \quad [\text{See Fig. } (a)]$$

$$R_0 = \text{Resistance looking in the open terminals in Fig. } (b)$$

$$= 100 \parallel 100 = \frac{100 \times 100}{100 + 100} = 50\ \Omega$$

Using the same procedure to the left of *ZZ*, we have,

$$E''_0 = \frac{20}{40 + 20 + 60} \times 60 = 10\text{V}$$

$$R''_0 = (40 + 20) \parallel 60 = \frac{60 \times 60}{60 + 60} = 30\ \Omega$$

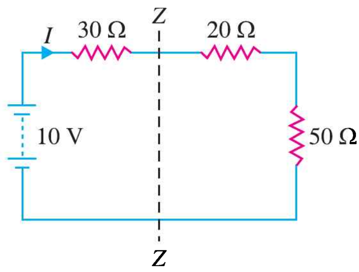


Fig. 1.37

The original circuit then reduces to that shown in Fig. 1.37. By Ohm's law, current *I* in 50 Ω resistor is

$$I = \frac{10}{30 + 20 + 50} = \mathbf{0.1\ A}$$

1.15 Norton's Theorem

Fig. 1.38 (*i*) shows a network enclosed in a box with two terminals *A* and *B* brought out. The network in the box may contain any number of resistors and e.m.f. sources connected in any manner. But according to Norton, the entire circuit behind terminals *A* and *B* can be replaced by a current source of output I_N in parallel with a single resistance R_N as shown in Fig. 1.38 (*ii*). The value of I_N is determined as mentioned in Norton's theorem. The resistance R_N is the same as Thevenin's resistance R_0. Once Norton's equivalent circuit is determined [See Fig. 1.38 (*ii*)], then current through any load R_L connected across terminals *AB* can be readily obtained.

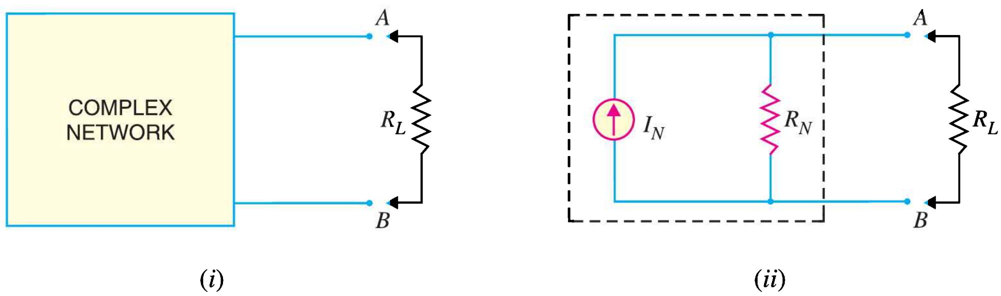


Fig. 1.38

Hence Norton's theorem as applied to d.c. circuits may be stated as under :

Any network having two terminals A and B can be replaced by a current source of output I_N in parallel with a resistance R_N.

(*i*) *The output I_N of the current source is equal to the current that would flow through AB when terminals A and B are short circuited.*

(*ii*) *The resistance R_N is the resistance of the network measured between terminals A and B with load (R_L) removed and sources of e.m.f. replaced by their internal resistances, if any.*

Norton's theorem is *converse* of Thevenin's theorem in that Norton equivalent circuit uses a current generator instead of voltage generator and resistance R_N (which is the same as R_0) in parallel with the generator instead of being in series with it.

Illustration. Fig. 1.39 illustrates the application of Norton's theorem. As far as circuit behind terminals *AB* is concerned [See Fig. 1.39 (*i*)], it can be replaced by a current source of output I_N in parallel with a resistance R_N as shown in Fig. 1.39 (*iv*). The output I_N of the current generator is equal to the current that would flow through *AB* when terminals *A* and *B* are short-circuited as shown in Fig. 1.39 (*ii*). The load *R′* on the source when terminals *AB* are short-circuited is given by :

$$R' = R_1 + \frac{R_2 R_3}{R_2 + R_3} = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_2 + R_3}$$

$$\text{Source current, } I' = \frac{V}{R'} = \frac{V(R_2 + R_3)}{R_1 R_2 + R_1 R_3 + R_2 R_3}$$

$$\text{Short-circuit current, } I_N = \text{Current in } R_2 \text{ in Fig. 1.39 } (ii)$$

$$= I' \times \frac{R_3}{R_2 + R_3} = \frac{V R_3}{R_1 R_2 + R_1 R_3 + R_2 R_3}$$

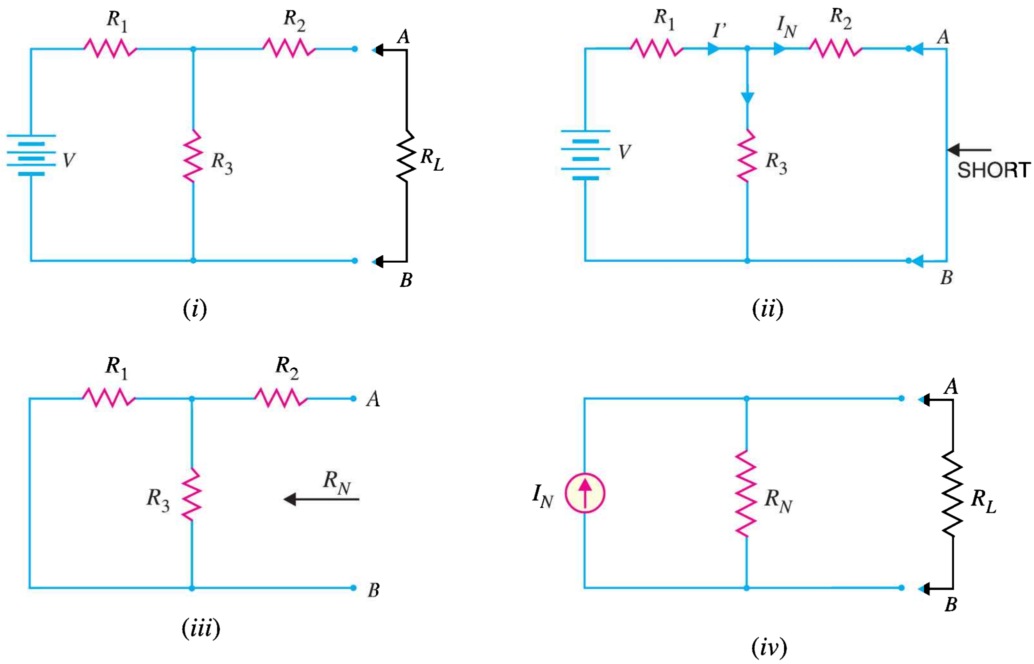


Fig. 1.39

To find R_N, remove the load R_L and replace the voltage source by a short circuit because its resistance is assumed zero [See Fig. 1.39 (*iii*)].

$$\therefore \quad R_N = \text{Resistance at terminals } AB \text{ in Fig. 1.39 } (iii).$$

$$= R_2 + \frac{R_1 R_3}{R_1 + R_3}$$

Thus the values of I_N and R_N are known. The Norton equivalent circuit will be as shown in Fig. 1.39 (*iv*).

1.16 Procedure for Finding Norton Equivalent Circuit

(i) Open the two terminals (*i.e.* remove any load) between which we want to find Norton equivalent circuit.

(ii) Put a short-circuit across the terminals under consideration. Find the short-circuit current flowing in the short circuit. It is called Norton current I_N.

(iii) Determine the resistance between the two open terminals with all ideal voltage sources shorted and all ideal current sources opened (a non-ideal source is replaced by its internal resistance). It is called Norton's resistance R_N. It is easy to see that $R_N = R_0$.

(iv) Connect I_N and R_N in parallel to produce Norton equivalent circuit between the two terminals under consideration.

(*v*) Place the load resistor removed in step (*i*) across the terminals of the Norton equivalent circuit. The load current can now be calculated by using current-divider rule. This load current will be the same as the load current in the original circuit.

Example 1.12. *Using Norton's theorem, find the current in 8 Ω resistor in the network shown in Fig. 1.40 (i).*

Solution. We shall reduce the network to the left of *AB* in Fig. 1.40 (*i*) to Norton's equivalent circuit. For this purpose, we are required to find I_N and R_N.

(*i*) With load (*i.e.*, 8 Ω) removed and terminals *AB* short circuited [See Fig. 1.40 (*ii*)], the current that flows through *AB* is equal to I_N. Referring to Fig. 1.40 (*ii*),

$$\text{Load on the source} = 4\ \Omega + 5\ \Omega \parallel 6\ \Omega = 4 + \frac{5 \times 6}{5 + 6} = 6.727\ \Omega$$

$$\text{Source current, } I' = 40/6.727 = 5.94\text{ A}$$

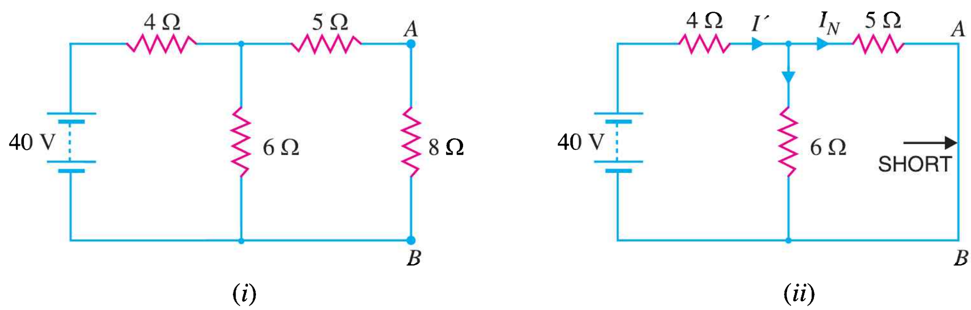


Fig. 1.40

$$\therefore \quad \text{Short-circuit current in } AB,\ I_N = I' \times \frac{6}{6+5} = 5.94 \times 6/11 = 3.24\text{ A}$$

(*ii*) With load (*i.e.*, 8 Ω) removed and battery replaced by a short (since its internal resistance is assumed zero), the resistance at terminals *AB* is equal to R_N as shown in Fig. 1.41 (*i*).

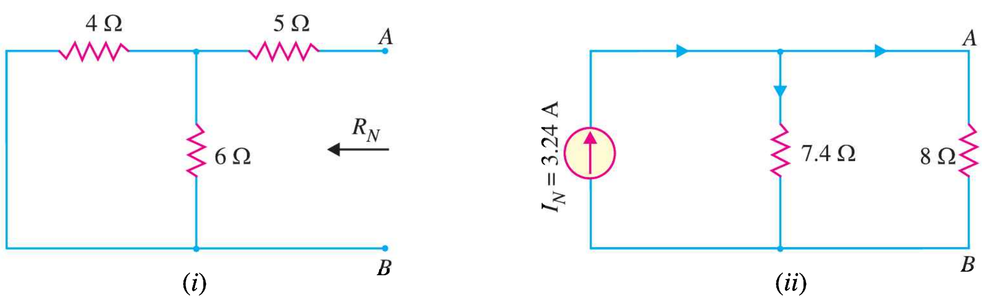


Fig. 1.41

$$R_N = 5\ \Omega + 4\ \Omega \parallel 6\ \Omega = 5 + \frac{4 \times 6}{4 + 6} = 7.4\ \Omega$$

The Norton's equivalent circuit behind terminals *AB* is I_N (= 3.24 A) in parallel with R_N (= 7.4 Ω). When load (*i.e.*, 8 Ω) is connected across terminals *AB*, the circuit becomes as shown in Fig. 1.41 (*ii*). The current source is supplying current to two resistors 7.4 Ω and 8 Ω in parallel.

$$\therefore \quad \text{Current in 8 }\Omega\text{ resistor} = 3.24 \times \frac{7.4}{8 + 7.4} = \mathbf{1.55\ A}$$

Example 1.13. *Find the Norton equivalent circuit at terminals X – Y in Fig. 1.42.*

Solution. We shall first find the Thevenin equivalent circuit and then convert it to an equivalent current source. This will then be Norton equivalent circuit.

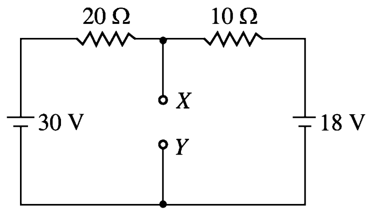


Fig. 1.42

Finding Thevenin Equivalent circuit. To find E_0, refer to Fig. 1.43 (*i*). Since 30 V and 18 V sources are in opposition, the circuit current I is given by :

$$I = \frac{30 - 18}{20 + 10} = \frac{12}{30} = 0.4 \text{ A}$$

Applying Kirchhoff's voltage law to loop *ABCDA*, we have,

$$30 - 20 \times 0.4 - E_0 = 0 \qquad \therefore \; E_0 = 30 - 8 = 22 \text{ V}$$

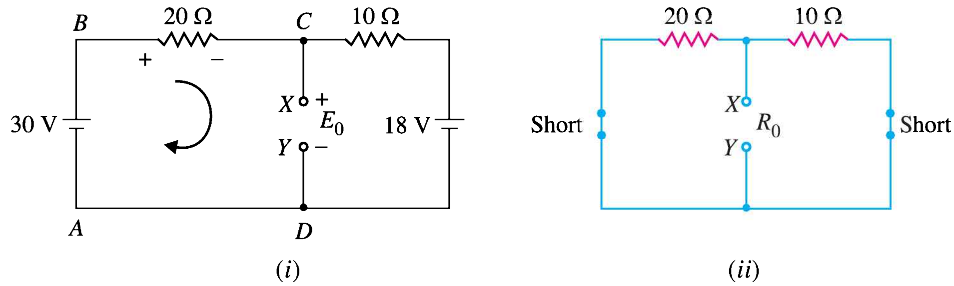


Fig. 1.43

To find R_0, we short both voltage sources as shown in Fig. 1.43 (*ii*). Notice that 10 Ω and 20 Ω resistors are then in parallel.

$$\therefore \qquad R_0 = 10\ \Omega \parallel 20\ \Omega = \frac{10 \times 20}{10 + 20} = 6.67\Omega$$

Therefore, Thevenin equivalent circuit will be as shown in Fig. 1.44 (*i*). Now it is quite easy to convert it into equivalent current source.

$$I_N = \frac{E_0}{R_0} = \frac{22}{6.67} = 3.3\text{A} \qquad \text{[See Fig. 1.44 (}ii\text{)]}$$

$$R_N = R_0 = 6.67\ \Omega$$

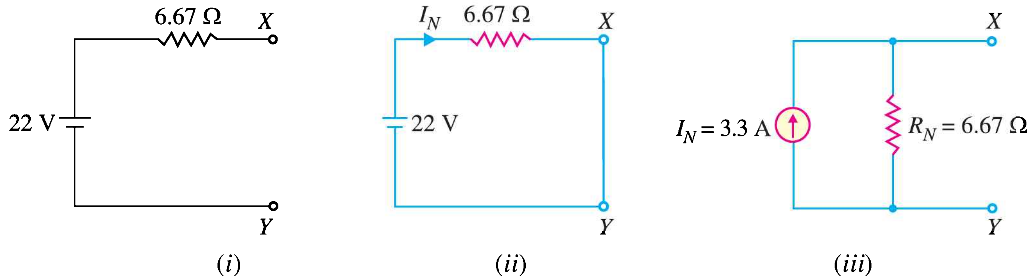


Fig. 1.44

Fig. 1.44 (*iii*) shows Norton equivalent circuit. Observe that the Norton equivalent resistance has the same value as the Thevenin equivalent resistance. Therefore, R_N is found exactly the same way.

Example 1.14. *Show that when Thevenin's equivalent circuit of a network is converted into Norton's equivalent circuit,* $I_N = E_0/R_0$ *and* $R_N = R_0$. *Here* E_0 *and* R_0 *are Thevenin voltage and Thevenin resistance respectively.*

Solution. Fig. 1.45 (*i*) shows a network enclosed in a box with two terminals *A* and *B* brought out. Thevenin's equivalent circuit of this network will be as shown in Fig. 1.45 (*ii*). To find Norton's equivalent circuit, we are to find I_N and R_N. Referring to Fig. 1.45 (*ii*),

I_N = Current flowing through short-circuited *AB* in Fig. 1.45 (*ii*)

$= E_0/R_0$

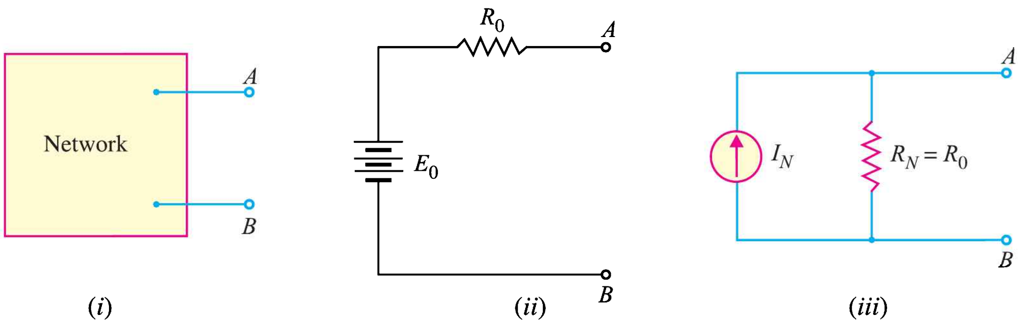
R_N = Resistance at terminals *AB* in Fig. 1.45 (*ii*)

$= R_0$

Fig. 1.45 (*iii*) shows Norton's equivalent circuit. Hence we arrive at the following two important conclusions :

(*i*) To convert Thevenin's equivalent circuit into Norton's equivalent circuit,

$$I_N = E_0/R_0 \quad ; \quad R_N = R_0$$



(*i*) (*ii*) (*iii*)

Fig. 1.45

(*ii*) To convert Norton's equivalent circuit into Thevenin's equivalent circuit,

$$E_0 = I_N R_N \quad ; \quad R_0 = R_N$$

1.17 Chassis and Ground

It is the usual practice to mount the electronic components on a metal base called *chassis*. For example, in Fig. 1.46, the voltage source and resistors are connected to the chassis. As the resistance of chassis is very low, therefore, it provides a conducting path and may be considered as a piece of wire.

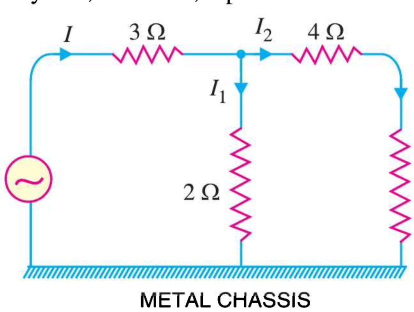


Fig. 1.46

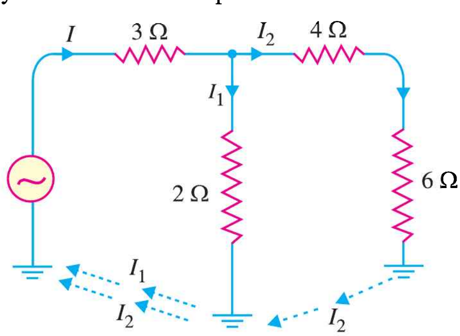


Fig. 1.47

It is customary to refer to the chassis as *ground*. Fig. 1.47 shows the symbol for chassis. It may be seen that all points connected to chassis are shown as grounded and represent the same potential. The adoption of this scheme (*i.e.* showing points of same potential as grounded) often simplifies the electronic circuits. In our further discussion, we shall frequently use this scheme.

MULTIPLE-CHOICE QUESTIONS

1. The outermost orbit of an atom can have a maximum of .... electrons.
(*i*) 8 (*ii*) 6
(*iii*) 4 (*iv*) 3

2. When the outermost orbit of an atom has less than 4 electrons, the material is generally a ....
(*i*) non-metal (*ii*) metal
(*iii*) semiconductor (*iv*) none of above

3. The valence electrons have ....
(*i*) very small energy
(*ii*) least energy
(*iii*) maximum energy
(*iv*) none of the above

4. A large number of free electrons exist in ....
(*i*) semiconductors (*ii*) metals
(*iii*) insulators (*iv*) non-metals

5. An ideal voltage source has .... internal resistance.
(*i*) small (*ii*) large
(*iii*) infinite (*iv*) zero

6. An ideal current source has .... internal resistance.
(*i*) infinite (*ii*) zero
(*iii*) small (*iv*) none of the above

7. Maximum power is transferred if load resistance is equal to .... of the source.
(*i*) half the internal resistance
(*ii*) internal resistance
(*iii*) twice the internal resistance
(*iv*) none of the above

8. Efficiency at maximum power transfer is ....
(*i*) 75% (*ii*) 25%
(*iii*) 90% (*iv*) 50%

9. When the outermost orbit of an atom has exactly 4 valence electrons, the material is generally ....
(*i*) a metal (*ii*) a non-metal
(*iii*) a semiconductor
(*iv*) an insulator

10. Thevenin's theorem replaces a complicated circuit facing a load by an ....
(*i*) ideal voltage source and parallel resistor
(*ii*) ideal current source and parallel resistor
(*iii*) ideal current source and series resistor
(*iv*) ideal voltage source and series resistor

11. The output voltage of an ideal voltage source is ....
(*i*) zero (*ii*) constant
(*iii*) dependent on load resistance
(*iv*) dependent on internal resistance

12. The current output of an ideal current source is ....
(*i*) zero (*ii*) constant
(*iii*) dependent on load resistance
(*iv*) dependent on internal resistance

13. Norton's theorem replaces a complicated circuit facing a load by an ....
(*i*) ideal voltage source and parallel resistor
(*ii*) ideal current source and parallel resistor
(*iii*) ideal voltage source and series resistor
(*iv*) ideal current source and series resistor

14. The practical example of ideal voltage source is ....
(*i*) lead-acid cell (*ii*) dry cell
(*iii*) Daniel cell (*iv*) none of the above

15. The speed of electrons in vacuum is .... than in a conductor.
(*i*) less (*ii*) much more
(*iii*) much less (*iv*) none of the above

16. Maximum power will be transferred from a source of 10 Ω resistance to a load of ....
(*i*) 5 Ω (*ii*) 20 Ω
(*iii*) 10 Ω (*iv*) 40 Ω

17. When the outermost orbit of an atom has more than 4 electrons, the material is generally a ....
(*i*) metal (*ii*) non-metal
(*iii*) semiconductor (*iv*) none of the above

18. An ideal source consists of 5 V in series with 10 kΩ resistance. The current magnitude of equivalent current source is ....
(*i*) 2 mA (*ii*) 3.5 mA
(*iii*) 0.5 mA (*iv*) none of the above

19. To get Thevenin voltage, you have to ....
(*i*) short the load resistor
(*ii*) open the load resistor
(*iii*) short the voltage source
(*iv*) open the voltage source

20. To get the Norton current, you have to ....
(*i*) short the load resistor
(*ii*) open the load resistor
(*iii*) short the voltage source
(*iv*) open the voltage source

21. The open-circuited voltage at the terminals of load R_L in a network is 30 V. Under the conditions of maximum power transfer, the load voltage will be ....
(*i*) 30 V (*ii*) 10 V
(*iii*) 5 V (*iv*) 15 V

22. Under the conditions of maximum power transfer, a voltage source is delivering a power of 30 W to the load. The power produced by the source is ....
(*i*) 45 W (*ii*) 60 W
(*iii*) 30 W (*iv*) 90 W

23. The maximum power transfer theorem is used in ....
(*i*) electronic circuits
(*ii*) power system
(*iii*) home lighting circuits
(*iv*) none of the above

24. The Norton resistance of a network is 20 Ω and the shorted-load current is 2 A. If the network is loaded by a resistance equal to 20 Ω, the current through the load will be ....
(*i*) 2 A (*ii*) 0.5 A
(*iii*) 4 A (*iv*) 1 A

25. The Norton current is sometimes called the ....
(*i*) shorted-load current
(*ii*) open-load current
(*iii*) Thevenin current
(*iv*) Thevenin voltage

Answers to Multiple-Choice Questions

1. (*i*)	**2.** (*ii*)	**3.** (*iii*)	**4.** (*ii*)	**5.** (*iv*)
6. (*i*)	**7.** (*ii*)	**8.** (*iv*)	**9.** (*iii*)	**10.** (*iv*)
11. (*ii*)	**12.** (*ii*)	**13.** (*ii*)	**14.** (*i*)	**15.** (*ii*)
16. (*iii*)	**17.** (*ii*)	**18.** (*iii*)	**19.** (*ii*)	**20.** (*i*)
21. (*iv*)	**22.** (*ii*)	**23.** (*i*)	**24.** (*iv*)	**25.** (*i*)

Chapter Review Topics

1. What is electronics ? Mention some important applications of electronics.
2. Describe briefly the structure of atom.
3. Explain how valence electrons determine the electrical properties of a material.
4. Explain constant voltage and current sources. What is their utility ?
5. Derive the condition for transfer of maximum power from a source to a load.
6. State and explain Thevenin's theorem.
7. Write short notes on the following :
(*i*) Atomic structure (*ii*) Valence electrons (*iii*) Free electrons

Problems

1. A dry battery developing 12 V has an internal resistance of 10 Ω. Find the output current if load is (*i*) 100 Ω (*ii*) 10 Ω (*iii*) 2 Ω and (*iv*) 1 Ω. **[(*i*) 0.1A (*ii*) 0.6A (*iii*) 1A (*iv*) 1.1A]**
2. Convert the current source in Fig. 1.48 into the equivalent voltage source.
[36 V in series with 900 Ω]

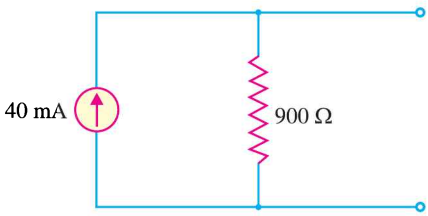


Fig. 1.48

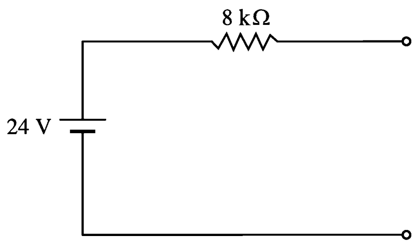


Fig. 1.49

3. Convert the voltage source in Fig. 1.49 into equivalent current source. **[3 mA in parallel with 8 kΩ]**
4. Using Norton's Theorem, find the current in branch *AB* containing 6 Ω resistor of the network shown in Fig. 1.50. **[0.466 A]**

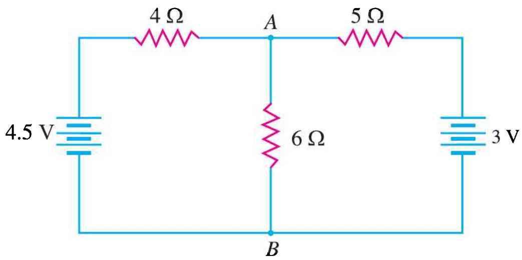


Fig. 1.50

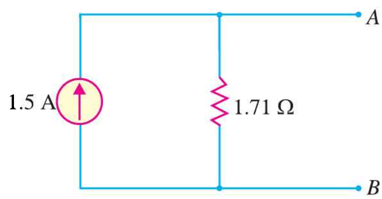


Fig. 1.51

5. Fig. 1.51 shows Norton's equivalent circuit of a network behind terminals *A* and *B*. Convert it into Thevenin's equivalent circuit. **[2.56 V in series with 1.71 Ω]**

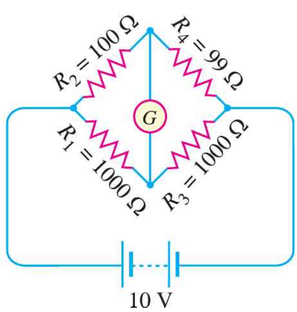


Fig. 1.52

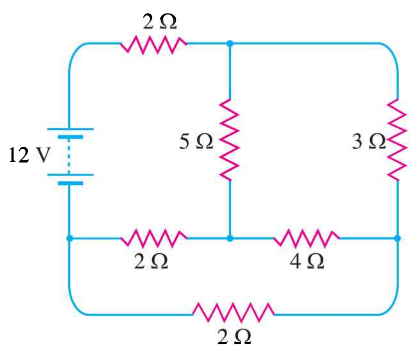


Fig. 1.53

6. A power amplifier has an internal resistance of 5 Ω and develops open circuited voltage of 12 V. Find the efficiency and power transferred to a load of (*i*) 20 Ω (*ii*) 5 Ω. **[(*i*) 80%, 4.6 W (*ii*) 50%, 7.2 W]**
7. Using Thevenin's theorem, find the current through the galvanometer in the Wheatstone bridge shown in Fig. 1.52. **[38.6 μA]**
8. Using Thevenin's theorem, find the current through 4 Ω resistor in the circuit of Fig. 1.53. **[0.305A[**

Discussion Questions

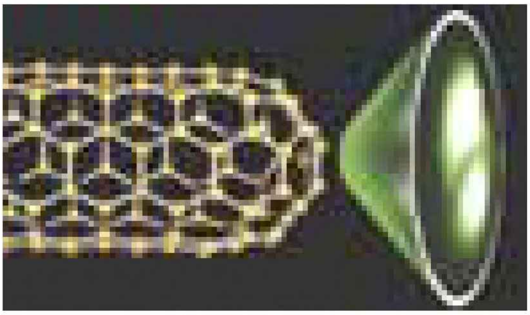
1. Why are free electrons most important for electronics ?
2. Why do insulators not have any free electrons ?
3. Where do you apply Thevenin's theorem ?
4. Why is maximum power transfer theorem important in electronic circuits ?
5. What are the practical applications of a constant current source ?

2

Electron Emission

INTRODUCTION

The reader is familiar with the current conduction (*i.e.* flow of electrons) through a conductor. The valence electrons of the conductor atoms are loosely bound to the atomic nuclei. At room temperature, the thermal energy in the conductor is adequate to break the bonds of the valence electrons and leave them unattached to any one nucleus. These unbound electrons move at random within the conductor and are known as *free electrons*. If an electric field is applied across the conductor, these free electrons move through the conductor in an orderly manner, thus constituting electric current. This is how these free electrons move through the conductor or electric current flows through a wire.



Electron Emission

Many electronic devices depend for their operation on the movement of electrons in an evacuated space. For this purpose, the free electrons must be ejected from the surface of metallic conductor by supplying sufficient energy from some external source. This is known as *electron emission*. The emitted electrons can be made to move in vacuum under the influence of an electric field, thus constituting electric current in vacuum. In this chapter, we shall confine our attention to the various aspects of electron emission.

2.1 Electron Emission

The liberation of electrons from the surface of a substance is known as **electron emission.**

For electron emission, metals are used because they have many free electrons. If a piece of metal is investigated at room temperature, the random motion of free electrons is as shown in Fig. 2.1. However, these electrons are free only to the extent that they may transfer from one atom to another within the metal but they cannot leave the metal surface to provide electron emission. It is because the free electrons that start at the surface of metal find behind them positive nuclei pulling them back and none pulling forward. Thus at the surface of a metal, a free electron encounters forces that prevent it to leave the metal. In other words, the metallic surface offers a barrier to free electrons and is known as *surface barrier*.

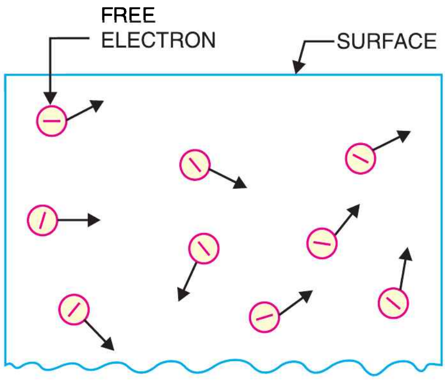


Fig. 2.1

However, if sufficient external energy is given to the free electron, its kinetic energy is increased and thus electron will cross over the surface barrier to leave the metal. This additional energy required by an electron to overcome the surface barrier of the metal is called *work function* of the metal.

The amount of additional energy required to emit an electron from a metallic surface is known as **work function** *of that metal.*

Thus, if the total energy required to liberate an electron from a metal is 4 eV* and the energy already possessed by the electron is 0.5 eV, then additional energy required (*i.e.*, work function) is 4.0– 0.5 = 3.5 eV. The work function of pure metals varies roughly from 2 to 6 eV. It depends upon the nature of metal, its purity and the conditions of its surface. It may be noted that it is desirable that metal used for electron emission should have low work function so that a small amount of energy is required to cause emission of electrons.

2.2 Types of Electron Emission

The electron emission from the surface of a metal is possible only if sufficient additional energy (*equal to the work function of the metal*) is supplied from some external source. This external energy may come from a variety of sources such as heat energy, energy stored in electric field, light energy or kinetic energy of the electric charges bombarding the metal surface. Accordingly, there are following four principal methods of obtaining electron emission from the surface of a metal :

* Work function is the additional energy required for the liberation of electrons. Therefore, it should have the conventional unit of energy *i.e.* joules. But this unit is very large for computing electronics work. Therefore, in practice, a smaller unit called *electron volt* (abbreviated as eV) is used.

One electron-volt is the amount of energy acquired by an electron when it is accelerated through a potential difference of 1 V.

Thus, if an electron moves from a point of 0 potential to a point of +10V, then amount of energy acquired by the electron is 10 eV.

Since charge on an electron $= 1.602 \times 10^{-19}$ C and voltage = 1 V,

$\therefore$ 1 electron-volt $= Q\,V = (1.602 \times 10^{-19}) \times 1$ J

or 1 eV $= 1.602 \times 10^{-19}$ J

(*i*) **Thermionic emission.** In this method, the metal is heated to sufficient temperature (about 2500°C) to enable the free electrons to leave the metal surface. The number of electrons emitted depends upon the temperature. The higher the temperature, the greater is the emission of electrons. This type of emission is employed in vacuum tubes.

(*ii*) **Field emission.** In this method, a strong electric field (*i.e.* a high positive voltage) is applied at the metal surface which pulls the free electrons out of metal because of the attraction of positive field. The stronger the electric field, the greater is the electron emission.

(*iii*) **Photo-electric emission.** In this method, the energy of light falling upon the metal surface is transferred to the free electrons within the metal to enable them to leave the surface. The greater the intensity (*i.e.* brightness) of light beam falling on the metal surface, the greater is the photo-electric emission.

(*iv*) **Secondary emission.** In this method, a high velocity beam of electrons strikes the metal surface and causes the free electrons of the metal to be knocked out from the surface.

2.3 Thermionic Emission

The process of electron emission from a metal surface by supplying thermal energy to it is known as **thermionic emission.**

At ordinary temperatures, the energy possessed by free electrons in the metal is inadequate to cause them to escape from the surface. When heat is applied to the metal, some of heat energy is converted into kinetic energy, causing accelerated motion of free electrons. When the temperature rises sufficiently, these electrons acquire additional energy equal to the work function of the metal. Consequently, they overcome the restraining surface barrier and leave the metal surface.

Metals *with lower work function* will require less additional energy and, therefore, will emit electrons at lower temperatures. The commonly used materials for electron emission are *tungsten, thoriated tungsten* and *metallic oxides of barium and strontium.* It may be added here that high temperatures are necessary to cause thermionic emission. For example, pure tungsten must be heated to about 2300°C to get electron emission. However, oxide coated emitters need only 750°C to cause thermionic emission.

Richardson-Dushman equation. The amount of thermionic emission increases rapidly as the emitter temperature is raised. The emission current density is given by Richardson-Dushman equation given below :

$$J_s = A\,T^2\,e^{-\frac{b}{T}} \text{ amp/m}^2 \qquad ...(i)$$

where J_s = emission current density *i.e.* current per square metre of the emitting surface

T = absolute temperature of emitter in K

A = constant, depending upon the type of emitter and is measured in amp/m^2/K^2

b = a constant for the emitter

e = natural logarithmic base

The value of b is constant for a metal and is given by :

$$b = \frac{\phi e}{k}$$

where ϕ = work function of emitter

e = electron charge = 1.602×10^{-19} coulomb

k = Boltzmann's constant = 1.38×10^{-23} J/K

$$\therefore \quad b = \frac{\phi \times 1.602 \times 10^{-19}}{1.38 \times 10^{-23}} = 11600\,\phi\text{ K}$$

Putting the value of b in exp. (i), we get,

$$J_s = AT^2 e^{-\frac{11600\,\phi}{T}} \quad ...(ii)$$

The following points may be noted from eqn. (ii) :

(i) The emission is markedly affected by temperature changes. Doubling the temperature of an emitter may increase electron emission by more than 10^7 times. For instance, emission from pure tungsten metal is about 10^{-6} ampere per sq. cm. at 1300°C but rises to enormous value of about 100 amperes when temperature is raised to 2900°C.

(ii) Small changes in the work function of the emitter can produce enormous effects on emission. Halving the work function has exactly the same effect as doubling the temperature.

Example 2.1. *A tungsten filament consists of a cylindrical cathode 5 cm long and 0.01 cm in diameter. If the operating temperature is 2500 K, find the emission current. Given that* $A = 60.2 \times 10^4\ A/m^2/K^2$, $\phi = 4.517\ eV$.

Solution. $A = 60.2 \times 10^4\ \text{amp/m}^2/\text{K}^2$, $T = 2500\ \text{K}$, $\phi = 4.517\ \text{eV}$

$\therefore$ $b = 11600\,\phi\,\text{K} = 11600 \times 4.517\ \text{K} = 52400\ \text{K}$

Using Richardson-Dushman equation, emission current density is given by :

$$J_s = AT^2 e^{-\frac{b}{T}}\ \text{amp/m}^2 = 60.2 \times 10^4 \times (2500)^2 \times (2.718)^{-\frac{52400}{2500}}$$

$$= 0.3 \times 10^4\ \text{amp/m}^2$$

Surface area of cathode, $a = \pi\, d\, l = 3.146 \times 0.01 \times 5 = 0.157\ \text{cm}^2 = 0.157 \times 10^{-4}\ \text{m}^2$

$\therefore$ Emission current $= J_s \times a = (0.3 \times 10^4) \times (0.157 \times 10^{-4}) = \mathbf{0.047\ A}$

Example 2.2. *A tungsten wire of unknown composition emits 0.1 amp/cm² at a temperature of 1900 K. Find the work function of tungsten filament. Determine whether the tungsten is pure or contaminated with substance of lower work function. Given that A = 60.2 amp/cm²/K².*

Solution. $J_s = 0.1\ \text{amp/cm}^2$; $A = 60.2\ \text{amp/cm}^2/\text{K}^2$; $T = 1900\ \text{K}$

Let ϕ electron-volt be the work function of the filament.

$\therefore$ $b = 11600\,\phi\ \text{K}$

Using Richardson-Dushman equation, emission current density is given by :

$$J_s = A\,T^2 e^{-\frac{b}{T}}\ \text{amp/cm}^2$$

or $$0.1 = 60.2 \times (1900)^2 \times e^{-\frac{11600\phi}{1900}}$$

or $$e^{-\frac{11600\phi}{1900}} = \frac{0.1}{60.2 \times (1900)^2} = 4.6 \times 10^{-10}$$

or $$e^{-6.1\,\phi} = 4.6 \times 10^{-10}$$

or $$-6.1\,\phi \log_e e = \log_e 4.6 - 10 \log_e 10$$

or $$-6.1\,\phi = 1.526 - 23.02$$

$\therefore$ $$\phi = \frac{1.526 - 23.02}{-6.1} = \mathbf{3.56\ eV}$$

Since the work function of pure tungsten is 4.52 eV, the sample must be contaminated. Thoriated tungsten has a work function ranging from 2.63 eV to 4.52 eV, depending upon the percentage of metallic thorium. Therefore, the sample is most likely to be thoriated tungsten.

2.4 Thermionic Emitter

The substance used for electron emission is known as an *emitter* or *cathode*. The cathode is heated in an evacuated space to emit electrons. If the cathode were heated to the required temperature in open air, it would burn up because of the presence of oxygen in the air. A cathode should have the following properties:

Thermionic Emitter

(*i*) **Low work function.** The substance selected as cathode should have low work function so that electron emission takes place by applying small amount of heat energy *i.e.* at low temperatures.

(*ii*) **High melting point.** As electron emission takes place at very high temperatures (>1500°C), therefore, the substance used as a cathode should have high melting point. For a material such as copper, which has the advantage of a low work function, it is seen that it cannot be used as a cathode because it melts at 810°C. Consequently, it will vaporise before it begins to emit electrons.

(*iii*) **High mechanical strength.** The emitter should have high mechanical strength to withstand the bombardment of positive ions. In any vacuum tube, no matter how careful the evacuation, there are always present some gas molecules which may form ions by impact with electrons when current flows. Under the influence of electric field, the positive ions strike the cathode. If high voltages are used, the cathode is subjected to considerable bombardment and may be damaged.

2.5 Commonly Used Thermionic Emitters

The high temperatures needed for satisfactory thermionic emission in vacuum tubes limit the number of suitable emitters to such substances as *tungsten, thoriated tungsten* and certain *oxide coated metals*.

Thoriated Tungsten

(*i*) **Tungsten.** It was the earliest material used as a cathode and has a slightly higher work function (4.52 eV). The important factors in its favour are : high melting point (3650 K), greater mechanical strength and longer life. The disadvantages are : high operating temperature (2500 K), high work function and low emission efficiency. Therefore, it is used in applications involving voltages exceeding 5 kV *e.g.* in X-ray tubes.

(*ii*) **Thoriated tungsten.** A mixture of two metals may have a lower work function than either of the pure metals alone. Thus, a tungsten emitter with a small quantity of thorium has a work function of 2.63 eV, compared with 3.4 eV for thorium and 4.52 eV for tungsten. At the same time, thoriated tungsten provides thermionic emission at lower temperature (1700°C) with consequent reduction in the heating power required.

In the manufacture of this type of cathode, tungsten filament is impregnated with thorium oxide and heated to a very high temperature (1850°C to 2500°C). The thorium oxide is reduced to metallic thorium and coats the filament surface with a thin layer of thorium. Thoriated tungsten cathodes are used for intermediate power tubes at voltages between 500 to 5000 volts.

(*iii*) **Oxide-coated cathode.** The cathode of this *type consists of a nickel ribbon coated with

* Oxides of any alkaline-earth metal (*e.g.* calcium, strontium, barium etc.) have very good emission characteristics. In the manufacture of this type of emitter, the base metal (*e.g.* nickel) is first coated with a mixture of strontium and barium carbonates. It is then heated to a high temperature in vacuum glass tube until the carbonates decompose into oxides. By proper heating, a layer of oxides of barium and strontium is coated over the cathode surface to give oxide-coated emitter.

barium and strontium oxides. The oxide-coated cathode has low work function (1.1 eV), operates at comparatively low temperature (750°C) and has high emission efficiency. However, the principal limitation of oxide-coated cathode is that it cannot withstand high voltages. Therefore, it is mostly used in receiving tubes or where voltages involved do not exceed 1000 volts.

S.No.	Emitter	Work Function	Operating temperature	Emission efficiency
1	*Tungsten*	4.52 eV	2327°C	4 mA/watt
2	*Thoriated tungsten*	2.63 eV	1700°C	60 mA/watt
3	*Oxide-coated*	1.1 eV	750°C	200 mA/watt

2.6 Cathode Construction

As cathode is sealed in vacuum, therefore, the most convenient way to heat it is electrically. On this basis, the thermionic cathodes are divided into two types viz directly heated cathode and indirectly heated cathode.

(*i*) **Directly heated cathode.** In this type, the cathode consists of oxide-coated nickel ribbon, called the *filament. The heating current is directly passed through this ribbon which emits the electrons. Fig. 2.2 (*i*) shows the structure of directly heated cathode whereas Fig. 2.2 (*ii*) shows its symbol.

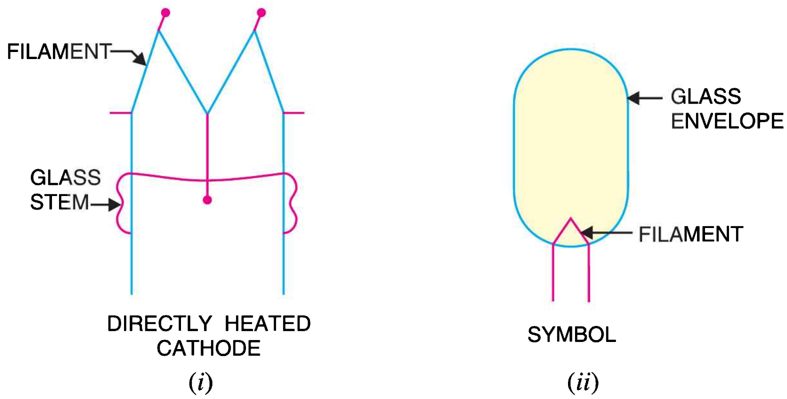
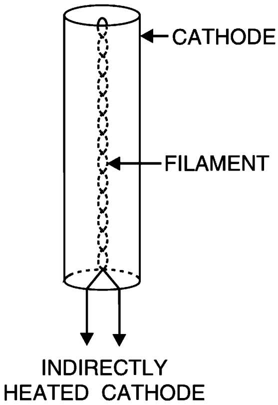


Fig. 2.2

The directly heated cathode is more efficient in converting heating power into thermionic emission. Therefore, it is generally used in power tubes that need large amounts of emission and in small tubes operated from batteries where efficiency and quick heating are important. The principal limitation of this type of cathode is that any variation in heater voltage affects the electron emission and thus produces *hum* in the circuit.

(*ii*) **Indirectly heated cathode.** In this type, the cathode consists of a thin metal sleeve coated with barium and strontium oxides. A filament or heater is enclosed within the sleeve and insulated from it. There is no electrical connection between the heater and the cathode. The heating current is passed through the heater and the cathode is heated indirectly through heat transfer from the heater element. Fig. 2.3 (*i*) shows the structure of indirectly heated cathode whereas Fig. 2.3 (*ii*) shows its symbol.

* ***Filament.*** The term filament (literally means a thin wire) denotes the element through which the cathode heating current flows. In case of directly heated, cathode is itself the filament. If indirectly heated, heater is the filament.



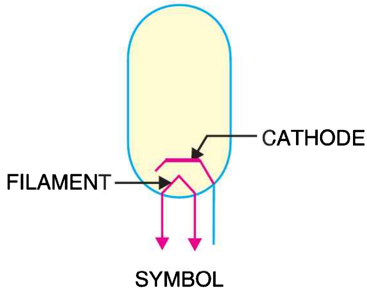


Fig. 2.3

Indirectly heated cathode has many advantages. As cathode is completely separated from the heating circuit, therefore, it can be readily connected to any desired potential as needed, independent of the heater potential. Furthermore, because of relatively large mass of cylindrical cathode, it takes time to heat or cool and as such does not introduce hum due to heater voltage fluctuations. Finally, a.c. can be used in the heater circuit to simplify the power requirements. Almost all modern receiving tubes use this type of cathode.

2.7 Field Emission

The process of electron emission by the application of strong electric field at the surface of a metal is known as **field emission.**

When a metal surface is placed close to a high voltage conductor which is positive *w.r.t.* the metal surface, the electric field exerts attractive force on the free electrons in the metal. If the positive potential is great enough, it succeeds in overcoming the restraining forces of the metal surface and the free electrons will be emitted from the metal surface as shown in Fig. 2.4.

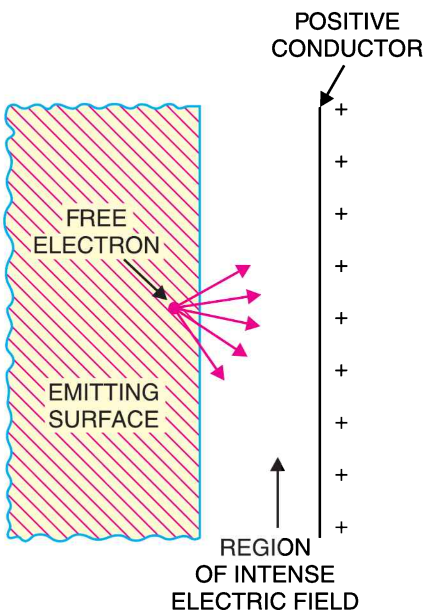


Fig. 2.4

Very intense electric field is required to produce field emission. Usually, a voltage of the order of a million volts per centimetre distance between the emitting surface and the positive conductor is necessary to cause field emission. Field emission can be obtained at temperatures much lower (*e.g.* room temperature) than required for thermionic emission and, therefore, it is also sometimes called *cold cathode emission* or *auto- electronic emission.*

2.8 Secondary Emission

Electron emission from a metallic surface by the bombardment of high-speed electrons or other particles is known as **secondary emission.**

When high-speed electrons suddenly strike a metallic surface, they may give some or all of their kinetic energy to the free electrons in the metal. If the energy of the striking electrons is sufficient, it may cause free electrons to escape from the metal surface. This phenomenon is called *secondary emission*. The electrons that strike the metal are called *primary electrons* while the emitted electrons are known as *secondary electrons*. The intensity of secondary emission depends upon the emitter material, mass and energy of the bombarding particles.

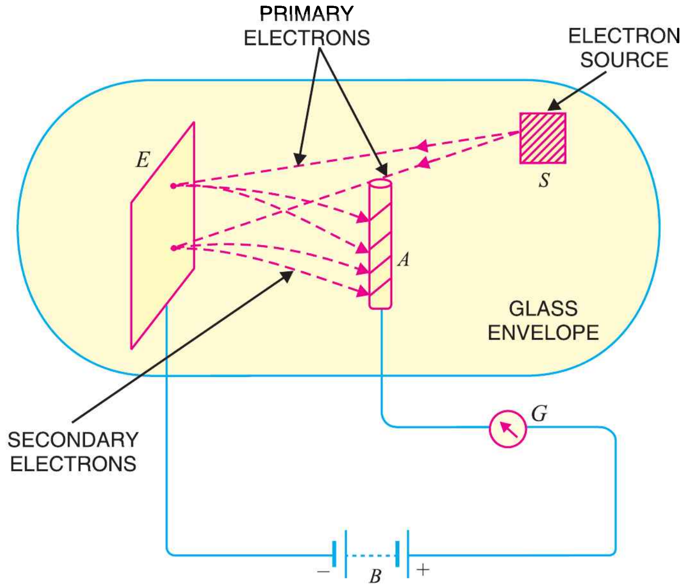


Fig. 2.5

The principle of secondary emission is illustrated in Fig. 2.5. An evacuated glass envelope contains an emitting surface *E*, the collecting anode *A* and a source of primary electrons *S*. The anode is maintained at positive potential *w.r.t.* the emitting surface by battery *B*. When the primary electrons strike the emitting surface *E*, they knock out secondary electrons which are attracted to the anode and constitute a flow of current. This current may be measured by connecting a sensitive galvanometer *G* in the anode circuit.

The effects of secondary emission are very undesirable in many electronic devices. For example, in a tetrode valve, secondary emission is responsible for the negative resistance. In some electronic devices, however, secondary emission effects are utilised *e.g.* *electron multiplier, cathode ray tube etc.

2.9 Photo Electric Emission

Electron emission from a metallic surface by the application of light is known as **photo electric emission.**

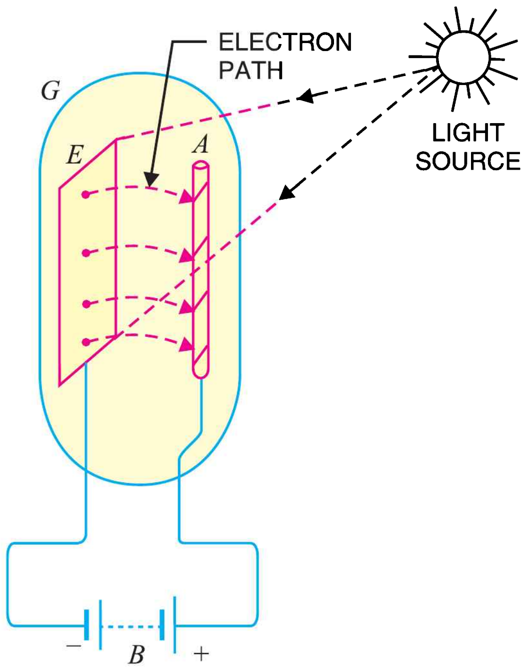


Fig. 2.6

When a beam of light strikes the surface of certain metals (*e.g.* potassium, sodium, cesium), the energy of photons of light is transferred to the free electrons within the metal. If the energy of the

* An interesting aspect of secondary emission is that a high-speed bombarding electron may liberate as many as 10 "secondary electrons". This amounts to a multiplication of electron flow by a ratio as great as 10 and is utilised in current multiplier devices.

striking photons is greater than the work function of the metal, then free electrons will be knocked out from the surface of the metal. The emitted electrons are known as *photo electrons* and the phenomenon is known as *photoelectric emission*. The amount of photoelectric emission depends upon the intensity of light falling upon the emitter and frequency of radiations. The greater the intensity and frequency of radiations, the greater is the photo electric emission. Photo-electric emission is utilised in photo tubes which form the basis of television and sound films.

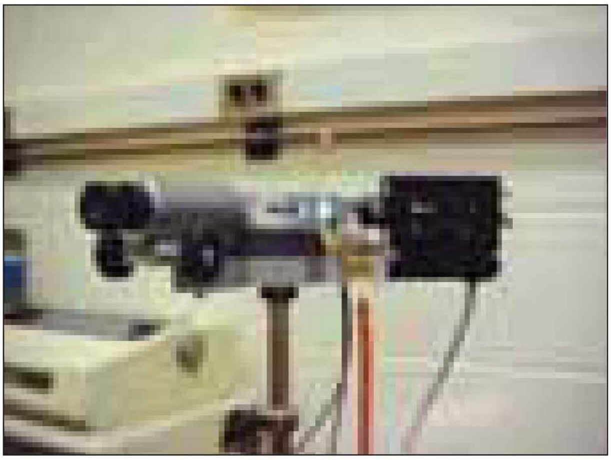


Photo Tube

Fig. 2.6 illustrates the phenomenon of photoelectric emission. The emitter *E* and anode *A* are enclosed in an evacuated glass envelope *G*. A battery *B* maintains the anode at positive potential w.r.t. emitter. When light of suitable intensity and frequency falls on the emitter, electrons are ejected from its surface. These electrons are attracted by the positive anode to constitute current in the circuit. It may be noted that current will exist in the circuit so long as illumination is maintained.

MULTIPLE-CHOICE QUESTIONS

1. Work function of metals is generally measured in ....

(*i*) joules (*ii*) electron-volt
(*iii*) watt-hour (*iv*) watt

2. The operating temperature of an oxide-coated emitter is about ....

(*i*) 750ºC (*ii*) 1200ºC
(*iii*) 2300ºC (*iv*) 3650ºC

3. .... is used in high voltage (> 10 kV) applications.

(*i*) tungsten emitter
(*ii*) oxide-coated emitter
(*iii*) thoriated-tungsten emitter
(*iv*) none of the above

4. A desirable characteristic of an emitter is that it should have .... work function.

(*i*) large (*ii*) very large
(*iii*) small (*iv*) none of the above

5. The thermionic emitter that has the highest operating temperature is ....

(*i*) oxide-coated (*ii*) thoriated-tungsten
(*iii*) tungsten (*iv*) none of the above

6. If the temperature of an emitter is increased two times, the electron emission is ....

(*i*) increased two times
(*ii*) increased four times
(*iii*) increased several million times
(*iv*) none of the above

7. In X-ray tubes, ... emitter is used.

(*i*) thoriated tungsten
(*ii*) tungsten
(*iii*) oxide-coated
(*iv*) none of the above

8. The life of an oxide-coated emitter is about ....

(*i*) 500 hours (*ii*) 1000 hours
(*iii*) 200 hours (*iv*) 10,000 hours

9. The electrons emitted by a thermionic emitter are called ....

(*i*) free electrons
(*ii*) loose electrons
(*iii*) thermionic electrons
(*iv*) bound electrons

10. The work function of an oxide-coated emitter is about ....
(*i*) 1.1 eV (*ii*) 4 eV
(*iii*) 2.63 eV (*iv*) 4.52 eV

11. The warm-up time of a directly heated cathode is .... that of indirectly heated cathode.
(*i*) more than (*ii*) less than
(*iii*) same as (*iv*) data incomplete

12. The most commonly used emitter in the tubes of a radio receiver is ....
(*i*) tungsten (*ii*) thoriated-tungsten
(*iii*) oxide-coated (*iv*) none of the above

13. Field emission is utilised in ....
(*i*) vacuum tubes
(*ii*) TV picture tubes
(*iii*) gas-filled tubes
(*iv*) mercury pool devices

14. Oxide-coated emitters have electron emission of .... per watt of heating power.
(*i*) 5-10 mA (*ii*) 40-90 mA
(*iii*) 50-100 mA (*iv*) 150-1000 mA

15. The oxide-coated cathodes can be used for voltages upto ....
(*i*) 1000 V (*ii*) 3000 V
(*iii*) 4000 V (*iv*) 10,000 V

Answers to Multiple-Choice Questions

1. (*ii*)	2. (*i*)	3. (*i*)	4. (*iii*)	5. (*iii*)
6. (*iii*)	7. (*ii*)	8. (*iv*)	9. (*iii*)	10. (*i*)
11. (*ii*)	12. (*iii*)	13. (*iv*)	14. (*iv*)	15. (*i*)

Chapter Review Topics

1. What is electron emission ? Explain the terms : surface barrier and work function.
2. What general conditions must be satisfied before an electron can escape from the surface of a material ?
3. Name and explain briefly four practical ways by which electron emission can occur.
4. What are the materials used for thermionic emitters ? Compare the relative merits of each.
5. Discuss briefly construction and relative advantages of directly and indirectly heated cathodes.

Problems

1. An oxide-coated emitter has a surface area of 0.157 cm^2. If the operating temperature is 110 K, find the emission current. Given A = 100 $A/m^2/K^2$, work function = 1.04 eV. **[0.0352 A]**
2. A tungsten filament of unknown composition emits 1000 A/m^2 at an operating temperature of 1900 K. Find the work function of tungsten filament. Given $A = 60.2 \times 10^4$ $A/m^2/K^2$. **[3.44 eV]**
3. Calculate the total emission available from barium-strontium oxide emitter, 10 cm long and 0.01 cm in diameter, operated at 1900 K. Given that $A = 10^{-12}$ $Amp/cm^2/K^2$ and b = 12,000. **[0.345 A]**

Discussion Questions

1. Why does electron emission not occur at room temperature ?
2. Why are high temperatures necessary for thermionic emission ?
3. Why are electron emitters heated electrically ?
4. Why are thermionic emitters heated in vacuum ?
5. Why are tungsten and thoriated tungsten cathodes always of directly heated type ?
6. Why cannot oxide-coated cathodes be used for voltages exceeding 1000 volts?
7. Why do directly heated cathodes introduce hum in the circuit ?
8. Why are directly heated cathodes used in high power applications ?

3

Gas-Filled Tubes

INTRODUCTION

In the vacuum tubes, the electrons flow from cathode to anode in vacuum. In such tubes, extreme care is taken to produce as perfect a vacuum as possible to prevent ionisation of gases and the resulting large uncontrolled currents. It may be mentioned here that the secret of triode is the fine control of free electrons within valve by the electrostatic fields of grid and anode. If gas is present even in small amount, the electrons flowing from cathode to anode will cause ionisation of the gas. The ionised molecules would interfere with the control and make the device useless as an amplifier.

In certain applications, fine control of electrons within the valve is of less importance than the efficient handling and turning on and off of heavy currents. In such situations, some inert gases (*e.g.* argon, neon, helium) at low pressures are purposely introduced into the valve envelope. Such tubes are known as *gas-filled tubes*. The gas-filled tubes are capable of doing various jobs that vacuum tubes cannot perform and which are especially useful in industrial and control circuits. In this chapter, we shall focus our attention on some important types of gas-filled tubes with special reference to their characteristic properties.

Gas-filled Tube

3.1 Gas-Filled Tubes

A **gas-filled** *tube is essentially a vacuum tube having a small amount of some inert gas at low pressure.*

The gas pressure in a gas-filled tube usually ranges from 10 mm of Hg to 50 mm. The construction of gas-filled tubes is similar to that of vacuum tubes, except that the cathodes, grids and anodes are usually larger in order to carry heavier current. However, the characteristic properties of the two are markedly different. Firstly, a gas-filled tube can conduct much *more current than the equivalent vacuum tube. It is because the electrons flowing from cathode to anode collide with gas molecules and ionise them *i.e.* knock out electrons from them. The additional electrons flow to the anode together with the original electrons, resulting in the increase in plate current. Secondly, a gas filled tube has far less control on the electrons in the tube than that of vacuum tube. Once the ionisation starts, the control of gas-filled tube is tremendously reduced.

Classification. Gas-filled tubes are usually classified according to the type of electron emission employed. On this basis, they may be classified into two types namely; *cold-cathode type* and *hot-cathode type.*

Cold-cathode type. In this type of gas-filled tubes, the cathode is not heated as in a vacuum tube. The ionisation of the gas is caused by the energy available from natural sources such as cosmic rays, sun rays or radioactive particles in air. These natural sources are the underlying reasons for the start of conduction in cold-cathode gas tubes. Most cold-cathode tubes are used as diodes.

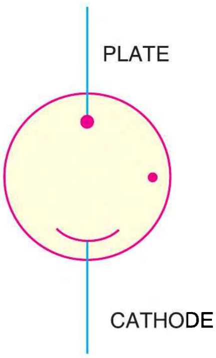


Fig. 3.1

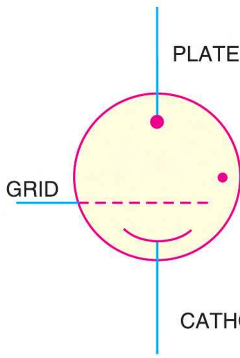


Fig. 3.2

Fig. 3.1 shows the schematic symbol for a cold-cathode gas diode, known as *glow tube*. The dot within the circle indicates the presence of gas. Fig. 3.2 shows the schematic symbol of cold-cathode gas triode, known as *grid glow tube*.

Hot-cathode type. In this type of gas-filled tubes, the cathode is heated just as in an ordinary vacuum tube. The electrons flowing from cathode to plate cause ionisation of the gas molecules. Such tubes are used as diodes, triodes and tetrodes.

* The ability of a gas-filled tube to carry large current is, of course, no recommendation in itself. A copper wire will do the same thing and with better efficiency. But a gas filled tube has one special ability which the wire does not possess ; the ability to carry current in one direction.

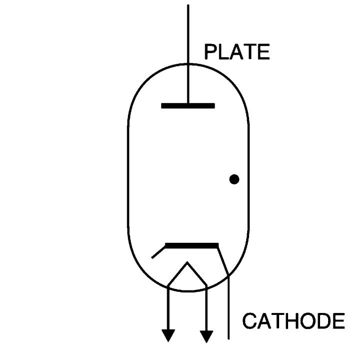


Fig. 3.3

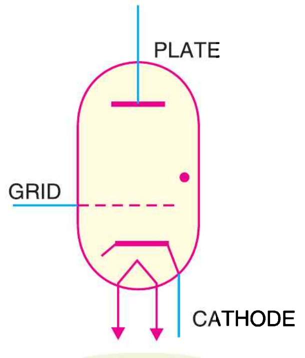


Fig. 3.4

Fig. 3.3 shows the schematic symbol of a hot-cathode gas diode, known as *phanotron* whereas Fig. 3.4 shows the symbol of hot-cathode gas triode, known as *thyratron*.

3.2 Conduction in a Gas

A gas under ordinary pressure is a perfect insulator and cannot conduct current. However, if the gas pressure is low, it is possible to produce a large number of free electrons in the gas by the process of ionisation and thus cause the gas to become a conductor. This is precisely what happens in gas-filled tubes. The current conduction in a gas at low pressure can be beautifully illustrated by referring to the hot-cathode gas diode shown in Fig. 3.5. The space between cathode and anode of the tube contains gas molecules. When cathode is heated, it emits a large number of electrons. These electrons form a cloud of electrons near the cathode, called **space charge.** If anode is made positive *w.r.t.* cathode, the electrons (magenta dots) from the space charge speed towards the anode and collide with gas molecules (cyan circles) in the tube.

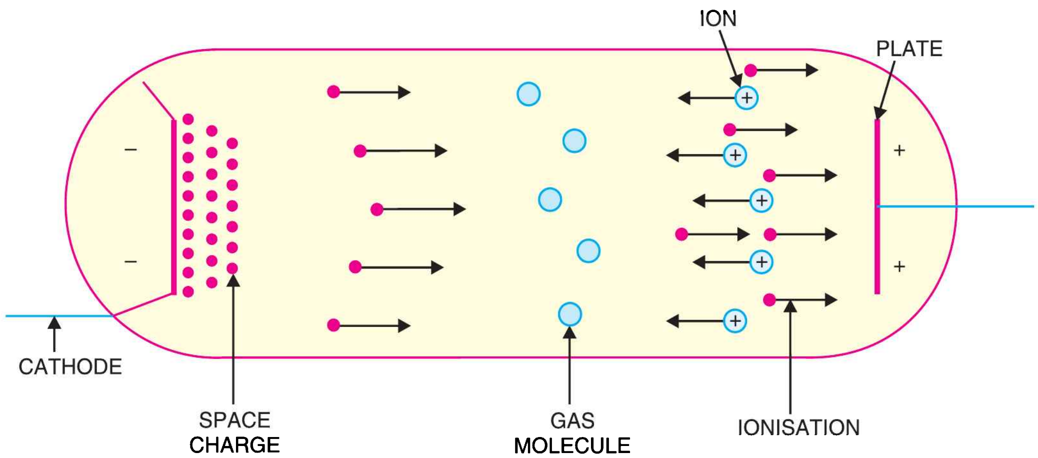


Fig. 3.5

If the anode-cathode voltage is low, the electrons do not possess the necessary energy to cause ionisation of the gas. Therefore, the plate current flow in the tube is only due to the electrons emitted by the cathode. As the anode-cathode voltage is increased, the electrons acquire more speed and energy and a point–called *ionisation voltage* is reached, where ionisation of the gas starts. The ionisation of gas produces free electrons and positive gas ions (cyan circles with +ve signs). The additional free electrons flow to the anode together with the original electrons, thus increasing plate current. However, the increase in plate current due to these added electrons is practically negligible. But the major effect is that the positive gas ions slowly drift towards the cathode and neutralise the space charge. Consequently, the resistance of the tube decreases, resulting in large plate current. *Hence, it is due to the neutralisation of space charge by the positive gas ions that plate current in a gas tube is too much increased.*

The following points may be noted regarding the conduction in a gas at low pressure :

(*i*) At low anode-cathode voltage, the ionisation of the gas does not occur and the plate current is about the same as for a vacuum tube at the same anode voltage.

(*ii*) At some anode-cathode voltage, called ionisation voltage, ionisation of the gas takes place. The plate current increases dramatically to a large value due to the neutralisation of space charge by the positive gas ions. The ionisation voltage depends upon the type and pressure of gas in the tube.

(*iii*) Once ionisation has started, it is maintained at anode-cathode voltage much lower than ionisation voltage. However, minimum anode-cathode voltage, called *deionising voltage*, exists below which ionisation cannot be maintained. Under such conditions, the positive gas ions combine with electrons to form neutral gas molecules and conduction stops. Because of this switching action, a gas-filled tube can be used as an electronic switch.

3.3 Cold-Cathode Gas Diode

Fig. 3.6 shows the cut-away view of cold-cathode gas diode. It essentially consists of two electrodes, cathode and anode, mounted fairly close together in an envelope filled with some inert gas at low pressure. The anode is in the form of a thin wire whereas cathode is of cylindrical metallic surface having oxide coating. The anode is always held at positive potential *w.r.t.* cathode.

Operation. Fig. 3.7 shows a circuit that can be used to investigate the operation of cold-cathode gas diode. Electric conduction through the tube passes through three successive discharge phases *viz.* Townsend discharge, the glow discharge and the arc discharge.

(*i*) **Townsend discharge.** At low anode-cathode voltage, the tube conducts an extremely small current (1mA). It is because the cathode is cold and as such no source of electrons is present. However, natural sources (*e.g.* cosmic rays *etc.*) cause some ionisation of the gas, creating a few free electrons. These electrons move towards the anode to constitute a small current. This stage of conduction is known as **Townsend discharge*. In this phase of conduction, no visible light is associated.

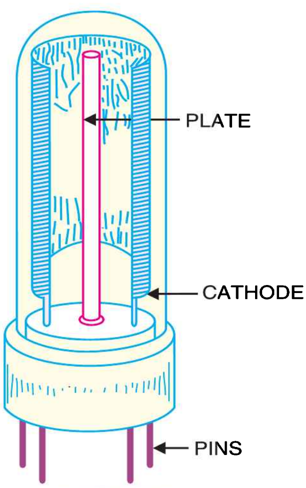


Fig. 3.6

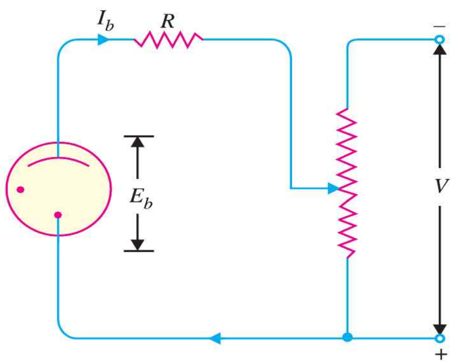


Fig. 3.7

* The volt-ampere characteristics of glow tube were first investigated by J. S. Townsend in 1901 and hence the name.

(*ii*) **Glow discharge.** As the anode-cathode voltage is increased, the electrons acquire more and more energy. At some voltage, known as *ionisation voltage*, ionisation of the gas starts and the tube current rises to a large value. The voltage across the tube drops to a low value, which remains constant regardless of the plate current. At the same time, glow is seen in the gas and on a portion of the cathode. This phase of conduction is known as *glow discharge*.

The fact that glow tube maintains constant voltage across it in the glow discharge region needs some explanation. In this region, any increase in supply voltage causes more current to flow; the drop across series resistance *R* increases but the voltage E_b across the tube remains constant. As the current increases, the degree of ionisation increases and the glow covers a greater part of cathode surface and hence the ionised gas path between cathode and anode has greater area of cross-section. As resistance is inversely proportional to the area of cross-section, therefore, resistance of the tube decreases. Hence, the voltage across the tube remains constant. Reverse is also true should the supply voltage decrease. Thus in the glow discharge region, the resistance of the tube changes so as to maintain constant voltage across it.

(*iii*) **Arc discharge.** Once the cathode glow covers the entire surface of the cathode, the *x*-sectional area of gas path cannot increase further. This region is known as *abnormal glow*. If the current density is further increased, the discharge becomes an arc.

3.4 Characteristics of Cold-Cathode Diode

The volt-ampere characteristic of a cold-cathode diode is shown in Fig. 3.8. At low anode-cathode voltage, the tube current is very small (1mA) and is due to the ionisation of gas molecules by the natural sources. This stage of conduction upto voltage *B* is known as *Townsend discharge* and is non-self maintained discharge because it requires an external source to cause ionisation. At some critical voltage such as *B*, the tube fires and the voltage across the tube drops (from *B* to *C*) and remains constant regardless of plate current. This is the start of second conduction and is known as *glow discharge*. In this region (*C* to *D*), voltage across the tube remains constant even if the plate current increases.

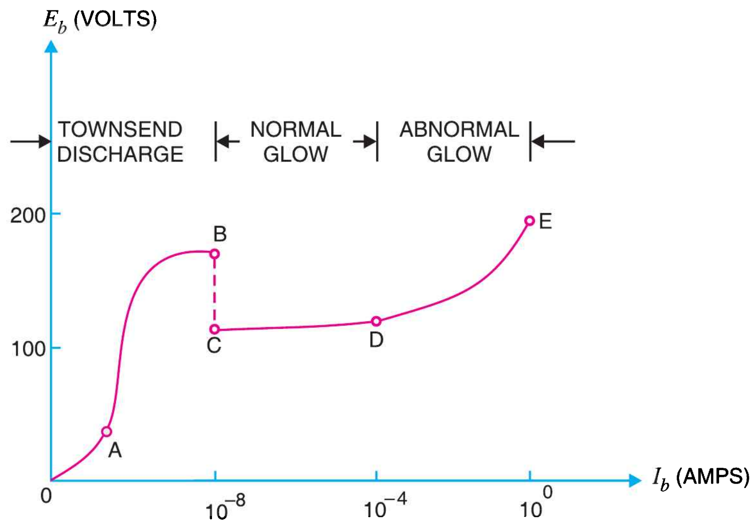


Fig. 3.8

After the glow discharge, the voltage across the tube no longer remains constant. Now, if the supply voltage is raised, not only will the circuit current increase but the voltage across the tube will start to rise again. This stage of conduction (*D* to *E*) is known as *abnormal glow*.

3.5 Applications of Glow Tubes

The outstanding characteristic of a cold-cathode gas diode (or glow tube) to maintain constant voltage across it in the glow discharge region renders it suitable for many industrial and control applications. A few of such applications are mentioned below :

(*i*) *As a voltage regulating tube.* A glow tube maintains constant voltage across it in the glow discharge region. This characteristic permits it to be used as a voltage regulating tube. Fig. 3.9 shows a simple circuit commonly used to maintain constant voltage across a load. The glow tube (*VR* tube) is connected in parallel with the load R_L across which constant voltage is desired. So long as the tube operates in the glow discharge region, it will maintain constant voltage (= 150V) across the load. The extra voltage is dropped across the series resistance *R*.


Voltage Regulating Tube

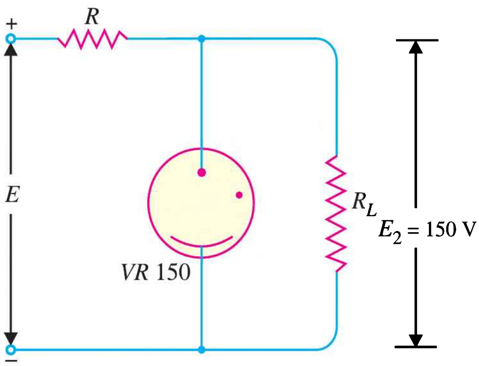


Fig. 3.9

(*ii*) *As a polarity indicator.* As the cathode is surrounded by a characteristic glow, therefore, it can be useful to indicate the polarity of a direct voltage.

(*iii*) *As an electronic switch*, which closes at ionisation potential, permitting a large current to flow, and opens at the deionising voltage, blocking the current flow.

(*iv*) *As a radio frequency field detector.* A strong radio-frequency field is capable of ionising the gas without direct connection to the tube. Therefore, the tube can indicate the presence of radio-frequency field.

3.6 Hot-Cathode Gas Diode

A hot-cathode gas diode is frequently used as a rectifier for moderate voltages because of high efficiency and better regulation. A hot-cathode gas diode consists of an oxide-coated cathode and a metallic anode enclosed in a glass envelope containing some inert gas under reduced pressure. For proper operation of the tube, anode is always held at a positive potential *w.r.t.* cathode.

Operation. Fig. 3.10 shows a circuit that can be used to investigate the operation of a hot-cathode gas diode. When cathode is heated, a large number of electrons are emitted. At low anode-cathode voltage, the tube conducts very small current. Under such conditions, the gas is not ionised and the tube acts similar to a vacuum diode — the voltage across the tube increases with plate current. This action continues until anode-cathode voltage becomes equal to the ionisation potential of the gas. Once this potential is reached, the gas begins to ionise, creating free electrons and positive gas ions. The positive gas ions move towards the cathode and tend to neutralise the space charge, thus decreasing the internal resistance of the tube. If now the plate voltage is increased, the plate current also increases due to increased degree of ionisation. This further reduces the tube resistance. As a result, increase in plate current is offset by the decrease in tube resistance *and the voltage across the tube remains constant.* Therefore, in a hot-cathode gas diode, not only the internal drop within the tube is small but also it remains constant. For this reason, a gas diode has better efficiency and regulation than for a vacuum diode when used as a rectifier.

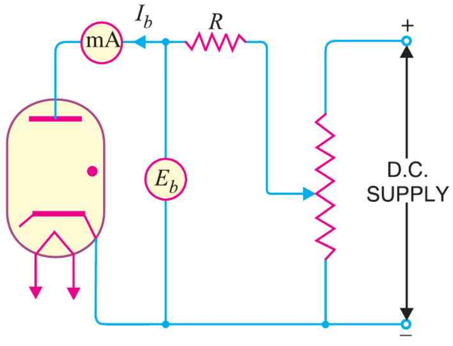


Fig. 3.10

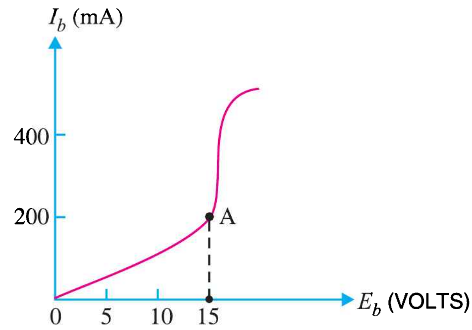


Fig. 3.11

Plate Characteristics. Fig. 3.10 shows the circuit that can be used to determine the volt-ampere (E_b/I_b) characteristics of a hot-cathode gas diode. The series resistance *R* is used to limit the current to reach a dangerously high value. Fig. 3.11 shows the plate characteristic of hot-cathode diode. It is clear that at first, plate current rises slowly with increase in anode-cathode voltage. However, at some voltage, known as ionisation voltage (point *A*), the plate current rises sharply and the voltage drop across the tube remains constant. The extra voltage is dropped across the series resistance *R*. Any attempt to raise the anode-cathode voltage above the ionising value is fruitless. Increasing the voltage E_b above point *A* results in higher plate current (I_b) and large drop across *R* but the voltage E_b across the tube remains constant.

3.7 Thyratron

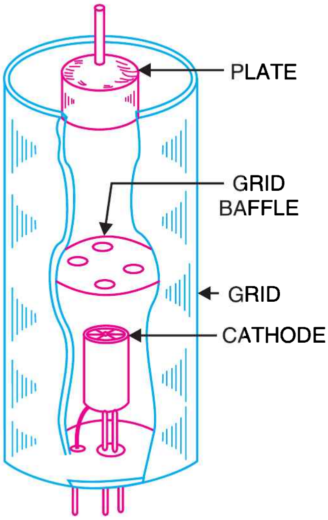
A hot-cathode gas triode is known by the trade name *thyratron*. As discussed before, a gas diode fires at a fixed plate potential, depending upon the type of gas used and gas pressure. Very often it is necessary to control the plate potential at which the tube is to fire. Such a control is obviously impossible with a gas diode. However, if a third electrode, known as *control grid* is introduced in a gas diode, this control is possible. The tube is then known as hot-cathode gas triode or thyratron. By controlling the negative potential on the control grid, the tube can be fired at any plate potential.

Construction. Figs. 3.12 (*i*) and 3.12 (*ii*) respectively show the cut-away view and schematic symbol of a thyratron. It consists of three electrodes, namely; *cathode, anode* and *control grid* enclosed in a glass envelope containing some inert gas at low pressure. The cathode and anode are approximately planar. The control grid of thyratron has a special structure quite different from that of a vacuum tube. It consists of a metal cylinder surrounding the cathode with one or more perforated discs known as *grid baffles* near the centre.

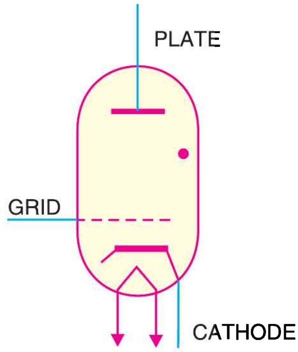
Operation. When cathode is heated, it emits plenty of electrons by thermionic emission. If the control grid is made sufficiently negative, the electrons do not have the necessary energy to ionise the gas and the plate current is substantially zero. As the negative grid voltage is reduced, the electrons acquire more speed and energy. At some grid voltage, called *critical grid voltage*, ionisation of the gas occurs and the plate current rises to a large value.

The negative grid voltage, for a given plate potential, at which ionisation of the gas starts is known as **critical grid voltage.**

At critical grid voltage, gas ionises, creating free electrons and positive gas ions. The positive ions tend to neutralise the space charge, resulting in large plate current. In addition, these positive ions are attracted by the negative grid and neutralise the normal negative field of the grid, thereby



(i)



(ii)

Fig. 3.12

preventing the grid from exerting any further control on the plate current of the tube. The grid now loses all control and the tube behaves as a diode. *Therefore, the function of control grid is only to start the conduction of anode current.* Once the conduction is started, the tube acts as a gas diode. It is important to realise the usefulness of control grid. We have seen that the ionisation does not start at low values of plate current. In a gas diode, it requires that the plate potential should be increased until sufficient plate current is flowing to cause ionisation. However, by adjusting the negative voltage on the grid, the desired plate current can be obtained to cause ionisation.

It may be mentioned here that once the thyratron fires, the only way to stop conduction is to reduce plate voltage to zero for a period *long enough for deionisation of the gas in the tube.

3.8 Applications of Thyratron

As the grid voltage has no control over the magnitude of plate current once the thyratron fires, therefore, it cannot be used as an amplifier like a vacuum triode. However, because of its triggering action, it is useful in switching and relay applications. Thyratrons are also used as controlled rectifiers for controlling the amount of d.c. power fed to the load. They are also used in motor control circuits.

MULTIPLE-CHOICE QUESTIONS

1. A gas diode can conduct .... the equivalent vacuum diode for the same plate voltage.
   (i) less current than
   (ii) more current than
   (iii) same current as
   (iv) none of the above
2. A gas-filled tube has .... resistance before ionisation.
   (i) very high (ii) small
   (iii) very small (iv) zero
3. The PIV of a hot cathode gas diode is .... the equivalent vacuum diode.
   (i) the same as that of
   (ii) more than that of
   (iii) less than that of
   (iv) none of the above

* 100 to 1000 μsec.

4. The anode-to-cathode potential of a gas-filled tube at which gas deionises and stops conduction is called ....potential.
   (*i*) extinction (*ii*) striking
   (*iii*) ionising (*iv*) none of the above
5. A thyratron can be used as ....
   (*i*) an oscillator (*ii*) an amplifier
   (*iii*) a controlled switch
   (*iv*) none of the above
6. The internal resistance of a gas-filled tube is .... that of a vacuum tube.
   (*i*) the same as (*ii*) more than
   (*iii*) less than (*iv*) none of the above
7. A cold cathode tube is generally used as a ....
   (*i*) diode (*ii*) triode
   (*iii*) tetrode (*iv*) pentode
8. Conduction in a cold cathode tube is started by ....
   (*i*) thermionic emission
   (*ii*) secondary emission
   (*iii*) natural sources
   (*iv*) none of the above
9. The cathode heating time of thermionic gas diode is .... that of a vacuum diode.
   (*i*) the same as (*ii*) much more than
   (*iii*) much less than (*iv*) none of the above
10. The solid state equivalent of thyratron is ....
    (*i*) FET (*ii*) transistor
    (*iii*) SCR (*iv*) crystal diode
11. The solid state equivalent of cold cathode diode is ....
    (*i*) zener diode (*ii*) crystal diode
    (*iii*) LED (*iv*) transistor
12. The noise in a gas-filled tube is .... that in a vacuum tube.
    (*i*) the same as (*ii*) more than
    (*iii*) less than (*iv*) none of the above
13. The ionisation potential in a gas diode depends upon ....
    (*i*) plate current
    (*ii*) cathode construction
    (*iii*) size of the tube
    (*iv*) type and pressure of gas
14. If the gas pressure in a gas-filled diode is increased, its PIV rating ....
    (*i*) remains the same
    (*ii*) is increased
    (*iii*) is decreased
    (*iv*) none of the above
15. Once a thyratron is fired, its control grid .... over the plate current.
    (*i*) loses all control
    (*ii*) exercises fine control
    (*iii*) exercises rough control
    (*iv*) none of the above
16. To stop conduction in a thyratron, the .... voltage should be reduced to zero.
    (*i*) grid (*ii*) plate
    (*iii*) filament (*iv*) none of the above
17. Ionisation of cold cathode diode takes place at .... plate potential compared to hot cathode gas diode.
    (*i*) the same (*ii*) much higher
    (*iii*) much lesser (*iv*) none of the above
18. A gas-filled tube has .... internal resistance after ionisation.
    (*i*) low (*ii*) high
    (*iii*) very high (*iv*) moderate
19. The gas-filled tubes can handle .... peak inverse voltage (PIV) as compared to equivalent vacuum tubes.
    (*i*) more (*ii*) less
    (*iii*) the same (*iv*) none of the above
20. A cold cathode diode is used as .... tube.
    (*i*) a rectifier
    (*ii*) a power-controlled
    (*iii*) a regulating
    (*iv*) an amplifier

Answers to Multiple-Choice Questions

1. (*ii*)	**2.** (*i*)	**3.** (*iii*)	**4.** (*i*)	**5.** (*iii*)
6. (*iii*)	**7.** (*i*)	**8.** (*iii*)	**9.** (*ii*)	**10.** (*iii*)
11. (*i*)	**12.** (*ii*)	**13.** (*iv*)	**14.** (*iii*)	**15.** (*i*)
16. (*ii*)	**17.** (*ii*)	**18.** (*i*)	**19.** (*ii*)	**20.** (*iii*)

Chapter Review Topics

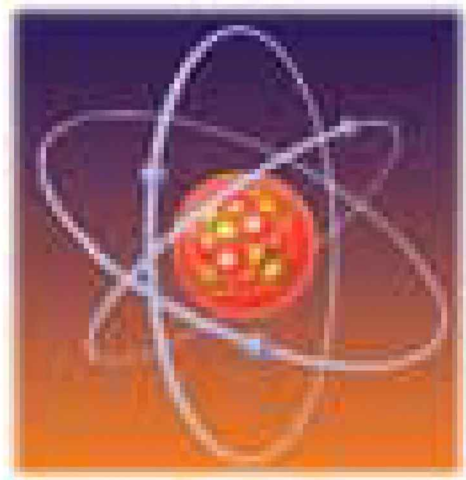
1. Explain the differences between a gas tube and equivalent vacuum tube.
2. Explain how ionisation takes place in a hot-gas diode. How does current conduction take place in such a tube ?
3. Give the schematic symbols of glow tube, hot-cathode gas diode and thyratron.
4. Explain the construction, operation and characteristics of a glow tube.
5. Discuss some applications of glow tubes.
6. What is a thyratron ? How does it differ from a vacuum triode ?
7. Write short notes on the following :
   (*i*) Characteristics of hot-cathode gas diode
   (*ii*) Applications of thyratrons

Discussion Questions

1. What are the advantages of gas tubes over vacuum tubes ?
2. What is the difference between the action of thyratron and vacuum triode ?
3. Why cannot thyratrons be used as rectifiers for high voltages ?
4. Can gas diodes be used as rectifiers for high voltages ?
5. What is the drawback of a gas diode compared to a thyratron ?

4

Atomic Structure



INTRODUCTION

The study of atomic structure is of considerable importance for electronics engineering. Unfortunately, the size of an atom is so small that it is virtually impossible to see it even with the most powerful microscope. Therefore, we have to employ indirect method for the study of its structure. The method consists of studying the properties of atom experimentally. After this, a *guess* is made regarding the possible structure of atom, which should satisfy the properties studied experimentally.

Various scientists have given different theories regarding the structure of atom. However, for the purpose of understanding electronics, the study of Bohr's atomic model is adequate. Although numerous refinements on Bohr's atomic model have since been made, we still believe in the laws that Bohr applied to the atomic world. In this chapter, we shall deal with Bohr's atomic model in order to understand the problems facing the electronic world.

4.1 Bohr's Atomic Model

In 1913, Neils Bohr, Danish Physicist gave clear explanation of atomic structure. According to Bohr:

(*i*) An atom consists of a positively charged nucleus around which negatively charged electrons revolve in different *circular orbits*.

(*ii*) The electrons can revolve around the nucleus only in certain permitted orbits *i.e.* orbits of certain radii are allowed.

(*iii*) The electrons in each permitted orbit have a certain fixed amount of energy. The larger the orbit (*i.e.* larger radius), the greater is the energy of electrons.

(*iv*) If an electron is given additional energy (*e.g.* heat, light etc.), it is lifted to the higher orbit. The atom is said to be in a state of *excitation*. This state does not last long, because the electron soon falls back to the original lower orbit. As it falls, it gives back the acquired energy in the form of heat, light or other radiations.

Neils Bohr (1885-1962)

Fig. 4.1 shows the structure of silicon atom. It has 14 electrons. Two electrons revolve in the first orbit, 8 in the second orbit and 4 in the third orbit. The first, second, third orbits etc. are also known as *K*, *L*, *M* orbits respectively.

These electrons can revolve only in permitted orbits (*i.e.* orbits of *radii r_1, r_2 and r_3) and not in any arbitrary orbit. Thus, all radii between r_1 and r_2 or between r_2 and r_3 are forbidden. Each orbit has fixed amount of energy associated with it. If an electron in the first orbit is to be lifted to the second orbit, just the **right amount of energy should be supplied to it. When this electron jumps from the second orbit to first, it will give back the acquired energy in the form of electromagnetic radiations.

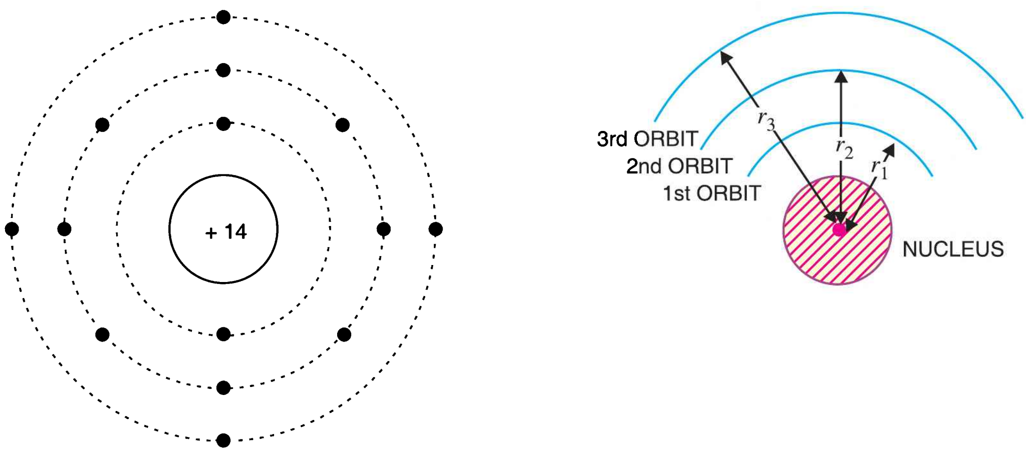


Fig. 4.1

4.2 Energy Levels

It has already been discussed that each orbit has fixed amount of energy associated with it. The electrons moving in a particular orbit possess the energy of that orbit. The larger the orbit, the greater is its energy. It becomes clear that outer orbit electrons possess more energy than the inner orbit electrons.

A convenient way of representing the energy of different orbits is shown in Fig. 4.2 (*ii*). This is known as energy level diagram. The first orbit represents the *first energy level*, the second orbit

* The values of radii are determined from quantum considerations.

** So that its total energy is equal to that of second orbit.

indicates the *second energy level* and so on. The larger the orbit of an electron, the greater is its energy and higher is the energy level.

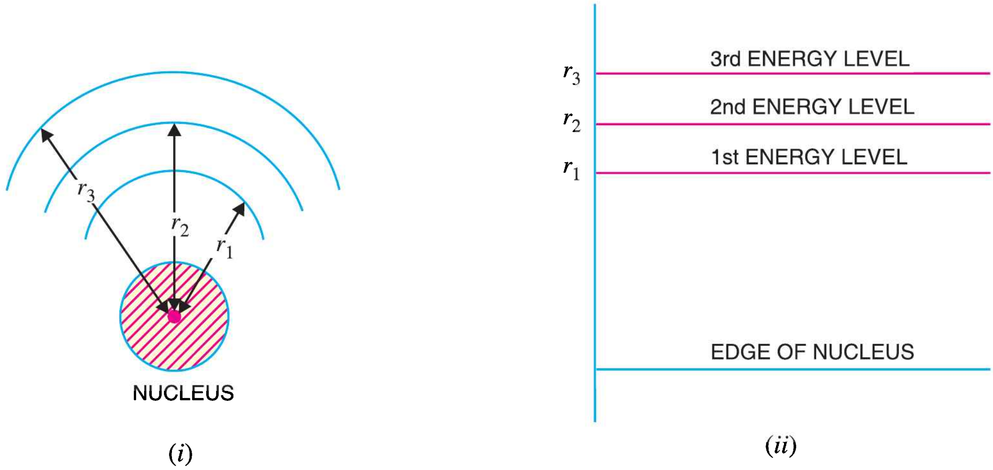


Fig. 4.2

4.3 Energy Bands

In case of a single isolated atom, the electrons in any orbit possess definite energy. However, an atom in a solid is greatly influenced by the closely-packed neighbouring atoms. The result is that the electron in any orbit of such an atom can have a range of energies rather than a single energy. This is known as *energy band.*

The range of energies possessed by an electron in a solid is known as **energy band.**

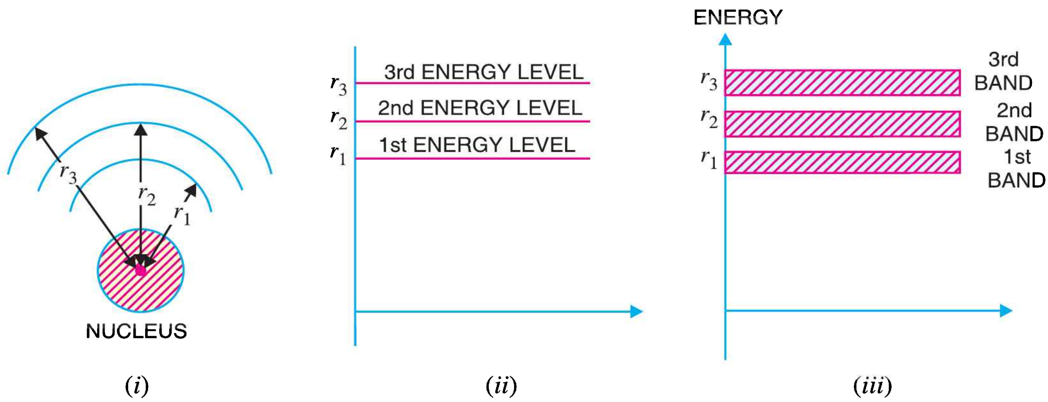


Fig. 4.3

The concept of energy band can be easily understood by referring to Fig. 4.3. Fig. 4.3 (*ii*) shows the energy levels of a single isolated atom of silicon. Each orbit of an atom has a single energy. Therefore, an electron can have only single energy corresponding to the orbit in which it exists. However, when the atom is in a solid, the electron in any orbit can have a range of energies. For instance, electrons in the first orbit have slightly different energies because no two electrons in this orbit see exactly the same charge environment. Since there are millions of first orbit electrons, the slightly different energy levels form a band, called 1st energy band [See Fig. 4.3 (*iii*)]. The electrons in the first orbit can have any energy range in this band. Similarly, second orbit electrons form second energy band and so on.

4.4 Important Energy Bands in Solids

As discussed before, individual *K*, *L*, *M* etc. energy levels of an isolated atom are converted into corresponding bands when the atom is in a solid. Though there are a number of energy bands in solids, the following are of particular importance [See Fig. 4.4] :

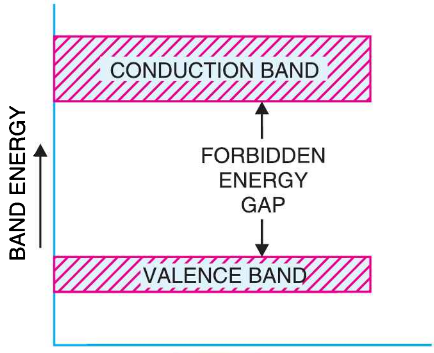


Fig. 4.4

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

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

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

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

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

(*iii*) **Forbidden energy gap.** *The separation between conduction band and valence band on the energy level diagram is known as* **forbidden energy gap.**

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

4.5 Classification of Solids and Energy Bands

We know that some solids are good conductors of electricity while others are insulators. There is also an intermediate class of semiconductors. The difference in the behaviour of solids as regards their electrical conductivity can be beautifully explained in terms of energy bands. The electrons in the lower energy band are tightly bound to the nucleus and play no part in the conduction process. However, the valence and conduction bands are of particular importance in ascertaining the electrical behaviour of various solids.

Insulators

(*i*) **Insulators.** Insulators (*e.g.* wood, glass etc.) are those substances which do not allow the passage of electric current through them. In terms of energy band, the valence band is full while the conduction band is empty. Further, the energy gap between valence and conduction bands is very large ($\simeq$ 15 eV) as shown in Fig. 4.5. Therefore, a very high electric field is required to push the valence electrons to the conduction band.

For this reason, the electrical conductivity of such materials is extremely small and may be regarded as nil under ordinary conditions.

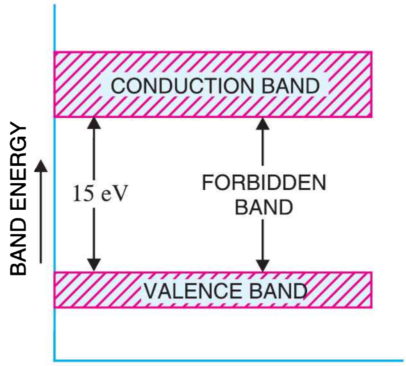


Fig. 4.5

At room temperature, the valence electrons of the insulators do not have enough energy to cross over to the conduction band. However, when the temperature is raised, some of the valence electrons may acquire enough energy to cross over to the conduction band. Hence, the resistance of an insulator decreases with the increase in temperature *i.e.* an insulator has negative temperature co-efficient of resistance.

(*ii*) Conductors. Conductors (*e.g.* copper, aluminium) are those substances which easily allow the passage of electric current through them. It is because there are a large number of free electrons available in a conductor. In terms of energy band, the valence and conduction bands overlap each other as shown in Fig. 4.6. Due to this overlapping, a slight potential difference across a conductor causes the free electrons to constitute electric current. Thus, the electrical behaviour of conductors can be satisfactorily explained by the band energy theory of materials.

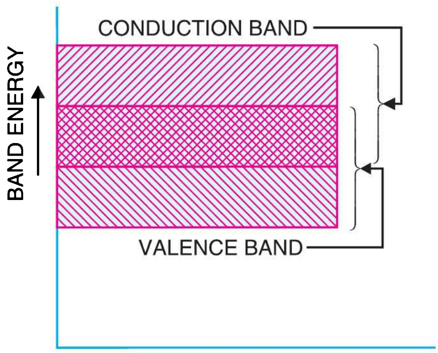


Fig. 4.6

(*iii*) Semicondutors. Semiconductors (*e.g.* germanium, silicon etc.) are those substances whose electrical conductivity lies inbetween conductors and insulators. In terms of energy band, the valence band is almost filled and conduction band is almost empty. Further, the energy gap between valence and conduction bands is very small as shown in Fig. 4.7. Therefore, comparatively smaller electric field (smaller than insulators but much greater than conductors) is required to push the electrons from the valence band to the conduction band. In short, a semiconductor has :

(*a*) almost full valence band

(*b*) almost empty conduction band

(*c*) small energy gap ($\simeq$ 1 eV) between valence and conduction bands.

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

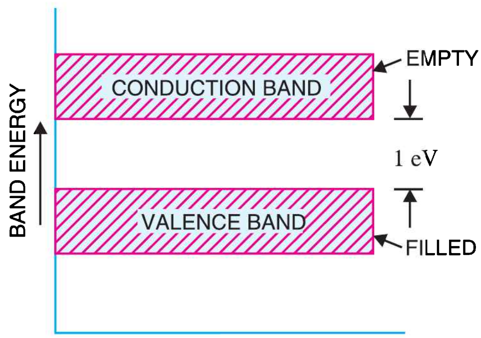


Fig. 4.7

4.6 Silicon

During the infancy of electronic industry, both germanium and silicon were used in the manufacture of semiconductor devices. As the electronic field advanced, it was realised that silicon was superior to germanium in many respects. Since silicon is the most widely used material in the manufacture of semiconductor devices, we shall continue to discuss the properties of this material (as compared to germanium) as and when we get the chance.

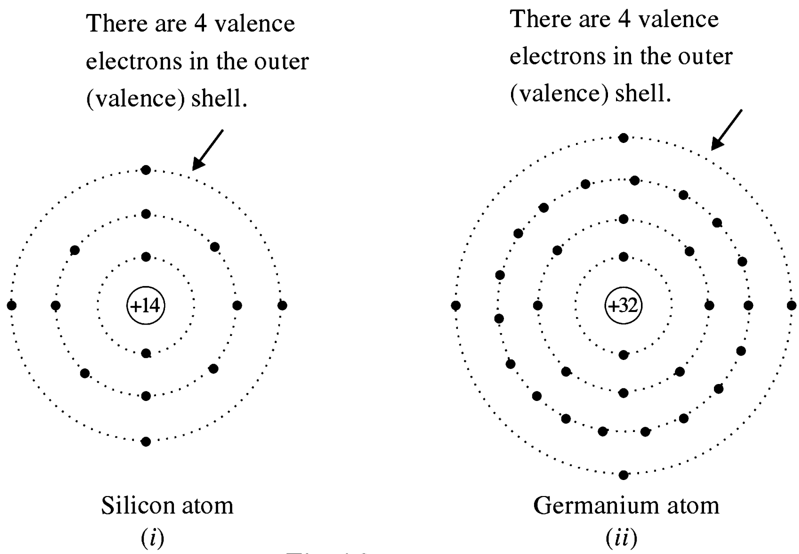


Fig. 4.8

(i) Note the atomic structure of germanium and silicon in Fig. 4.8 carefully. The valence electrons in germanium are in the fourth orbit while the valence electrons in silicon are in the third orbit; closer to the nucleus. Therefore, the germanium valence electrons are at higher energy level than those in silicon. This means that germanium valence electrons require smaller amount of additional energy to escape from the atom and become free electron. What is the effect of this property? This property makes germanium more unstable at high temperatures. This is the basic reason why silicon is widely used as semiconductor material.

(ii) Fig. 4.9 shows the energy level/band of silicon atom. The atomic number of silicon is 14 so that its 14 electrons are distributed in 3 orbits. Each energy level/band is associated with certain amount of energy and is separated from the adjacent bands by energy gap. ***No electron can exist in the energy gap.*** For an electron to jump from one orbit to the next higher orbit, external energy

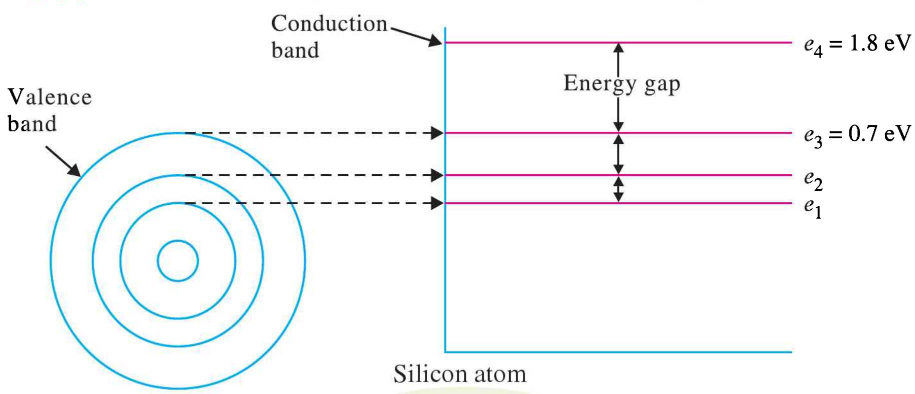


Fig. 4.9

(e.g. heat) equal to the energy difference of the orbits must be supplied. For example, the valence band is shown to have an energy level of 0.7 eV. The conduction band is shown to have an energy level of 1.8 eV. Thus for an electron to jump from the valence band to the conduction band, an energy = 1.8 – 0.7 = 1.1 eV must be supplied. As you will see, the energy band description of substances is very important in understanding many fields of science and engineering including electronics.

MULTIPLE-CHOICE QUESTIONS

1. The electrons in the third orbit of an atom have .... energy than the electrons in the second orbit.
(*i*) more (*ii*) less
(*iii*) the same (*iv*) none of the above

2. When an electron jumps from higher orbit to a lower orbit, it .... energy.
(*i*) absorbs (*ii*) emits
(*iii*) sometimes emits, sometimes absorbs
(*iv*) none of the above

3. Which of the following is quantized according to Bohr's theory of atom ?
(*i*) linear momentum of electron
(*ii*) linear velocity of electron
(*iii*) angular momentum of electron
(*iv*) angular velocity of electron

4. A semiconductor has .... band.
(*i*) almost empty valence
(*ii*) almost empty conduction
(*iii*) almost full conduction
(*iv*) none of the above

5. The electrons in the conduction band are known as ....
(*i*) bound electrons (*ii*) valence electrons
(*iii*) free electrons (*iv*) none of the above

6. In insulators, the energy gap between valence and conduction bands is ....
(*i*) very large (*ii*) zero
(*iii*) very small (*iv*) none of the above

7. In a conductor, the energy gap between valence and conduction bands is ....
(*i*) large (*ii*) very large
(*iii*) very small (*iv*) none of the above

8. According to Bohr's theory of atom, an electron can move in an orbit of ....
(*i*) any radius
(*ii*) certain radius
(*iii*) some range of radii
(*iv*) none of the above

9. In a semiconductor, the energy gap between valence and conduction bands is about ....
(*i*) 15 eV (*ii*) 100 eV
(*iii*) 50 eV (*iv*) 1 eV

10. The energy gap between valence and conduction bands in insulators is about ....
(*i*) 15 eV (*ii*) 1.5 eV
(*iii*) zero (*iv*) 0.5 eV

Answers to Multiple-Choice Questions

1. (*i*)	**2.** (*ii*)	**3.** (*iii*)	**4.** (*ii*)	**5.** (*iii*)
6. (*i*)	**7.** (*iii*)	**8.** (*ii*)	**9.** (*iv*)	**10.** (*i*)

Chapter Review Topics

1. Explain the salient features of Bohr's atomic model.
2. Explain the concept of energy bands in solids.
3. Describe the valence band, conduction band and forbidden energy gap with the help of energy level diagram.
4. Give the energy band description of conductors, semiconductors and insulators.

Discussion Questions

1. Why is the energy of an electron more in higher orbits ?
2. What is the concept of energy band ?
3. Why do conduction band electrons possess very high energy ?
4. Why are valence electrons of a material so important ?
5. What is the difference between energy level and energy band ?

5

Semiconductor Physics

INTRODUCTION

Certain substances like germanium, silicon, carbon etc. are neither good conductors like copper nor insulators like glass. In other words, the resistivity of these materials lies inbetween conductors and insulators. Such substances are classified as *semiconductors*. Semiconductors have some useful properties and are being extensively used in electronic circuits. For instance, *transistor*—a semiconductor device is fast replacing bulky vacuum tubes in almost all applications. Transistors are only one of the family of semiconductor devices ; many other semiconductor devices are becoming increasingly popular. In this chapter, we shall focus our attention on the different aspects of semiconductors.

5.1 Semiconductor

It is not easy to define a semiconductor if we want to take into account all its physical characteristics. However, generally, a semiconductor is defined on the basis of electrical conductivity as under :

A **semiconductor** *is a substance which has resistivity (10^{-4} to 0.5 Ωm) inbetween conductors and insulators e.g. germanium, silicon, selenium, carbon etc.*

The reader may wonder, when a semiconductor is neither a good conductor nor an insulator, then why not to classify it as a *resistance material* ? The answer shall be readily available if we study the following table :

S.No.	Substance	Nature	Resistivity
1	Copper	good conductor	1.7×10^{-8} Ω m
2	Germanium	semiconductor	0.6 Ω m
3	Glass	insulator	9×10^{11} Ω m
4	Nichrome	resistance material	10^{-4} Ω m

Comparing the resistivities of above materials, it is apparent that the resistivity of germanium (semiconductor) is quite high as compared to copper (conductor) but it is quite low when compared with glass (insulator). This shows that resistivity of a semiconductor lies inbetween conductors and insulators. However, it will be wrong to consider the semiconductor as a resistance material. For example, nichrome, which is one of the highest resistance material, has resistivity much lower than germanium. This shows that electrically germanium cannot be regarded as a conductor or insulator or a resistance material. This gave such substances like germanium the name of semiconductors.

It is interesting to note that it is not the resistivity alone that decides whether a substance is semiconductor or not. For example, it is just possible to prepare an alloy whose resistivity falls within the range of semiconductors but the alloy cannot be regarded as a semiconductor. In fact, semiconductors have a number of peculiar properties which distinguish them from conductors, insulators and resistance materials.

Properties of Semiconductors

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

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

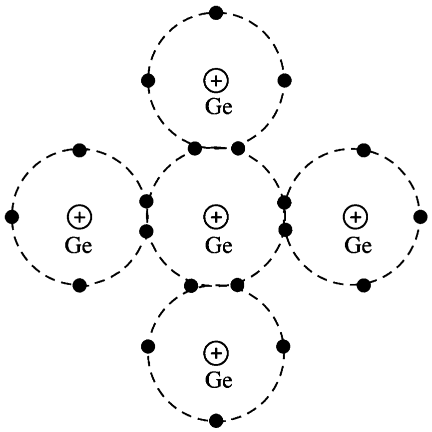
(*iii*) When a suitable metallic impurity (*e.g.* arsenic, gallium etc.) is added to a semiconductor, its current conducting properties change appreciably. This property is most important and is discussed later in detail.

5.2 Bonds in Semiconductors

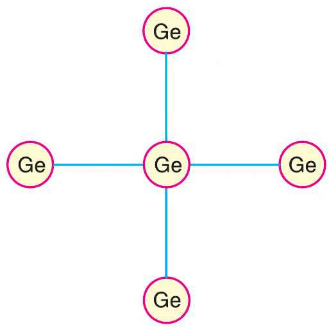
The atoms of every element are held together by the bonding action of valence electrons. This bonding is due to the fact that it is the tendency of each atom to complete its last orbit by acquiring 8 electrons in it. However, in most of the substances, the last orbit is incomplete *i.e.* the last orbit does not have 8 electrons. This makes the atom active to enter into bargain with other atoms to acquire 8 electrons in the last orbit. To do so, the atom may lose, gain or share valence electrons with other atoms. In semiconductors, bonds are formed by sharing of valence electrons. Such bonds are called *co-valent bonds*. In the formation of a co-valent bond, each atom contributes equal number of valence electrons and the contributed electrons are shared by the atoms engaged in the formation of the bond.

Fig. 5.1 shows the co-valent bonds among germanium atoms. A germanium atom has *4 valence electrons. It is the tendency of each germanium atom to have 8 electrons in the last orbit. To do so, each germanium atom positions itself between four other germanium atoms as shown in Fig. 5.1 (*i*). Each neighbouring atom shares one valence electron with the central atom. In this business of sharing, the central atom completes its last orbit by having 8 electrons revolving around the nucleus. In this way, the central atom sets up co-valent bonds. Fig. 5.1 (*ii*) shows the bonding diagram.

The following points may be noted regarding the co-valent bonds :



COVALENT BONDS AMONG Ge ATOMS
(*i*)

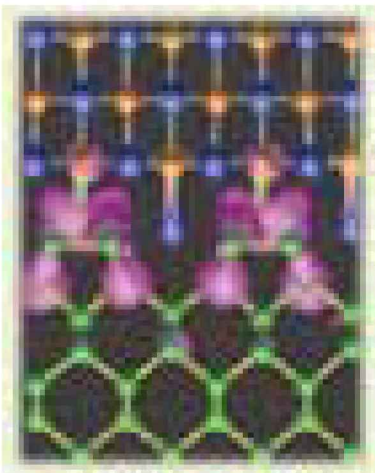


BONDING DIAGRAM
(*ii*)

Fig. 5.1

(*i*) Co-valent bonds are formed by sharing of valence electrons.

(*ii*) In the formation of co-valent bond, each valence electron of an atom forms direct bond with the valence electron of an adjacent atom. In other words, valence electrons are associated with particular atoms. For this reason, valence electrons in a semiconductor are not free.



Bonds in Semiconductor

5.3 Crystals

A substance in which the atoms or molecules are arranged in an orderly pattern is known as a *crystal*. All semi-conductors have crystalline structure. For example, referring to Fig. 5.1, it is clear that each atom is surrounded by neighbouring atoms in a repetitive manner. Therefore, a piece of germanium is generally called germanium crystal.

5.4 Commonly Used Semiconductors

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

* A germanium atom has 32 electrons. First orbit has 2 electrons, second 8 electrons, third 18 electrons and the fourth orbit has 4 electrons.

(*i*) **Germanium.** Germanium has become the model substance among the semiconductors; the main reason being that it can be purified relatively well and crystallised easily. Germanium is an earth element and was discovered in 1886. It is recovered from the ash of certain coals or from the flue dust of zinc smelters. Generally, recovered germanium is in the form of germanium dioxide powder which is then reduced to pure germanium.

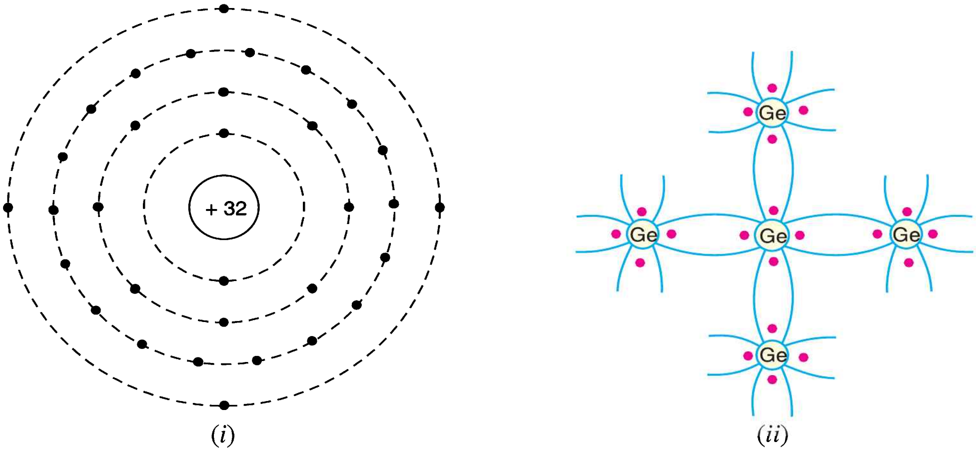


Fig. 5.2

The atomic number of germanium is 32. Therefore, it has 32 protons and 32 electrons. Two electrons are in the first orbit, eight electrons in the second, eighteen electrons in the third and four electrons in the outer or valence orbit [See Fig. 5.2 (*i*)]. It is clear that germanium atom has four valence electrons *i.e.*, it is a tetravalent element. Fig. 5.2 (*ii*) shows how the various germanium atoms are held through co-valent bonds. As the atoms are arranged in an orderly pattern, therefore, germanium has crystalline structure.

(*ii*) **Silicon.** Silicon is an element in most of the common rocks. Actually, sand is silicon dioxide. The silicon compounds are chemically reduced to silicon which is 100% pure for use as a semiconductor.

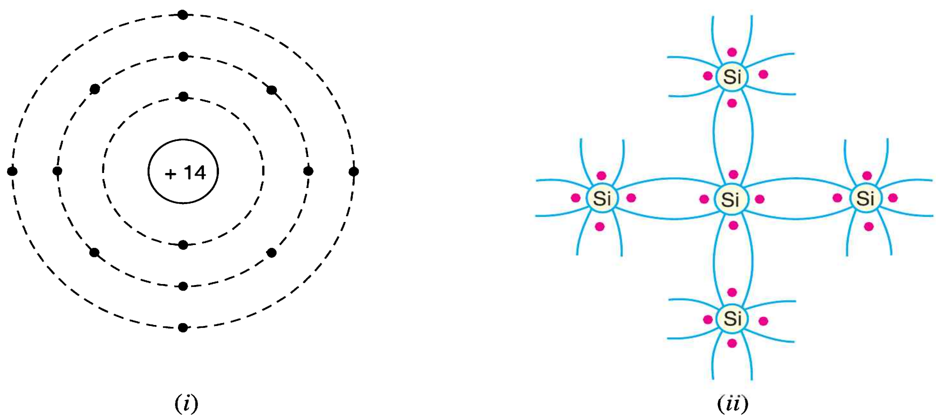


Fig. 5.3

The atomic number of silicon is 14. Therefore, it has 14 protons and 14 electrons. Two electrons are in the first orbit, eight electrons in the second orbit and four electrons in the third orbit [See Fig. 5.3 (*i*)]. It is clear that silicon atom has four valence electrons *i.e.* it is a tetravalent element. Fig. 5.3 (*ii*) shows how various silicon atoms are held through co-valent bonds. Like germanium, silicon atoms are also arranged in an orderly manner. Therefore, silicon has crystalline structure.

5.5 Energy Band Description of Semiconductors

It has already been discussed that a semiconductor is a substance whose resistivity lies between conductors and insulators. The resistivity is of the order of 10^{-4} to 0.5 ohm metre. However, a semiconductor can be defined much more comprehensively on the basis of energy bands as under :

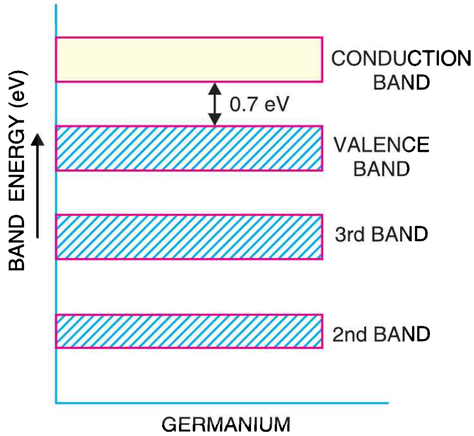


Fig. 5.4

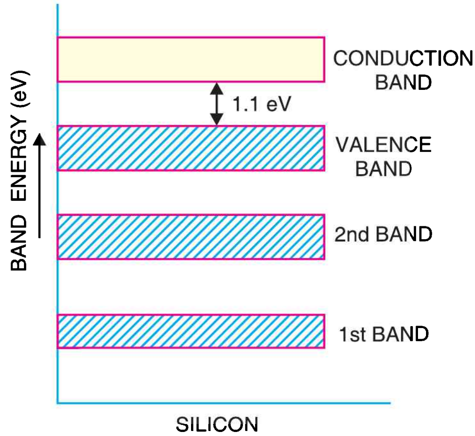


Fig. 5.5

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

Figs. 5.4 and 5.5 show the energy band diagrams of germanium and silicon respectively. It may be seen that forbidden energy gap is very small; being 1.1 eV for silicon and 0.7 eV for germanium. Therefore, relatively small energy is needed by their valence electrons to cross over to the conduction band. Even at room temperature, some of the valence electrons may acquire sufficient energy to enter into the conduction band and thus become free electrons. However, at this temperature, the number of free electrons available is very *small. Therefore, at room temperature, a piece of germanium or silicon is neither a good conductor nor an insulator. For this reason, such substances are called *semiconductors*.

The energy band description is extremely helpful in understanding the current flow through a semiconductor. Therefore, we shall frequently use this concept in our further discussion.

5.6 Effect of Temperature on Semiconductors

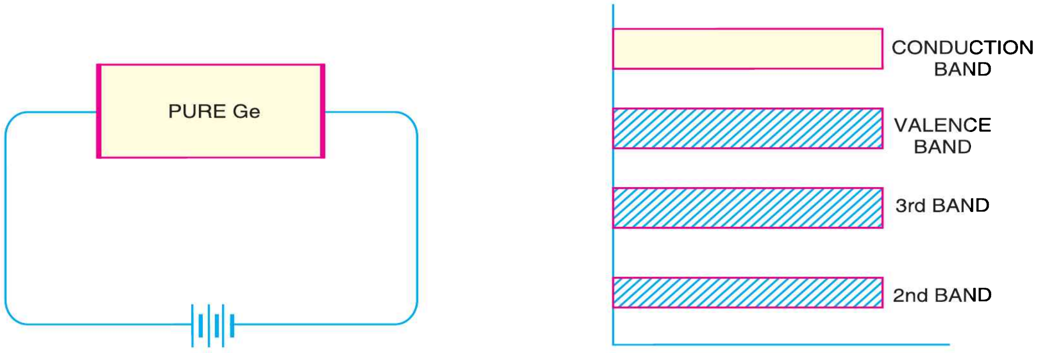
The electrical conductivity of a semiconductor changes appreciably with temperature variations. This is a very important point to keep in mind.

(*i*) At absolute zero. At absolute zero temperature, all the electrons are tightly held by the semiconductor atoms. The inner orbit electrons are bound whereas the valence electrons are engaged in co-valent bonding. At this temperature, the co-valent bonds are very strong and there are no free electrons. Therefore, the semiconductor crystal behaves as a perfect insulator [See Fig. 5.6 (*i*)].

In terms of energy band description, the valence band is filled and there is a large energy gap between valence band and conduction band. Therefore, no valence electron can reach the conduction band to become free electron. It is due to the non-availability of free electrons that a semiconductor behaves as an insulator.

* Out of 10^{10} semiconductor atoms, one atom provides a free electron.

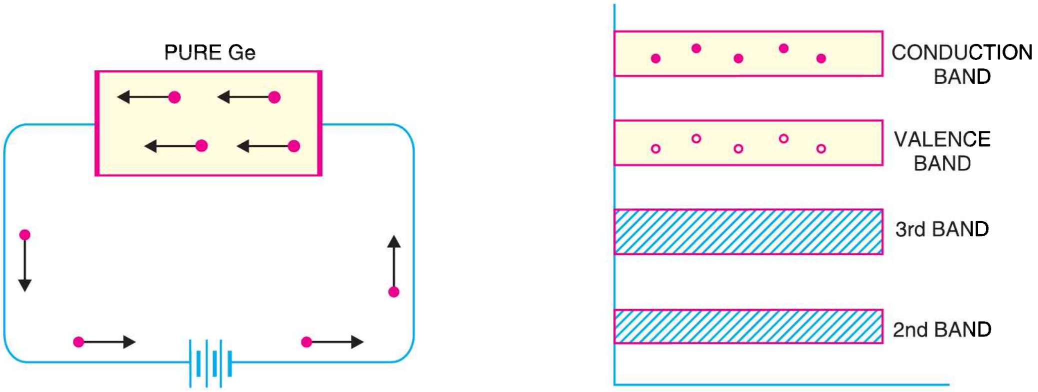
(*ii*) Above absolute zero. When the temperature is raised, some of the covalent bonds in the semiconductor break due to the thermal energy supplied. The breaking of bonds sets those electrons *free* which are engaged in the formation of these bonds. The result is that a few free electrons exist in the semiconductor. These free electrons can constitute a tiny electric current if potential difference is



(*i*) (*ii*)

Fig. 5.6

applied across the semiconductor crystal [See Fig. 5.7 (*i*)]. *This shows that the resistance of a semiconductor decreases with the rise in temperature i.e.* it has negative temperature coefficient of resistance. It may be added that at room temperature, current through a semiconductor is too small to be of any practical value.



(*i*) (*ii*)

Fig. 5.7

Fig. 5.7 (*ii*) shows the energy band diagram. As the temperature is raised, some of the valence electrons acquire sufficient energy to enter into the conduction band and thus become free electrons. Under the influence of electric field, these free electrons will constitute electric current. It may be noted that each time a valence electron enters into the conduction band, a *hole* is created in the valence band. As we shall see in the next article, holes also contribute to current. In fact, hole current is the most significant concept in semiconductors.

5.7 Hole Current

At room temperature, some of the co-valent bonds in pure semiconductor break, setting up free electrons. Under the influence of electric field, these free electrons constitute electric current. At the same time, another current – the hole current – also flows in the semiconductor. When a covalent bond is broken due to thermal energy, the removal of one electron leaves a vacancy *i.e.* a missing electron in the covalent bond. This missing electron is called a *hole which acts as a positive charge. For one electron set free, one hole is created. Therefore, thermal energy creates *hole-electron pairs*; there being as many holes as the free electrons. The current conduction by holes can be explained as follows :

The hole shows a missing electron. Suppose the valence electron at *L* (See Fig. 5.8) has become free electron due to thermal energy. This creates a hole in the co-valent bond at *L*. The hole is a strong centre of attraction **for the electron. A valence electron (say at *M*) from nearby co-valent bond comes to fill in the hole at *L*. This results in the creation of hole at *M*. Another valence electron (say at *N*) in turn may leave its bond to fill the hole at *M*, thus creating a hole at *N*. Thus the hole having a positive charge has moved from *L* to *N* *i.e.* towards the negative terminal of supply. This constitutes *hole current*.

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

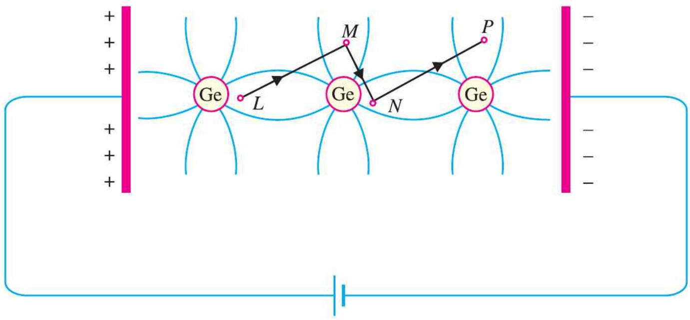


Fig. 5.8

Energy band description. The hole current can be beautifully explained in terms of energy bands. Suppose due to thermal energy, an electron leaves the valence band to enter into the conduction band as shown in Fig. 5.9.

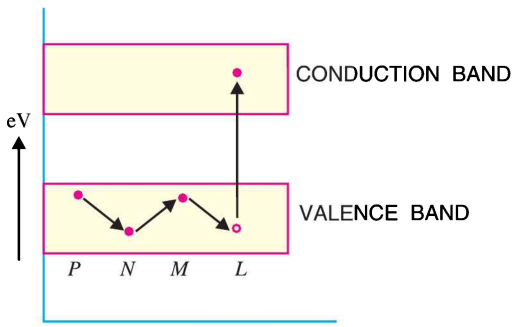


Fig. 5.9

This leaves a vacancy at *L*. Now the valence electron at *M* comes to fill the hole at *L*. The result is that hole disappears at *L* and appears at *M*. Next, the valence electron at *N* moves into the hole at *M*. Consequently, hole is created at *N*. It is clear that valence electrons move along the path *PNML* whereas holes move in the opposite direction *i.e.* along the path *LMNP*.

* Note that hole acts as a virtual charge, although there is no physical charge on it.

** There is a strong tendency of semiconductor crystal to form co-valent bonds. Therefore, a hole attracts an electron from the neighbouring atom.

*** Unlike the normal current which is by free electrons.

5.8 Intrinsic Semiconductor

A semiconductor in an extremely pure form is known as an **intrinsic semiconductor**.

In an intrinsic semiconductor, even at room temperature, hole-electron pairs are created. When electric field is applied across an intrinsic semiconductor, the current conduction takes place by two processes, namely ; by *free electrons* and *holes* as shown in Fig. 5.10. The free electrons are produced due to the breaking up of some covalent bonds by thermal energy. At the same time, holes are created in the covalent bonds. Under the influence of electric field, conduction through the semiconductor is by both free electrons and holes. Therefore, the total current inside the semiconductor is the sum of currents due to free electrons and holes.

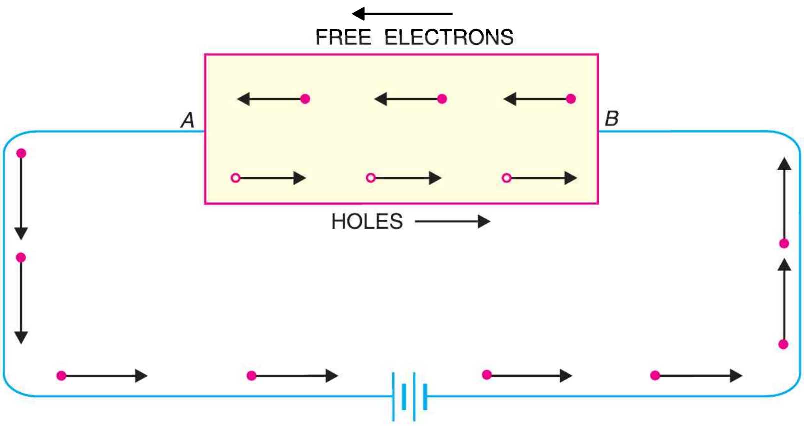


Fig. 5.10

It may be noted that current in the external wires is fully electronic *i.e.* by electrons. What about the holes ? Referring to Fig. 5.10, holes being positively charged move towards the negative terminal of supply. As the holes reach the negative terminal *B*, electrons enter the semiconductor crystal near the terminal and combine with holes, thus cancelling them. At the same time, the loosely held electrons near the positive terminal *A* are attracted away from their atoms into the positive terminal. This creates new holes near the positive terminal which again drift towards the negative terminal.

5.9 Extrinsic Semiconductor

The intrinsic semiconductor has little current conduction capability at room temperature. To be useful in electronic devices, the pure semiconductor must be altered so as to significantly increase its conducting properties. This is achieved by adding a small amount of suitable impurity to a semiconductor. It is then called *impurity* or *extrinsic semiconductor*. The process of adding impurities to a semiconductor is known as *doping*. The amount and type of such impurities have to be closely controlled during the preparation of extrinsic semiconductor. Generally, for 10^8 atoms of semiconductor, one impurity atom is added.

The purpose of adding impurity is to increase either the number of free electrons or holes in the semiconductor crystal. As we shall see, if a pentavalent impurity (having 5 valence electrons) is added to the semiconductor, a large number of free electrons are produced in the semiconductor. On the other hand, addition of trivalent impurity (having 3 valence electrons) creates a large number of holes in the semiconductor crystal. Depending upon the type of impurity added, extrinsic semiconductors are classified into:

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

5.10 *n*-type Semiconductor

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

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

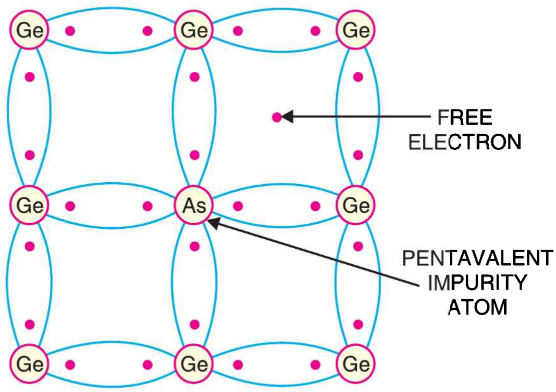


Fig. 5.11

To explain the formation of *n*-type semiconductor, consider a pure germanium crystal. We know that germanium atom has four valence electrons. When a small amount of pentavalent impurity like arsenic is added to germanium crystal, a large number of free electrons become available in the crystal. The reason is simple. Arsenic is pentavalent *i.e.* its atom has five valence electrons. An arsenic atom fits in the germanium crystal in such a way that its four valence electrons form covalent bonds with four germanium atoms. The *fifth* valence electron of arsenic atom finds no place in co-valent bonds and is thus free as shown in Fig. 5.11. Therefore, for each arsenic atom added, one free electron will be available in the germanium crystal. Though each arsenic atom provides one free electron, yet an extremely small amount of arsenic impurity provides enough atoms to supply millions of free electrons.

Fig. 5.12 shows the energy band description of *n*-type semi-conductor. The addition of pentavalent impurity has produced a number of conduction band electrons *i.e.*, free electrons. The four valence electrons of pentavalent atom form covalent bonds with four neighbouring germanium atoms. The fifth left over valence electron of the pentavalent atom cannot be accommodated in the valence band and travels to the conduction band. The following points may be noted carefully :

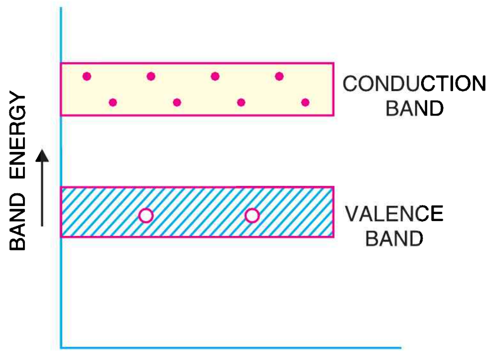


Fig. 5.12

(*i*) Many new free electrons are produced by the addition of pentavalent impurity.

(*ii*) Thermal energy of room temperature still generates a few hole-electron pairs. However, the number of free electrons provided by the pentavalent impurity far exceeds the number of holes. It is due to this predominance of electrons over holes that it is called *n*-type semiconductor (*n* stands for negative).

n-type conductivity. The current conduction in an *n*-type semiconductor is *predominantly* by free electrons *i.e.* negative charges and is called *n-type* or *electron type conductivity*. To understand *n*-type conductivity, refer to Fig. 5.13. When p.d. is applied across the *n*-type semiconductor, the free electrons (donated by impurity) in the crystal will be directed towards the positive terminal, constituting electric current. As the current flow through the crystal is by free electrons which are carriers of negative charge, therefore, this type of conductivity is called negative or *n*-type conductivity. It may be noted that conduction is just as in ordinary metals like copper.

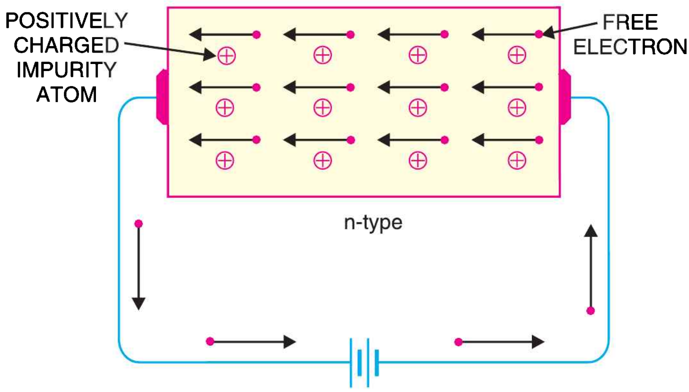


Fig. 5.13

5.11 *p*-type Semiconductor

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

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

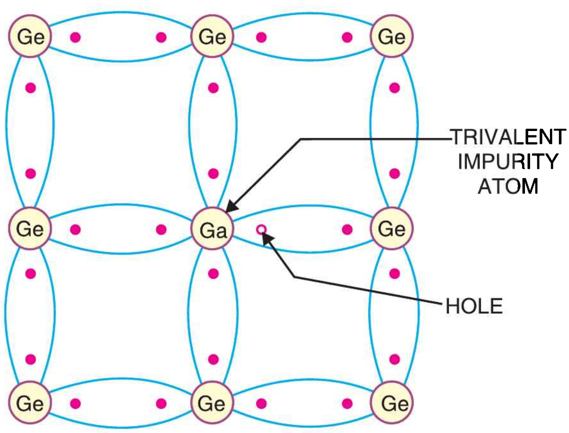


Fig. 5.14

To explain the formation of *p*-type semiconductor, consider a pure germanium crystal. When a small amount of trivalent impurity like gallium is added to germanium crystal, there exists a large number of holes in the crystal. The reason is simple. Gallium is trivalent *i.e.* its atom has three valence electrons. Each atom of gallium fits into the germanium crystal but now only three co-valent bonds can be formed. It is because three valence electrons of gallium atom can form only three single co-valent bonds with three germanium atoms as shown in Fig. 5.14. In the fourth co-valent bond, only germanium atom contributes one valence electron while gallium has no valence electron to contribute as all its three valence electrons are already engaged in the co-valent bonds with neighbouring germanium atoms. In other words, fourth bond is incomplete; being short of one electron. This missing electron is called a *hole*. Therefore, for each gallium atom added, one hole is created. A small amount of gallium provides millions of holes.

Fig. 5.15 shows the energy band description of the *p*-type semiconductor. The addition of trivalent impurity has produced a large number of holes. However, there are a few conduction band electrons due to thermal energy associated with room temperature. But the holes far outnumber the conduction band electrons. It is due to the predominance of holes over free electrons that it is called *p*-type semiconductor (*p* stands for positive).

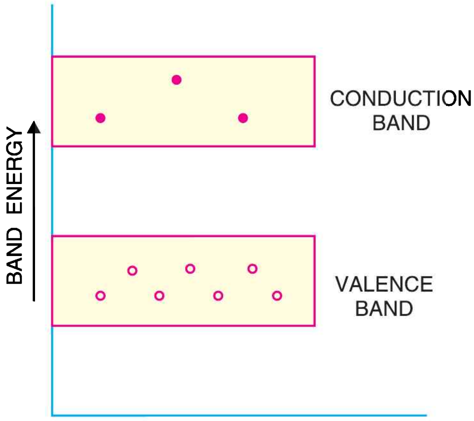


Fig. 5.15

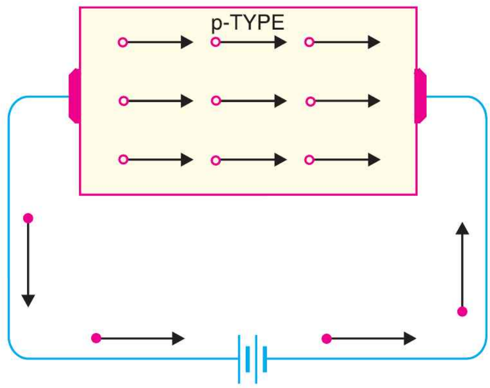


Fig. 5.16

p-type conductivity. The current conduction in *p*-type semiconductor is predominantly by holes *i.e.* positive charges and is called *p-type* or *hole-type conductivity*. To understand *p*-type conductivity, refer to Fig. 5.16. When *p.d.* is applied to the *p*-type semiconductor, the holes (donated by the impurity) are shifted from one co-valent bond to another. As the holes are positively charged, therefore, they are directed towards the negative terminal, constituting what is known as hole current. It may be noted that in *p*-type conductivity, the valence electrons move from one co-valent bond to another unlike the *n*-type where current conduction is by free electrons.

5.12 Charge on *n*-type and *p*-type Semiconductors

As discussed before, in *n*-type semiconductor, current conduction is due to excess of electrons whereas in a *p*-type semiconductor, conduction is by holes. The reader may think that *n*-type material has a net negative charge and *p*-type a net positive charge. But this conclusion is wrong. It is true that *n*-type semiconductor has excess of electrons but these extra electrons were supplied by the atoms of donor impurity and each atom of donor impurity is electrically neutral. When the impurity atom is added, the term "excess electrons" refers to an excess with regard to the number of electrons needed to fill the co-valent bonds in the semiconductor crystal. The extra electrons are free electrons and increase the conductivity of the semiconductor. The situation with regard to *p*-type semiconductor is also similar. *It follows, therefore, that n-type as well as p-type semiconductor is electrically neutral.*

5.13 Majority and Minority Carriers

It has already been discussed that due to the effect of impurity, *n*-type material has a large number of free electrons whereas *p*-type material has a large number of holes. However, it may be recalled that even at room temperature, some of the co-valent bonds break, thus releasing an equal number of free electrons and holes. An *n*-type material has its share of electron-hole pairs (released due to breaking of bonds at room temperature) but in addition has a much larger quantity of free electrons due to the effect of impurity. These impurity-caused free electrons are not associated with holes. Consequently, an *n*-type material has a large number of free electrons and a small number of holes as shown in Fig. 5.17 (*i*). The free electrons in this case are considered *majority carriers* — since the majority portion of current in *n*-type material is by the flow of free electrons — and the holes are the *minority carriers*.

Similarly, in a *p*-type material, holes outnumber the free electrons as shown in Fig. 5.17 (*ii*). Therefore, holes are the majority carriers and free electrons are the minority carriers.

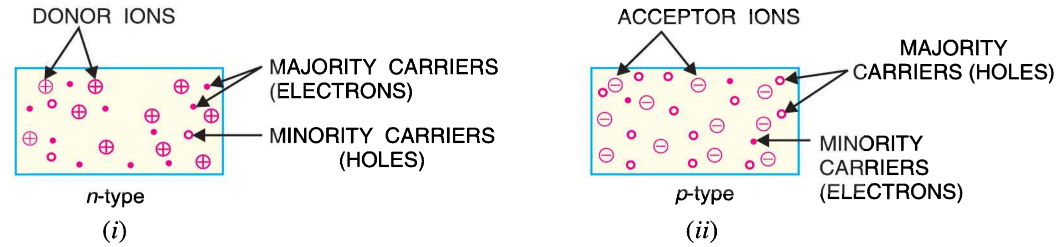


Fig. 5.17

5.14 *pn* Junction

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

Most semiconductor devices contain one or more *pn* junctions. The *pn* junction is of great importance because it is in effect, the *control element* for semiconductor devices. A thorough knowledge of the formation and properties of *pn* junction can enable the reader to understand the semiconductor devices.

Formation of *pn* junction. In actual practice, the characteristic properties of *pn* junction will not be apparent if a *p*-type block is just brought in contact with *n*-type block. In fact, *pn* junction is fabricated by special techniques. One common method of making *pn* junction is called *alloying*. In this method, a small block of indium (trivalent impurity) is placed on an *n*-type germanium slab as shown in Fig. 5.18 (*i*). The system is then heated to a temperature of about 500ºC. The indium and some of the germanium melt to form a small puddle of molten germanium-indium mixture as shown in Fig. 5.18 (*ii*). The temperature is then lowered and puddle begins to solidify. Under proper conditions, the atoms of indium impurity will be suitably adjusted in the germanium slab to form a single crystal. The addition of indium overcomes the excess of electrons in the *n*-type germanium to such an extent that it creates a *p*-type region.

As the process goes on, the remaining molten mixture becomes increasingly rich in indium. When all germanium has been redeposited, the remaining material appears as indium button which is frozen on to the outer surface of the crystallised portion as shown in Fig. 5.18 (*iii*). This button serves as a suitable base for soldering on leads.

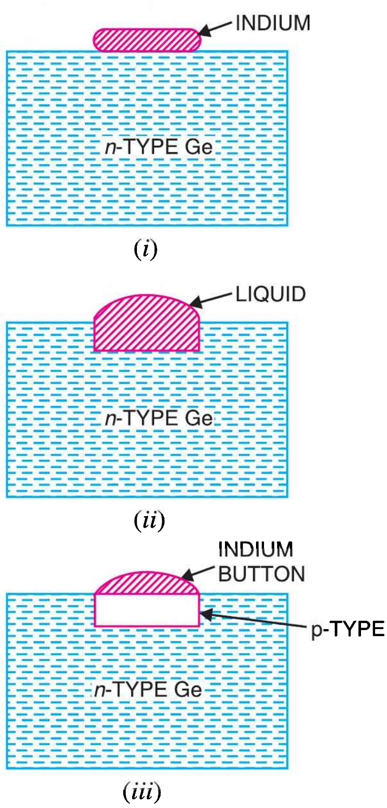
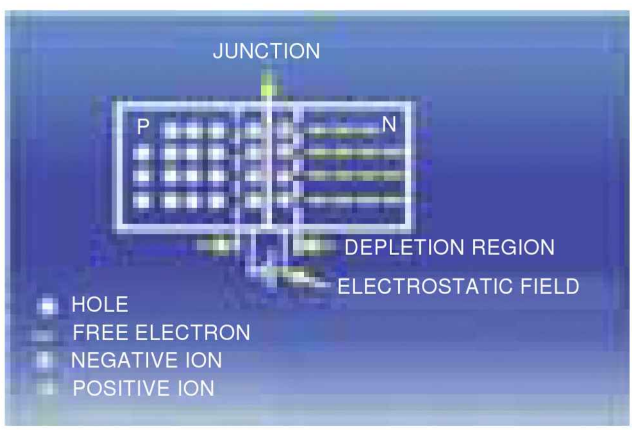


Fig. 5.18



5.15 Properties of *pn* Junction

At the instant of *pn*-junction formation, the free electrons near the junction in the *n* region begin to diffuse across the junction into the *p* region where they combine with holes near the junction. The result is that *n* region loses free electrons as they diffuse into the junction. This creates a layer of positive charges (pentavalent ions) near the junction. As the electrons move across the junction, the *p* region loses holes as the electrons and holes combine. The result is that there is a layer of negative charges (trivalent ions) near the junction. These two layers of positive and negative charges form the *depletion region* (or *depletion layer*). The term depletion is due to the fact that near the junction, the region is depleted (*i.e.* emptied) of *charge carries* (free electrons and holes) due to diffusion across the junction. It may be noted that depletion layer is formed very quickly and is very thin compared to the *n* region and the *p* region. For clarity, the width of the depletion layer is shown exaggerated.

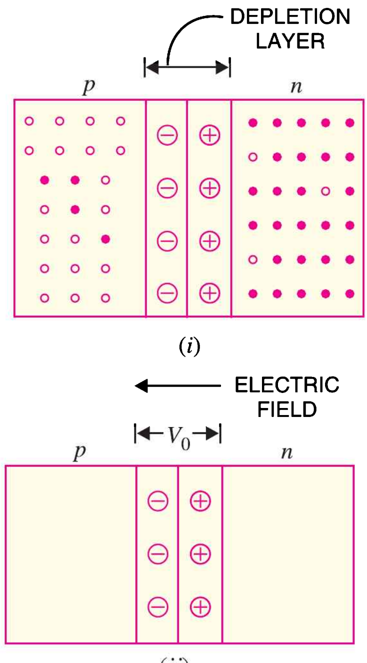


Fig. 5.19

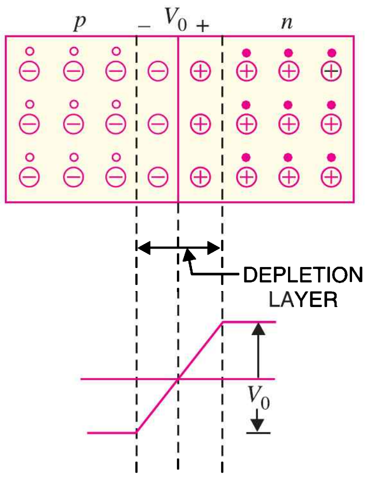


Fig. 5.20

Once *pn* junction is formed and depletion layer created, the diffusion of free electrons stops. In other words, the depletion region acts as a barrier to the further movement of free electrons across the junction. The positive and negative charges set up an electric field. This is shown by a black arrow in Fig. 5.19 (*i*). The electric field is a barrier to the free electrons in the *n*-region. There exists a potential difference across the depletion layer and is called **barrier potential** (V_0). The barrier potential of a *pn* junction depends upon several factors including the type of semiconductor material, the amount of doping and temperature. The typical barrier potential is approximately:

For silicon, $V_0 = 0.7$ V ; For germanium, $V_0 = 0.3$ V

Fig. 5.20 shows the potential (V_0) distribution curve.

5.16 Applying D.C. Voltage Across *pn* Junction or Biasing a pn Junction

In electronics, the term bias refers to the use of d.c. voltage to establish certain operating conditions for an electronic device. In relation to a *pn* junction, there are following two bias conditions :

1. **Forward biasing** 2. **Reverse biasing**

1. **Forward biasing.** *When external d.c. voltage applied to the junction is in such a direction that it cancels the potential barrier, thus permitting current flow, it is called* **forward biasing.**

To apply forward bias, connect positive terminal of the battery to *p*-type and negative terminal to *n*-type as shown in Fig. 5.21. The applied forward potential establishes an electric field which acts against the field due to potential barrier. Therefore, the resultant field is weakened and the barrier height is reduced at the junction as shown in Fig. 5.21. As potential barrier voltage is very small (0.1 to 0.3 V), therefore, a small forward voltage is sufficient to completely eliminate the barrier. Once the potential barrier is eliminated by the forward voltage, junction resistance becomes almost zero and a low resistance path is established for the entire circuit. Therefore, current flows in the circuit. This is called *forward current*. With forward bias to *pn* junction, the following points are worth noting :

(*i*) The potential barrier is reduced and at some forward voltage (0.1 to 0.3 V), it is eliminated altogether.

(*ii*) The junction offers low resistance (called *forward resistance*, R_f) to current flow.

(*iii*) Current flows in the circuit due to the establishment of low resistance path. The magnitude of current depends upon the applied forward voltage.

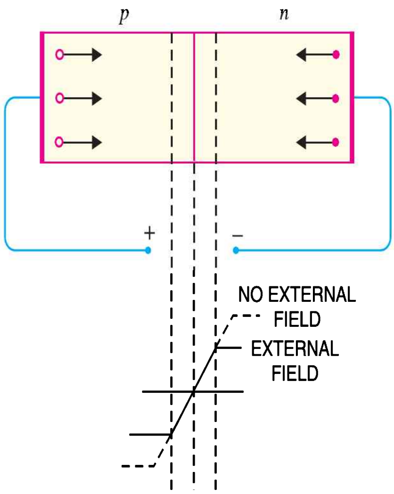


Fig. 5.21

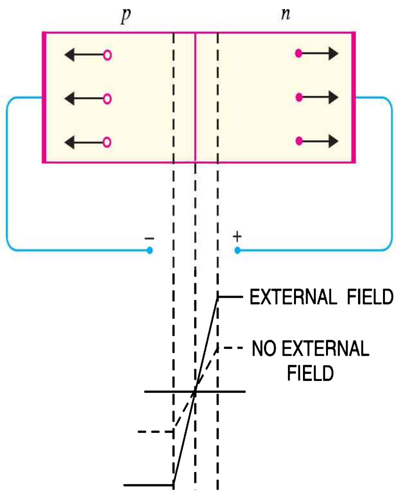


Fig. 5.22

2. **Reverse biasing.** *When the external d.c. voltage applied to the junction is in such a direction that potential barrier is increased, it is called* **reverse biasing.**

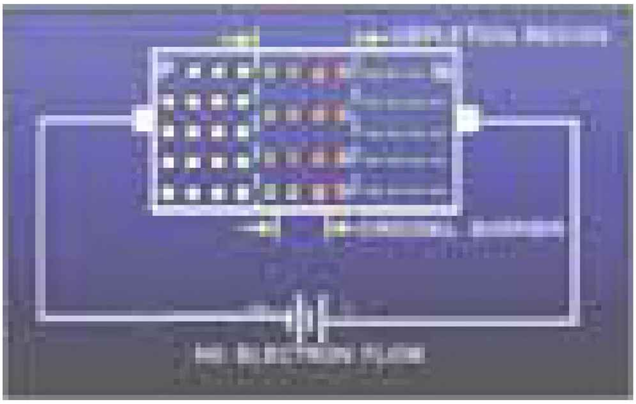
To apply reverse bias, connect negative terminal of the battery to *p*-type and positive terminal to *n*-type as shown in Fig. 5.22. It is clear that applied reverse voltage establishes an electric field which acts in the same direction as the field due to potential barrier. Therefore, the resultant field at the junction is strengthened and the barrier height is increased as shown in Fig. 5.22. The increased potential barrier prevents the flow of charge carriers across the junction. Thus, a high resistance path is established for the entire circuit and hence the current does not flow. With reverse bias to *pn* junction, the following points are worth noting :

(*i*) The potential barrier is increased.

(*ii*) The junction offers very high resistance (called *reverse resistance*, R_r) to current flow.

(*iii*) No current flows in the circuit due to the establishment of high resistance path.

Conclusion. From the above discussion, it follows that with reverse bias to the junction, a high resistance path is established and hence no current flow occurs. On the other hand, with forward bias to the junction, a low resistance path is set up and hence current flows in the circuit.



5.17 Current Flow in a Forward Biased pn Junction

We shall now see how current flows across *pn* junction when it is forward biased. Fig. 5.23 shows a forward biased *pn* junction. Under the influence of forward voltage, the free electrons in *n*-type move *towards the junction, leaving behind positively charged atoms. However, more electrons arrive from the negative battery terminal and enter the *n*-region to take up their places. As the free electrons reach the junction, they become **valence electrons. As valence electrons, they move through the holes in the *p*-region. The valence electrons move towards left in the *p*-region which is equivalent to the holes moving to right. When the valence electrons reach the left end of the crystal, they flow into the positive terminal of the battery.

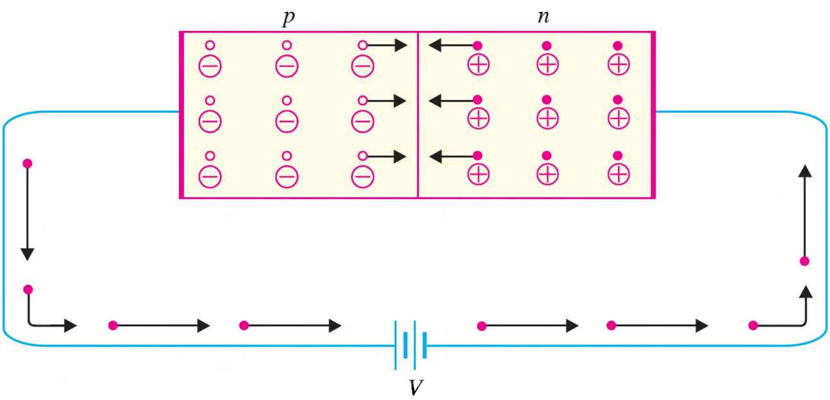


Fig. 5.23

The mechanism of current flow in a forward biased *pn* junction can be summed up as under :

(*i*) The free electrons from the negative terminal continue to pour into the *n*-region while the free electrons in the *n*-region move towards the junction.

(*ii*) The electrons travel through the *n*-region as free-electrons *i.e.* current in *n*-region is by free electrons.

* Note that negative terminal of battery is connected to *n*-type. It repels the free electrons in *n*-type towards the junction.

** A hole is in the co-valent bond. When a free electron combines with a hole, it becomes a valence electron.

(*iii*) When these electrons reach the junction, they combine with holes and become valence electrons.

(*iv*) The electrons travel through *p*-region as valence electrons *i.e.* current in the *p*-region is by holes.

(*v*) When these valence electrons reach the left end of crystal, they flow into the positive terminal of the battery.

From the above discussion, it is concluded that in *n*-type region, current is carried by free electrons whereas in *p*-type region, it is carried by holes. However, in the external connecting wires, the current is carried by free electrons.

5.18 Volt-Ampere Characteristics of *pn* Junction

Volt-ampere or *V-I* characteristic of a *pn* junction (also called a *crystal or semiconductor diode*) is the curve between voltage across the junction and the circuit current. Usually, voltage is taken along *x*-axis and current along *y*-axis. Fig. 5.24 shows the *circuit arrangement for determining the *V-I* characteristics of a *pn* junction. The characteristics can be studied under three heads, namely; *zero external voltage*, *forward bias* and *reverse bias*.

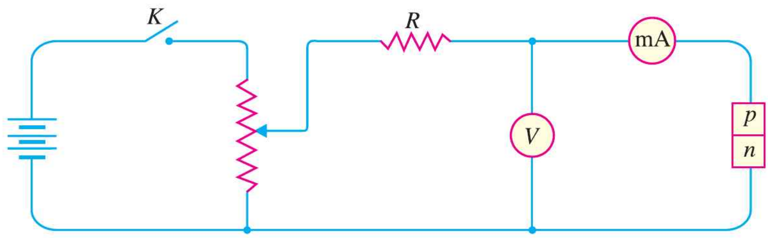


Fig. 5.24

(*i*) **Zero external voltage.** When the external voltage is zero, *i.e.* circuit is open at *K*, the potential barrier at the junction does not permit current flow. Therefore, the circuit current is zero as indicated by point *O* in Fig. 5.25.

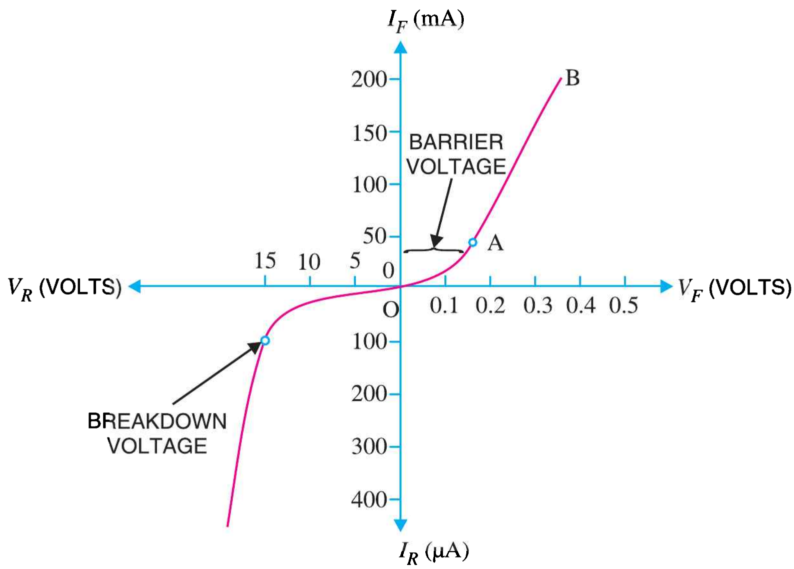


Fig. 5.25

---

* *R* is the current limiting resistance. It prevents the forward current from exceeding the permitted value.

(*ii*) **Forward bias.** With forward bias to the *pn* junction *i.e.* *p*-type connected to positive terminal and *n*-type connected to negative terminal, the potential barrier is reduced. At some forward voltage (0.7 V for Si and 0.3 V for Ge), the potential barrier is altogether eliminated and current starts flowing in the circuit. From now onwards, the current increases with the increase in forward voltage. Thus, a rising curve *OB* is obtained with forward bias as shown in Fig. 5.25. From the forward characteristic, it is seen that at first (*region OA*),the current increases very slowly and the curve is non-linear. It is because the external applied voltage is used up in overcoming the potential barrier. However, once the external voltage exceeds the potential barrier voltage, the *pn* junction behaves like an ordinary conductor. Therefore, the current rises very sharply with increase in external voltage (*region AB on the curve*). The curve is almost linear.

(*iii*) **Reverse bias.** With reverse bias to the *pn* junction *i.e.* *p*-type connected to negative terminal and *n*-type connected to positive terminal, potential barrier at the junction is increased. Therefore, the junction resistance becomes very high and practically no current flows through the circuit. However, in practice, a very small current (of the order of μ*A*) flows in the circuit with reverse bias as shown in the reverse characteristic. This is called *reverse *saturation current* (I_s) and is due to the minority carriers. It may be recalled that there are a few free electrons in *p*-type material and a few holes in *n*-type material. These undesirable free electrons in *p*-type and holes in *n*-type are called *minority carriers*. As shown in Fig. 5.26, to these minority carriers, the applied reverse bias appears as forward bias. Therefore, a **small current flows in the reverse direction.

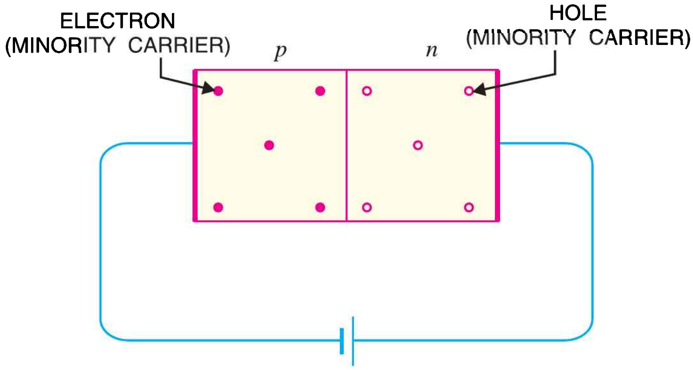


Fig. 5.26

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

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

5.19 Important Terms

Two important terms often used with *pn* junction (*i.e.* crystal diode) are *breakdown voltage* and *knee voltage*. We shall now explain these two terms in detail.

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

Under normal reverse voltage, a very little reverse current flows through a *pn* junction. However, if the reverse voltage attains a high value, the junction may break down with sudden rise in

* The term saturation comes from the fact that it reaches its maximum level quickly and does not significantly change with the increase in reverse voltage.

** Reverse current increases with reverse voltage but can generally be regarded as negligible over the working range of voltages.

reverse current. For understanding this point, refer to Fig. 5.27. Even at room temperature, some hole-electron pairs (minority carriers) are produced in the depletion layer as shown in Fig. 5.27 (*i*). With reverse bias, the electrons move towards the positive terminal of supply. At large reverse voltage, these electrons acquire high enough velocities to dislodge valence electrons from semiconductor atoms as shown in Fig. 5.27 (*ii*). The newly liberated electrons in turn free other valence electrons. In this way, we get an *avalanche* of free electrons. Therefore, the *pn* junction conducts a very large reverse current.

Once the breakdown voltage is reached, the high reverse current may damage the junction. Therefore, care should be taken that reverse voltage across a *pn* junction is always less than the breakdown voltage.

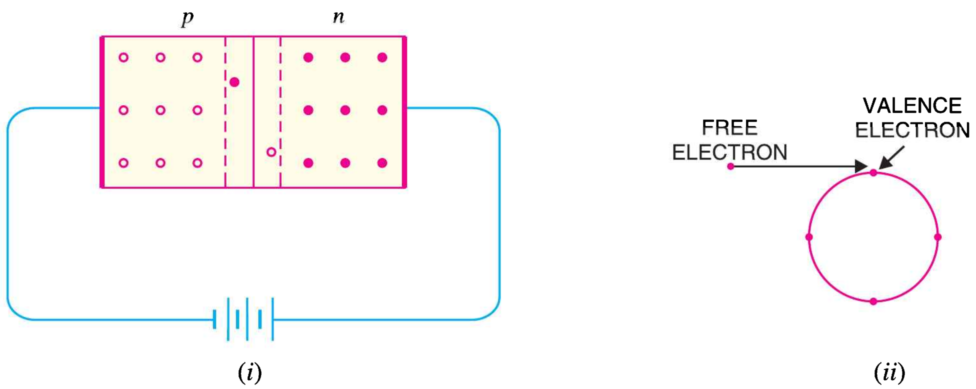


Fig. 5.27

(*ii*) Knee voltage. *It is the forward voltage at which the current through the junction starts to increase rapidly.*

When a diode is forward biased, it conducts current very slowly until we overcome the potential barrier. For silicon *pn* junction, potential barrier is 0.7 V whereas it is 0.3 V for germanium junction. It is clear from Fig. 5.28 that knee voltage for silicon diode is 0.7 V and 0.3 V for germanium diode.

Once the applied forward voltage exceeds the knee voltage, the current starts increasing rapidly. It may be added here that in order to get useful current through a *pn* junction, the applied voltage must be more than the knee voltage.

Note. The potential barrier voltage is also known as *turn-on voltage*. This is obtained by taking the straight line portion of the forward characteristic and extending it back to the horizontal axis.

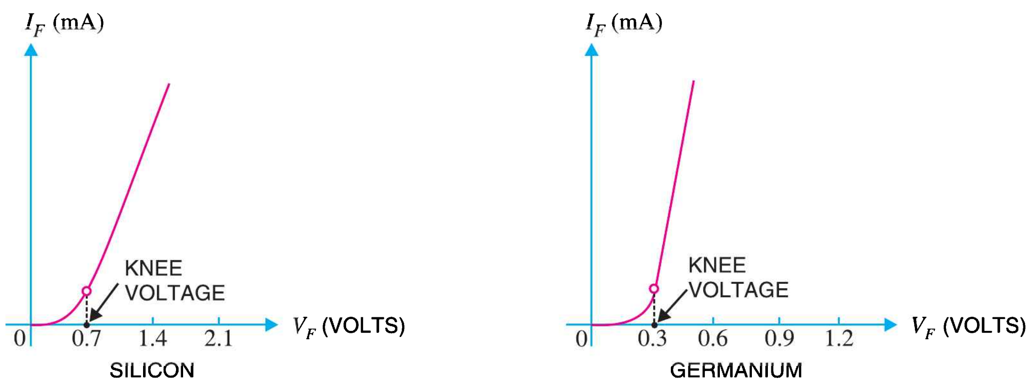


Fig. 5.28

5.20 Limitations in the Operating Conditions of *pn* Junction

Every *pn* junction has limiting values of *maximum forward current, peak inverse voltage* and *maximum power rating*. The *pn* junction will give satisfactory performance if it is operated within these limiting values. However, if these values are exceeded, the *pn* junction may be destroyed due to excessive heat.

(*i*) *Maximum forward current.* It is the highest instantaneous forward current that a *pn* junction can conduct without damage to the junction. Manufacturer's data sheet usually specifies this rating. If the forward current in a *pn* junction is more than this rating, the junction will be destroyed due to overheating.

(*ii*) *Peak inverse voltage (PIV).* It is the maximum reverse voltage that can be applied to the *pn* junction without damage to the junction. If the reverse voltage across the junction exceeds its PIV, the junction may be destroyed due to excessive heat. The peak inverse voltage is of particular importance in rectifier service. A *pn* junction *i.e.* a crystal diode is used as a rectifier to change alternating current into direct current. In such applications, care should be taken that reverse voltage across the diode during negative half-cycle of a.c. does not exceed the PIV of diode.

(*iii*) *Maximum power rating.* It is the maximum power that can be dissipated at the junction without damaging it. The power dissipated at the junction is equal to the product of junction current and the voltage across the junction. This is a very important consideration and is invariably specified by the manufacturer in the data sheet.

MULTIPLE-CHOICE QUESTIONS

1. A semiconductor is formed by .... bonds.
(*i*) covalent (*ii*) electrovalent
(*iii*) co-ordinate (*iv*) none of the above

2. A semiconductor has .... temperature coefficient of resistance.
(*i*) positive (*ii*) zero
(*iii*) negative (*iv*) none of the above

3. The most commonly used semiconductor is ....
(*i*) germanium (*ii*) silicon
(*iii*) carbon (*iv*) sulphur

4. A semiconductor has generally .... valence electrons.
(*i*) 2 (*ii*) 3
(*iii*) 6 (*iv*) 4

5. The resistivity of pure germanium under standard conditions is about ....
(*i*) $6 \times 10^4\ \Omega$ cm (*ii*) $60\ \Omega$ cm
(*iii*) $3 \times 10^6\ \Omega$ cm (*iv*) $6 \times 10^{-4}\ \Omega$ cm

6. The resistivity of pure silicon is about ....
(*i*) $100\ \Omega$ cm (*ii*) $6000\ \Omega$ cm
(*iii*) $3 \times 10^5\ \Omega$ cm (*iv*) $1.6 \times 10^{-8}\ \Omega$ cm

7. When a pure semiconductor is heated, its resistance ....
(*i*) goes up (*ii*) goes down
(*iii*) remains the same (*iv*) cannot say

8. The strength of a semiconductor crystal comes from ....
(*i*) forces between nuclei
(*ii*) forces between protons
(*iii*) electron-pair bonds
(*iv*) none of the above

9. When a pentavalent impurity is added to a pure semiconductor, it becomes ....
(*i*) an insulator
(*ii*) an intrinsic semiconductor
(*iii*) *p*-type semiconductor
(*iv*) *n*-type semiconductor

10. Addition of pentavalent impurity to a semiconductor creates many ....
(*i*) free electrons (*ii*) holes
(*iii*) valence electrons
(*iv*) bound electrons

11. A pentavalent impurity has .... valence electrons.
(*i*) 3 (*ii*) 5
(*iii*) 4 (*iv*) 6

12. An *n*-type semiconductor is ....
(*i*) positively charged
(*ii*) negatively charged

(*iii*) electrically neutral
(*iv*) none of the above

13. A trivalent impurity has .... valence electrons.
(*i*) 4 (*ii*) 5
(*iii*) 6 (*iv*) 3

14. Addition of trivalent impurity to a semiconductor creates many ....
(*i*) holes (*ii*) free electrons
(*iii*) valence electrons
(*iv*) bound electrons

15. A hole in a semiconductor is defined as ....
(*i*) a free electron
(*ii*) the incomplete part of an electron pair bond
(*iii*) a free proton
(*iv*) a free neutron

16. The impurity level in an extrinsic semiconductor is about .... of pure semiconductor.
(*i*) 10 atoms for 10^8 atoms
(*ii*) 1 atom for 10^8 atoms
(*iii*) 1 atom for 10^4 atoms
(*iv*) 1 atom for 100 atoms

17. As the doping to a pure semiconductor increases, the bulk resistance of the semiconductor ....
(*i*) remains the same
(*ii*) increases
(*iii*) decreases
(*iv*) none of the above

18. A hole and electron in close proximity would tend to ....
(*i*) repel each other
(*ii*) attract each other
(*iii*) have no effect on each other
(*iv*) none of the above

19. In a semiconductor, current conduction is due ....
(*i*) only to holes
(*ii*) only to free electrons
(*iii*) to holes and free electrons
(*iv*) none of the above

20. The random motion of holes and free electrons due to thermal agitation is called ....
(*i*) diffusion (*ii*) pressure
(*iii*) ionisation (*iv*) none of the above

21. A forward biased *pn* junction has a resistance of the ....
(*i*) order of Ω (*ii*) order of kΩ
(*iii*) order of MΩ (*iv*) none of the above

22. The battery connections required to forward bias a *pn* junction are ....
(*i*) +ve terminal to *p* and −ve terminal to *n*
(*ii*) −ve terminal to *p* and +ve terminal to *n*
(*iii*) −ve terminal to *p* and −ve terminal to *n*
(*iv*) none of the above

23. The barrier voltage at a *pn* junction for germanium is about ....
(*i*) 3.5 V (*ii*) 3V
(*iii*) zero (*iv*) 0.3 V

24. In the depletion region of a *pn* junction, there is a shortage of ....
(*i*) acceptor ions (*ii*) holes and electrons
(*iii*) donor ions (*iv*) none of the above

25. A reverse biased *pn* junction has ....
(*i*) very narrow depletion layer
(*ii*) almost no current
(*iii*) very low resistance
(*iv*) large current flow

26. A *pn* junction acts as a ....
(*i*) controlled switch
(*ii*) bidirectional switch
(*iii*) unidirectional switch
(*iv*) none of the above

27. A reverse biased *pn* junction has resistance of the........
(*i*) order of Ω (*ii*) order of kΩ
(*iii*) order of MΩ (*iv*) none of the above

28. The leakage current across a *pn* junction is due to ....
(*i*) minority carriers
(*ii*) majority carriers
(*iii*) junction capacitance
(*iv*) none of the above

29. When the temperature of an extrinsic semiconductor is increased, the pronounced effect is on ....
(*i*) junction capacitance

(*ii*) minority carriers
(*iii*) majority carriers
(*iv*) none of the above

30. With forward bias to a *pn* junction, the width of depletion layer ....
(*i*) decreases (*ii*) increases
(*iii*) remains the same
(*iv*) none of the above

31. The leakage current in a *pn* junction is of the order of ....
(*i*) A (*ii*) mA
(*iii*) kA (*iv*) μA

32. In an intrinsic semiconductor, the number of free electrons ....
(*i*) equals the number of holes
(*ii*) is greater than the number of holes
(*iii*) is less than the number of holes
(*iv*) none of the above

33. At room temperature, an intrinsic semiconductor has ....
(*i*) many holes only
(*ii*) a few free electrons and holes
(*iii*) many free electrons only
(*iv*) no holes or free electrons

34. At absolute temperature, an intrinsic semiconductor has ....
(*i*) a few free electrons
(*ii*) many holes
(*iii*) many free electrons
(*iv*) no holes or free electrons

35. At room temperature, an intrinsic silicon crystal acts approximately as ....
(*i*) a battery
(*ii*) a conductor
(*iii*) an insulator
(*iv*) a piece of copper wire

Answers to Multiple-Choice Questions

1. (*i*)	**2.** (*iii*)	**3.** (*ii*)	**4.** (*iv*)	**5.** (*ii*)
6. (*ii*)	**7.** (*ii*)	**8.** (*iii*)	**9.** (*iv*)	**10.** (*i*)
11. (*ii*)	**12.** (*iii*)	**13.** (*iv*)	**14.** (*i*)	**15.** (*ii*)
16. (*ii*)	**17.** (*iii*)	**18.** (*ii*)	**19.** (*iii*)	**20.** (*i*)
21. (*i*)	**22.** (*i*)	**23.** (*iv*)	**24.** (*ii*)	**25.** (*ii*)
26. (*iii*)	**27.** (*iii*)	**28.** (*i*)	**29.** (*ii*)	**30.** (*i*)
31. (*iv*)	**32.** (*i*)	**33.** (*ii*)	**34.** (*iv*)	**35.** (*iii*)

Chapter Review Topics

1. What do you understand by a semi-conductor ? Discuss some important properties of semiconductors.
2. Which are the most commonly used semiconductors and why ?
3. Give the energy band description of semiconductors.
4. Discuss the effect of temperature on semiconductors.
5. Give the mechanism of hole current flow in a semiconductor.
6. What do you understand by intrinsic and extrinsic semiconductors ?
7. What is a *pn* junction ? Explain the formation of potential barrier in a *pn* junction.
8. Discuss the behaviour of a *pn* junction under forward and reverse biasing.
9. Draw and explain the *V-I* characteristics of a *pn* junction.
10. Write short notes on the following :
(*i*) Breakdown voltage
(*ii*) Knee voltage
(*iii*) Limitations in the operating conditions of *pn* junction

Discussion Questions

1. Why is a semiconductor an insulator at ordinary temperature ?
2. Why are electron carriers present in *p*-type semiconductor ?
3. Why is silicon preferred to germanium in the manufacture of semiconductor devices ?
4. What is the importance of peak inverse voltage ?

6

Semiconductor Diode

INTRODUCTION

It has already been discussed in the previous chapter that a *pn* junction conducts current easily when forward biased and practically no current flows when it is reverse biased. This unilateral conduction characteristic of *pn* junction (*i.e.* semiconductor diode) is similar to that of a vacuum diode. Therefore, like a vacuum diode, a semiconductor diode can also accomplish the job of *rectification i.e.* change alternating current to direct current. However, semiconductor diodes have become more *popular as they are smaller in size, cheaper and robust and usually operate with greater efficiency. In this chapter, we shall focus our attention on the circuit performance and applications of semiconductor diodes.

* On the other hand, vacuum diodes can withstand high reverse voltages and can operate at fairly high temperatures.

6.1 Semiconductor Diode

A pn junction is known as a **semi-conductor** *or* ***crystal diode**.

The outstanding property of a crystal diode to conduct current in one direction only permits it to be used as a rectifier. A crystal diode is usually represented by the schematic symbol shown in Fig. 6.1. The arrow in the symbol indicates the direction of easier conventional current flow.

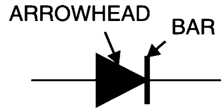
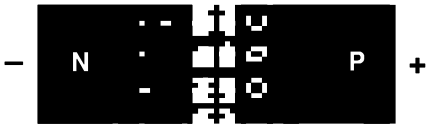


Fig. 6.1



A crystal diode has two terminals. When it is connected in a circuit, one thing to decide is whether the diode is forward or reverse biased. There is an easy rule to ascertain it. If the external circuit is trying to push the conventional current in the direction of arrow, the diode is forward biased. On the other hand, if the conventional current is trying to flow opposite to arrowhead, the diode is reverse biased. Putting in simple words :

(*i*) If *arrowhead* of diode symbol is *positive w.r.t. bar* of the symbol, the diode is forward biased.

(*ii*) If the *arrowhead* of diode symbol is *negative w.r.t. bar*, the diode is reverse biased.

Identification of crystal diode terminals. While using a crystal diode, it is often necessary to know which end is arrowhead and which end is bar. For this purpose, the following methods are available :

(*i*) Some manufacturers actually paint the symbol on the body of the diode *e.g. BY*127, *BY*114 crystal diodes manufactured by *BEL* [See Fig. 6.2 (*i*)].

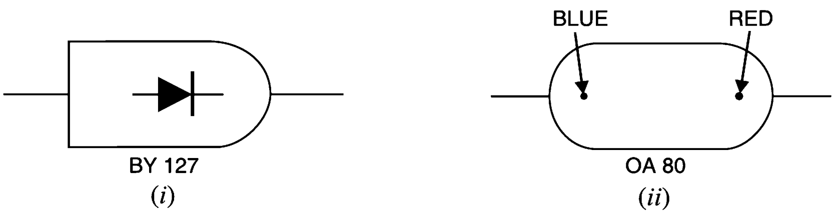


Fig. 6.2

(*ii*) Sometimes, red and blue marks are used on the body of the crystal diode. Red mark denotes arrow whereas blue mark indicates bar *e.g. OA80* crystal diode [See Fig. 6.2 (*ii*)].

6.2 Crystal Diode as a Rectifier

Fig. 6.3 illustrates the rectifying action of a crystal diode. The a.c. input voltage to be rectified, the diode and load R_L are connected in series. The d.c. output is obtained across the load as explained in the following discussion. During the positive half-cycle of a.c. input voltage, the arrowhead becomes positive *w.r.t.* bar. Therefore, diode is forward biased and conducts current in the circuit. The result is that positive half-cycle of input voltage appears across R_L as shown. However, during the negative half-cycle of input a.c. voltage, the diode becomes reverse biased because now the arrowhead is negative *w.r.t.* bar. Therefore, diode does not conduct and no voltage appears across load R_L. The result is that output consists of positive half-cycles of input a.c. voltage while the negative half-cycles are suppressed. In this way, crystal diode has been able to do rectification i.e. change a.c. into d.c. It may be seen that output across R_L is pulsating d.c.

* So called because *pn* junction is grown out of a crystal.

It is interesting to see that behaviour of diode is like a *switch*. When the diode is forward biased, it behaves like a closed switch and connects the a.c. supply to the load R_L. However, when the diode is reverse biased, it behaves like an open switch and disconnects the a.c. supply from the load R_L. This switching action of diode permits only the positive half-cycles of input a.c. voltage to appear across R_L.

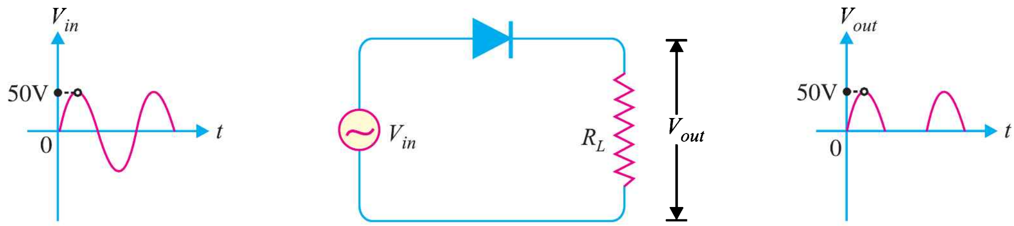


Fig. 6.3

Example 6.1. *In each diode circuit of Fig. 6.4, find whether the diodes are forward or reverse biased.*

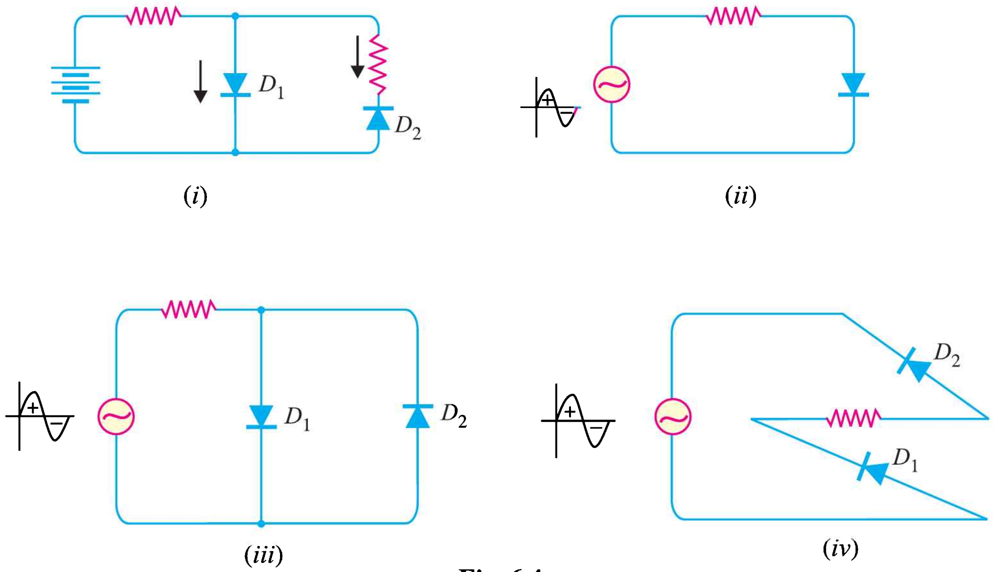


Fig. 6.4

Solution.

(i) Refer to Fig. 6.4 (*i*). The conventional current coming out of battery flows in the branch circuits. In diode D_1, the conventional current flows in the direction of arrowhead and hence this diode is forward biased. However, in diode D_2,the conventional current flows opposite to arrowhead and hence this diode is reverse biased.

(ii) Refer to Fig. 6.4 (*ii*). During the positive half-cycle of input a.c. voltage, the conventional current flows in the direction of arrowhead and hence diode is forward biased. However, during the negative half-cycle of input a.c. voltage, the diode is reverse biased.

(iii) Refer to Fig. 6.4 (*iii*). During the positive half-cycle of input a.c. voltage, conventional current flows in the direction of arrowhead in D_1 but it flows opposite to arrowhead in D_2. Therefore, during positive half-cycle, diode D_1 is forward biased and diode D_2 reverse biased. However, during the negative half-cycle of input a.c. voltage, diode D_2 is forward biased and D_1 is reverse biased.

(*iv*) Refer to Fig. 6.4 (*iv*). During the positive half-cycle of input a.c. voltage, both the diodes are reverse biased. However, during the negative half-cycle of input a.c. voltage, both the diodes are forward biased.

6.3 Resistance of Crystal Diode

It has already been discussed that a forward biased diode conducts easily whereas a reverse biased diode practically conducts no current. It means that *forward resistance* of a diode is quite small as compared with its *reverse resistance.*

1. Forward resistance. The resistance offered by the diode to forward bias is known as *forward resistance.* This resistance is not the same for the flow of direct current as for the changing current. Accordingly; this resistance is of two types, namely; *d.c. forward resistance* and *a.c. forward resistance.*

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

$$\therefore \quad \text{d.c. forward resistance, } R_f = \frac{OA}{OB}$$

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

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

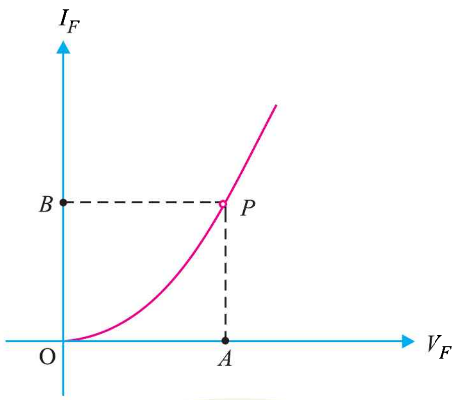


Fig. 6.5

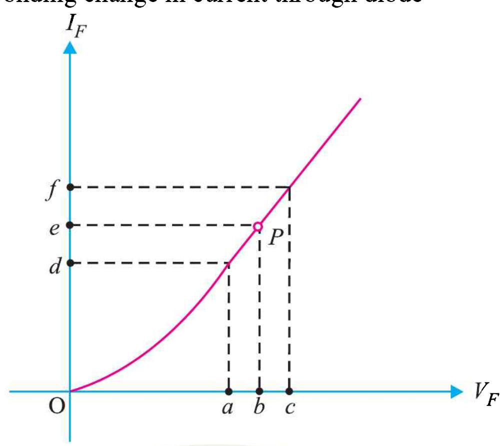


Fig. 6.6

The a.c. forward resistance is more significant as the diodes are generally used with alternating voltages. The a.c. forward resistance can be determined from the forward characteristic as shown in Fig. 6.6. If *P* is the operating point at any instant, then forward voltage is *ob* and forward current is *oe*. To find the a.c. forward resistance, vary the forward voltage on both sides of the operating point equally as shown in Fig. 6.6 where *ab* = *bc*. It is clear from this figure that :

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

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

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

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

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

6.4 Equivalent Circuit of Crystal Diode

It is generally profitable to replace a device or system by its equivalent circuit. An equivalent circuit of a device (*e.g.* crystal diode, transistor etc.) is a combination of electric elements, which when connected in a circuit, acts exactly as does the device when connected in the same circuit. Once the device is replaced by its equivalent circuit, the resulting network can be solved by traditional circuit analysis techniques. We shall now find the equivalent circuit of a crystal diode.

(*i*) *Approximate Equivalent circuit. When the forward voltage V_F is applied across a diode, it will not conduct till the potential barrier V_0 at the junction is overcome. When the forward voltage exceeds the potential barrier voltage, the diode starts conducting as shown in Fig. 6.7 (*i*). The forward current I_f flowing through the diode causes a voltage drop in its internal resistance r_f. Therefore, the forward voltage V_F applied across the *actual* diode has to overcome :

(*a*) potential barrier V_0

(*b*) internal drop $I_f r_f$

$$\therefore \quad V_F = V_0 + I_f r_f$$

For a silicon diode, V_0 = 0.7 V whereas for a germanium diode, V_0 = 0.3 V.

Therefore, approximate equivalent circuit for a crystal diode is a switch in series with a battery V_0 and internal resistance r_f as shown in Fig. 6.7 (*ii*). This approximate equivalent circuit of a diode is very helpful in studying the performance of the diode in a circuit.

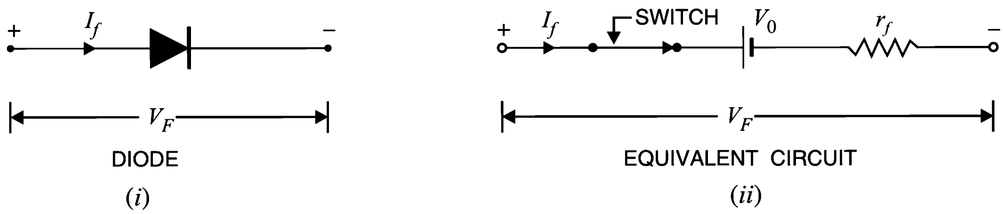


Fig. 6.7

(*ii*) Simplified Equivalent circuit. For most applications, the internal resistance r_f of the crystal diode can be ignored in comparison to other elements in the equivalent circuit. The equivalent circuit then reduces to the one shown in Fig. 6.8 (*ii*). This simplified equivalent circuit of the crystal diode is frequently used in diode-circuit analysis.

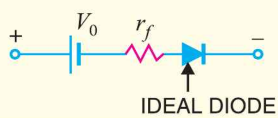
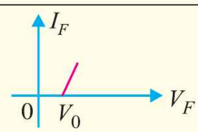
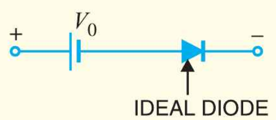
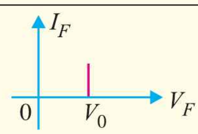
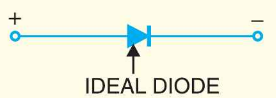
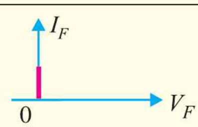
Fig. 6.8

* We assume here that *V*/*I* characteristic of crystal diode is linear.

***(iii)* Ideal diode model.** An ideal diode is one which behaves as a perfect conductor when forward biased and as a perfect insulator when reverse biased. Obviously, in such a hypothetical situation, forward resistance $r_f = 0$ and potential barrier V_0 is considered negligible. It may be mentioned here that although ideal diode is never found in practice, yet diode circuit analysis is made on this basis. *Therefore, while discussing diode circuits, the diode will be assumed ideal unless and until stated otherwise.*

6.5 Crystal Diode Equivalent Circuits

It is desirable to sum up the various models of crystal diode equivalent circuit in the tabular form given below:

S.No.	Type	Model	Characteristic
1.	Approximate model		
2.	Simplified model		
3.	Ideal Model		

Example 6.2. *An a.c. voltage of peak value 20 V is connected in series with a silicon diode and load resistance of 500 Ω. If the forward resistance of diode is 10 Ω, find :*

(i) peak current through diode *(ii) peak output voltage*

What will be these values if the diode is assumed to be ideal ?

Solution.

Peak input voltage = 20 V

Forward resistance, r_f = 10 Ω

Load resistance, R_L = 500 Ω

Potential barrier voltage, V_0 = 0.7 V

The diode will conduct during the positive half-cycles of a.c. input voltage only. The equivalent circuit is shown in Fig. 6.9 (*ii*).

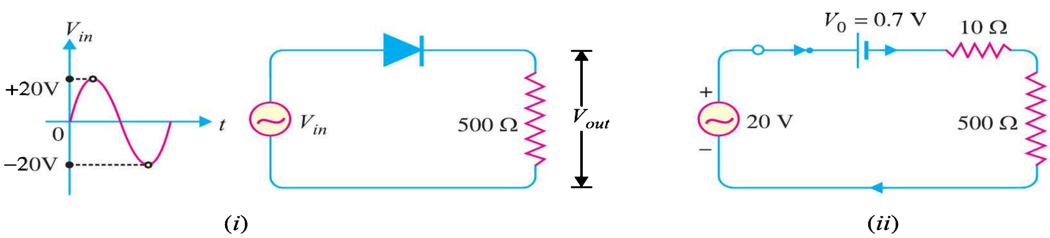


Fig. 6.9

(*i*) The peak current through the diode will occur at the instant when the input voltage reaches positive peak *i.e.* $V_{in} = V_F = 20$ V.

$$\therefore \quad V_F = V_0 + (I_f)_{peak}\,[r_f + R_L] \qquad ...(i)$$

$$\text{or} \quad (I_f)_{peak} = \frac{V_F - V_0}{r_f + R_L} = \frac{20 - 0.7}{10 + 500} = \frac{19.3}{510}\text{ A} = \mathbf{37.8\ mA}$$

(*ii*) Peak output voltage $= (I_f)_{peak} \times R_L = 37.8\text{ mA} \times 500\ \Omega = \mathbf{18.9\ V}$

Ideal diode. For an ideal diode, put $V_0 = 0$ and $r_f = 0$ in equation (*i*).

$$\therefore \quad V_F = (I_f)_{peak} \times R_L$$

$$\text{or} \quad (I_f)_{peak} = \frac{V_F}{R_L} = \frac{20\text{ V}}{500\ \Omega} = \mathbf{40\ mA}$$

Peak output voltage $= (I_f)_{peak} \times R_L = 40\text{ mA} \times 500\ \Omega = \mathbf{20\ V}$

Comments. It is clear from the above example that output voltage is *nearly* the same whether the actual diode is used or the diode is considered ideal. This is due to the fact that input voltage is quite large as compared with V_0 and voltage drop in r_f. Therefore, nearly the whole input forward voltage appears across the load. For this reason, diode circuit analysis is generally made on the ideal diode basis.

Example 6.3. *Find the current through the diode in the circuit shown in Fig. 6.10 (i). Assume the diode to be ideal.*

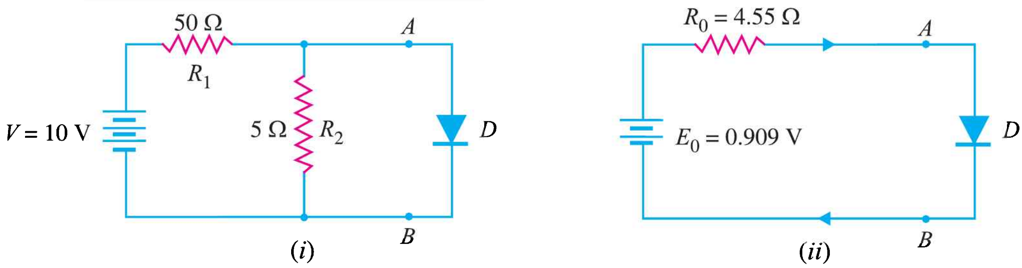


Fig. 6.10

Solution. We shall use Thevenin's theorem to find current in the diode. Referring to Fig. 6.10(*i*),

E_0 = Thevenin's voltage

= Open circuited voltage across *AB* with diode removed

$$= \frac{R_2}{R_1 + R_2} \times V = \frac{5}{50 + 5} \times 10 = 0.909\text{ V}$$

R_0 = Thevenin's resistance

= Resistance at terminals *AB* with diode removed and battery replaced by a short circuit

$$= \frac{R_1 R_2}{R_1 + R_2} = \frac{50 \times 5}{50 + 5} = 4.55\ \Omega$$

Fig. 6.10 (*ii*) shows Thevenin's equivalent circuit. Since the diode is ideal, it has zero resistance.

$$\therefore \quad \text{Current through diode} = \frac{E_0}{R_0} = \frac{0.909}{4.55} = 0.2\text{ A} = \mathbf{200\ mA}$$

Example 6.4. *Calculate the current through 48 Ω resistor in the circuit shown in Fig. 6.11 (i). Assume the diodes to be of silicon and forward resistance of each diode is 1 Ω.*

Solution. Diodes D_1 and D_3 are forward biased while diodes D_2 and D_4 are reverse biased. We can, therefore, consider the branches containing diodes D_2 and D_4 as "open". Replacing diodes D_1 and D_3 by their equivalent circuits and making the branches containing diodes D_2 and D_4 open, we get the circuit shown in Fig. 6.11 (*ii*). Note that for a silicon diode, the barrier voltage is 0.7 V.

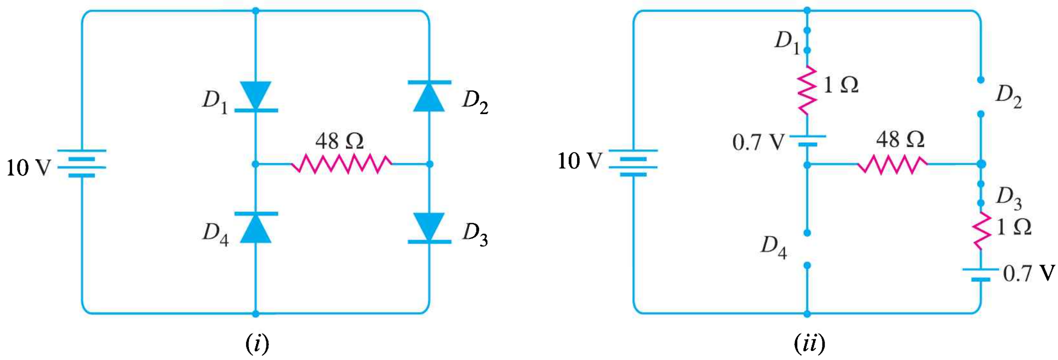


Fig. 6.11

Net circuit voltage = 10 – 0.7 – 0.7 = 8.6 V

Total circuit resistance = 1 + 48 + 1 = 50 Ω

∴ Circuit current = 8.6/50 = 0.172 A = **172 mA**

Example 6.5. *Determine the current I in the circuit shown in Fig. 6.12 (i). Assume the diodes to be of silicon and forward resistance of diodes to be zero.*

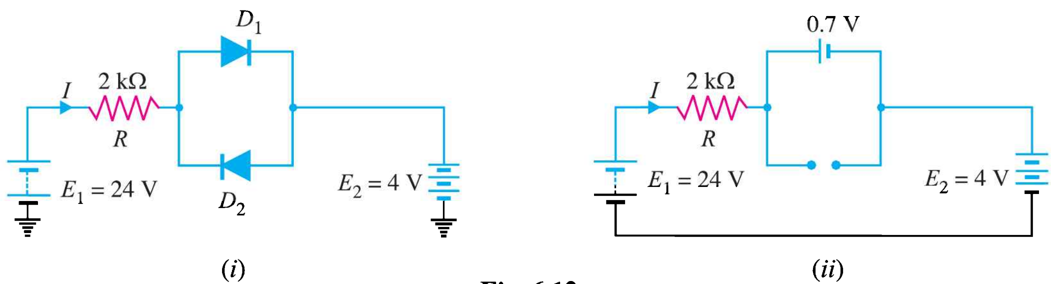


Fig. 6.12

Solution. The conditions of the problem suggest that diode D_1 is forward biased and diode D_2 is reverse biased. We can, therefore, consider the branch containing diode D_2 as open as shown in Fig. 6.12 (*ii*). Further, diode D_1 can be replaced by its simplified equivalent circuit.

$$\therefore \quad I = \frac{E_1 - E_2 - V_0}{R} = \frac{24 - 4 - 0.7}{2\ \text{k}\Omega} = \frac{19.3\ \text{V}}{2\ \text{k}\Omega} = \mathbf{9.65\ mA}$$

Example 6.6. *Find the voltage V_A in the circuit shown in Fig. 6.13 (i). Use simplified model.*

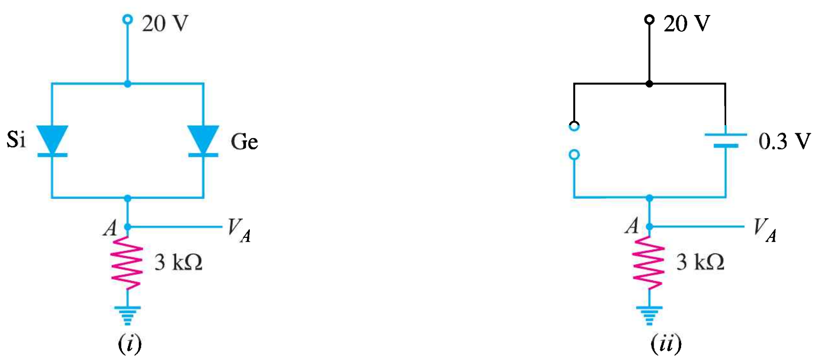


Fig. 6.13

Solution. It appears that when the applied voltage is switched on, both the diodes will turn "on". But that is not so. When voltage is applied, germanium diode ($V_0 = 0.3$ V) will turn on first and a level of 0.3 V is maintained across the parallel circuit. The silicon diode never gets the opportunity to have 0.7 V across it and, therefore, remains in open-circuit state as shown in Fig. 6.13 (*ii*).

$$\therefore \quad V_A = 20 - 0.3 = \mathbf{19.7\ V}$$

Example 6.7. *Find V_Q and I_D in the network shown in Fig. 6.14 (i). Use simplified model.*

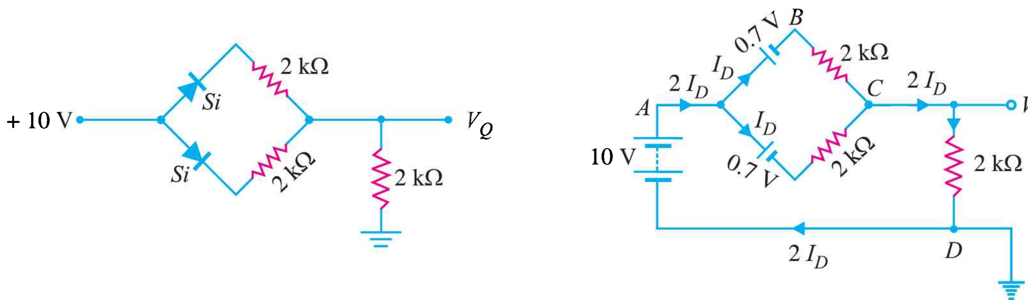
Solution. Replace the diodes by their simplified models. The resulting circuit will be as shown in Fig. 6.14 (*ii*). By symmetry, current in each branch is I_D so that current in branch *CD* is $2I_D$. Applying Kirchhoff's voltage law to the closed circuit *ABCDA*, we have,

$$-0.7 - I_D \times 2 - 2\,I_D \times 2 + 10 = 0 \qquad (I_D \text{ in mA})$$

or $$6\,I_D = 9.3$$

$$\therefore \quad I_D = \frac{9.3}{6} = \mathbf{1.55\ mA}$$

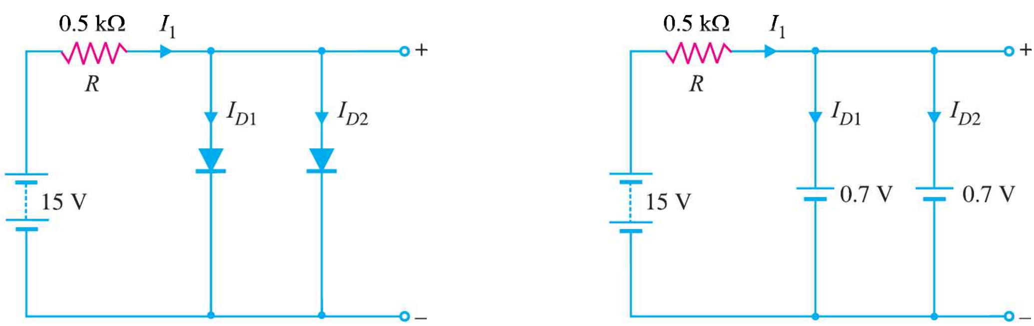
Also $$V_Q = (2\,I_D) \times 2\text{ k}\Omega = (2 \times 1.55\text{ mA}) \times 2\text{ k}\Omega = \mathbf{6.2\ V}$$



(*i*) (*ii*)

Fig. 6.14

Example 6.8. *Determine current through each diode in the circuit shown in Fig. 6.15 (i). Use simplified model. Assume diodes to be similar.*



(*i*) (*ii*)

Fig. 6.15

Solution. The applied voltage forward biases each diode so that they conduct current in the same direction. Fig. 6.15 (*ii*) shows the equivalent circuit using simplified model. Referring to Fig. 6.15 (*ii*),

$$I_1 = \frac{\text{Voltage across } R}{R} = \frac{15 - 0.7}{0.5\text{ k}\Omega} = 28.6\text{ mA}$$

Since the diodes are similar, $I_{D1} = I_{D2} = \frac{I_1}{2} = \frac{28.6}{2} = \mathbf{14.3\ mA}$

Comments. Note the use of placing the diodes in parallel. If the current rating of each diode is

20 mA and a single diode is used in this circuit, a current of 28.6 mA would flow through the diode, thus damaging the device. By placing them in parallel, the current is limited to a safe value of 14.3 mA for the same terminal voltage.

Example 6.9. *Determine the currents I_1, I_2 and I_3 for the network shown in Fig. 6.16(i). Use simplified model for the diodes.*

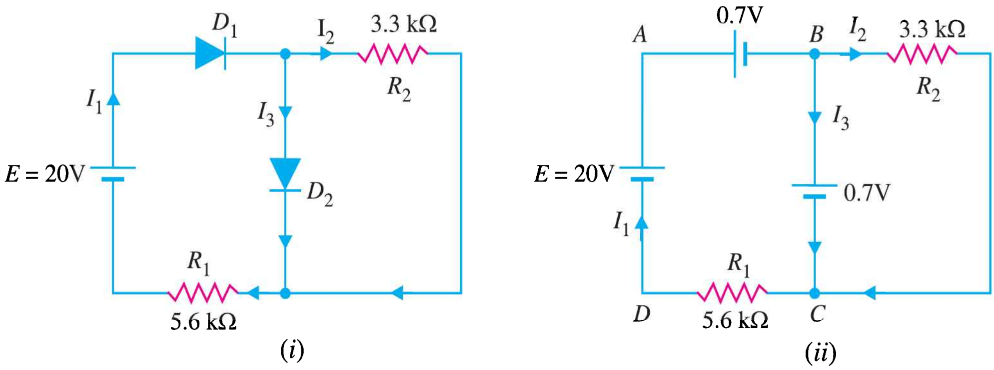


Fig. 6.16

Solution. An inspection of the circuit shown in Fig. 6.16 (*i*) shows that both diodes D_1 and D_2 are forward biased. Using simplified model for the diodes, the circuit shown in Fig. 6.16 (*i*) becomes the one shown in Fig. 6.16 (*ii*). The voltage across R_2 (= 3.3 k Ω) is 0.7V.

$$\therefore \quad I_2 = \frac{0.7 \text{ V}}{3.3 \text{ k}\Omega} = \mathbf{0.212 \text{ mA}}$$

Applying Kirchhoff's voltage law to loop *ABCDA* in Fig. 6.16 (*ii*), we have,

$$-0.7 - 0.7 - I_1 R_1 + 20 = 0$$

$$\therefore \quad I_1 = \frac{20 - 0.7 - 0.7}{R_1} = \frac{18.6 \text{ V}}{5.6 \text{ k}\Omega} = \mathbf{3.32 \text{ mA}}$$

Now $\quad I_1 = I_2 + I_3$

$$\therefore \quad I_3 = I_1 - I_2 = 3.32 - 0.212 = \mathbf{3.108 \text{ mA}}$$

Example 6.10. *Determine if the diode (ideal) in Fig. 6.17 (i) is forward biased or reverse biased.*

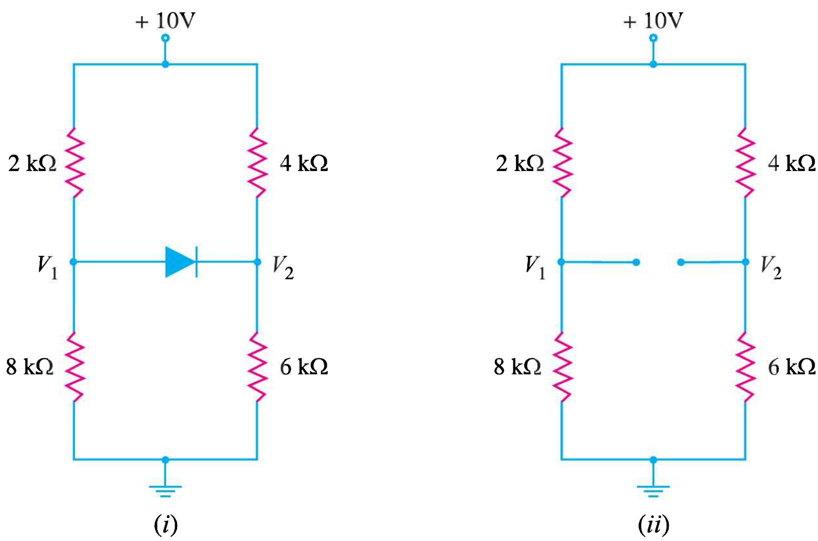


Fig. 6.17

Solution. Let us assume that diode in Fig. 6.17 (*i*) is *OFF i.e.* it is reverse biased. The circuit then becomes as shown in Fig. 6.17 (*ii*). Referring to Fig. 6.17 (*ii*), we have,

$$V_1 = \frac{10\text{ V}}{2\text{ k}\Omega + 8\text{ k}\Omega} \times 8\text{ k}\Omega = 8\text{V}$$

$$V_2 = \frac{10\text{ V}}{4\text{ k}\Omega + 6\text{ k}\Omega} \times 6\text{ k}\Omega = 6\text{V}$$

∴ Voltage across diode = $V_1 - V_2 = 8 - 6 = 2\text{V}$

Now $V_1 - V_2 = 2\text{V}$ is enough voltage to make the diode *forward biased*. Therefore, our initial assumption was wrong.

Example 6.11. *Determine the state of diode for the circuit shown in Fig. 6.18 (i) and find I_D and V_D. Assume simplified model for the diode.*

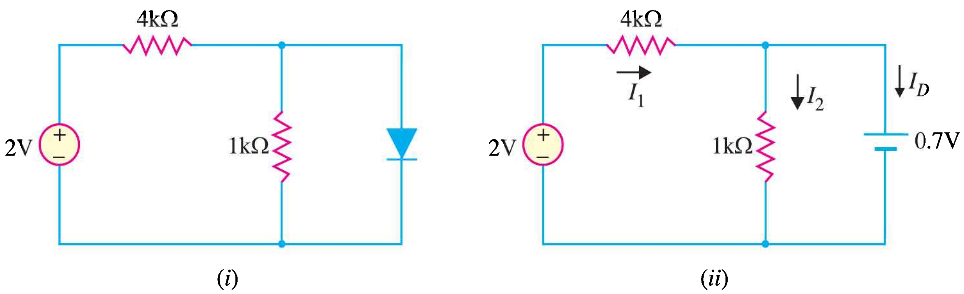


Fig. 6.18

Solution. Let us assume that the diode is *ON*. Therefore, we can replace the diode with a 0.7V battery as shown in Fig. 6.18 (*ii*). Referring to Fig. 6.18 (*ii*), we have,

$$I_1 = \frac{(2 - 0.7)\text{ V}}{4\text{ k}\Omega} = \frac{1.3\text{ V}}{4\text{ k}\Omega} = 0.325\text{ mA}$$

$$I_2 = \frac{0.7\text{ V}}{1\text{ k}\Omega} = 0.7\text{ mA}$$

Now $\quad I_D = I_1 - I_2 = 0.325 - 0.7 = -0.375\text{ mA}$

Since the diode current is negative, the diode must be **OFF** and the true value of diode current is I_D = **0 mA**. Our initial assumption was wrong. In order to analyse the circuit properly, we should replace the diode in Fig. 6.18 (*i*) with an open circuit as shown in Fig. 6.19. The voltage V_D across the diode is

$$V_D = \frac{2\text{ V}}{1\text{ k}\Omega + 4\text{ k}\Omega} \times 1\text{ k}\Omega = \mathbf{0.4V}$$

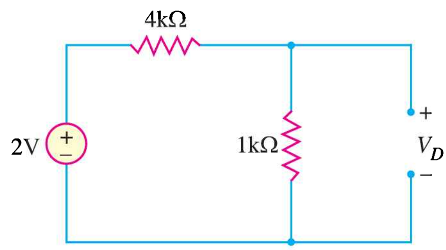


Fig. 6.19

We know that 0.7V is required to turn *ON* the diode. Since V_D is only 0.4V, the answer confirms that the diode is *OFF*.

6.6 Important Terms

While discussing the diode circuits, the reader will generally come across the following terms :

(*i*) **Forward current.** It is the current flowing through a forward biased diode. Every diode has a maximum value of forward current which it can safely carry. If this value is exceeded, the diode may be destroyed due to excessive heat. For this reason, the manufacturers' data sheet specifies the maximum forward current that a diode can handle safely.

(*ii*) Peak inverse voltage. *It is the maximum reverse voltage that a diode can withstand without destroying the junction.*

If the reverse voltage across a diode exceeds this value, the reverse current increases sharply and breaks down the junction due to excessive heat. Peak inverse voltage is extremely important when diode is used as a rectifier. In rectifier service, it has to be ensured that reverse voltage across the diode does not exceed its PIV during the negative half-cycle of input a.c. voltage. As a matter of fact, PIV consideration is generally the deciding factor in diode rectifier circuits. The peak inverse voltage may be between 10V and 10 kV depending upon the type of diode.

(*iii*) Reverse current or leakage current. It is the current that flows through a reverse biased diode. This current is due to the minority carriers. Under normal operating voltages, the reverse current is quite small. Its value is extremely small ($< 1\mu$ A) for silicon diodes but it is appreciable ($\simeq$ 100 μA) for germanium diodes.

It may be noted that the reverse current is usually very small as compared with forward current. For example, the forward current for a typical diode might range upto 100 mA while the reverse current might be only a few μA—a ratio of many thousands between forward and reverse currents.

6.7 Crystal Diode Rectifiers

For reasons associated with economics of generation and transmission, the electric power available is usually an a.c. supply. The supply voltage varies sinusoidally and has a frequency of 50 Hz. It is used for lighting, heating and electric motors. But there are many applications (*e.g.* electronic circuits) where d.c. supply is needed. When such a d.c. supply is required, the mains a.c. supply is rectified by using crystal diodes. The following two rectifier circuits can be used :

(*i*) Half-wave rectifier (*ii*) Full-wave rectifier

6.8 Half-Wave Rectifier

In half-wave rectification, the rectifier conducts current only during the positive half-cycles of input a.c. supply. The negative half-cycles of a.c. supply are suppressed *i.e.* during negative half-cycles, no current is conducted and hence no voltage appears across the load. Therefore, current always flows in one direction (*i.e.* d.c.) through the load though after every half-cycle.

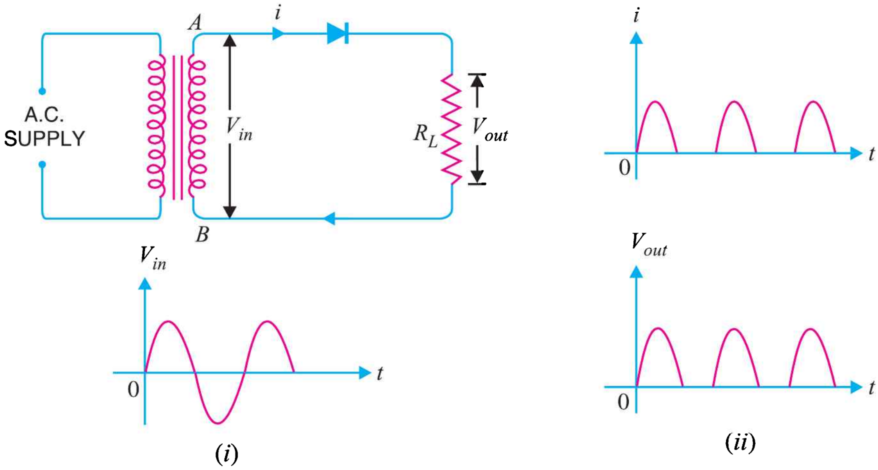


Fig. 6.20

Circuit details. Fig. 6.20 shows the circuit where a single crystal diode acts as a half-wave rectifier. The a.c. supply to be rectified is applied in series with the diode and load resistance R_L. Generally, a.c. supply is given through a transformer. The use of transformer permits two advantages. Firstly, it allows us to step up or step down the a.c. input voltage as the situation demands. Secondly, the transformer isolates the rectifier circuit from power line and thus reduces the risk of electric shock.

Operation. The a.c. voltage across the secondary winding *AB* changes polarities after every half-cycle. During the positive half-cycle of input a.c. voltage, end *A* becomes positive *w.r.t.* end *B*. This makes the diode forward biased and hence it conducts current. During the negative half-cycle, end *A* is negative *w.r.t.* end *B*. Under this condition, the diode is reverse biased and it conducts no current. Therefore, current flows through the diode during positive half-cycles of input a.c. voltage only ; it is blocked during the negative half-cycles [See Fig. 6.20 (*ii*)]. In this way, current flows through load R_L always in the same direction. Hence d.c. output is obtained across R_L. It may be noted that output across the load is pulsating d.c. These pulsations in the output are further smoothened with the help of *filter circuits* discussed later.

Disadvantages : The main disadvantages of a half-wave rectifier are :

(*i*) The pulsating current in the load contains alternating component whose basic frequency is equal to the supply frequency. Therefore, an elaborate filtering is required to produce steady direct current.

(*ii*) The a.c. supply delivers power only half the time. Therefore, the output is low.

6.9 Output Frequency of Half-Wave Rectifier

The output frequency of a half-wave rectifier is equal to the input frequency (50 Hz). Recall how a complete cycle is defined. A waveform has a complete cycle when it repeats the same wave pattern over a given time. Thus in Fig. 6.21 (*i*), the a.c. input voltage repeats the same wave pattern over 0° – 360°, 360° – 720° and so on. In Fig. 6.21 (*ii*), the output waveform also repeats the same wave pattern over 0° – 360°, 360° – 720° and so on. This means that when input a.c. completes one cycle, the output half-wave rectified wave also completes one cycle. In other words, the output frequency is equal to the input frequency *i.e.*

$$f_{out} = f_{in}$$

For example, if the input frequency of sine wave applied to a half-wave rectifier is 100 Hz, then frequency of the output wave will also be 100 Hz.

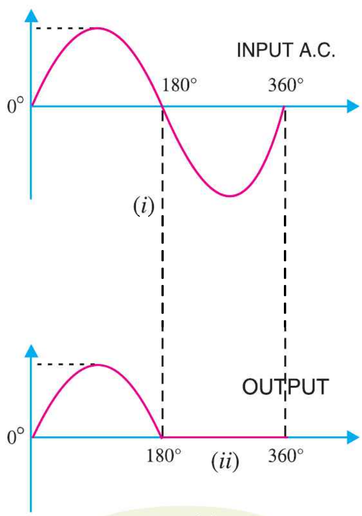


Fig. 6.21

6.10 Efficiency of Half-Wave Rectifier

The ratio of d.c. power output to the applied input a.c. power is known as **rectifier efficiency** *i.e.*

$$\text{Rectifier efficiency, } \eta = \frac{\text{d.c. power output}}{\text{Input a.c. power}}$$

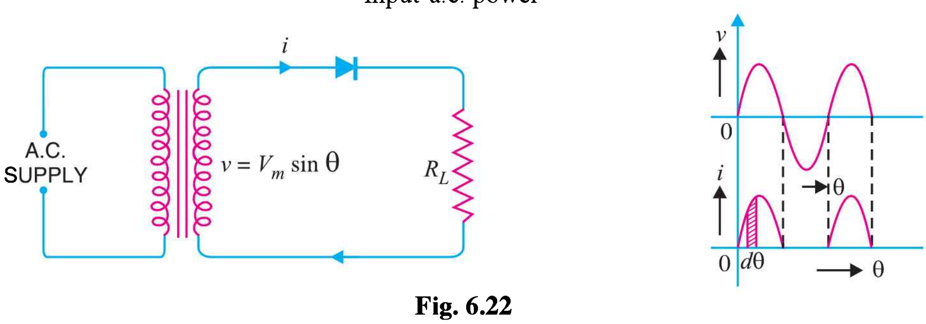


Fig. 6.22

Consider a half-wave rectifier shown in Fig. 6.22. Let $v = V_m \sin \theta$ be the alternating voltage that appears across the secondary winding. Let r_f and R_L be the diode resistance and load resistance respectively. The diode conducts during positive half-cycles of a.c. supply while no current conduction takes place during negative half-cycles.

d.c. power. The output current is pulsating direct current. Therefore, in order to find d.c. power, average current has to be found out.

$$*I_{av} = I_{dc} = \frac{1}{2\pi}\int_0^{\pi} i\, d\theta = \frac{1}{2\pi}\int_0^{\pi} \frac{V_m \sin\theta}{r_f + R_L}\, d\theta$$

$$= \frac{V_m}{2\pi(r_f + R_L)}\int_0^{\pi} \sin\theta\, d\theta = \frac{V_m}{2\pi(r_f + R_L)}[-\cos\theta]_0^{\pi}$$

$$= \frac{V_m}{2\pi(r_f + R_L)} \times 2 = \frac{V_m}{(r_f + R_L)} \times \frac{1}{\pi}$$

$$= \frac{^{**}I_m}{\pi} \qquad \left[\because I_m = \frac{V_m}{(r_f + R_L)}\right]$$

$$\therefore \quad \text{d.c. power, } P_{dc} = I_{dc}^2 \times R_L = \left(\frac{I_m}{\pi}\right)^2 \times R_L \qquad ...(i)$$

a.c. power input : The a.c. power input is given by :

$$P_{ac} = I_{rms}^2 (r_f + R_L)$$

For a half-wave rectified wave, $I_{rms} = I_m/2$

$$\therefore \quad P_{ac} = \left(\frac{I_m}{2}\right)^2 \times (r_f + R_L) \qquad ...(ii)$$

$$\therefore \quad \text{Rectifier efficiency} = \frac{\text{d.c. output power}}{\text{a.c. input power}} = \frac{(I_m/\pi)^2 \times R_L}{(I_m/2)^2\,(r_f + R_L)}$$

$$= \frac{0.406\, R_L}{r_f + R_L} = \frac{0.406}{1 + \frac{r_f}{R_L}}$$

The efficiency will be maximum if r_f is negligible as compared to R_L.

∴ Max. rectifier efficiency = 40.6%

This shows that in half-wave rectification, a maximum of 40.6% of a.c. power is converted into d.c. power.

Example 6.12. *The applied input a.c. power to a half-wave rectifier is 100 watts. The d.c. output power obtained is 40 watts.*

(*i*) *What is the rectification efficiency ?*

(*ii*) *What happens to remaining 60 watts ?*

Solution.

(***i***) Rectification efficiency $= \frac{\text{d.c. output power}}{\text{a.c. input power}} = \frac{40}{100} = 0.4 = \mathbf{40\%}$

(***ii***) 40% efficiency of rectification does not mean that 60% of power is lost in the rectifier circuit. In fact, a crystal diode consumes little power due to its small internal resistance. The 100 W

---

* Average value $= \frac{\text{Area under the curve over a cycle}}{\text{Base}} = \frac{\int_0^{\pi} i\, d\theta}{2\pi}$

** It may be remembered that the *area of one-half cycle of a sinusoidal wave* is twice the peak value. Thus in this case, peak value is I_m and, therefore, area of one-half cycle is 2 I_m.

$$\therefore \quad I_{av} = I_{dc} = \frac{2\, I_m}{2\pi} = \frac{I_m}{\pi}$$

a.c. power is contained as 50 watts in positive half-cycles and 50 watts in negative half-cycles. The 50 watts in the negative half-cycles are not supplied at all. Only 50 watts in the positive half-cycles are converted into 40 watts.

$$\therefore \quad \text{Power efficiency} = \frac{40}{50} \times 100 = 80\%$$

Although 100 watts of a.c. power was supplied, the half-wave rectifier accepted only 50 watts and converted it into 40 watts d.c. power. Therefore, it is appropriate to say that efficiency of rectification is 40% and *not* 80% which is power efficiency.

Example 6.13. *An a.c. supply of 230 V is applied to a half-wave rectifier circuit through a transformer of turn ratio 10 : 1. Find (i) the output d.c. voltage and (ii) the peak inverse voltage. Assume the diode to be ideal.*

Solution.

Primary to secondary turns is

$$\frac{N_1}{N_2} = 10$$

R.M.S. primary voltage = 230 V

∴ Max. primary voltage is

$$V_{pm} = (\sqrt{2}) \times \text{r.m.s. primary voltage} = (\sqrt{2}) \times 230 = 325.3 \text{ V}$$

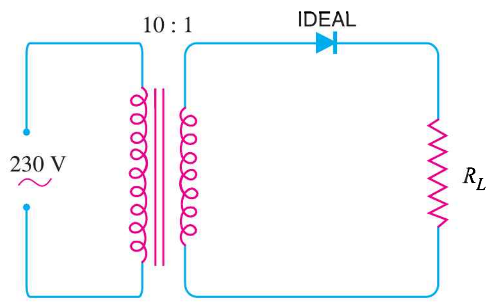


Fig. 6.23

Max. secondary voltage is

$$V_{sm} = V_{pm} \times \frac{N_2}{N_1} = 325.3 \times \frac{1}{10} = 32.53 \text{ V}$$

(i)

$$I_{d.c.} = \frac{I_m}{\pi}$$

$$\therefore \quad V_{dc} = \frac{I_m}{\pi} \times R_L = \frac{V_{sm}}{\pi} = \frac{32.53}{\pi} = \mathbf{10.36 \text{ V}}$$

(ii) During the negative half-cycle of a.c. supply, the diode is reverse biased and hence conducts no current. Therefore, the maximum secondary voltage appears across the diode.

∴ Peak inverse voltage = **32.53 V**

Example 6.14. *A crystal diode having internal resistance $r_f = 20\Omega$ is used for half-wave rectification. If the applied voltage $v = 50 \sin \omega t$ and load resistance $R_L = 800\ \Omega$, find :*

(i) I_m, I_{dc}, I_{rms} *(ii) a.c. power input and d.c. power output*

(iii) d.c. output voltage *(iv) efficiency of rectification.*

Solution.

$$v = 50 \sin \omega t$$

$$\therefore \quad \text{Maximum voltage, } V_m = 50 \text{ V}$$

(i)

$$I_m = \frac{V_m}{r_f + R_L} = \frac{50}{20 + 800} = 0.061 \text{ A} = \mathbf{61 \text{ mA}}$$

$$I_{dc} = I_m/\pi = 61/\pi = \mathbf{19.4 \text{ mA}}$$

$$I_{rms} = I_m/2 = 61/2 = \mathbf{30.5 \text{ mA}}$$

(ii)

$$\text{a.c. power input} = (I_{rms})^2 \times (r_f + R_L) = \left(\frac{30.5}{1000}\right)^2 \times (20 + 800) = \mathbf{0.763 \text{ watt}}$$

$$\text{d.c. power output} = I_{dc}^2 \times R_L = \left(\frac{19.4}{1000}\right)^2 \times 800 = \mathbf{0.301\ watt}$$

(*iii*) $$\text{d.c. output voltage} = I_{dc} R_L = 19.4 \text{ mA} \times 800\ \Omega = \mathbf{15.52\ volts}$$

(*iv*) $$\text{Efficiency of rectification} = \frac{0.301}{0.763} \times 100 = \mathbf{39.5\%}$$

Example 6.15. *A half-wave rectifier is used to supply 50V d.c. to a resistive load of 800 Ω. The diode has a resistance of 25 Ω. Calculate a.c. voltage required.*

Solution.

Output d.c. voltage, $V_{dc} = 50$ V

Diode resistance, $r_f = 25\ \Omega$

Load resistance, $R_L = 800\ \Omega$

Let V_m be the maximum value of a.c. voltage required.

$$\therefore \quad V_{dc} = I_{dc} \times R_L$$

$$= \frac{I_m}{\pi} \times R_L = \frac{V_m}{\pi(r_f + R_L)} \times R_L \qquad \left[\because I_m = \frac{V_m}{r_f + R_L}\right]$$

or $$50 = \frac{V_m}{\pi(25 + 800)} \times 800$$

$$\therefore \quad V_m = \frac{\pi \times 825 \times 50}{800} = \mathbf{162\ V}$$

Hence a.c. voltage of maximum value 162 V is required.

6.11 Full-Wave Rectifier

In full-wave rectification, current flows through the load in the same direction for both half-cycles of input a.c. voltage. This can be achieved with two diodes working alternately. For the positive half-cycle of input voltage, one diode supplies current to the load and for the negative half-cycle, the other diode does so ; current being always in the same direction through the load. Therefore, a full-wave rectifier utilises both half-cycles of input a.c. voltage to produce the d.c. output. The following two circuits are commonly used for full-wave rectification :

(*i*) Centre-tap full-wave rectifier (*ii*) Full-wave bridge rectifier

6.12 Centre-Tap Full-Wave Rectifier

The circuit employs two diodes D_1 and D_2 as shown in Fig. 6.24. A centre tapped secondary winding *AB* is used with two diodes connected so that each uses one half-cycle of input a.c. voltage. In other words, diode D_1 utilises the a.c. voltage appearing across the upper half (*OA*) of secondary winding for rectification while diode D_2 uses the lower half winding *OB*.

Operation. During the positive half-cycle of secondary voltage, the end *A* of the secondary winding becomes positive and end *B* negative. This makes the diode D_1 forward biased and diode D_2 reverse biased. Therefore, diode D_1 conducts while diode D_2 does not. The conventional current flow is through diode D_1, load resistor R_L and the upper half of secondary winding as shown by the dotted arrows. During the negative half-cycle, end *A* of the secondary winding becomes negative and end *B* positive. Therefore, diode D_2 conducts while diode D_1 does not. The conventional current flow is through diode D_2, load R_L and lower half winding as shown by solid arrows. Referring to Fig. 6.24, it may be seen that current in the load R_L is *in the same direction* for both half-cycles of input a.c. voltage. Therefore, d.c. is obtained across the load R_L. Also, the polarities of the d.c. output across the load should be noted.

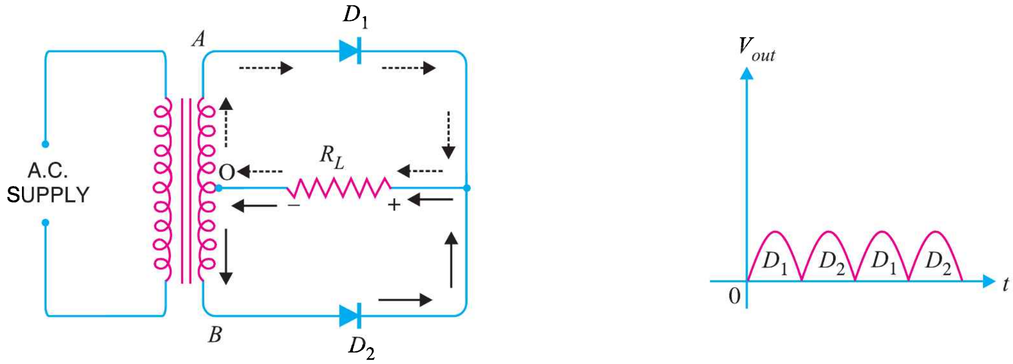


Fig. 6.24

Peak inverse voltage. Suppose V_m is the maximum voltage across the half secondary winding. Fig. 6.25 shows the circuit at the instant secondary voltage reaches its maximum value in the positive direction. At this instant, diode D_1 is conducting while diode D_2 is non-conducting. Therefore, whole of the secondary voltage appears across the non-conducting diode. Consequently, the peak inverse voltage is twice the maximum voltage across the half-secondary winding *i.e.*

$$\text{PIV} = 2\,V_m$$

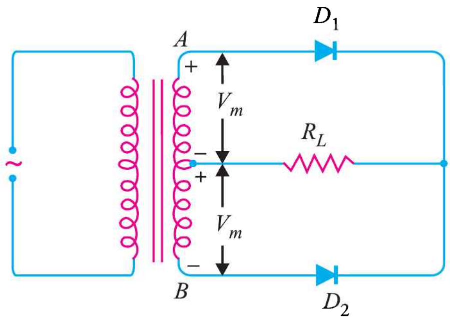


Fig. 6.25

Disadvantages

(i) It is difficult to locate the centre tap on the secondary winding.

(ii) The d.c. output is small as each diode utilises only one-half of the transformer secondary voltage.

(iii) The diodes used must have high peak inverse voltage.

6.13 Full-Wave Bridge Rectifier

The need for a centre tapped power transformer is eliminated in the bridge rectifier. It contains four diodes D_1, D_2, D_3 and D_4 connected to form bridge as shown in Fig. 6.26. The a.c. supply to be rectified is applied to the diagonally opposite ends of the bridge through the transformer. Between other two ends of the bridge, the load resistance R_L is connected.

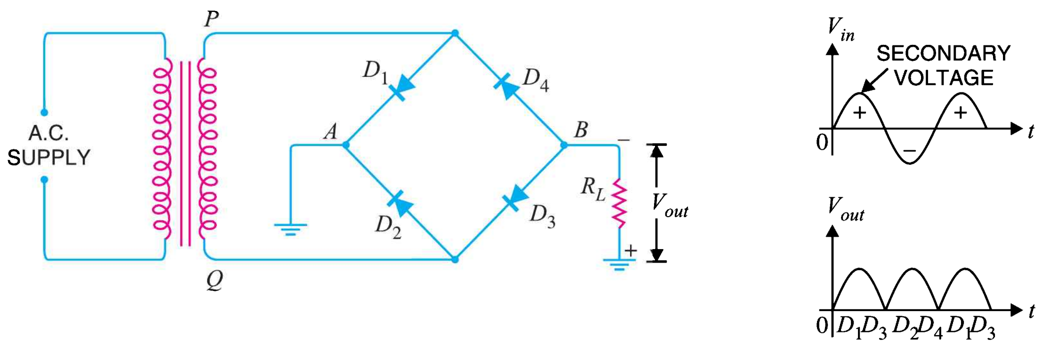


Fig. 6.26

Operation. During the positive half-cycle of secondary voltage, the end *P* of the secondary winding becomes positive and end *Q* negative. This makes diodes D_1 and D_3 forward biased while diodes D_2 and D_4 are reverse biased. Therefore, only diodes D_1 and D_3 conduct. These two diodes will be in series through the load R_L as shown in Fig. 6.27 (*i*). The conventional current flow is shown by dotted arrows. It may be seen that current flows from *A* to *B* through the load R_L.

During the negative half-cycle of secondary voltage, end *P* becomes negative and end *Q* positive. This makes diodes D_2 and D_4 forward biased whereas diodes D_1 and D_3 are reverse biased. Therefore, only diodes D_2 and D_4 conduct. These two diodes will be in series through the load R_L as shown in Fig. 6.27 (*ii*). The current flow is shown by the solid arrows. It may be seen that again current flows from *A* to *B* through the load *i.e.* in the same direction as for the positive half-cycle. Therefore, d.c. output is obtained across load R_L.

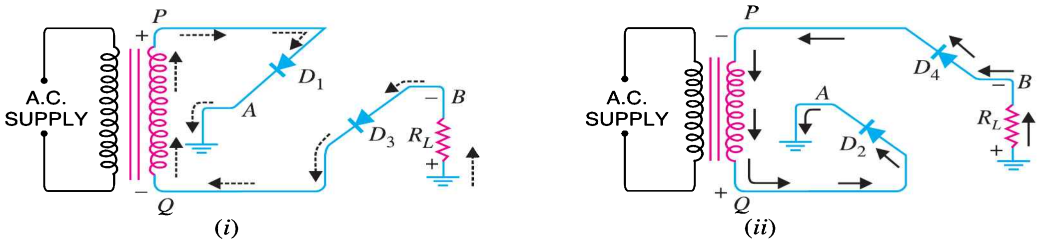


Fig. 6.27

Peak inverse voltage. The peak inverse voltage (PIV) of each diode is equal to the maximum secondary voltage of transformer. Suppose during positive half cycle of input a.c., end *P* of secondary is positive and end *Q* negative. Under such conditions, diodes D_1 and D_3 are forward biased while diodes D_2 and D_4 are reverse biased. Since the diodes are considered ideal, diodes D_1 and D_3 can be replaced by wires as shown in Fig. 6.28 (*i*). This circuit is the same as shown in Fig. 6.28 (*ii*).

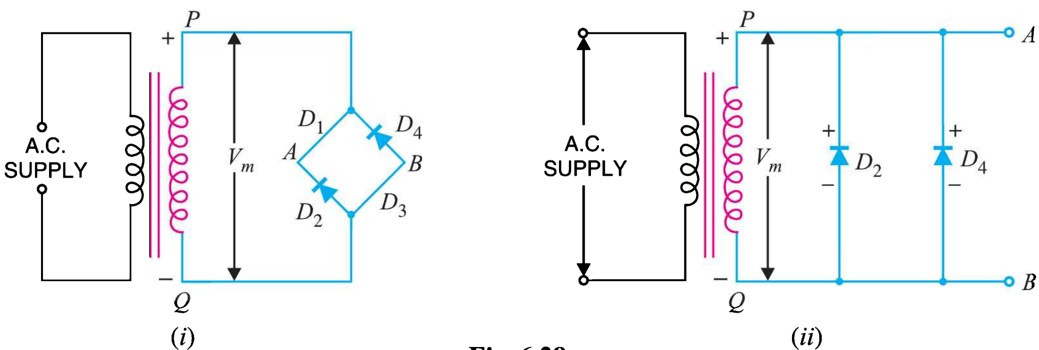


Fig. 6.28

Referring to Fig. 6.28 (*ii*), it is clear that two reverse biased diodes (*i.e.*, D_2 and D_4) and the secondary of transformer are in parallel. Hence PIV of each diode (D_2 and D_4) is equal to the maximum voltage (V_m) across the secondary. Similarly, during the next half cycle, D_2 and D_4 are forward biased while D_1 and D_3 will be reverse biased. It is easy to see that reverse voltage across D_1 and D_3 is equal to V_m.

Advantages

(*i*) The need for centre-tapped transformer is eliminated.

(*ii*) The output is twice that of the centre-tap circuit for the same secondary voltage.

(*iii*) The PIV is one-half that of the centre-tap circuit (for same d.c. output).

Disadvantages

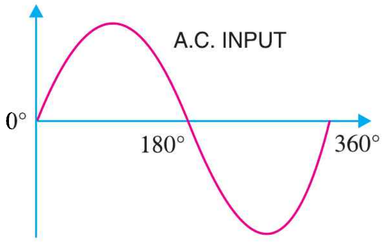
(*i*) It requires four diodes.

(*ii*) As during each half-cycle of a.c. input two diodes that conduct are in series, therefore, voltage drop in the internal resistance of the rectifying unit will be twice as great as in the centre tap circuit. This is objectionable when secondary voltage is small.

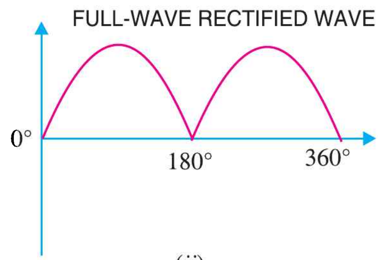
6.14 Output Frequency of Full-Wave Rectifier

The output frequency of a full-wave rectifier is double the input frequency. Remember that a wave has a complete cycle when it repeats the same pattern. In Fig. 6.29 (*i*), the input a.c. completes one cycle from 0° – 360°. However, the full-wave rectified wave completes 2 cycles in this period [See Fig. 6.29 (*ii*)]. Therefore, output frequency is twice the input frequency *i.e.*

$$f_{out} = 2 f_{in}$$



(*i*)



(*ii*)

Fig. 6.29

For example, if the input frequency to a full-wave rectifier is 100 Hz, then the output frequency will be 200 Hz.

6.15 Efficiency of Full-Wave Rectifier

Fig. 6.30 shows the process of full-wave rectification. Let $v = V_m \sin \theta$ be the a.c. voltage to be rectified. Let r_f and R_L be the diode resistance and load resistance respectively. Obviously, the rectifier will conduct current through the load in the same direction for both half-cycles of input a.c. voltage. The instantaneous current i is given by :

$$i = \frac{v}{r_f + R_L} = \frac{V_m \sin \theta}{r_f + R_L}$$

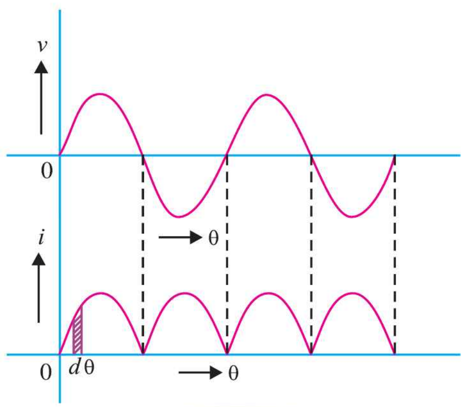


Fig. 6.30

d.c. output power. The output current is pulsating direct current. Therefore, in order to find the d.c. power, average current has to be found out. From the elementary knowledge of electrical engineering,

$$I_{dc} = \frac{2I_m}{\pi}$$

$\therefore$ d.c. power output, $P_{dc} = I_{dc}^2 \times R_L = \left(\frac{2I_m}{\pi}\right)^2 \times R_L$ ...(*i*)

a.c. input power. The a.c. input power is given by :

$$P_{ac} = I_{rms}^2 (r_f + R_L)$$

For a full-wave rectified wave, we have,

$$I_{rms} = I_m/\sqrt{2}$$

$\therefore$ $P_{ac} = \left(\frac{I_m}{\sqrt{2}}\right)^2 (r_f + R_L)$ ... (*ii*)

$\therefore$ Full-wave rectification efficiency is

$$\eta = \frac{P_{dc}}{P_{ac}} = \frac{(2I_m/\pi)^2 R_L}{\left(\frac{I_m}{\sqrt{2}}\right)^2 (r_f + R_L)}$$

$$= \frac{8}{\pi^2} \times \frac{R_L}{(r_f + R_L)} = \frac{0.812\, R_L}{r_f + R_L} = \frac{0.812}{1 + \frac{r_f}{R_L}}$$

The efficiency will be maximum if r_f is negligible as compared to R_L.

$\therefore$ Maximum efficiency = 81.2%

This is double the efficiency due to half-wave rectifier. Therefore, a full-wave rectifier is twice as effective as a half-wave rectifier.

Example 6.16. *A full-wave rectifier uses two diodes, the internal resistance of each diode may be assumed constant at 20 Ω. The transformer r.m.s. secondary voltage from centre tap to each end of secondary is 50 V and load resistance is 980 Ω. Find :*

(*i*) *the mean load current* (*ii*) *the r.m.s. value of load current*

Solution.

$$r_f = 20\ \Omega, \quad R_L = 980\ \Omega$$

Max. a.c. voltage, $V_m = 50 \times \sqrt{2} = 70.7$ V

Max. load current, $I_m = \frac{V_m}{r_f + R_L} = \frac{70.7 \text{ V}}{(20 + 980)\ \Omega} = 70.7$ mA

(*i*) Mean load current, $I_{dc} = \frac{2I_m}{\pi} = \frac{2 \times 70.7}{\pi} =$ **45 mA**

(*ii*) R.M.S. value of load current is

$$I_{rms} = \frac{I_m}{\sqrt{2}} = \frac{70.7}{\sqrt{2}} = \mathbf{50\ mA}$$

Example 6.17. *In the centre-tap circuit shown in Fig. 6.31, the diodes are assumed to be ideal i.e. having zero internal resistance. Find* :

(*i*) *d.c. output voltage* (*ii*) *peak inverse voltage* (*iii*) *rectification efficiency.*

Solution.

Primary to secondary turns, $N_1/N_2 = 5$

R.M.S. primary voltage = 230 V

∴ R.M.S. secondary voltage

$= 230 \times (1/5) = 46$ V

Maximum voltage across secondary

$= 46 \times \sqrt{2} = 65$V

Maximum voltage across half secondary winding is

$V_m = 65/2 = 32.5$ V

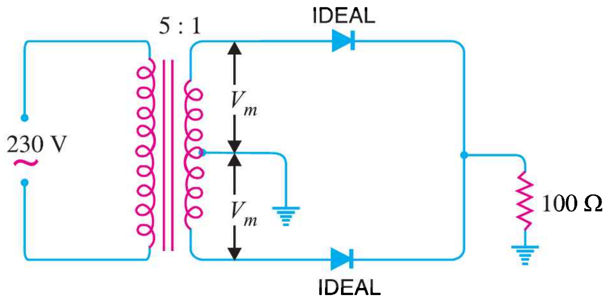


Fig. 6.31

(*i*) Average current, $I_{dc} = \dfrac{2V_m}{\pi R_L} = \dfrac{2 \times 32.5}{\pi \times 100} = 0.207$ A

∴ d.c. output voltage, $V_{dc} = I_{dc} \times R_L = 0.207 \times 100 =$ **20.7 V**

(*ii*) The peak inverse voltage is equal to the maximum secondary voltage, *i.e.*

$PIV =$ **65 V**

(*iii*) Rectification efficiency $= \dfrac{0.812}{1 + \dfrac{r_f}{R_L}}$

Since $r_f = 0$

∴ Rectification efficiency = **81.2 %**

Example 6.18. *In the bridge type circuit shown in Fig. 6.32, the diodes are assumed to be ideal. Find :* (*i*) *d.c. output voltage* (*ii*) *peak inverse voltage* (*iii*) *output frequency.*
Assume primary to secondary turns to be 4.

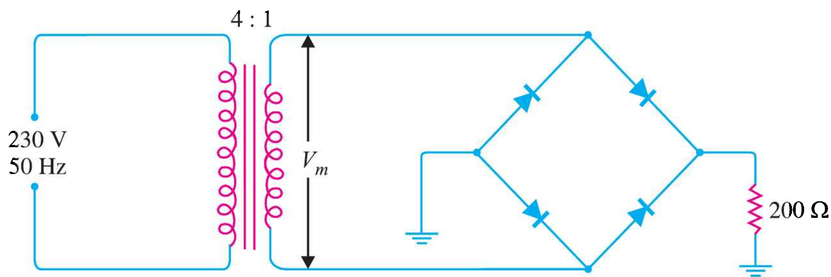


Fig. 6.32

Solution.

Primary/secondary turns, $N_1/N_2 = 4$

R.M.S. primary voltage = 230 V

∴ R.M.S. secondary voltage $= 230\ (N_2/N_1) = 230 \times (1/4) = 57.5$ V

Maximum voltage across secondary is

$$V_m = 57.5 \times \sqrt{2} = 81.3 \text{ V}$$

(*i*) Average current, $I_{dc} = \dfrac{2V_m}{\pi R_L} = \dfrac{2 \times 81.3}{\pi \times 200} = 0.26$ A

∴ d.c. output voltage, $V_{dc} = I_{dc} \times R_L = 0.26 \times 200 =$ **52 V** (*ii*) The peak inverse voltage is equal to the maximum secondary voltage *i.e.*

$$PIV = \mathbf{81.3\ V}$$

(*iii*) In full-wave rectification, there are two output pulses for each complete cycle of the input a.c. voltage. Therefore, the output frequency is twice that of the a.c. supply frequency *i.e.*

$$f_{out} = 2 \times f_{in} = 2 \times 50 = \mathbf{100\ Hz}$$

Example 6.19. *Fig. 6.33 (i) and Fig. 6.33 (ii) show the centre-tap and bridge type circuits having the same load resistance and transformer turn ratio. The primary of each is connected to 230V, 50 Hz supply.*

(i) Find the d.c. voltage in each case.

(ii) PIV for each case for the same d.c. output. Assume the diodes to be ideal.

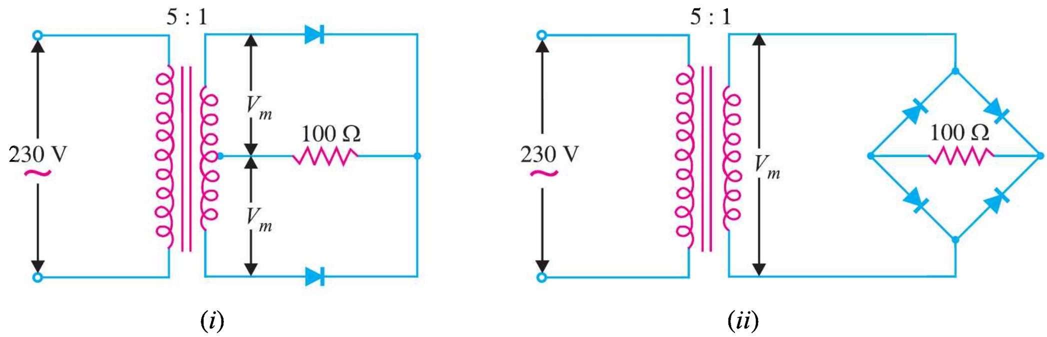


Fig. 6.33

Solution.

(*i*) D.C. output voltage

Centre-tap circuit

$$\text{R.M.S. secondary voltage} = 230 \times 1/5 = 46 \text{ V}$$

$$\text{Max. voltage across secondary} = 46 \times \sqrt{2} = 65 \text{ V}$$

Max. voltage appearing across half secondary winding is

$$V_m = 65/2 = 32.5 \text{ V}$$

$$\text{Average current, } I_{dc} = \frac{2V_m}{\pi R_L}$$

$$\text{D.C. output voltage, } V_{dc} = I_{dc} \times R_L = \frac{2V_m}{\pi R_L} \times R_L$$

$$= \frac{2V_m}{\pi} = \frac{2 \times 32.5}{\pi} = \mathbf{20.7\ V}$$

Bridge Circuit

$$\text{Max. voltage across secondary, } V_m = 65 \text{ V}$$

$$\text{D.C. output voltage, } V_{dc} = I_{dc} R_L = \frac{2V_m}{\pi R_L} \times R_L = \frac{2\,V_m}{\pi} = \frac{2 \times 65}{\pi} = \mathbf{41.4\ V}$$

This shows that for the same secondary voltage, the d.c. output voltage of bridge circuit is twice that of the centre-tap circuit.

(*ii*) PIV for same d.c. output voltage

The d.c. output voltage of the two circuits will be the same if V_m (*i.e.* max. voltage utilised by each circuit for conversion into d.c.) is the same. For this to happen, the turn ratio of the transformers should be as shown in Fig. 6.34.

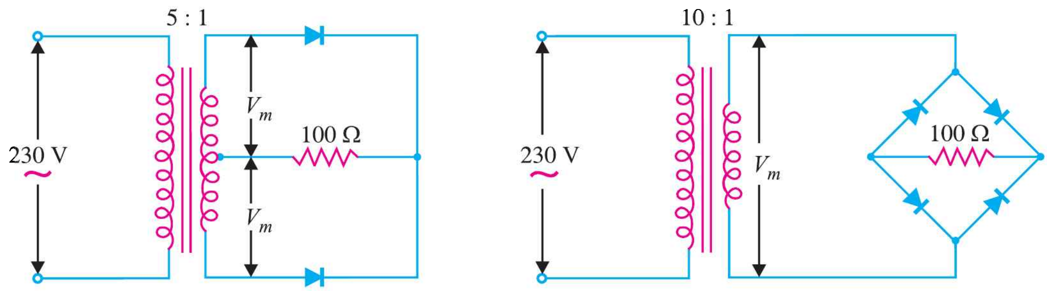
Centre-tap circuit

R.M.S. secondary voltage $= 230 \times 1/5 = 46$ V

Max. voltage across secondary $= 46 \times \sqrt{2} = 65$ V

Max. voltage across half secondary winding is

$$V_m = 65/2 = 32.5 \text{ V}$$



(*i*) (*ii*)

Fig. 6.34

$\therefore$ $PIV = 2\,V_m = 2 \times 32.5 = \mathbf{65\ V}$

Bridge type circuit

R.M.S. secondary voltage $= 230 \times 1/10 = 23$ V

Max. voltage across secondary, $V_m = 23 \times \sqrt{2} = 32.5$ V

$\therefore$ $PIV = V_m = \mathbf{32.5\ V}$

This shows that for the same d.c. output voltage, *PIV* of bridge circuit is half that of centre-tap circuit. This is a distinct advantage of bridge circuit.

Example 6.20. *The four diodes used in a bridge rectifier circuit have forward resistances which may be considered constant at 1Ω and infinite reverse resistance. The alternating supply voltage is 240 V r.m.s. and load resistance is 480 Ω. Calculate (i) mean load current and (ii) power dissipated in each diode.*

Solution.

Max. a.c. voltage, $V_m = 240 \times \sqrt{2}$ V

(***i***) At any instant in the bridge rectifier, two diodes in series are conducting. Therefore, total circuit resistance = $2\,r_f + R_L$.

$$\text{Max. load current, } I_m = \frac{V_m}{2\,r_f + R_L} = \frac{240 \times \sqrt{2}}{2 \times 1 + 480} = 0.7 \text{ A}$$

$$\therefore \quad \text{Mean load current, } I_{dc} = \frac{2\,I_m}{\pi} = \frac{2 \times 0.7}{\pi} = \mathbf{0.45\ A}$$

(***ii***) Since each diode conducts only half a cycle, diode r.m.s. current is :

$$I_{r.m.s.} = I_m/2 = 0.7/2 = 0.35 \text{ A}$$

$$\text{Power dissipated in each diode} = I^2_{r.m.s.} \times r_f = (0.35)^2 \times 1 = \mathbf{0.123\ W}$$

Example 6.21. *The bridge rectifier shown in Fig. 6.35 uses silicon diodes. Find (i) d.c. output* *voltage (ii) d.c. output current. Use simplified model for the diodes.*

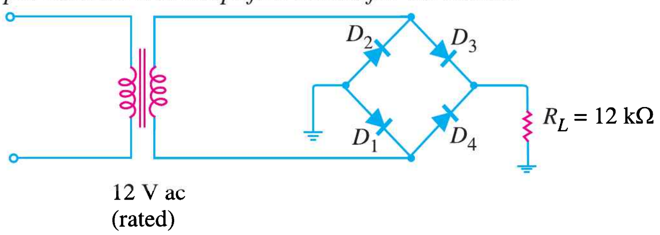


Fig. 6.35

Solution. The conditions of the problem suggest that the a.c voltage across transformer secondary is 12V r.m.s.

∴ Peak secondary voltage is

$$V_{s\,(pk)} = 12 \times \sqrt{2} = 16.97 \text{ V}$$

(*i*) At any instant in the bridge rectifier, two diodes in series are conducting.

∴ Peak output voltage is

$$V_{out\,(pk)} = 16.97 - 2\,(0.7) = 15.57 \text{ V}$$

∴ Average (or d.c.) output voltage is

$$V_{av} = V_{dc} = \frac{2\,V_{out\,(pk)}}{\pi} = \frac{2 \times 15.57}{\pi} = \mathbf{9.91V}$$

(*ii*) Average (or d.c.) output current is

$$I_{av} = \frac{V_{av}}{R_L} = \frac{9.91\text{V}}{12 \text{ k}\Omega} = \mathbf{825.8\ \mu A}$$

6.16 Faults in Centre-Tap Full-Wave Rectifier

The faults in a centre-tap full-wave rectifier may occur in the transformer or rectifier diodes. Fig. 6.36 shows the circuit of a centre-tap full-wave rectifier. A fuse is connected in the primary of the transformer for protection purposes.

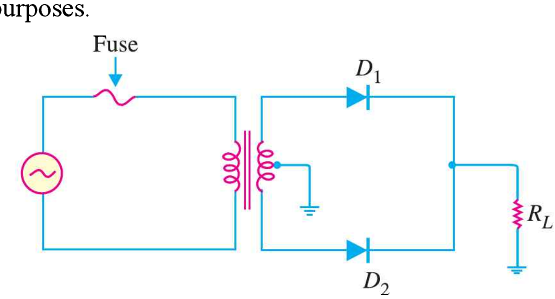


Fig. 6.36

We can divide the rectifier faults into two classes viz.

1. Faults in transformer **2.** Faults in rectifier diodes

1. Faults in Transformer. The transformer in a rectifier circuit can develop the following faults :

(*i*) A shorted primary or secondary winding.

(*ii*) An open primary or secondary winding.

(*iii*) A short between the primary or secondary winding and the transformer frame.

(*i*) In most cases, a **shorted primary** or **shorted secondary** will cause the fuse in the primary to blow. If the fuse does not blow, the d.c. output from the rectifier will be extremely low and the transformer itself will be very hot.

(*ii*) When the **primary** or **secondary winding of the transformer opens**, the output from the rectifier will drop to zero. In this case, the primary fuse will not blow. If you believe that either transformer winding is open, a simple resistance check will verify your doubt. If either winding reads a very high resistance, the winding is open.

(*iii*) If **either winding shorts to the transformer casing**, the primary fuse will blow. This fault can be checked by measuring the resistances from the winding leads to the transformer casing. A low resistance measurement indicates that a winding-to-case short exists.

2. Faults in Rectifier Diodes. If a fault occurs in a rectifier diode, the circuit conditions will indicate the type of fault.

(*i*) If **one diode in the centre-tap full-wave rectifier is shorted**, the primary fuse will blow. The reason is simple. Suppose diode D_2 in Fig. 6.36 is shorted. Then diode D_2 will behave as a wire. When diode D_1 is forward biased, the transformer secondary will be shorted through D_1. This will cause excessive current to flow in the secondary (and hence in the primary), causing the primary fuse to blow.

(*ii*) If **one diode in the centre-tap full-wave rectifier opens**, the output from the rectifier will resemble the output from a half-wave rectifier. The remedy is to replace the diode.

Bridge Rectifier Faults. The transformer faults and their remedies for bridge rectifier circuits are the same as for centre-tap full-wave rectifier. Again symptoms for shorted and open diodes in the bridge rectifier are the same as those for the centre-tap circuit. In the case of bridge circuit, you simply have more diodes that need to be tested.

6.17 Nature of Rectifier Output

It has already been discussed that the output of a rectifier is pulsating d.c. as shown in Fig. 6.37. In fact, if such a waveform is carefully analysed, it will be found that it contains a d.c. component and an a.c. component. The a.c. component is responsible for the *pulsations in the wave. The reader may wonder how a pulsating d.c. voltage can have an a.c. component when the voltage never becomes negative. The answer is that any wave which varies in a regular manner has an a.c. component.

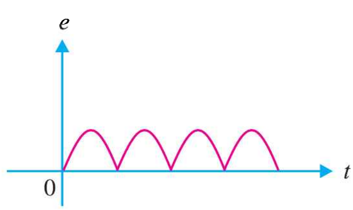


Fig. 6.37

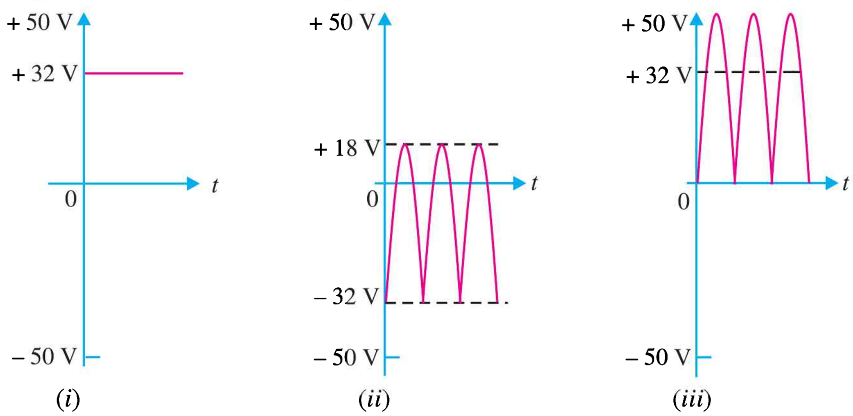


Fig. 6.38

* Means changing output voltage.

The fact that a pulsating d.c. contains both d.c. and a.c. components can be beautifully illustrated by referring to Fig. 6.38. Fig. 6.38 (*i*) shows a pure d.c. component, whereas Fig. 6.38 (*ii*) shows the *a.c. component. If these two waves are added together, the resulting wave will be as shown in Fig. 6.38 (*iii*). It is clear that the wave shown in Fig. 6.38 (*iii*) never becomes negative, although it contains both a.c. and d.c. components. The striking resemblance between the rectifier output wave shown in Fig. 6.37 and the wave shown in Fig. 6.38 (*iii*) may be noted.

Rectifier

It follows, therefore, that a pulsating output of a rectifier contains a d.c. component and an a.c. component.

6.18 Ripple Factor

The output of a rectifier consists of a d.c. component and an a.c. component (also known as *ripple*). The a.c. component is undesirable and accounts for the pulsations in the rectifier output. The effectiveness of a rectifier depends upon the magnitude of a.c. component in the output ; the smaller this component, the more effective is the rectifier.

The ratio of r.m.s. value of a.c. component to the d.c. component in the rectifier output is known as **ripple factor** *i.e.*

$$\text{Ripple factor} = \frac{\text{r.m.s. value of a.c component}}{\text{value of d.c. component}} = \frac{I_{ac}}{I_{dc}}$$

Therefore, ripple factor is very important in deciding the effectiveness of a rectifier. The smaller the ripple factor, the lesser the effective a.c. component and hence more effective is the rectifier.

Mathematical analysis. The output current of a rectifier contains d.c. as well as a.c. component. The undesired a.c. component has a frequency of 100 Hz (*i.e.* double the supply frequency 50 Hz) and is called the *ripple* (See Fig. 6.39). It is a fluctuation superimposed on the d.c. component.

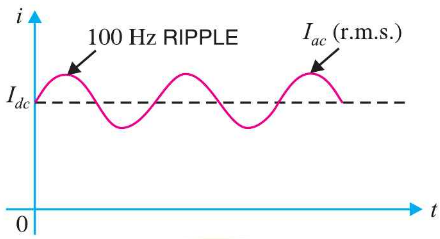


Fig. 6.39

By definition, the effective (*i.e.* r.m.s.) value of total load current is given by :

$$I_{rms} = \sqrt{I_{dc}^2 + I_{ac}^2}$$

or

$$I_{ac} = \sqrt{I_{rms}^2 - I_{dc}^2}$$

Dividing throughout by I_{dc}, we get,

$$\frac{I_{ac}}{I_{dc}} = \frac{1}{I_{dc}}\sqrt{I_{rms}^2 - I_{dc}^2}$$

But I_{ac}/I_{dc} is the ripple factor.

$$\therefore \quad \text{Ripple factor} = \frac{1}{I_{dc}}\sqrt{I_{rms}^2 - I_{dc}^2} = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1}$$

(*i*) **For half-wave rectification.** In half-wave rectification,

$$I_{rms} = I_m/2 \quad ; \quad I_{dc} = I_m/\pi$$

* Although the a.c. component is not a sine-wave, yet it is alternating one.

$$\therefore \qquad \text{Ripple factor} = \sqrt{\left(\frac{I_m/2}{I_m/\pi}\right)^2 - 1} = 1.21$$

It is clear that a.c. component exceeds the d.c. component in the output of a half-wave rectifier. This results in greater pulsations in the output. Therefore, half-wave rectifier is ineffective for conversion of a.c. into d.c.

(*ii*) **For full-wave rectification.** In full-wave rectification,

$$I_{rms} = \frac{I_m}{\sqrt{2}} \quad ; \quad I_{dc} = \frac{2\,I_m}{\pi}$$

$$\therefore \qquad \text{Ripple factor} = \sqrt{\left(\frac{I_m/\sqrt{2}}{2\,I_m/\pi}\right)^2 - 1} = 0.48$$

$$i.e. \qquad \frac{\text{effective a.c. component}}{\text{d.c. component}} = 0.48$$

This shows that in the output of a full-wave rectifier, the d.c. component is more than the a.c. component. Consequently, the pulsations in the output will be less than in half-wave rectifier. For this reason, full-wave rectification is invariably used for conversion of a.c. into d.c.

Example 6.22. *A power supply A delivers 10 V dc with a ripple of 0.5 V r.m.s. while the power supply B delivers 25 V dc with a ripple of 1 mV r.m.s. Which is better power supply ?*

Solution. The lower the ripple factor of a power supply, the better it is.

For power supply A

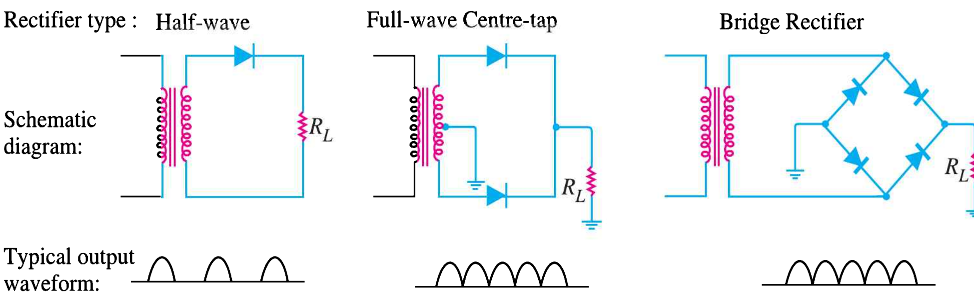
$$\text{Ripple factor} = \frac{V_{ac(r.m.s.)}}{V_{dc}} = \frac{0.5}{10} \times 100 = 5\%$$

For power supply B

$$\text{Ripple factor} = \frac{V_{ac(r.m.s.)}}{V_{dc}} = \frac{0.001}{25} \times 100 = 0.004\%$$

Clearly, power supply *B* is better.

6.19 Comparison of Rectifiers



S. No.	Particulars	Half-wave	Centre-tap	Bridge type
1	No. of diodes	1	2	4
2	Transformer necessary	no	yes	no
3	Max. efficiency	40.6%	81.2%	81.2%
4	Ripple factor	1.21	0.48	0.48
5	Output frequency	f_{in}	$2\,f_{in}$	$2\,f_{in}$
6	Peak inverse voltage	V_m	$2\,V_m$	V_m

A comparison among the three rectifier circuits must be made very judiciously. Although bridge circuit has some disadvantages, it is the best circuit from the viewpoint of overall performance. When cost of the transformer is the main consideration in a rectifier assembly, we invariably use the bridge circuit. This is particularly true for large rectifiers which have a low-voltage and a high-current rating.

6.20 Filter Circuits

Generally, a rectifier is required to produce pure d.c. supply for using at various places in the electronic circuits. However, the output of a rectifier has pulsating *character *i.e.* it contains a.c. and d.c. components. The a.c. component is undesirable and must be kept away from the load. To do so, a *filter circuit* is used which removes (or *filters out*) the a.c. component and allows only the d.c. component to reach the load.

*A **filter circuit** is a device which removes the a.c. component of rectifier output but allows the d.c. component to reach the load.*

Obviously, a filter circuit should be installed between the rectifier and the load as shown in Fig. 6.40. A filter circuit is generally a combination of inductors (L) and capacitors (C). The filtering action of L and C depends upon the basic electrical principles. A capacitor passes a.c. readily but does not **pass d.c. at all. On the other hand, an inductor †opposes a.c. but allows d.c. to pass through it. It then becomes clear that suitable network of L and C can effectively remove the a.c. component, allowing the d.c. component to reach the load.

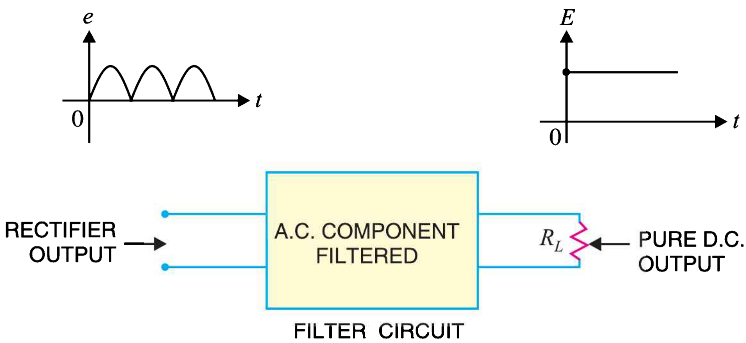


Fig. 6.40

6.21 Types of Filter Circuits

The most commonly used filter circuits are *capacitor filter*, *choke input filter* and *capacitor input filter or π-filter*. We shall discuss these filters in turn.

(*i*) **Capacitor filter.** Fig. 6.41 (*ii*) shows a typical capacitor filter circuit. It consists of a capacitor C placed across the rectifier output in parallel with load R_L. The pulsating direct voltage of the rectifier is applied across the capacitor. As the rectifier voltage increases, it charges the capacitor and also supplies current to the load. At the end of quarter cycle [Point A in Fig. 6.41 (*iii*)], the

* If such a d.c. is applied in an electronic circuit, it will produce a *hum*.

** A capacitor offers infinite reactance to d.c. For d.c., $f = 0$.

$$\therefore \quad X_C = \frac{1}{2\pi fC} = \frac{1}{2\pi \times 0 \times C} = \infty$$

Hence, a capacitor does not allow d.c. to pass through it.

† We know $X_L = 2\pi fL$. For d.c., $f = 0$ and, therefore, $X_L = 0$. Hence inductor passes d.c. quite readily. For a.c., it offers opposition and drops a part of it.

capacitor is charged to the peak value V_m of the rectifier voltage. Now, the rectifier voltage starts to decrease. As this occurs, the capacitor discharges through the load and voltage across it (*i.e.* across parallel combination of *R*-*C*) decreases as shown by the line *AB* in Fig. 6.41 (*iii*). The voltage across load will decrease only slightly because immediately the next voltage peak comes and recharges the capacitor. This process is repeated again and again and the output voltage waveform becomes *ABCDEFG*. It may be seen that very little ripple is left in the output. Moreover, output voltage is higher as it remains substantially near the peak value of rectifier output voltage.

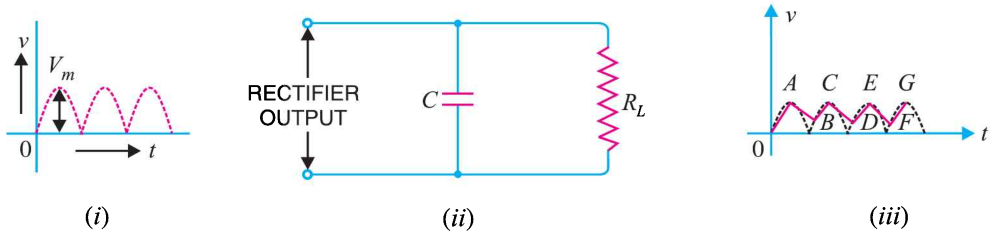


Fig. 6.41

The capacitor filter circuit is extremely popular because of its low cost, small size, little weight and good characteristics. For small load currents (say upto 50 mA), this type of filter is preferred. It is commonly used in transistor radio battery eliminators.

(*ii*) **Choke input filter.** Fig. 6.42 shows a typical choke input filter circuit. It consists of a *choke *L* connected in series with the rectifier output and a filter capacitor *C* across the load. Only a single filter section is shown, but several identical sections are often used to reduce the pulsations as effectively as possible.

The pulsating output of the rectifier is applied across terminals 1 and 2 of the filter circuit. As discussed before, the pulsating output of rectifier contains a.c. and d.c. components. The choke offers high opposition to the passage of a.c. component but negligible opposition to the d.c. component. The result is that most of the a.c. component appears across the choke while whole of d.c. component passes through the choke on its way to load. This results in the reduced pulsations at terminal 3.

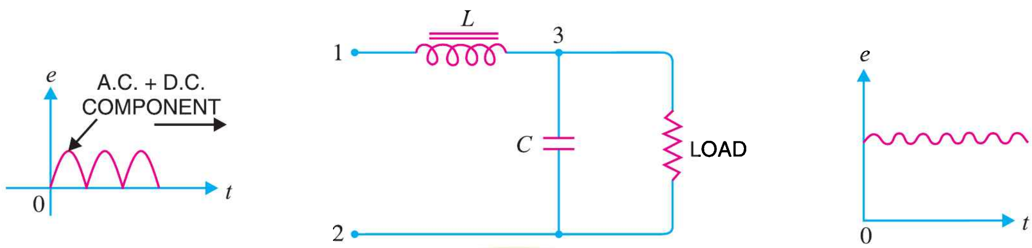


Fig. 6.42

At terminal 3, the rectifier output contains d.c. component and the remaining part of a.c. component which has managed to pass through the choke. Now, the low reactance of filter capacitor bypasses the a.c. component but prevents the d.c. component to flow through it. Therefore, only d.c. component reaches the load. In this way, the filter circuit has filtered out the a.c. component from the rectifier output, allowing d.c. component to reach the load.

(*iii*) **Capacitor input filter** *or* **π-filter.** Fig. 6.43 shows a typical capacitor input filter or **π-filter. It consists of a filter capacitor C_1 connected across the rectifier output, a choke *L* in series and

* The shorthand name of inductor coil is choke.

** The shape of the circuit diagram of this filter circuit appears like Greek letter π (pi) and hence the name π-filter.

another filter capacitor C_2 connected across the load. Only one filter section is shown but several identical sections are often used to improve the smoothing action.

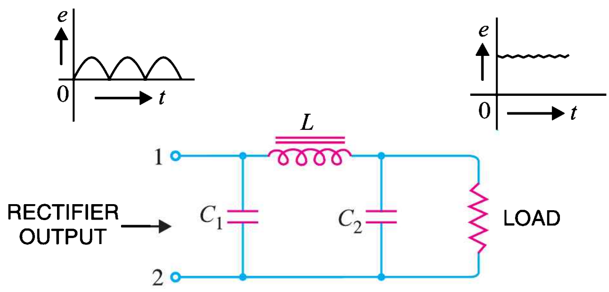


Fig. 6.43

The pulsating output from the rectifier is applied across the input terminals (*i.e.* terminals 1 and 2) of the filter. The filtering action of the three components *viz* C_1, L and C_2 of this filter is described below :

(*a*) The *filter capacitor* C_1 offers low reactance to a.c. component of rectifier output while it offers infinite reactance to the d.c. component. Therefore, capacitor C_1 bypasses an appreciable amount of a.c. component while the d.c. component continues its journey to the choke L.

(*b*) The *choke* L offers high reactance to the a.c. component but it offers almost zero reactance to the d.c. component. Therefore, it allows the d.c. component to flow through it, while the *unbypassed a.c. component is blocked.

(*c*) The *filter capacitor* C_2 bypasses the a.c. component which the choke has failed to block. Therefore, only d.c. component appears across the load and that is what we desire.

Example 6.23. *For the circuit shown in Fig. 6.44, find the output d.c. voltage.*

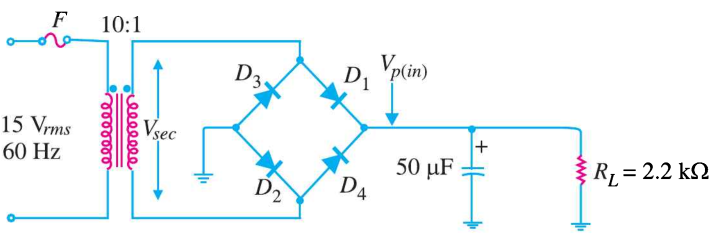


Fig. 6.44

Solution. It can be proved that output d.c. voltage is given by :

$$V_{dc} = V_{p(in)}\left(1 - \frac{1}{2f\ R_L C}\right)$$

Here $V_{p(in)}$ = Peak rectified full-wave voltage applied to the filter

f = Output frequency

Peak primary voltage, $V_{p(prim)} = \sqrt{2} \times 115 = 163\text{V}$

Peak secondary voltage, $V_{p(sec)} = \left(\frac{1}{10}\right) \times 163 = 16.3\text{V}$

Peak full-wave rectified voltage at the filter input is

$$V_{p\,(in)} = V_{p\,(sec)} - 2\times0.7 = 16.3 - 1.4 = 14.9\text{V}$$

For full-wave rectification, $f = 2\ f_{in} = 2\times60 = 120$ Hz

Now $$\frac{1}{2f\ R_L C} = \frac{1}{2\times120\times(2.2\times10^3)\times(50\times10^{-6})} = 0.038$$

* That part of a.c. component which could not be bypassed by capacitor C_1.

$$\therefore \quad V_{dc} = V_{p(in)}\left(1 - \frac{1}{2f\ R_L C}\right) = 14.9\ (1 - 0.038) = \mathbf{14.3V}$$

Example 6.24. *The choke of Fig. 6.45 has a d.c. resistance of 25 Ω. What is the d.c. voltage if the full-wave signal into the choke has a peak value of 25.7 V ?*

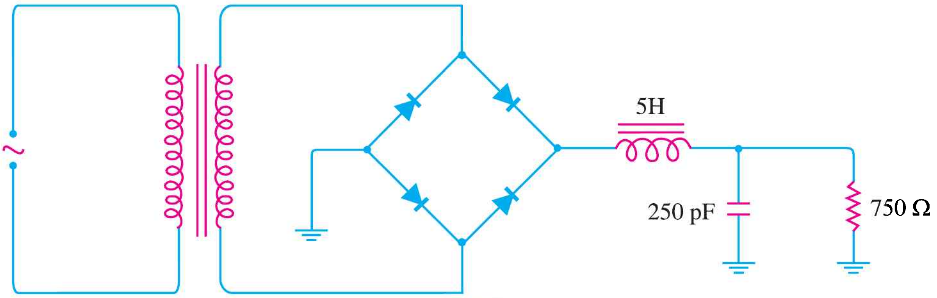


Fig. 6.45

Solution. The output of a full-wave rectifier has a d.c. component and an a.c. component. Due to the presence of a.c. component, the rectifier output has a pulsating character as shown in Fig. 6.46. The maximum value of the pulsating output is V_m and d.c. component is $V'_{dc} = 2\ V_m/\pi$.

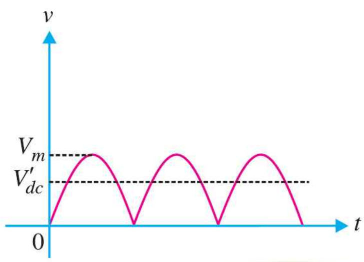


Fig. 6.46

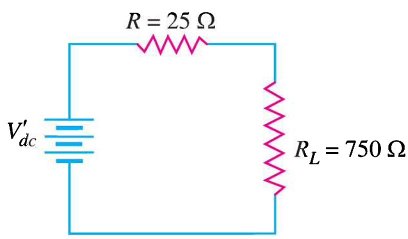


Fig. 6.47

For d.c. component V'_{dc}, the choke resistance is in series with the load as shown in Fig. 6.47.

$$\therefore \quad \text{Voltage across load, } V_{dc} = \frac{V'_{dc}}{R + R_L} \times R_L$$

$$\text{In our example, } V'_{dc} = \frac{2V_m}{\pi} = \frac{2 \times 25.7}{\pi} = 16.4 \text{ V}$$

$$\therefore \quad \text{Voltage across load, } V_{dc} = \frac{V'_{dc}}{R + R_L} \times R_L = \frac{16.4}{25 + 750} \times 750 = \mathbf{15.9\ V}$$

The voltage across the load is 15.9 V dc *plus* a small ripple.

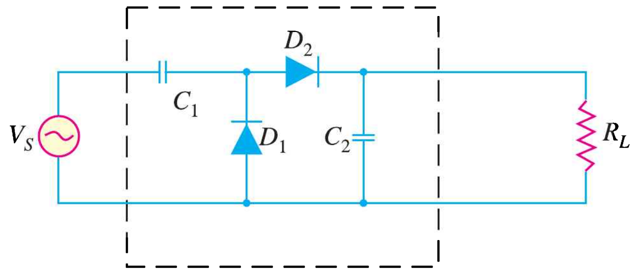
6.22 Voltage Multipliers

With a diode, we can build a rectifier to produce a d.c. voltage that is nearly equal to the peak value of input a.c. voltage. We can also use diodes and capacitors to build a circuit that will provide a *d.c output* that is multiple of the *peak input a.c. voltage*. Such a circuit is called a voltage multiplier. For example, a *voltage doubler* will provide a d.c. output that is *twice* the peak input a.c. voltage, a *voltage tripler* will provide a d.c. output that is three times the peak input a.c. voltage and so on.

While voltage multipliers provide d.c. output that is much greater than the peak input a.c. voltage, there is no power amplification and law of conservation of energy holds good. When a voltage multiplier *increases* the peak input voltage by a factor *n*, the peak input current is *decreased* by approximately the same factor. Thus the actual power output from a voltage multiplier will *never* be greater than the input power. In fact, there are losses in the circuit (e.g. in diodes, capacitors etc.) so that the output power will actually be *less than* the input power.

6.23 Half-Wave Voltage Doubler

A half-wave voltage doubler consists of two diodes and two capacitors connected in a manner as shown in Fig. 6.48. It will be shown that if the peak input a.c. voltage is $V_{S(pk)}$, the d.c. output voltage will be 2 $V_{S\,(pk)}$ provided the diodes are ideal (this assumption is fairly reasonable). The basic idea in a voltage multiplier is to charge each capacitor to the peak input a.c. voltage and to arrange the capacitors so that their stored voltages will add.



Half-wave voltage doubler

Fig. 6.48

Circuit action. We now discuss the working of a half-wave voltage doubler.

(*i*) During the negative half-cycle of a.c. input voltage [See Fig. 6.49 (*i*)], diode D_1 is forward biased and diode D_2 is reverse biased [See Fig. 6.49 (*i*)]. Therefore, diode D_1 can be represented by a *short* and diode D_2 as an *open*. The equivalent circuit then becomes as shown in Fig. 6.49 (*ii*).

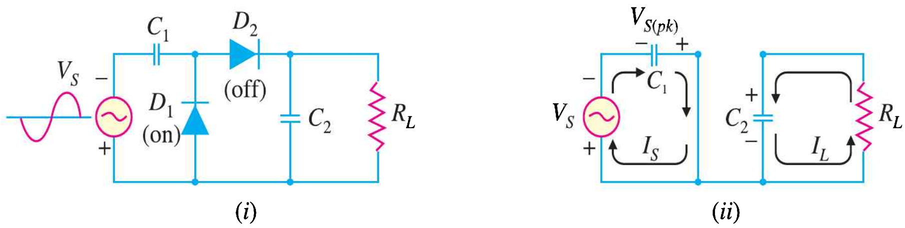


Fig. 6.49

As you can see [See Fig.6.49 (*ii*)], C_1 will charge until voltage across it becomes equal to peak value of source voltage [$V_{S\,(pk)}$]. At the same time, C_2 will be in the process of discharging through the load R_L (The source of this charge on C_2 will be explained in a moment). *Note that in all figures electron flow is shown.*

(*ii*) When the polarity of the input a.c. voltage reverses (*i.e.* during positive half-cycle), the circuit conditions become as shown in Fig. 6.50 (*i*). Now D_1 is reverse biased and D_2 is forward biased and the equivalent circuit becomes as shown in Fig. 6.50 (*ii*).

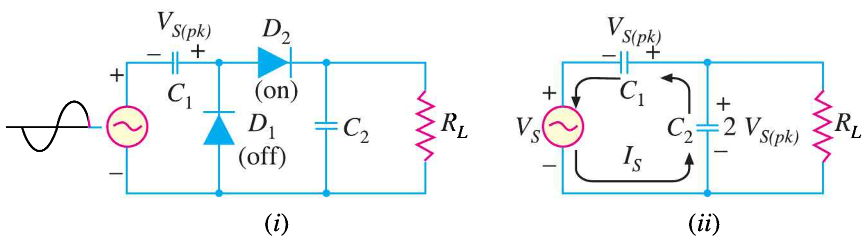


Fig. 6.50

Referring to Fig. 6.50 (*ii*), it is easy to see that C_1 (charged to $V_{S\,(pk)}$)and the source voltage (V_S) now act as *series-aiding* voltage sources. Thus C_2 will be charged to the sum of the series peak voltages *i.e.* 2 $V_{S(pk)}$.

(*iii*) When V_S returns to its original polarity (*i.e.* negative half-cycle), D_2 is again turned off (*i.e.* reverse biased). With D_2 turned off, the only discharge path for C_2 is through the load resistance R_L. The time constant (= $R_L\,C_2$) of this circuit is so adjusted that C_2 has little time to lose any of its charge before the input polarity reverses again. During the positive half-cycle, D_2 is turned on and C_2 recharges until voltage across it is again equal to 2 $V_{S\,(pk)}$.

$\therefore$ D.C. output voltage, $V_{dc} = 2\ V_{S\,(pk)}$

Since C_2 barely discharges between input cycles, the output waveform of the half-wave voltage doubler closely resembles that of a filtered half-wave rectifier. Fig. 6.51 shows the input and output waveforms for a half-wave voltage doubler.

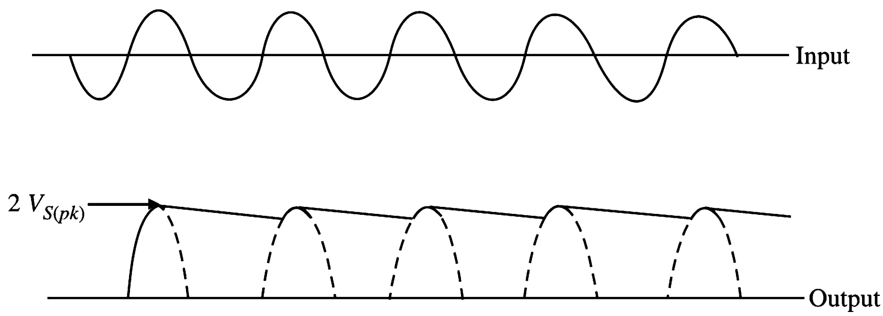


Fig. 6.51

The voltage multipliers have the disadvantage of poor voltage regulation. This means that d.c. output voltage drops considerably as the load current increases. Large filter capacitors are needed to help maintain the output voltage.

6.24 Voltage Stabilisation

A rectifier with an appropriate filter serves as a good source of d.c. output. However, the major disadvantage of such a power supply is that the output voltage changes with the variations in the input voltage or load. Thus, if the input voltage increases, the d.c. output voltage of the rectifier also increases. Similarly, if the load current increases, the output voltage falls due to the voltage drop in the rectifying element, filter chokes, transformer winding etc. In many electronic applications, it is desired that the output voltage should remain constant regardless of the variations in the input voltage or load. In order to ensure this, a voltage stabilising device, called voltage stabiliser is used. Several stabilising circuits have been designed but only *zener diode* as a voltage stabiliser will be discussed.

6.25 Zener Diode

It has already been discussed that when the reverse bias on a crystal diode is increased, a critical voltage, called *breakdown voltage* is reached where the reverse current increases sharply to a high value. The breakdown region is the knee of the reverse characteristic as shown in Fig. 6.52. The satisfactory explanation of this breakdown of the junction was first given by the American scientist *C.* Zener. Therefore, the breakdown voltage is sometimes called *zener voltage* and the sudden increase in current is known as zener current.

The breakdown or zener voltage depends upon the amount of doping. If the diode is heavily doped, depletion layer will be thin and consequently the breakdown of the junction will occur at a lower reverse voltage. On the other hand, a lightly doped diode has a higher breakdown voltage. When an ordinary crystal diode is properly doped so that it has a sharp breakdown voltage, it is called

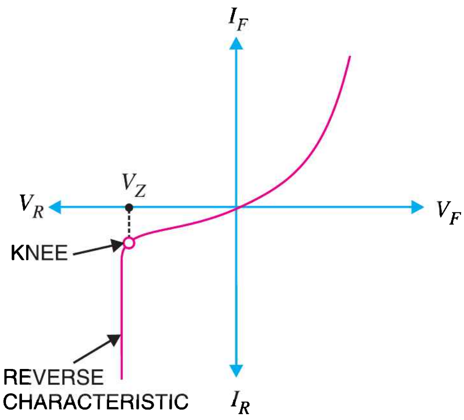


Fig. 6.52

a zener diode.

A properly doped crystal diode which has a sharp breakdown voltage is known as a **zener diode.**

Fig. 6.53 shows the symbol of a zener diode. It may be seen that it is just like an ordinary diode except that the bar is turned into *z*-shape. The following points may be noted about the zener diode:

(*i*) A zener diode is like an ordinary diode except that it is properly doped so as to have a sharp breakdown voltage.

(*ii*) A zener diode is always reverse connected *i.e.* it is always reverse biased.

(*iii*) A zener diode has sharp breakdown voltage, called zener voltage V_Z.

(*iv*) When forward biased, its characteristics are just those of ordinary diode.

(*v*) The zener diode is not immediately burnt just because it has entered the *breakdown region. As long as the external circuit connected to the diode limits the diode current to less than *burn out* value, the diode will not burn out.

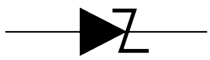


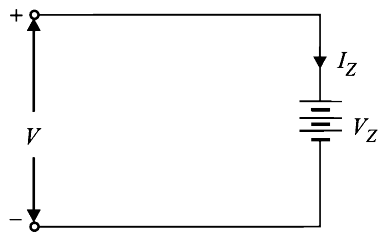
Fig. 6.53

6.26 Equivalent Circuit of Zener Diode

The analysis of circuits using zener diodes can be made quite easily by replacing the zener diode by its equivalent circuit.

(*i*) "On" state. When reverse voltage across a zener diode is equal to or more than break down voltage V_Z, the current increases very sharply. In this region, the curve is almost vertical. It means that voltage across zener diode is constant at V_Z even though the current through it changes. Therefore, in the breakdown region, an **ideal zener diode can be represented by a battery of voltage V_Z as shown in Fig. 6.54 (*ii*). Under such conditions, the zener diode is said to be in the "ON" state.

(*i*)
$V \geq V_Z$



(*ii*)
Equivalent circuit of zener for ''on'' state

Fig. 6.54

(*ii*) "OFF" state. When the reverse voltage across the zener diode is less than V_Z but greater than 0 V, the zener diode is in the "OFF" state. Under such conditions, the zener diode can be represented by an open-circuit as shown in Fig. 6.55 (*ii*).

* The current is limited only by both external resistance and the power dissipation of zener diode.

** This assumption is fairly reasonable as the impedance of zener diode is quite small in the breakdown region.