G.C.E. (Advanced Level)

Physics

Grade 13
Resource Book
Unit 09

Electronics

Department of Science
Faculty of Science and Technology
National Institute of Education
Maharagama
www.nie.lk

© 2020 National Institute of Education. All Rights Reserved.

G.C.E. (Advanced Level)
Physics
Grade 13
Resource Book

Electronics

1st Printing - 2022

Department of Science
Faculty of Science and Technology
National Institute of Education
Maharagama
www.nie.lk

Printed by : Printing and Publication Department
National Institute of Education
Maharagama

© 2020 National Institute of Education. All Rights Reserved.

Message from the Director General

The National Institute of Education takes opportune steps from time to time for the development of quality in education. Preparation of supplementary resource books for respective subjects is one such initiative.

Supplementary resource books have been composed by a team of curriculum developers of the National Institute of Education, subject experts from the national universities and experienced teachers from the school system. Because these resource books have been written so that they are in line with the G. C. E. (A/L) new syllabus implemented in 2017, students can broaden their understanding of the subject matter by referring these books while teachers can refer them in order to plan more effective learning teaching activities.

I wish to express my sincere gratitude to the staff members of the National Institute of Education and external subject experts who made their academic contribution to make this material available to you.

Dr. Sunil Jayantha Nawarathna
Director General
National Institute of Education
Maharagama.

© 2020 National Institute of Education. All Rights Reserved.

Message from the Director

Since 2017, a rationalized curriculum, which is an updated version of the previous curriculum is in effect for the G.C.E (A/L) in the general education system of Sri Lanka. In this new curriculum cycle, revisions were made in the subject content, mode of delivery and curricular materials of the G.C.E. (A/L) Physics, Chemistry and Biology. Several alterations in the learning teaching sequence were also made. A new Teachers' Guide was introduced in place of the previous Teacher's Instruction Manual. In concurrence to that, certain changes in the learning teaching methodology, evaluation and assessment are expected. The newly introduced Teachers' Guide provides learning outcomes, a guideline for teachers to mould the learning events, assessment and evaluation.

When implementing the previous curricula, the use of internationally recognized standard textbooks published in English was imperative for the Advanced Level science subjects. Due to the contradictions of facts related to the subject matter between different textbooks and inclusion of the content beyond the limits of the local curriculum, the usage of those books was not convenient for both teachers and students. This book comes to you as an attempt to overcome that issue.

As this book is available in Sinhala, Tamil, and English, the book offers students an opportunity to refer the relevant subject content in their mother tongue as well as in English within the limits of the local curriculum. It also provides both students and teachers a source of reliable information expected by the curriculum instead of various information gathered from the other sources.

This book authored by subject experts from the universities and experienced subject teachers is presented to you followed by the approval of the Academic Affairs Board and the Council of the National Institute of Education. Thus, it can be recommended as material of a high standard.

Dr. A. D. A. De Silva

Director

Department of Science

© 2020 National Institute of Education. All Rights Reserved.

Guidance
Mr. K. R. Pathmasiri
Faculty of Science and Technology
Deputy Director General - National Institute of Education

Supervision
Dr. A.D.A. De Silva
Director, Department of Science
National Institute of Education (NIE)

Editing :

P. Malavipathirana - Senior Lecturer, National Institute of Education
Dr. M.L.S. Piyatissa - Assistant Lecturer, National Institute of Education
Miss. R.A. Amarasinghe - Assistant Lecturer, National Institute of Education
Mrs. R.N.N. Weerasinghe - Assistant Lecturer, National Institute of Education

Subject Guidance

Prof. S.R.D. Rosa - Department of Physics, University of Colombo
Prof. L.R.A.K. Bandara - Department of Physics, University of Peradeniya

Writers

H. S. K. Wijayathilaka - (Retired) Sri Lanka Education Administrative Service - I
P. Wickramasekara - (Retired) Sri Lanka Teacher Service - I

Translation (Sinhala to English)

P. Wickramasekara - (Retired) Sri Lanka Teacher Service - I

English Language Editing

Dr. (Mrs.) C. L. Amarasekara - Consultant, National Institute of Education

Technical Page Setting

Mrs. D. M. I. R. Dissanayaka - National Institute of Education

Computer Type setting

Mrs. R. K. Fernando - National Institute of Education
Miss. S. A. S. N. Subasingha - National Institute of Education

Diagrams

J. R. Lankapura - Sri Lanka Teacher Service - I,
Wickramashila National School, Giriulla

Cover Page

Mrs. D. M. I. R. Dissanayaka - National Institute of Education

Other Assistance

Mangala Welipitiya - National Institute of Education
Mrs. W. P. P. Weerawardhana - National Institute of Education

© 2020 National Institute of Education. All Rights Reserved.

© 2020 National Institute of Education. All Rights Reserved.

© 2020 National Institute of Education. All Rights Reserved.

© 2020 National Institute of Education. All Rights Reserved.

Content

© 2020 National Institute of Education. All Rights Reserved.

© 2020 National Institute of Education. All Rights Reserved.

© 2020 National Institute of Education. All Rights Reserved.

Chapter One

Use of Semiconductors and Diodes

1.1 Introduction

We have studied under current electricity the effects (eg. heating, lighting, mechanical and magnetic) related to the flow of electrons through a conductor. In electronics it is expected to study the outcomes due to controlling or amplifying the flow of electrons, or other change carriers, in a vacuum or solid medium.

Many appliances used in day-to-day life are produced based on electronics. Most of them are devices made, using semiconductors coming in solid - state electronics. The starting point of electronics came up with the discovery of the ability of electrons to travel through a vacuum (nevertheless it was not known at the beginning that, what travels in this manner were negatively charged electrons).

The first experience of solid state electronics came up with a certain discovery made in 1940 by the scientist Russell Ohl who was attached to the Bell Laboratory. The discovery was that a germanium semiconductor, when added with a small amount of impurity showed photoelectric properties. Afterwards he discovered the rectification property of a p - n junction. with that, an interest to control the travel of electron was arisen. In 1941 the first detector diode was made using a p - n junction. The first transistor made using germanium semiconductors was introduced in 1947 by three American scientists John Bardeen, William Shookly and Walter Brattain. With the development of solid-state electronics the thermionic valve gradually went out and, instead the transistor came in.

1.2 Semiconductors

Depending on their electrical conductivity or resistivity materials are divided basically as conductors and insulators. In general the resistivity (ρ) of metal conductors is of $10^{-8}\,\Omega$ m, and that of non metal insulators is greater than $10^{12}\,\Omega$ m.

The resistivity is the reciprocal of the conductivity.

$$\text{Resistivity } (\rho) = \frac{1}{\text{Conductivity}(\sigma)}$$

The resistivity and conductivity of some materials at 20 °C are given in the following table.

Table 1.1 Resistivity and conductivity of some materials at 20 °C

Material	Resistivity (ρ) Ω m	Conductivity (σ) Moho
Silver	1.6×10^{-8}	6.25×10^{7}
Copper	1.7×10^{-8}	5.9×10^{7}
Gold	2.42×10^{-8}	4.2×10^{7}
Aluminium	2.8×10^{-8}	3.6×10^{7}
Iron	10×10^{-8}	1×10^{7}
Lead	20×10^{-8}	0.5×10^{7}
Manganin	44.5×10^{-8}	0.23×10^{7}
Constantan	49×10^{-8}	0.20×10^{7}
Nichrome	110×10^{-8}	0.09×10^{7}
Silicon	2.3×10^{3}	4.35×10^{-4}
Germanium	6.5×10^{-1}	1.54
Pyrex (glass)	1×10^{12}	1×10^{-12}
Ebonite	2×10^{13}	0.5×10^{-13}
Paraffin	3×10^{13}	0.33×10^{-13}
Mica	9×10^{13}	0.11×10^{-13}
Molten quartz	7×10^{16}	0.14×10^{-13}

Some materials in the above table have the resistivity which lies between that of very low resistivity (conductors) and very high resistivity (insulators). Those materials are categorized as "Semiconductors".

Accordingly, materials having resistivity less than $10^{-3}\,\Omega$ m are considered as conductors, materials having resistivity greater than $10^{5}\,\Omega$ m as insulators and materials having resistivity between $10^{-3}\,\Omega$ m and $10^{5}\,\Omega$ m as semiconductors.

Table 1.2 Conductors semiconductors and insulators based on resistivity

Conductors	**Semiconductors**	**Insulators**
$\rho < 10^{-3}\,\Omega$ m	$10^{-3}\,\Omega$ m $< \rho < 10^{5}\,\Omega$ m	$\rho > 10^{5}\,\Omega$ m

These semiconductors are the materials which are important in solid - state electronics.

The cause of electrical conduction of a material is the free electrons which act as charge carriers. The electrons in atoms of the material are bound to positively charged nuclei because of the electrical attractive forces. Therefore, it obstructs to the mobility of electrons. Because of the nature of bonds among atoms of the material the electrons can get opportunities to become free. The bonds in the atoms of solid materials can be divided into three major types.

(i) metalic bonds (ii) ionic bonds (iii) covalent bond

The bonds in the metal elements are known as metalic bonds.

It has been found practically that the distance between two nuclei atoms in the atomic lattice is less than twice the radius of the metal atom. So, it is clear that the outermost electron orbits (electron clouds) of the atoms overlap.

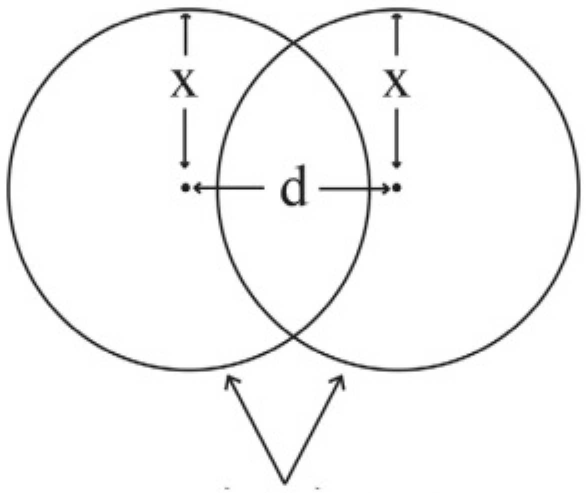


Figure 1.1

The valence electrons in the outermost orbit (valence orbit) are of less bound to the nucleus. Therefore, those electrons can move at random in the overlapped electron cloud. The atomic density of the lattice of a metal is very high. For example, the atomic density of copper is of the order 10^{29} m^{-3} while the free electron density of copper at room temperature (300 K) is about 8×10^{28} m^{-3} . Likewise almost every metal has a huge number of free electrons and those free electrons will lead to the high conductivity of metals.

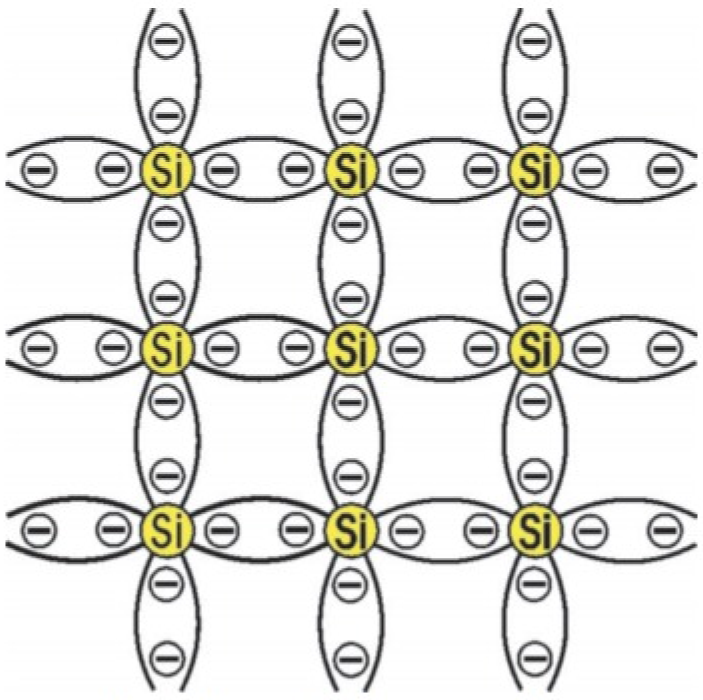
The electrical conductivity of materials such as sodium chloride (NaCl) is determined by the ionic bonds among their atoms. In sodium chloride a valence electron in the outermost orbit of the sodium atom gets out of it and then gets into the valence orbit of the chloride atom. With this the sodium becomes positive ions (Na^+) and chloride becomes negative ions (Cl^-). The bonds in sodium chloride molecules are electrical bonds among these ions. Since these bonds are very strong the materials these do not conduct electricity in the solid state (As there are no charge carriers). Nevertheless in an ionizing solvent like water these ions are able to move as change carriers because the electrical attractive forces among ions are much decreased. This decrease is due to the arrival of water having high permitting, among the ions. Aqueous sodium chloride has become a good conductor because of these ions. However in the dry solid state ionic materials behave as insulators because they don't have free ions. When these materials get molten as heated to a high temperature they conduct electricity. There also it should be stated that the conduction occurs due to positive and negative ions (charge carriers) and not due to free electrons.

© 2020 National Institute of Education. All Rights Reserved.
© 2020 National Institute of Education. All Rights Reserved.

Except for metals and salts other non - metallic materials have covalent bonds. In covalent bonds atoms of the material keep electrons in their valence orbits in common with neighboring atoms. By doing so, they can become stable atoms keeping the number of electrons in their valence orbits as eight. In many materials those covalent bonds are much stronger and therefore electrons with those covalent bonds are much stronger and therefore electrons can't move freely. Only a very few free electrons are formed by breaking a very few of those bonds. Insulators are the materials which have strong covalent bonds.

In semiconductors also, the atoms are bound due to covalent bonds. However those covalent bonds are not as strong as the bonds in insulators. Therefore semiconductors can produce free electrons by breaking their bonds even with the absorption of low amount of energy.

1.3 Intrinsic semiconductors

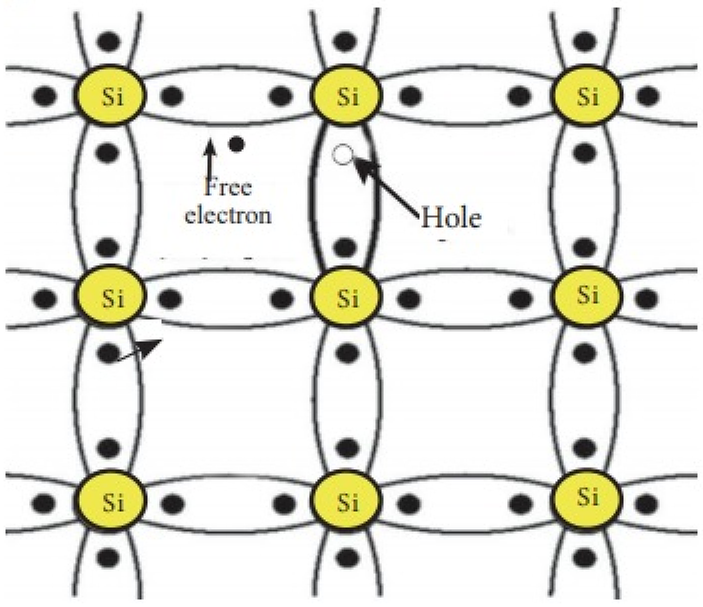


Single-crystal semiconductor
Figure 1.2

Elements in group iv of the periodic table have four (valence) electrons in the outermost orbit. When forming a crystal lattice each atom shares those four electrons with the neighboring four atoms and gets the valence orbit completed with eight electrons. Figure 1.2 shows how the lattice is formed. Although the crystal lattice is in the form of a three dimensional tetrahedron structure, here it has been used a plane structure to explain its properties.

When a semiconductor is in a pure state it is called an **intrinsic semiconductor.** Unlike insulators, semiconductors have weaker covalent bonds and hence some bonds get broken even at room temperature (due to heat) producing free electrons. From each bond thus broken there adds a **hole** and a **free electron** to the lattice.

The free electors created by the broken bonds can recombine with holes and make bonds again. These electron - hole pair generation and recombination processes come to a dynamic equilibrium, corresponding to the temperature of the material. Therefore, a particular number of holes and free electrons remain in the lattice.



The formation of holes and free - electrons is shown in Figure 1.3.

In a conductor, electrical is conducted by free electrons only. However in semiconductors both holes and free electrons contribute towards the conduction of electricity. Therefore, these holes and free - electrons are called "carriers" in general.

Figure 1.3

The concentration of carriers of the above mentioned intrinsic semiconductors at room temperature (300 K) are given below.

Table 1.3 Concentration of carries of Ge and Si

Semiconductors	Intrinsic carrier conductors
Ge - Germanium	2.5×10^{13} cm^{-3}
Si - Silicon	1.5×10^{10} cm^{-3}

When the temperature rises, above concentrations get changed considerably as a result of further breaking of bonds in the lattice. As the binding energy is lower in germanium, its carrier concentration increase heavily with the temperature rise. The effect of temperature on silicon is lower than on germanium. The breaking up of bonds with the temperature rise is called the "thermal agitation".

1.4 Mechanism of flow of current in semiconductors

We know that there are some number of free electrons and an equal number of holes occur in a semiconductor, because of thermal agitation. This process is known as **electron - hole pair generation.** Also, some of these free electrons, in recombination with holes, make bonds. At the end of these free electrons, in recombination with holes, make bonds. At the end of these two processes a dynamic equilibrium is reached, so that there are certain number of free electrons and an equal number of holes in the crystal lattice. Let us denote the concentration of free electrons by n_e and the concentration of holes by n_h , at some temperature. These two concentrations are always equal $(n_e = n_h)$. What is known as carrier concentration of the semiconductor is the concentration of free electrons or holes thus created. Let us denote the carrier concentration by n_i. Then,

$$\therefore \quad n_i = n_e = n_h$$

When a voltage is applied across an intrinsic semiconductor an electric field (E) creates in it. The free electrons move with a drift velocity (v) against the electric field because of the force exerted by the filed. /therefore there is a flow of current in the direction opposite to that of the flow of electrons. Let us denote this current by I_e. Since the holes in the semiconductor behave as if they are positive charges, there is a hole current in the direction of the field. The way the holes move in the semiconductor lattice is shown in Figure 1.4.

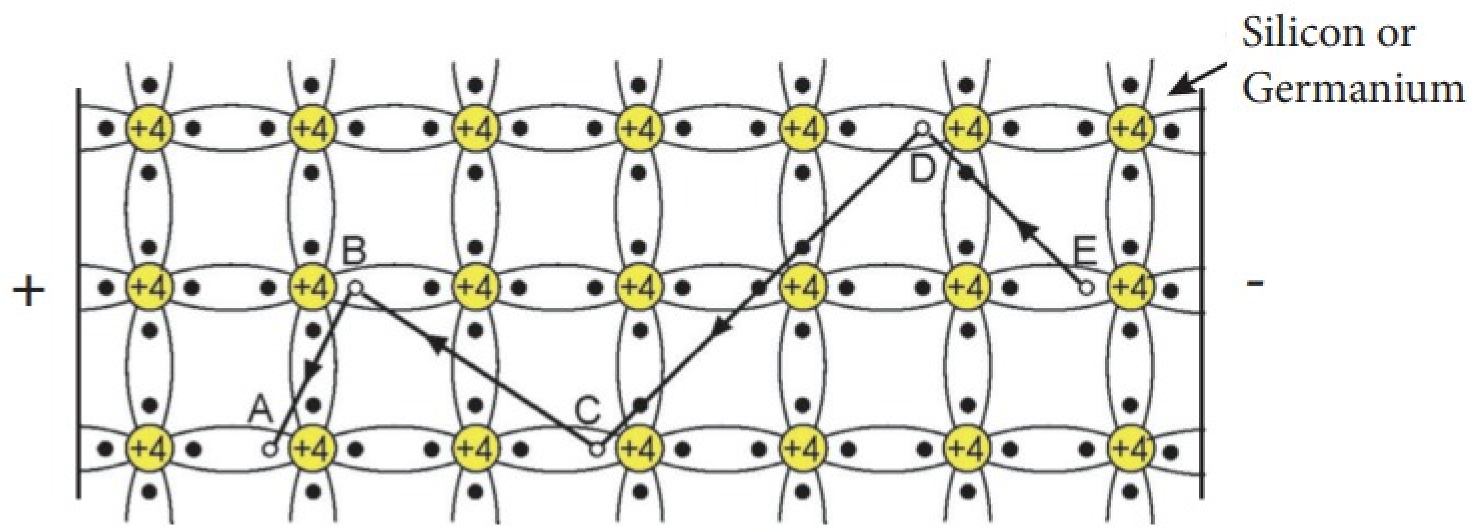


Figure 1.4
——— Motion of holes ———>
<——— Motion of electrons ———

The drift velocity of holes (v') is less than that of the electrons (v), for the reason that the motion of holes occurs in exchange with the bonds. Let us denote the hole current by I_h. The total current through the semiconductor is the sum of these two currents I_e and I_h. If we denote the total current by I_t then,

$$\boxed{I_i = I_e + I_h}$$

$$\text{Also, } v > v' \text{ as } I_e > I_h$$

Some important properties of semiconductors are given below.

(a) When the temperature of an intrinsic semiconductor is increased the intrinsic carrier concentration of it (n_i) also increases, as more bonds are broken. Therefore the resistivity of it decreases and hence the temperature coefficient of resistance of the semiconductor is negative (-).

(b) When an electromagnetic wave with an appropriate wavelength is incident on a semiconductor, it absorbs energy from the wave and as a result some bonds are broken. Therefore electron - hole pair generation occurs and the resistivity of the semiconductor decreases.

(c) When a very small amount of an element in the group III or V of the periodic table (having 3 or 5 valence electrons) is added to an intrinsic semiconductor lattice its carrier concentration increases very heavily. Therefore its resistivity decreases very much. This process of adding the element is known as **doping**. The element added is known as the **impurity element**. This will be discussed further under extrinsic semiconductors.

1.5 Extrinsic semiconductors

If a very small amount of an element in the group iii or v of the periodic table is added to an intrinsic semiconductor, its electrical conductivity increases very largely. This process is known as **"doping"**. The added element is called the **"impurity element"**. When an intrinsic semiconductor is doped with an impurity element, that semiconductor is now called an "extrinsic semiconductor".

Depending on whether the impurity element used for doping is in the group III (having three valence electrons) or in the group V (having five valance electrons) the extrinsic semiconductors are divided into two types.

(a) n – Type semiconductors
(b) p – Type semiconductors

1.5.1 n – Type semiconductors

Let us consider how the bonds are formed in the crystal lattice of an intrinsic semiconductor when a very little amount of one of the elements phosphorus (P), arsenic (As), antimony (Sb) or bismuth (Bi) which are in the group V, is added for an example when an arsenic atom is added to the lattice of a silicon intrinsic semiconductor, the arsenic atom makes bonds with the neighboring valence electron of the arsenic atom remains in the lattice without taking part in any bond. In this manner the arsenic atom becomes stable by completing the eight electrons in its valence orbit. The electron that the lattice received from the arsenic atom without participating to make any bond remains in the lattice as a free electron.

Since the electrons thus formed can move freely in the lattice, the conductivity of the lattice increases largely. Figure 1.5 shows how the bonds are formed in such an extrinsic silicon lattice.

Since this extra electron or 'carrier' can contribute to conduction, has a negative charge, this type of extrinsic semiconductors are called n - type semiconductors. Here, the arsenic atom is taken as the impurity, and since arsenic donates 'carriers' to the lattice, the arsenic atoms are known as 'donor atoms'. Every donor atom provides one carrier each, to the lattice.

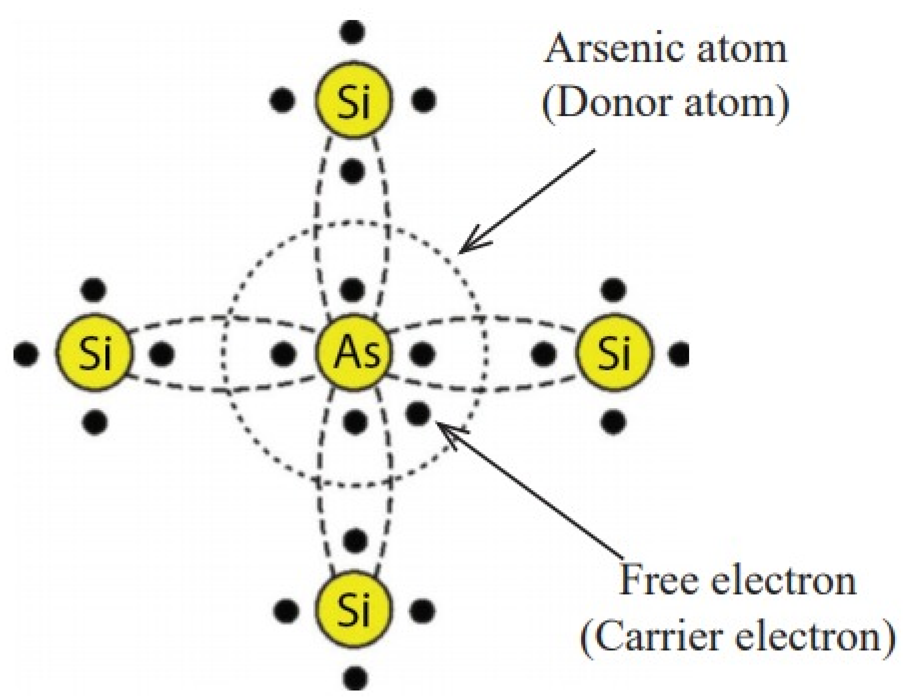


Figure 1.5
Entering an arsenic atom into silicon crystal lattice

The concentration of extra free electrons donated by the donor atoms (N_D) under normal doping conditions is very much higher than the concentration of free electrons, in the intrinsic semiconductor (n_e) due to thermal agitation.

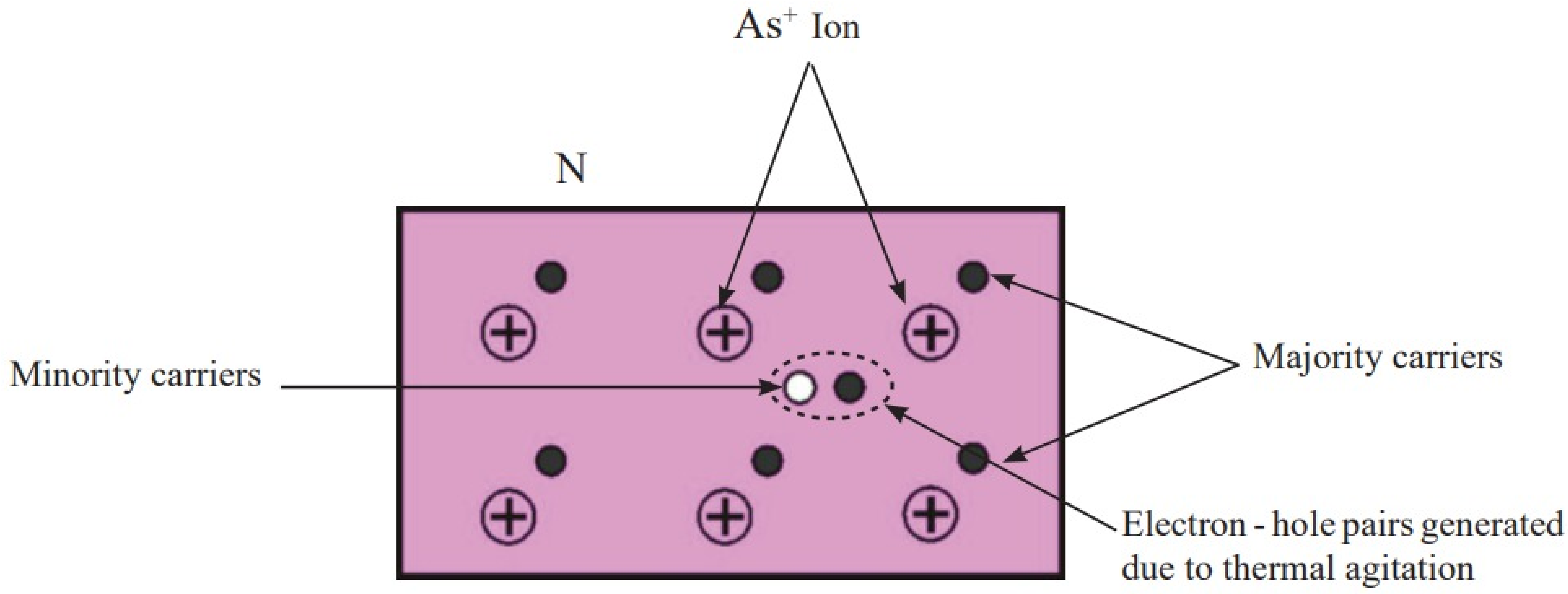
Now, since $N_D >>> n_e$ the total concentration of free electrons is taken to be,

$$N_D + n_e \simeq N_D$$

The concentration of holes in this n - type extrinsic semiconductor is nearly equal to the intrinsic concentration of holes n_h which is also equal to n_e.

So, this lattice of n - type extrinsic semiconductor has a very low concentration of holes compared its the concentration of free electrons. Therefore, 'free electrons' are know as the **majority carriers** and 'holes' as the **minority carriers** in an n - type extrinsic semiconductor.

Although, both these, the majority carriers and the minority carriers contribute towards the conduction of electricity, since only a few minority carriers are there, the contribution given by the minority carriers can be neglected. (Here, the as contributing arsenic atom acts as an As^+ ion)



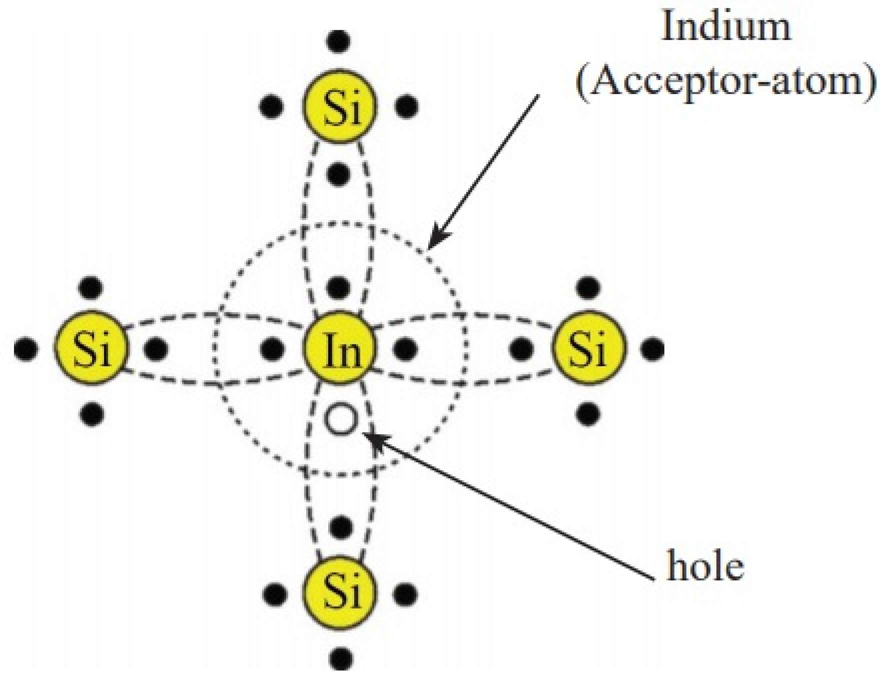
Representation of n - type extrinsic semiconductor
Figure 1.6

The total number of protons in extrinsic semiconductors, and the total number of electrons in the lattice are equal. Therefore the lattice is electrically neutral. Arsenic atoms from which the electrons have removed are in the lattice as positive ions. (Ions oscillating only about their center positions only dependent on the temperature of the lattice).

1.5.2 p – Type Semiconductors

If an intrinsic silicon (Si) lattice is doped with an impurity element having three valence electrons and in group III of the periodic table, then each impurity atom lacks an electron in making bonds. (to get its outermost orbit completed with eight electrons).

© 2020 National Institute of Education. All Rights Reserved.



Entering an Indium atom into silicon crystle lattice
Figure 1.7

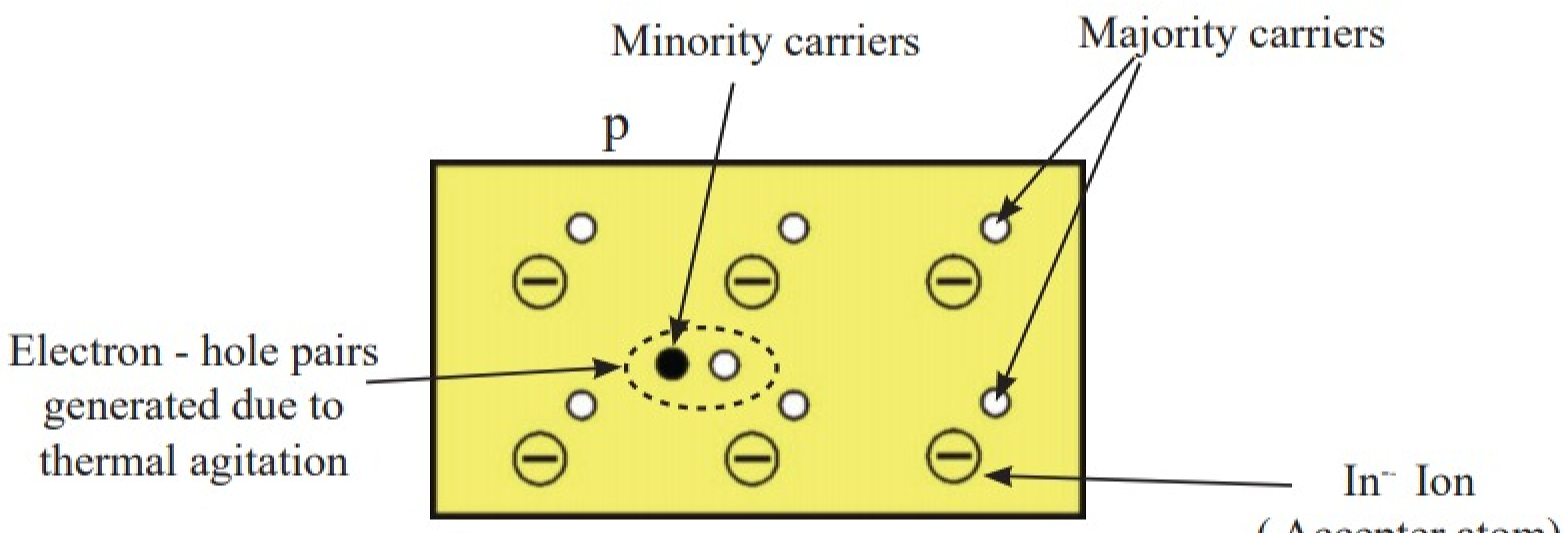
Elements in group III

Aluminium	- Al
Gallium	- Ga
Boron	- B
Indium	- In

When a silicon lattice is doped with indium as the impurity, the formation of bonds in the lattice is shown in Figure 1.7. One bond of each indium atom lacks an electron and the place where there is a vacancy for an electron remain as a hole.

Therefore, from each indium atom added to the lattice there is created a hole (carrier) in it. For the reason that they can accept electrons in this situation, the indium atoms are called acceptor atoms. The holes created in this process behave as if they were positive charges. Therefore this type of extrinsic semiconductors are called p - type semiconductors. Here the indium acceptor atoms act as In- ions.

With a similar type of explanation it can be shown that a p - type semiconductor gets a very high concentration of holes due to the acceptor atoms. The concentration of free electrons in it is relatively very much low. Therefore, the holes in p - type extrinsic semiconductor are identified as majority carriers and 'free electrons' as minority carriers.



Representation of p - type extrinsic semiconductor
Figure 1.8

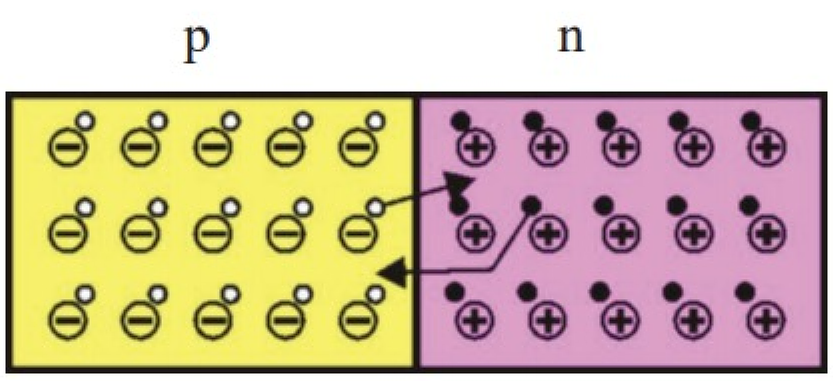
General properties of extrinsic semiconductors

- In extrinsic semiconductors, when the doping level is increased as the majority carrier concentration increases conductivity increases (resistivity decreases).
- When the temperature of an extrinsic semiconductor is increased its minority carrier concentration increases, as more bonds are broken. However the increase of conductivity due to this is little. (The percentage increase of majority carriers due to doping process is very much greater than that of minority carriers due to breaking of bonds).

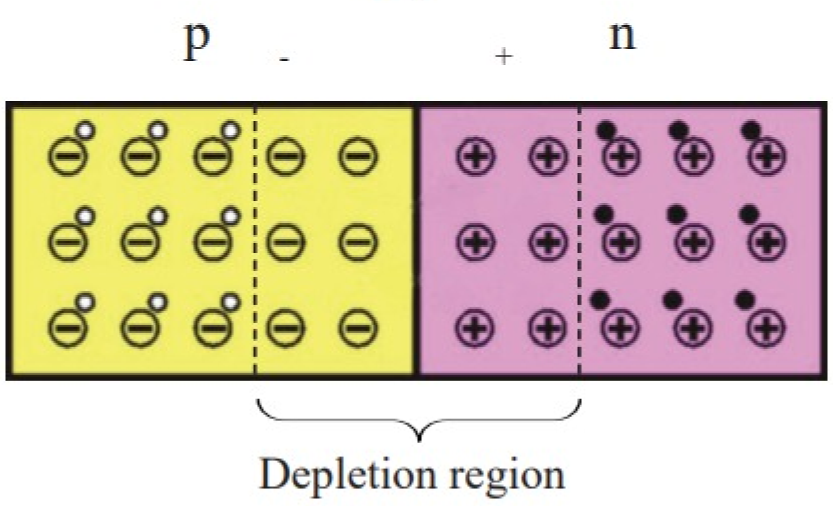
Action of p – n junction

When a junction is made using two extrinsic semiconductors, one is of p - type and the other is of n - type, it is known as a p-n junction. This type of a junction cannot be made by keeping the p - type and n - type semiconductors in contact or by soldering them. A special doping process should be followed to make a p - n junction by using the atoms of group (III) and group (V) on the two sides of a piece of intrinsic semiconductor.

As soon as the p - n junction is formed, the majority carriers in the neighborhood of the junction tend to recombine with holes and free electrons in the lattice of the opposite sides. /holes (majority carriers) in the p - type semiconductor diffuse to the n - type semiconductor across the junction and free electrons (majority carriers) in the n - type semiconductor diffuse to the p - type semiconductor leading to recombination of holes and the electrons. This process is explained by the following Figures 1.9 (a) and 1.9 (b).



(a)



(b)

Figure 1.9

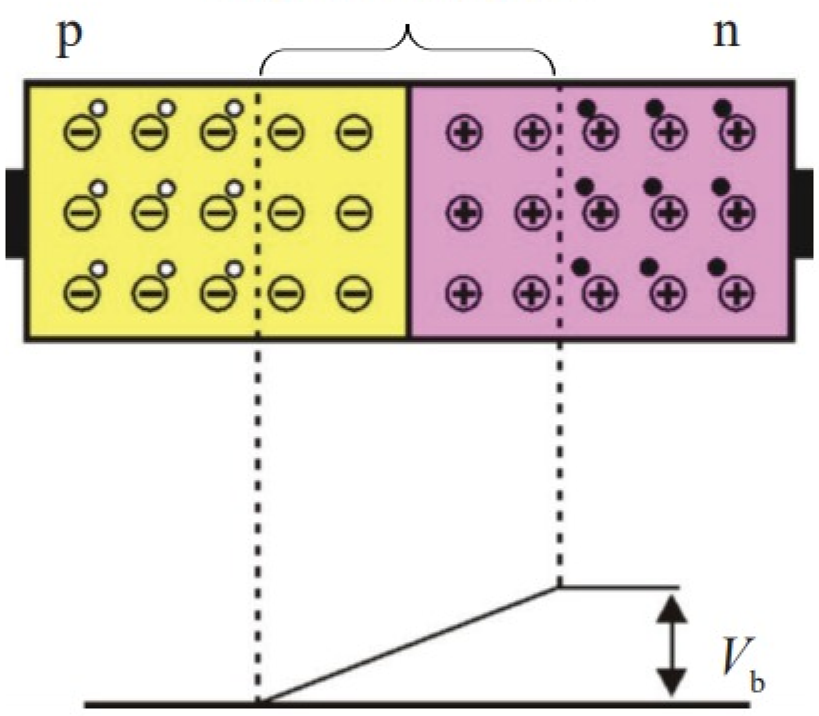
Because of the diffusion of majority carriers across the junction the neutral atoms in the neighborhood of junction become ions.

Because of the arrival of extra electrons to the p - region it gets a negative charge. Due to departure of electrons from the n - region it gets a positive charge. This process occurs instantly and in the end it will come to an equilibrium because the positive charge developed in the n - region opposes further arrival of electrons and the negative charge in the p - region opposes further arrival of holes. In this equilibrium p - type semiconductor gains a negative potential and n - type semiconductor gains a positive potential. This potential difference prevailing across the junction is known as the **"internal potential barrier"**. This is shown by Figure 1.10. On this occation there is a region without charge carriers in p - type and n - type semiconductors close to the junction. This region is known as the **attenuated region** or **depletion layer**. The above mentioned internal potential barrier (V_b) can be regarded as having an imaginary cell in the junction. The depletion layer is an above as thin as 10^{-6} m. The thickness of this layer depends on the level of doping. Rough value of internal potential barriers in different extrinsic semiconductors at room temperature is given below.

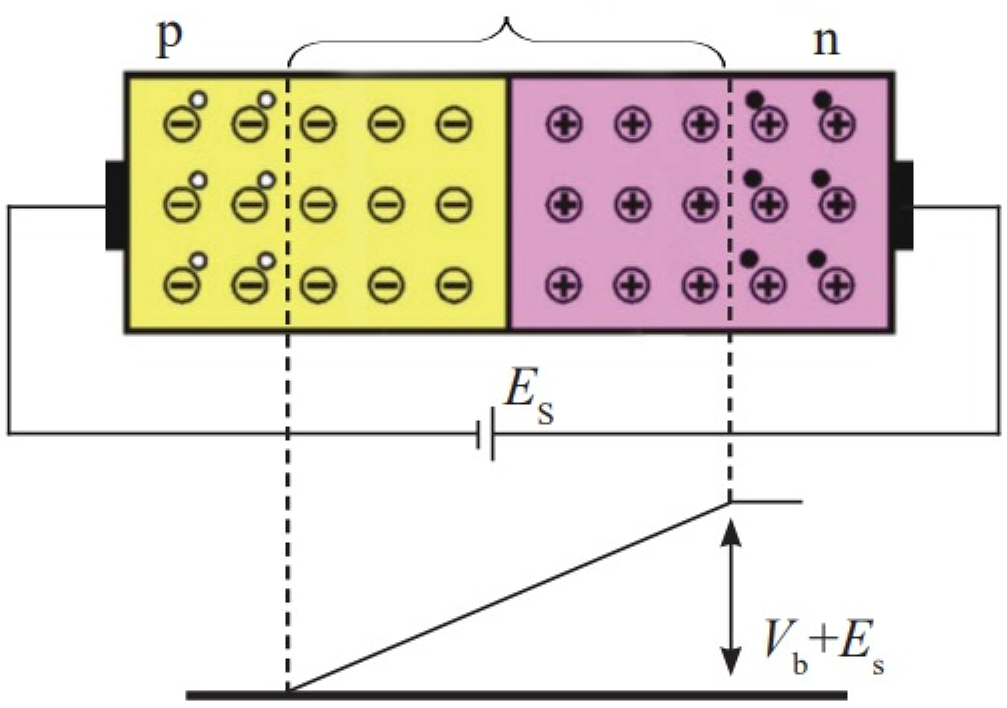
Silicon	0.7 V (0.6 ~ 0.7 V)
Germanium	0.3 V (0.2 ~ 0.3 V)

1.7 Reverse biasing p - n junction

Let us consider the action of a p - n junction when a potential difference is applied across it so that the negative terminal is connected to p - region of the junction.



(a)



(b)

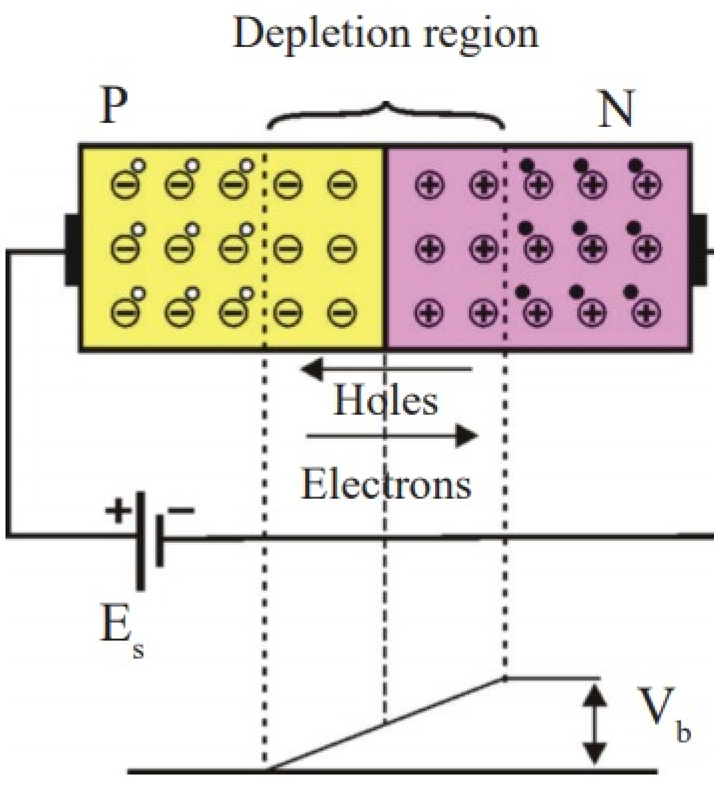
Figure 1.10

Because of the external potential difference (E_s), holes (behaving as positive charges) in the p - region get diffused towards the negative terminal and free electrons in the n - region get diffused towards the positive terminal of the external potential difference, and hence the depletion layer widens further. Therefore, internal potential barrier increases up to $V_b + E_s$ and the junction comes to an equilibrium again. As there is no flow of charges across the junction in this process, the flow of current also does not take place. Supplying a biasing potential difference to a p - n junction in this manner is known as reverse biasing the junction.

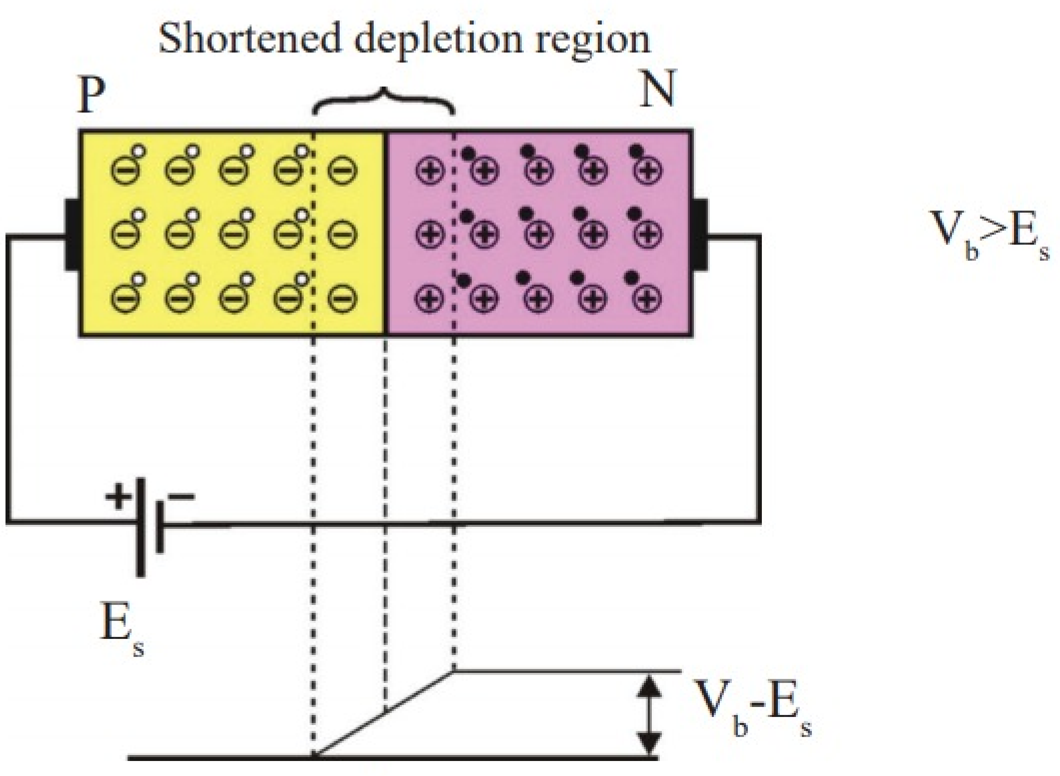
Although there is an internal potential difference for majority carriers, it is not a barrier for minority carriers. Therefore, electrons (minority carriers) in the p - region and holes (minority carriers) in the n - region) flow across the junction leading to a small current in the biasing circuit. This small current is known as "Saturation leakage current" and it depends on the rate at which the minority carriers are generated by thermal agitation. The charge generated per unit time due to temperature are flowing across the junction current. This leakage current (I_s) increases slightly with the temperature rise. Since this is a very small current of the order μA, in many cases the leakage current is ignored.

1.8 Forward biasing p – n junction

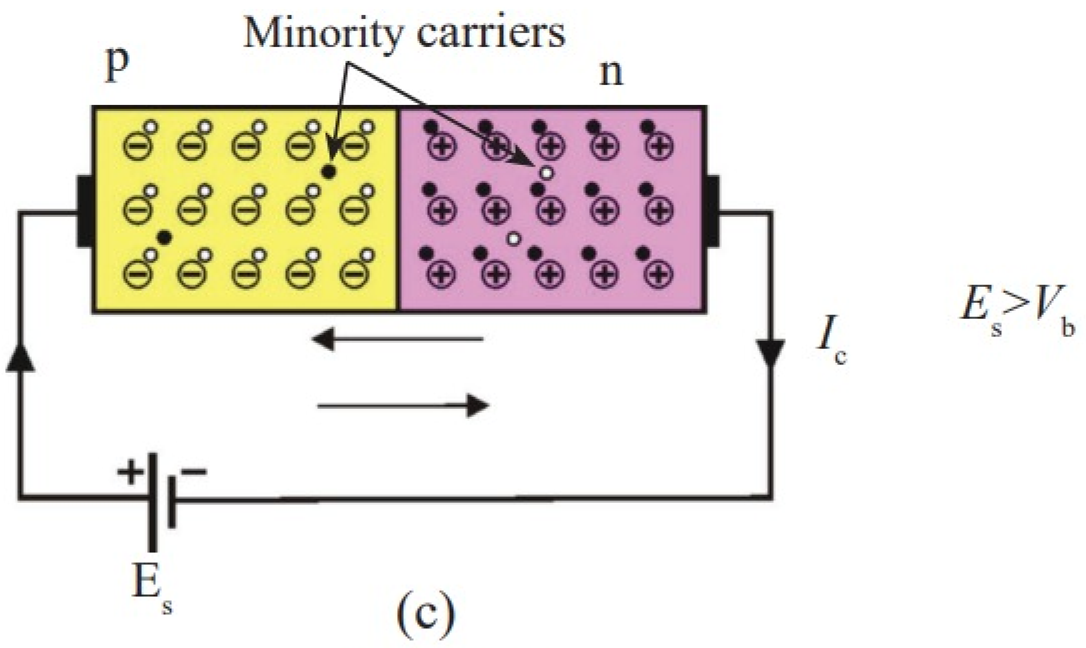
To make a p - n junction forward biased, an external potential difference (E_s) should be connected so that the p - region gets the positive potential and n - region gets the negative potential. Then, holes in p - region and free electrons in n - region are pushed towards the junction and hence the depletion layer gets narrower. If the internal potential difference (E_s) is greater than the internal potential barrier (V_b) then the charge carriers flow in the biasing circuit. The actions taken place in this process are shown in Figures 1.11 (a), (b) and (c).



(a)



(b)



(c)

Figure 1.11

On this occasion, the majority carriers are contributing towards the flow of current. Therefore, this biasing is known as the forward biasing of the p - n junction. Since the junction is reverse biased for minority carriers there is no current due to minority carriers.

Unlike a normal metal junction, p - n junction allows a current to flow when it is forward biased and does not allow a current to flow when it is reverse biased. Therefore, p - n junction acts as a **"valve"**.

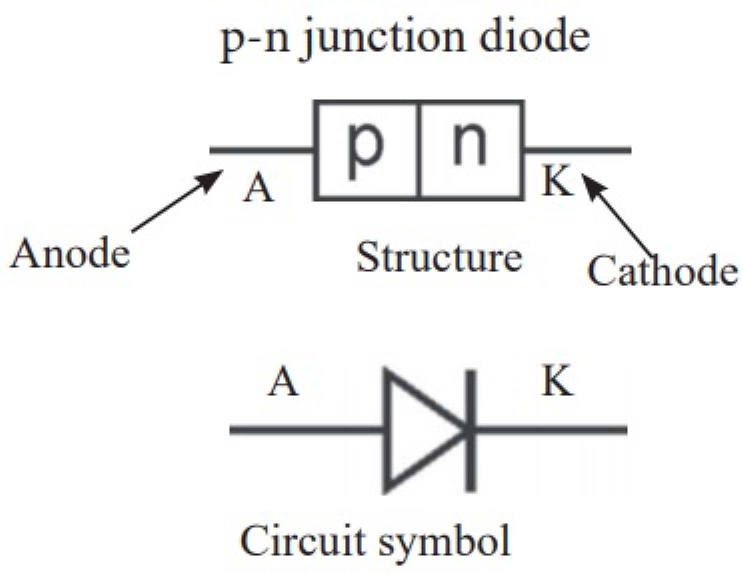


Figure 1.12

The way the p and n semiconductors are arranged in a diode and the circuit symbol of it are shown in Figure 1.12. The direction of flow of current (forward current) in the forward bias is indicated by the arrowhead in the circuit symbol of the diode. For the current to flow through it, the diode should be forward biased by applying a positive potential to p - region and negative potential to n - region. The terminal connected to p - region is named as anode (A) and that connected to n - region is named as cathode (K). Current through the external circuit flows from cathode to anode while that through the diode flows from anode to cathode. In general, the diodes made using p - n junctions are called junction diodes.

1.9 V - I characteristic curve of a p - n junction diode

Let us now consider the variation of current though a p - n junction diode under forward bias and reverse bias. We should use two circuits for this. The variable resistor (VR) shown in Figure 1.13 (a) acts as a potential divider. By adjusting VR, the forward bias voltage can be varied appropriately. Consequently the forward current can be measured using the milli - ammeter.

© 2020 National Institute of Education. All Rights Reserved.

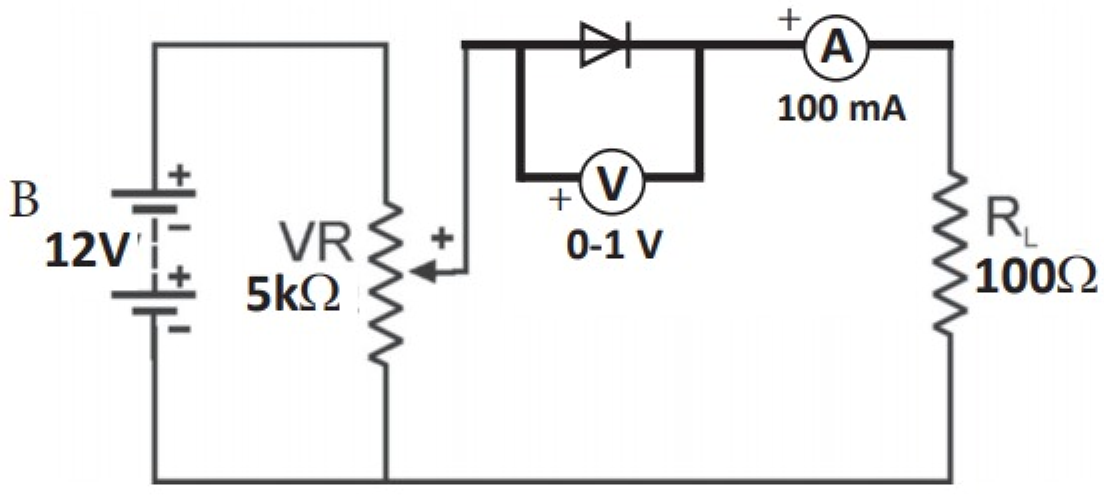


Figure 1.13 (a) Forward bias circuit

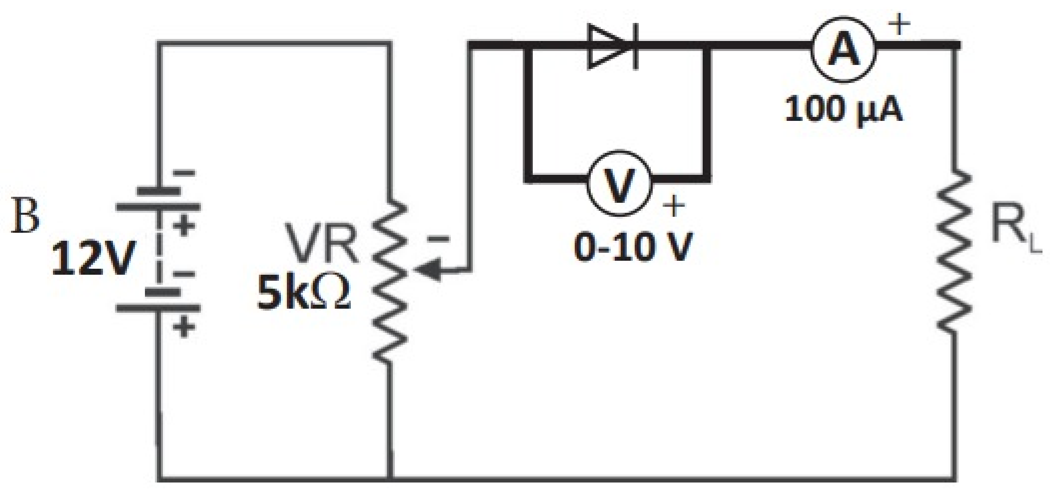


Figure 1.13 (b) Reverse bias circuit

The circuit shown in Figure 1.13 (a) is used to study reverse bass. In that circuit, the terminal connections of cell B have been interchanged and accordingly the terminal connections of voltmeter and ammeter are also interchanged. Under reverse bias the reverse current is of the order μA. Therefore, a micro - ammeter in the range 0 - 100 μA is used the measure it.

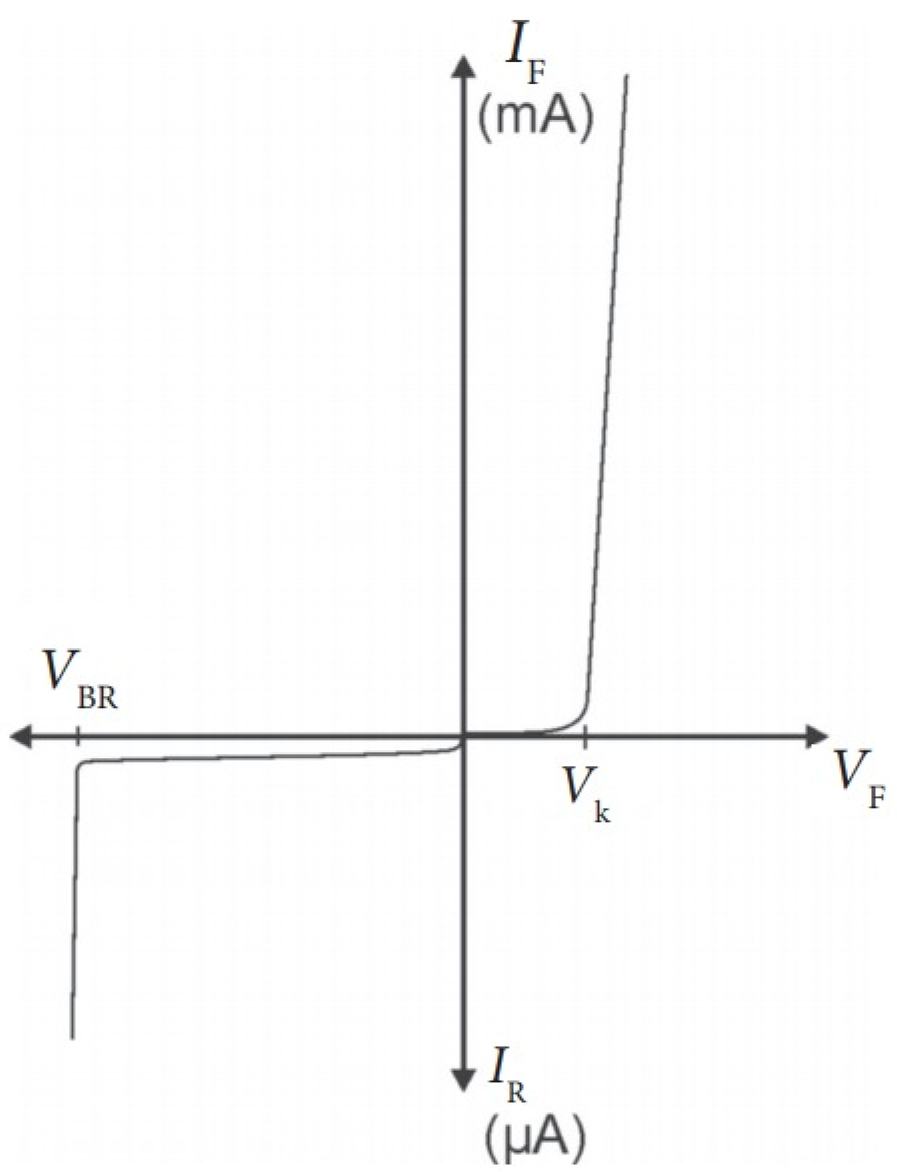


Figure 1.13 (c)

A roughly plotted *V* - *I* graph based on readings obtained in an experiment as described above, is shown in Figure 1.13 (c). It can be seen from the graph that, until a certain voltage is reached there is no measurable current and afterwards the current begins to increase very sensitively even for little increase of voltage.

Let us consider the special features regarding the curve, separately.

Cut-in Voltage or Knee Voltage - V_k)

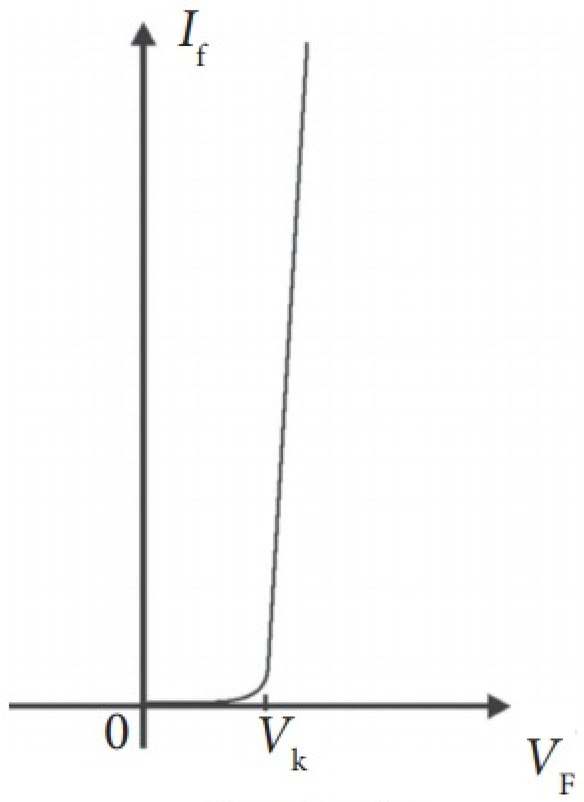


Figure 1.14

The voltage at which the forward current starts to increase rapidly when the forward bias voltage is increased, is known as the cutin voltage or knee voltage (V_k). It is the voltage at the point where the extrapolated linear part of the graph cuts the voltage axis (marked as V_k). The reason for this behaviour of the diode is the potential barrier of the p - n junction. When the potential barrier is overcome the forward current (I_c) varies almost linearly with the forward bias voltage (V_f). This potential barrier is about 0.2 V for germanium diodes whereas it is about 1.2 V for gallium arsenide diodes.

Resistance of a diode

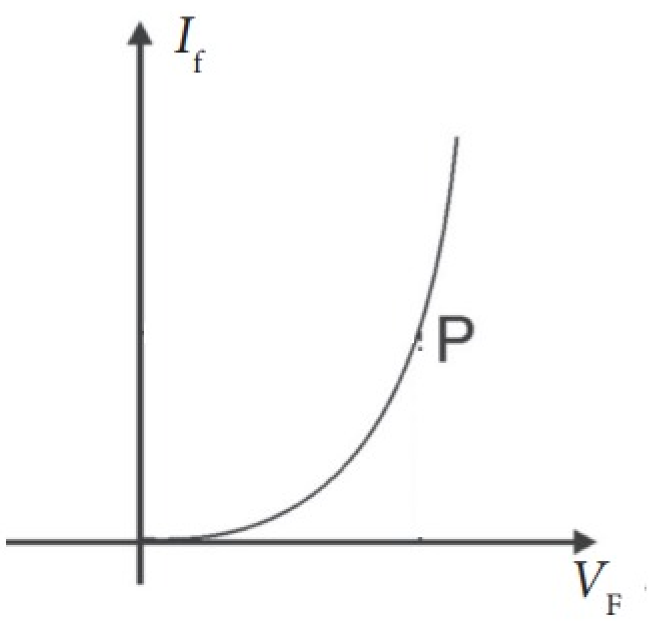


Figure 1.15

It can be seen from the graph in Figure 1.15 that the diode does not behave as an ohmic conductor. The resistance of the diode differs at different forward bias voltages. The diode has a very high resistance until it reaches the cut - in voltage and at voltages higher than that, the forward current varies linearly (roughly) with the forward bias voltage. On this occasion, the diode shows a low resistance.

1.9.2 Reverse Saturation Current (I_s)

It is clear from the characteristic curve, when the diode is reverse biased until the reverse bias voltage reaches a particular high voltage there is only a very small reverse current flowing. In the case of a germanium diode this current is of the order of μA and for a silicon diode it is of the order of pA.

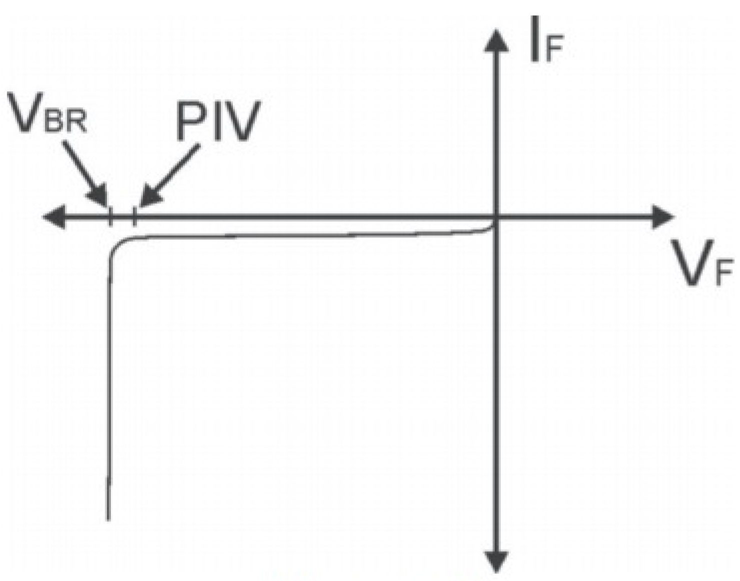


Figure 1.16

Reverse current flows because of the minority carriers created by thermal agitation. For these minority carriers the external reverse bias voltage acts as a forward bias voltage. As the reverse voltage is high, all the minority carriers created by thermal agitation contributes to the flow of current. Therefore, this is a saturation current and is known as **reverse saturation current** or **leakage current**.

The number of minority carriers created within unit time is constant at constant temperature. Therefore, the reverse saturation current is also constant.

1.9.3 Breakdown Voltage (V_{BR})

When the reverse bias voltage is increased the electric field across the p - n junction increases. Then there will be a large force on minority carriers due to this electric field (according to $F = Eq$). This force makes the minority carriers accelerated and hence their kinetic energy ($\frac{1}{2} mv^2$) increases. The carriers having this kinetic energy can break bonds in the lattice by colliding with them. This becomes a chain reaction and then a large number of minority carriers are generated instantly making a large reverse bias current.

Along with this due to heat emitted in the collisions the junction is subject to thermal breakdown. Then the p - n junction gets damaged and cannot be used again. In practical usage the diode should be used in such a way that the reverse bias voltage does not reach the breakdown voltage. This process which makes the diode breakdown, is known as avalanche breakdown.

1.9.4 Peak Inverse Voltage (PIV)

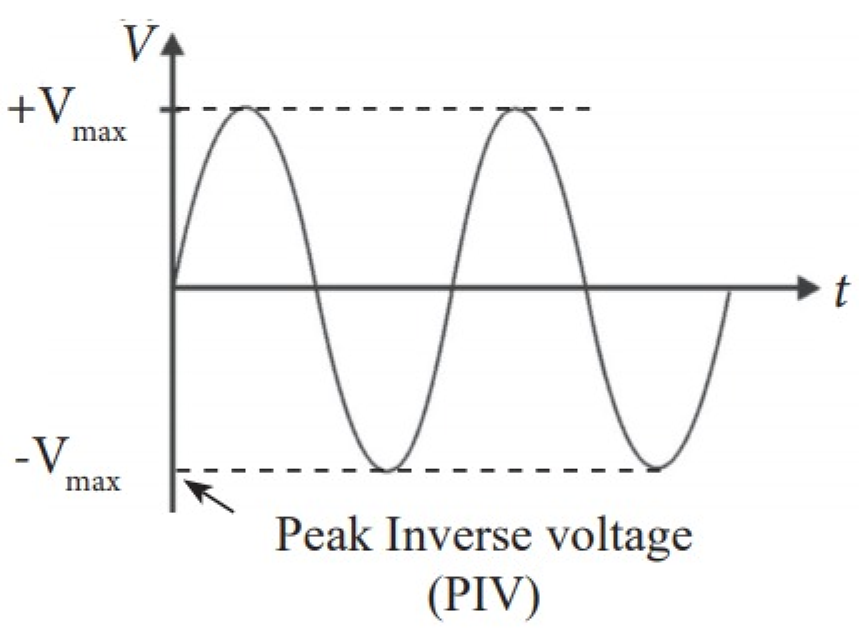


Figure 1.17

When an alternating voltage is supplied to a diode, the peak inverse voltage of that supply acts as a reverse bias voltage across the diode. Therefore, in the diodes when used with alternating voltages, the magnitude of breakdown voltage of the diode should be greater than the magnitude of peak inverse value of the alternating voltage used. Therefore, in data books for diodes, it is given the maximum value of peak inverse voltage (PIV) that the diode can be used safely. (This data is given in addition to the breakdown voltage). The voltage value given under this PIV is sufficiently lower than the V_{BR} voltage value and, the data which is more important in practical usage of diodes is this PIV value.

© 2020 National Institute of Education. All Rights Reserved.

For supplementary knowledge

Characteristic curves of silicon, germanium and gallium arsenide diodes.

Shown below are the characteristic curves of several diodes.

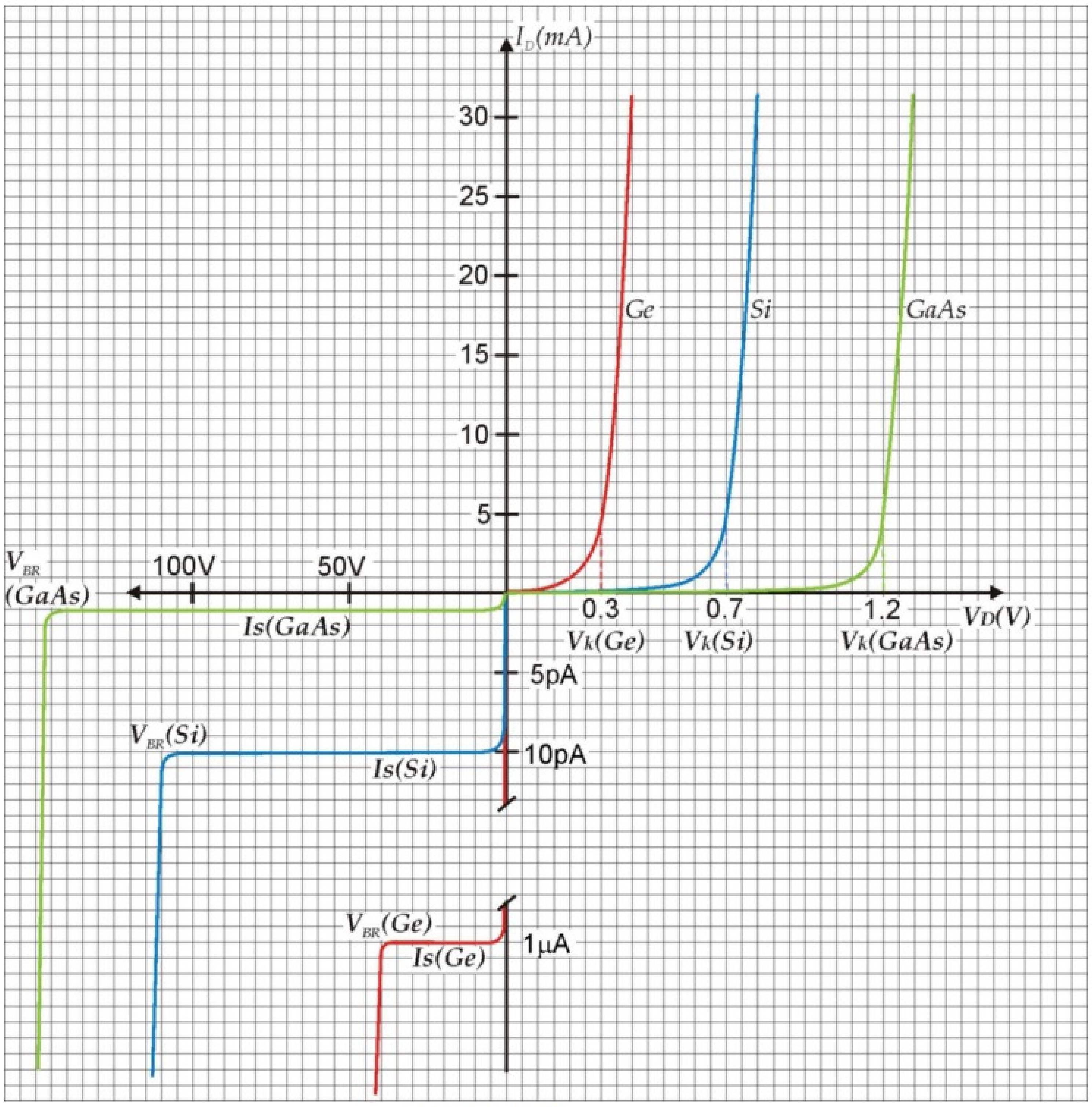


Figure 1.18

Accordingly, of the curves it can be seen that germanium diode has the lowest and gallium arsenide diode has the highest value of cut - in voltage.

Diode	V_k (V)
Ge	0.3
Si	0.7
GaAs	1.2

Also, the curves show that GaAs diode has the lowest reverse bias leakage current ($I_S \bumpeq$ 1pA) and germanium diode has the highest (1µA), and silicon diode has a value of about $I_S \bumpeq$ 10 pA. Since leakage current creates unnecessary conditions on many occasions, Si - diodes are widely used now instead of Ge - diodes. As the cost of production is high for GaAs - diodes, their usage is confined to special purposes only.

Temperature sensitivity of diodes

The graphs given below show how a silicon diode responds to temperature. We see that cut - in voltage (V_K) and breakdown voltage (V_{BR}) get changed depending on the temperature of it.

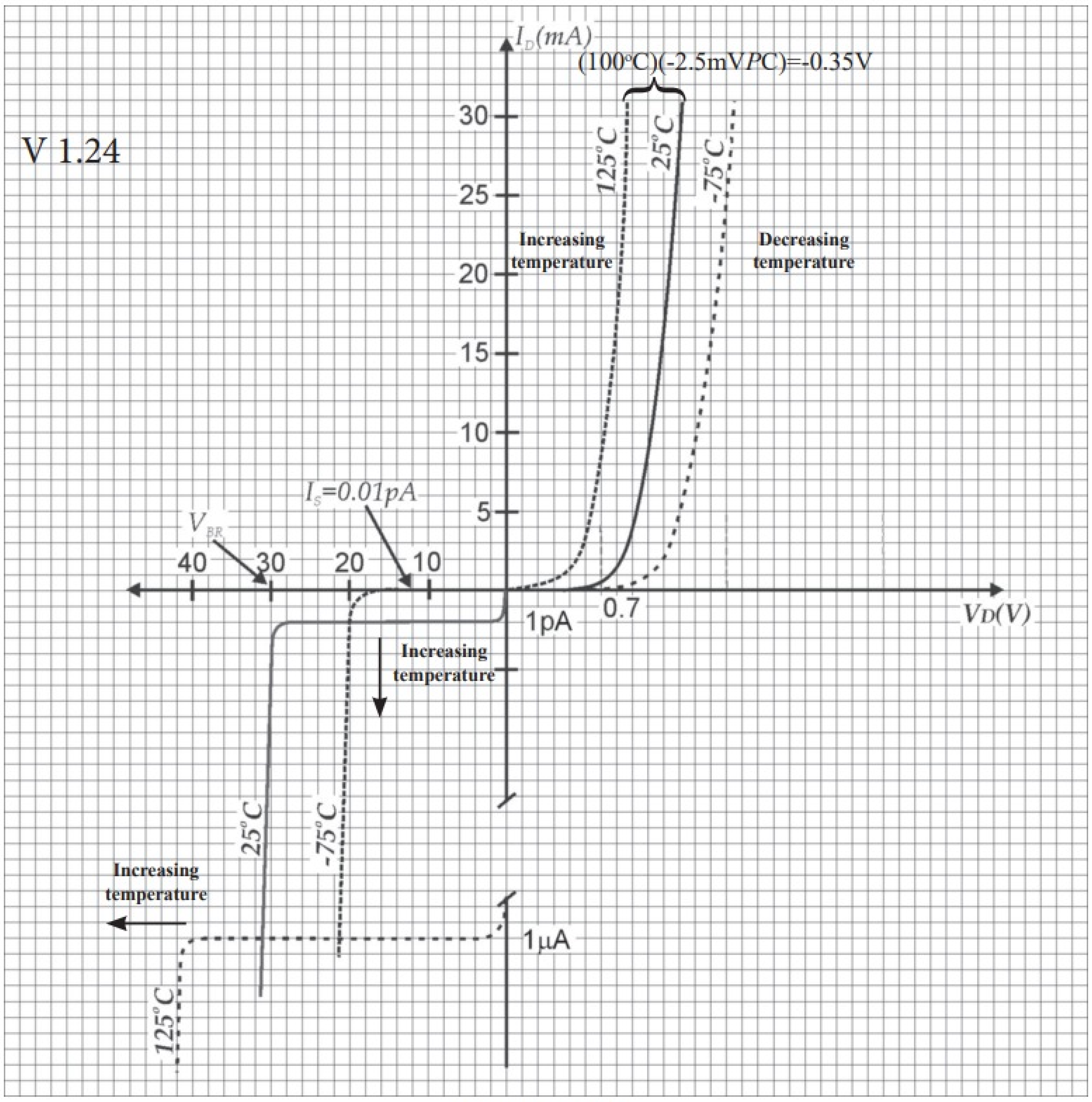


Figure 1.19

This variation can be described in short as follows.

- The cut - in voltage (V_K) decreases as temperature increases. (for all semiconductor diodes this is about 2 mV °C^{-1})
- The breakdown voltage (V_{BR}) increases with temperature rise. (for every 10 °C of temperature rise the breaking voltage increases by 1V)
- The leakage current or reverse bias saturation current (I_S) increases with temperature rise. For germanium diodes, I_S gets doubled per every temperature rise of 9 °C and for silicon diodes, per every 11°C.

  (In general it can be considered I_S gets doubled per every temperature rise of 10 °C for all diodes.)

© 2020 National Institute of Education. All Rights Reserved.

1.10 Maximum power of a diode (P_{DMax})

In the forward bias condition of a diode some amount of work is done when the forward current is flowing through the diode under a potential difference V_D. This work generates heat in the diode. Let us consider I_F as the maximum current that can be flown through the diode without damaging it due to heat. This potential difference V_D is approximately equal to V_K of the diode.

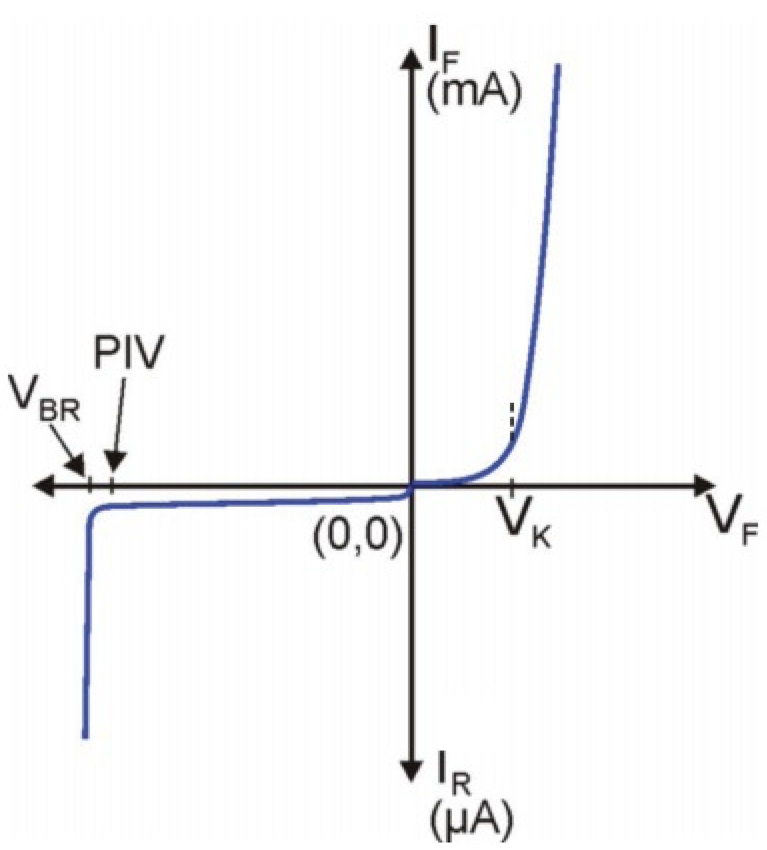
$$\therefore \; P_{DMax} = V_D . I_{FMax} = V_K . I_{FMax}$$

As there is a possibility of a breakdown due to the generation of heat, the maximum power of the diode is important. P_{DMax} can be calculated when I_F is known, taking the values $V_K = 0.7$ V for silicon diodes and $V_K = 0.3$ V for germanium diodes. If this power value is exceeded when operating, then the junction may get molten and become open or short.

Most of the diode data sheets provide the values of I_{FMax} and P_{DMax}.

1.11 Application of diodes

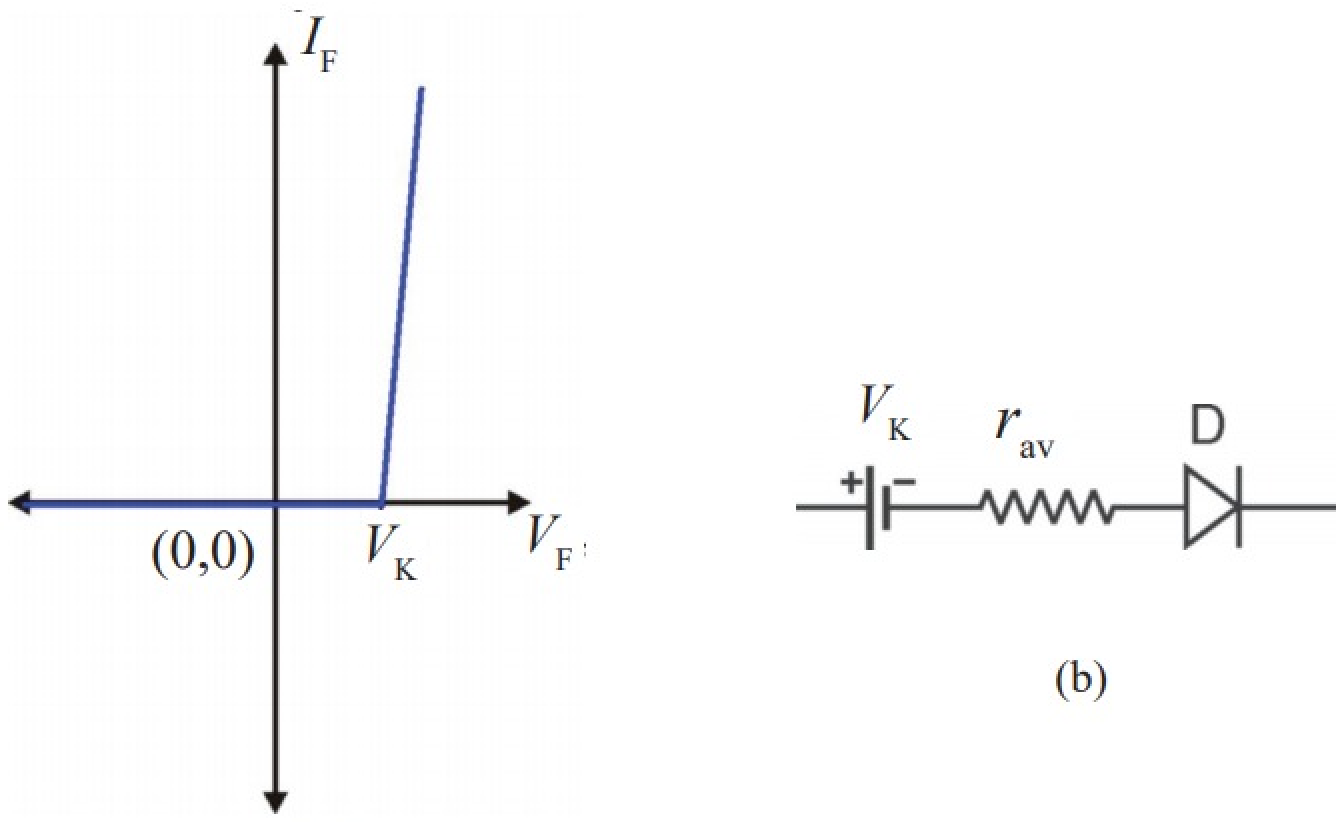
When diodes are used in circuits, the properties shown by its characteristic curve are made use of. Therefore, let us consider the characteristic curve of a practical diode.



Curve of real diode
Figure 1.20

When a diode is provided with a forward voltage until it reaches the knee (or cut - in) value the diode does not conduct a current. If the voltage is increased beyond the knee value, the forward current I_F increases rapidly. This is approximately a linear variation and the graph is not exactly a straight line (Figure 1.20). Therefore, in this region the dynamic resistance of the diode differs slightly at different points. Let us take the average dynamic resistance as r_{av}. Although a very small current (of the order pA or µA) flows through the diode when it is reverse biased, this current can be neglected and regarded as not a flowing current. In the reverse bias, until the diode reaches the breakdown voltage (V_{BR}) it does not conduct a current. When the reverse voltage value exceeds V_{BR} value, the diode undergoes a breakdown. In diode data sheets, the peak inverse voltage (PIV) that the diode can practically withstand, is given instead of V_{BR}. Therefore, in practical applications the PIV value given in data sheets is used instead of V_{BR}. Figure 1.21 (a) shows how to represent a practical diode in a linear manner for simplicity. The relevant equivalent circuit is given in Figure 1.20 (b).

© 2020 National Institute of Education. All Rights Reserved.



(a)

Figure 1.21

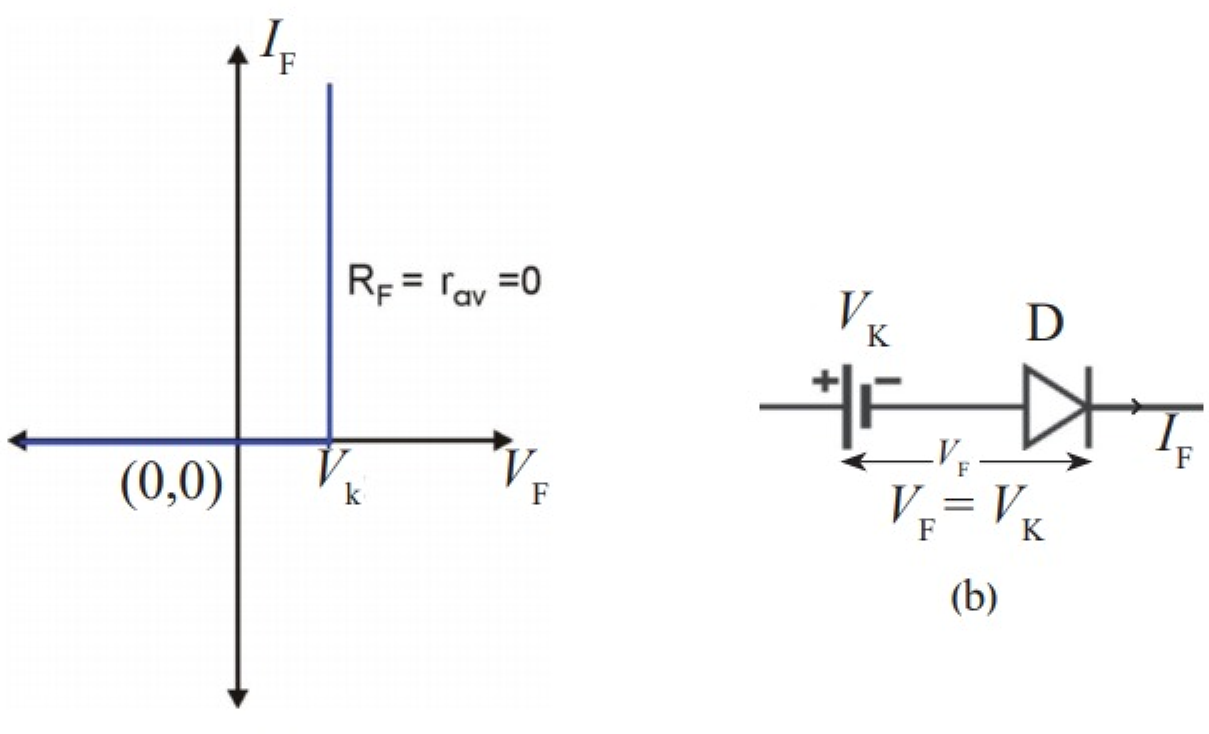
Accordingly, the relationship between the forward bias voltage (V_F) and forward bias current (I_F) can be stated as follows, (D above represents an ideal diode)

$$V_F = V_K + I_F r_{av}$$

where V_K is the knee voltage, r_{av} is the average forward bias dynamic resistance.

1.11.1 Near ideal diode

It is considered that the forward bias resistance is zero for the near ideal diode. ($R_F = r_{av} = 0$)



(a)

Figure 1.22

The characteristics curve, equivalent circuit and voltage relationship relevant to near ideal diode are shown above.

© 2020 National Institute of Education. All Rights Reserved.

Ideal diode

For the ideal diode not only the forward bias resistance (R_F) but also the potential barrier (V_K) is considered to be zero. (Figure 1.22 - page 21)

The V - I curve and the equivalent circuit corresponding by an ideal diode is shown in Figure 1.23.

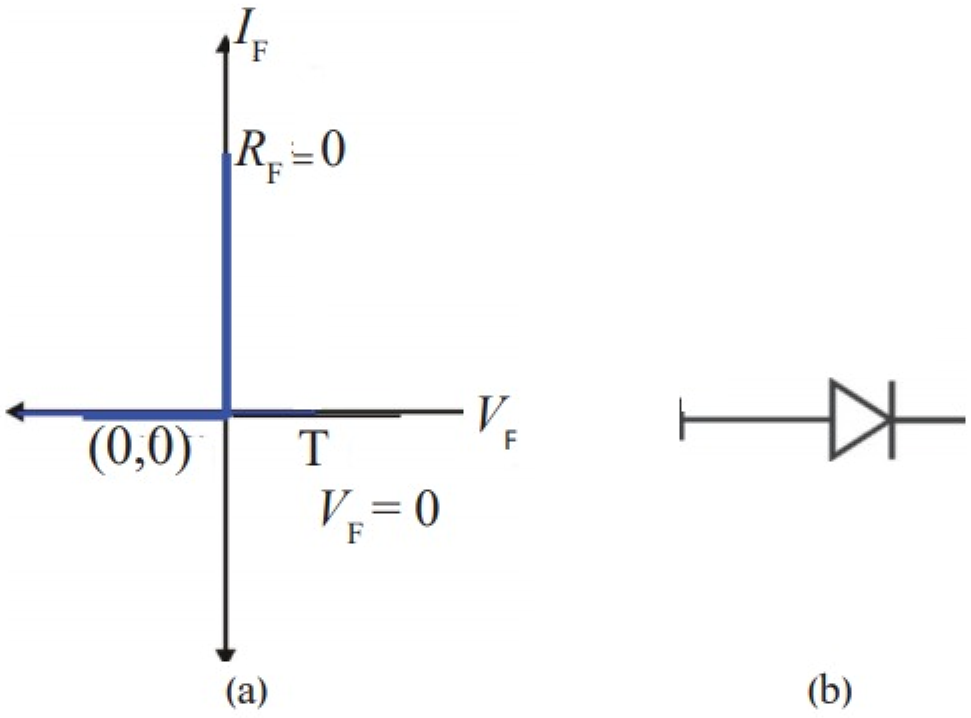


Figure 1.23

What we have discussed on practical and ideal diodes can be summarized as follows.

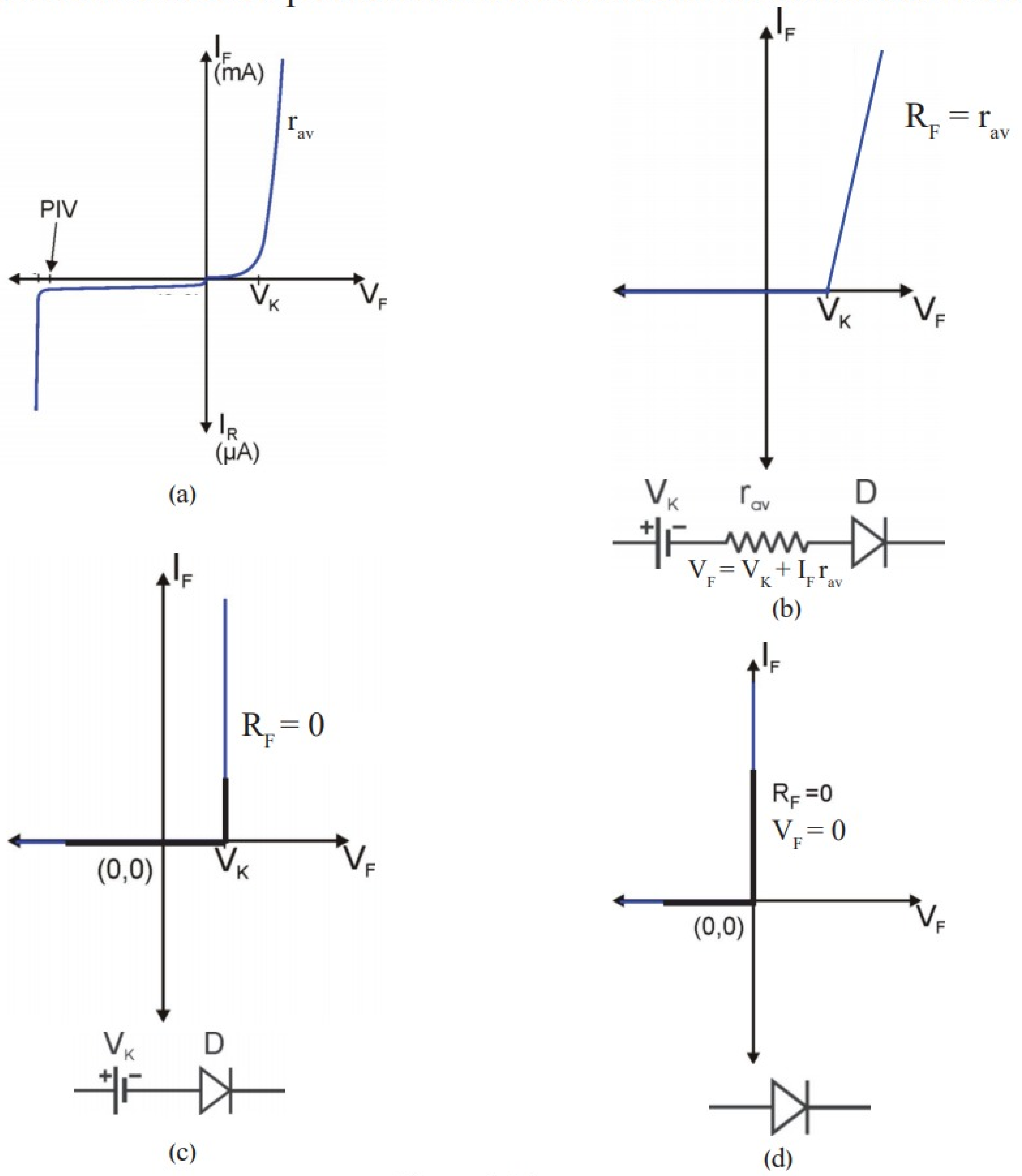


Figure 1.24

For the practical diode, the mathematical relationship between V_F and I_F is in fact an exponential function. In simple applications, mostly the diode is considered to be ideal and only the diode action is taken into consideration. When a small range of V_F is concerned the diode is taken to be a near ideal one with cut - in voltage V_K and calculations are carried out accordingly. For more accurate calculations both r_{av} and V_K are considered.

1.11.3 Diode as a switch

Let us consider the characteristic curves of an ideal and a mechanical switch.

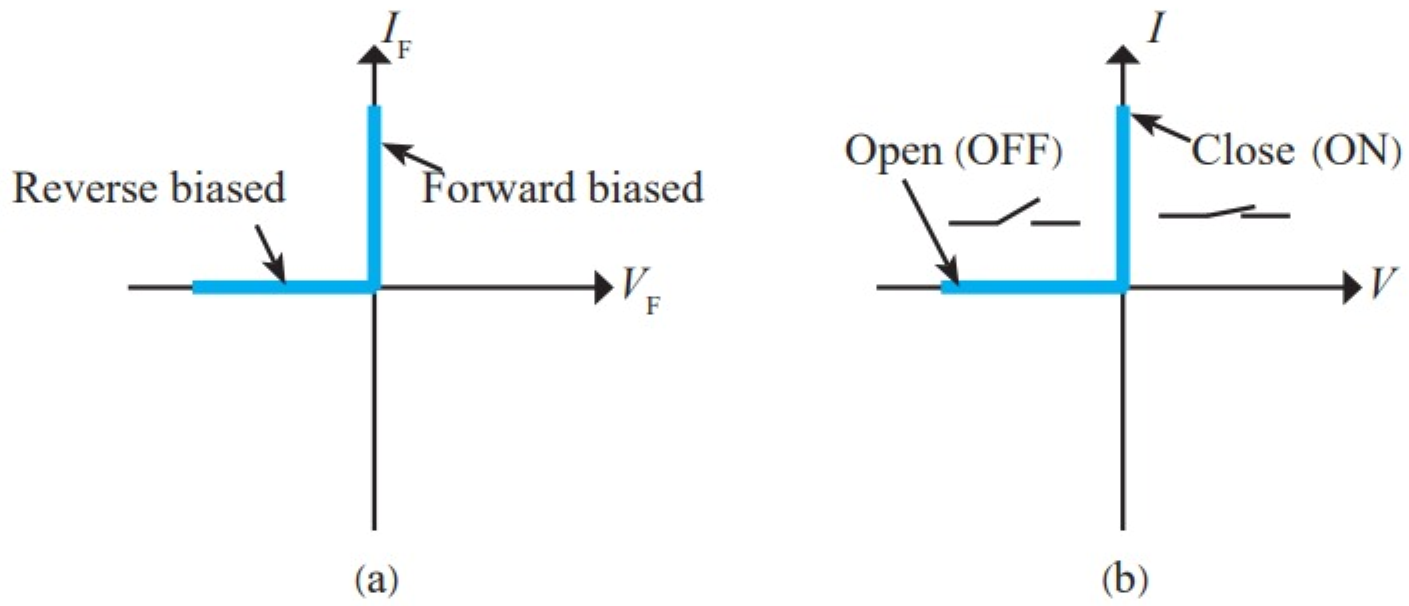


Figure 1.25

By comparing the two characteristics it can be seen that the reverse biased ideal diode does not conduct a current like in the open state of a mechanical switch and also when the diode is forward biased it allows a current to flow like in the closed state of a mechanical switch. Therefore, by making it forward biased or reverse biased a diode in an electronic circuit can be operated as a switch.

Eg : (i) Circuit which enables an electric supply to use a standby battery.

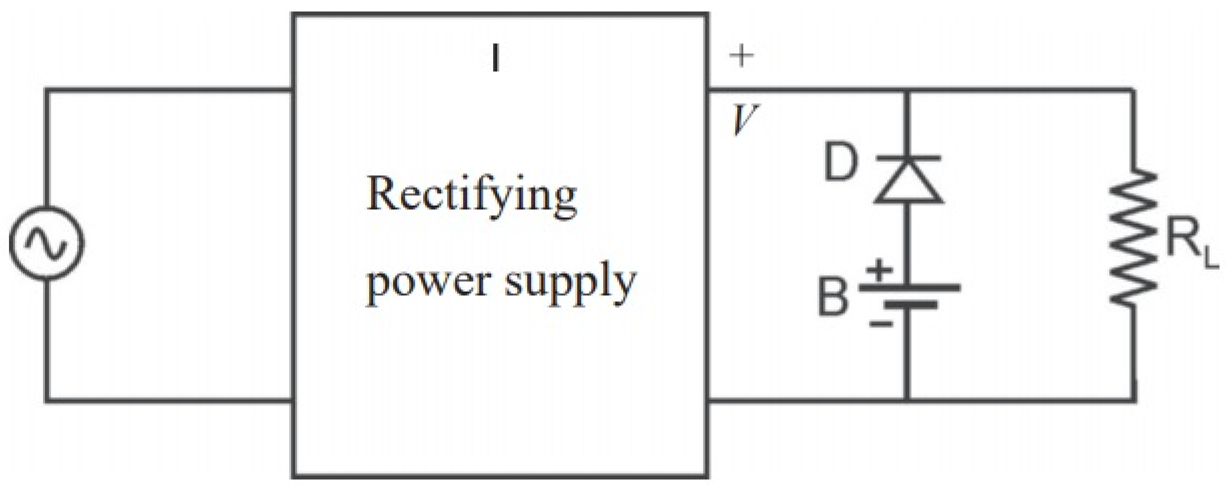


Figure 1.26

Let us suppose that the device (load) R_L is supplied with a d.c voltage V using a power pack working with an a.c supply voltage. If the a.c supply voltage fails suddenly the device R_L will not work. To overcome this, a battery with the same voltage V and a diode are connected to the power pack as shown in Figure 1.26. When the power supply is there, both sides of the diode (anode and cathode) get the same potential V making sure the diode does not get forward biased and therefore, no current is drawn from the battery. If the power supply fails, the diode will be forward biased due to potential at the positive terminal of the battery B, and the battery will supply the required voltage V to the load R_L. Here, the diode acts as a switch.

© 2020 National Institute of Education. All Rights Reserved.

(ii) Protecting an electronic device when batteries are connected with incorrect polarity.

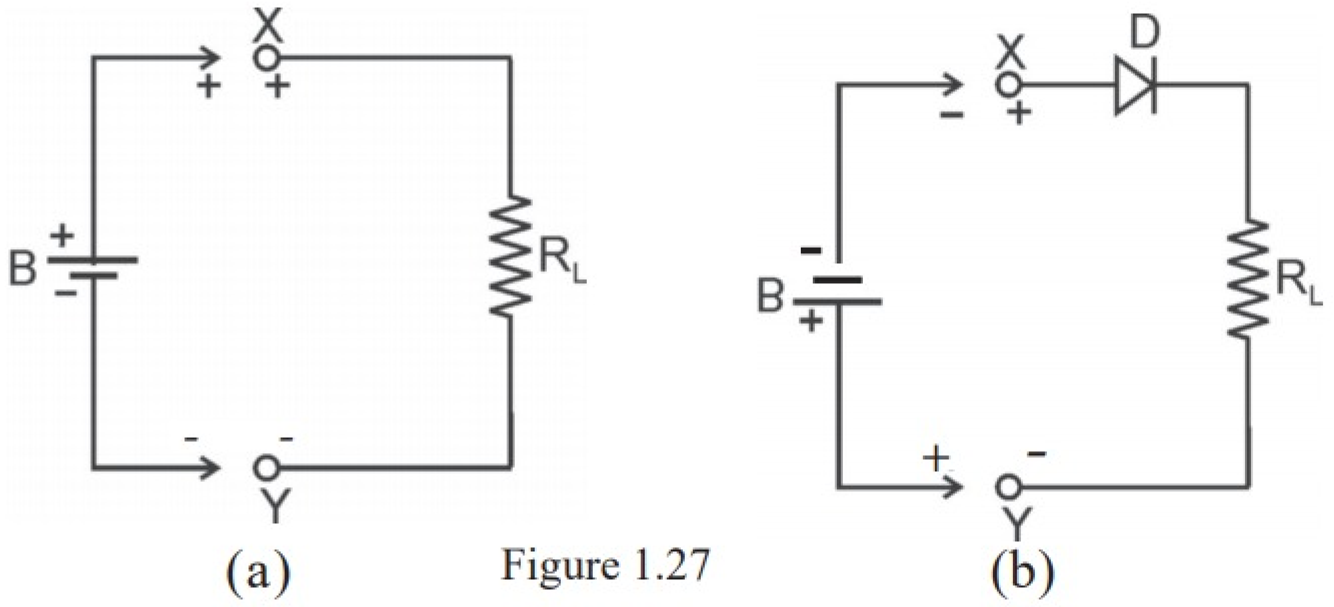


Figure 1.27

The correct way of connecting a battery (B) to an electronic device is shown by Figure 1.26 (a). (R_L represents the load resistance of the electronic circuit. If the device is connected to the battery with incorrect polarity then the device my get damaged. To prevent this, a diode D can be connected as shown in Figure 1.27 (b) so that it behaves as a switch. There the battery has been connected with the circuit, therefore a current does not flow through R_L. Only when the battery is connected correctly (as in Figure 1.27 (a)), does the diode get forward biased and pass a current to R_L.

1.11.4 Application of diode as a rectifier

The conversion of an alternating current into a direct current is known as rectification. Most of the electronic devices are working with direct current voltage supplies. The domestic mains electricity supply is an alternating voltage of 230 V and frequency 50 Hz. Therefore, to operate electronic devices using the mains electricity supply the rectification is essential. To do that, there is a **rectifier circuit** in the electronic device. The main component of that circuit is a diode which acts as a rectifier.

Rectifier circuits can be divided into two types depending on their operation.

(a) Half wave rectifier circuit
(b) Full wave rectifier circuit

Let us discuss the operation of these circuits separately.

(a) Half-wave rectification circuit

In half - wave rectification only one diode is connected in a series with the alternating voltage which is to be rectified.

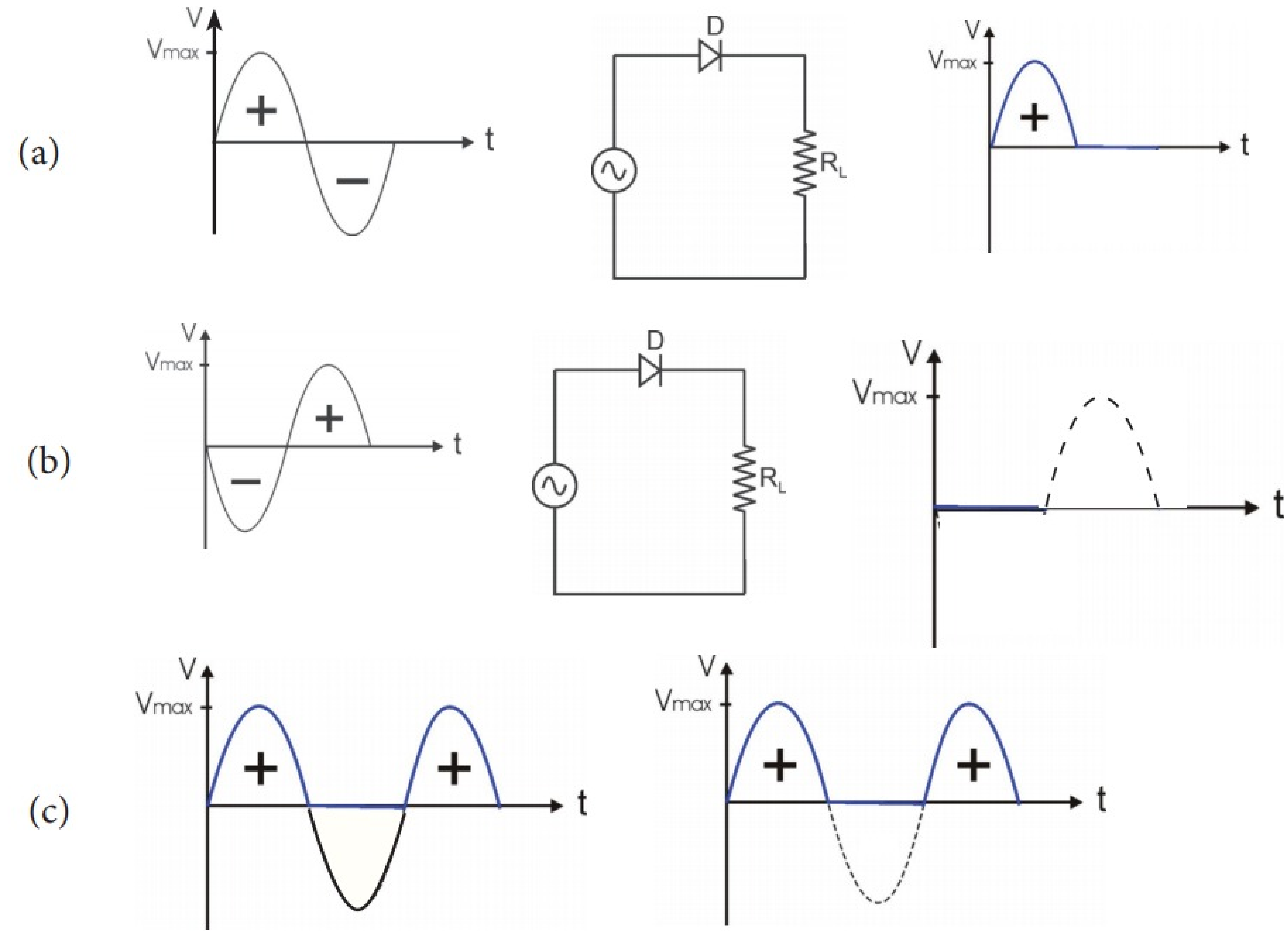


Figure 1.28

To explain the rectification process in Figure 1.28 (a) it is shown the instant where a positive half cycle of the alternating voltage is applied (Let us consider the diode to be an ideal diode). Now, since the diode is forward biased, the positive voltage appears across R_L as the output. Figure 1.28 (b) shows the instant where a negative half cycle of the alternating voltage is applied. Now the diode is reverse biased. Therefore, a voltage does not appear across R_L and the output is zero.

Figure 1.28 (c) shows the full wave of the input alternating voltage and the relevant output voltage wave.

In this rectification only the positive half cycles of the input alternating voltage wave is appearing as the output. Since, only one half of each cycle of the voltage wave is given as the output, this circuit is called the **half - wave rectifier** circuit.

(b) Full-wave rectification circuit

In this rectification the full wave of the input alternating voltage is subjected to rectification. Here, we consider only the bridge rectifier circuit. In the bridge rectifier, instead of a single diode, a bridge consisting of four diodes are connected as shown in Figure 1.29, is used.

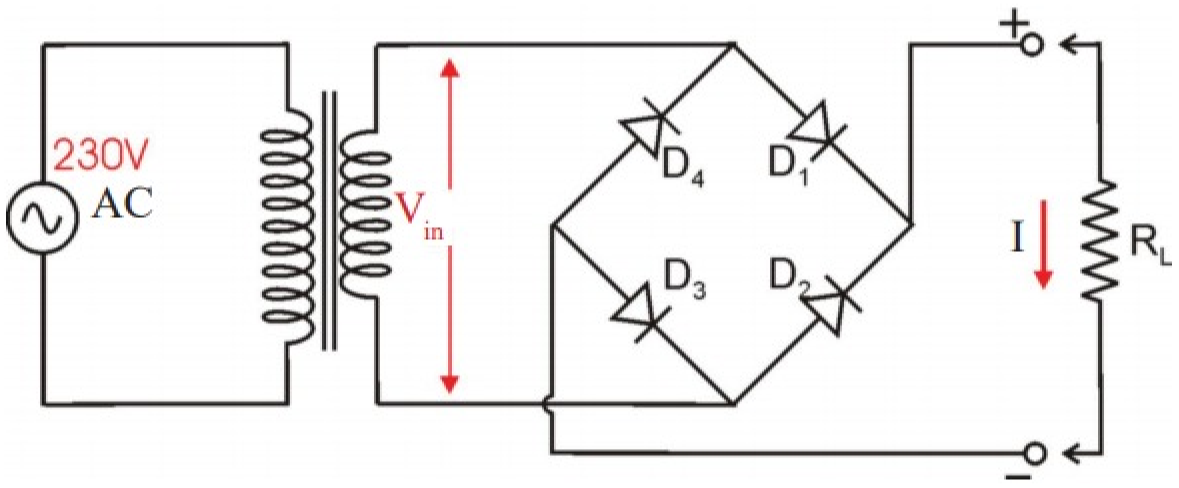


Figure 1.29

Let us consider how the diode bridge rectifies the alternating voltage which has been obtained by reducing the AC mains voltage to an appropriate value using a transformer.

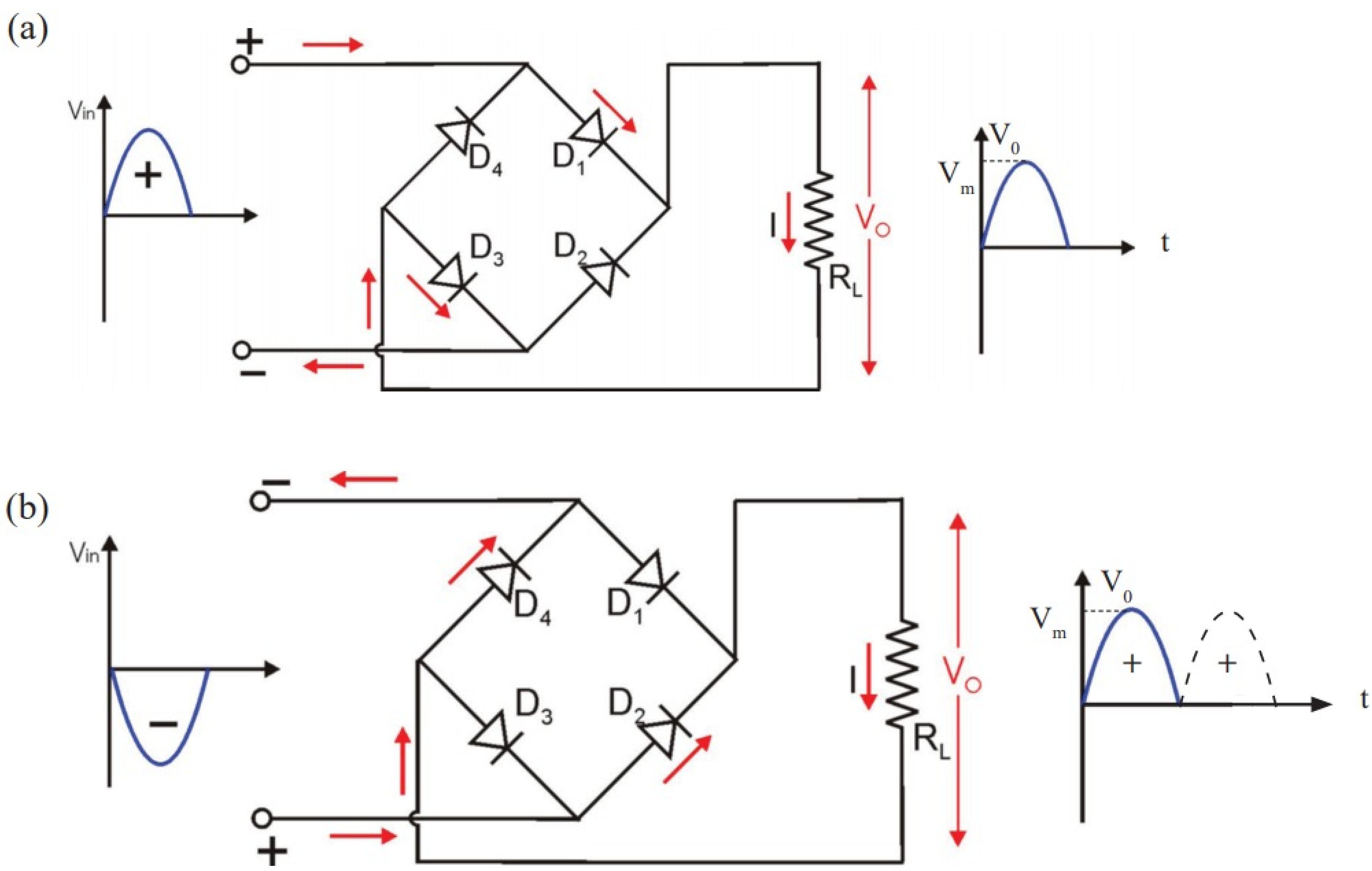


Figure 1.30

As shown by Figure 1.30 when a positive half of the alternating voltage is input to the bridge, the diodes D_1 and D_2 get forward biased and allow a current to flow through R_L. Figure 1.29 (b) shows how a current flows through R_L when a negative half cycle is input to the bridge. There the diodes D_2 and D_4 get forward biased and allow a current to flow. On both of the above occasions, it can be seen that the current flows through R_L in the same direction.

On both occasions the remaining two diodes (other than the two diodes which conduct) are reverse biased and therefore, do not contribute towards the flow of current. In full - wave rectification the output wave from is as shown by Figure 1.30 (b).

© 2020 National Institute of Education. All Rights Reserved.

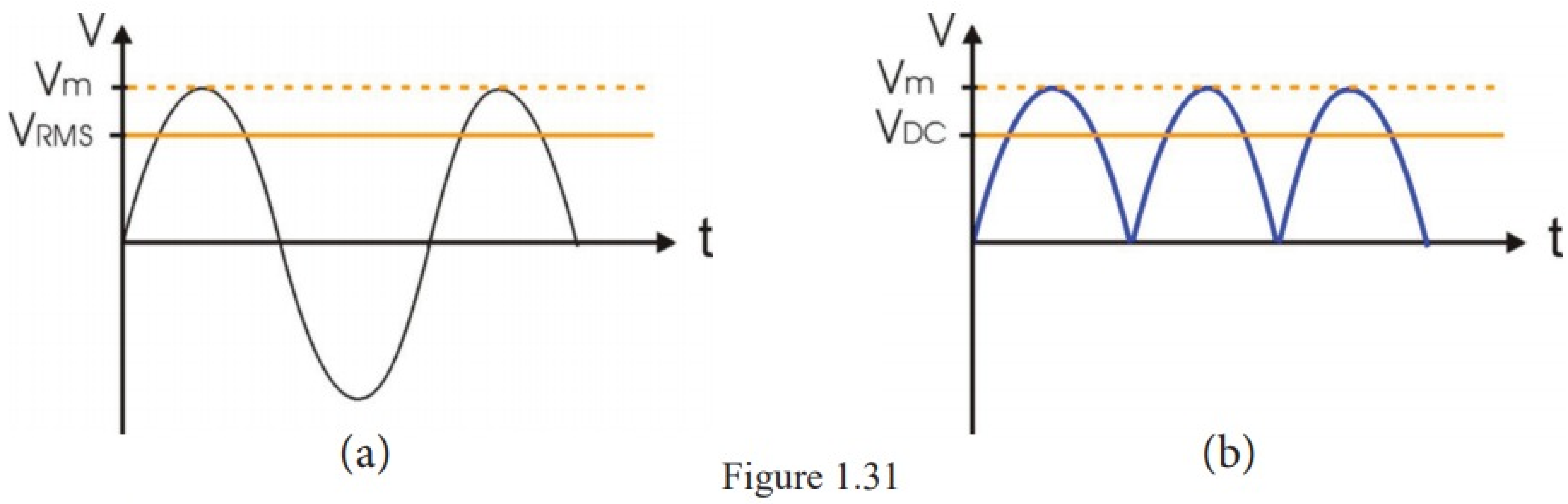


Figure 1.31

1.11.5 Smoothing

In both of the above rectifier circuits the output current is flowing along only one direction (DC) but the relevant output voltage is a series of pulses varying from zero to V_m. Most of the electronic circuits working with direct current, require a constant voltage as if the voltage is supplied by the batteries, for their correct operation. To achieve this, smoothing components are added to the rectifier circuit. The simplest component is a capacitor having a high capacitance, connected in parallel with the output. Let us first consider how to apply smoothing for a half - wave rectifier circuit.

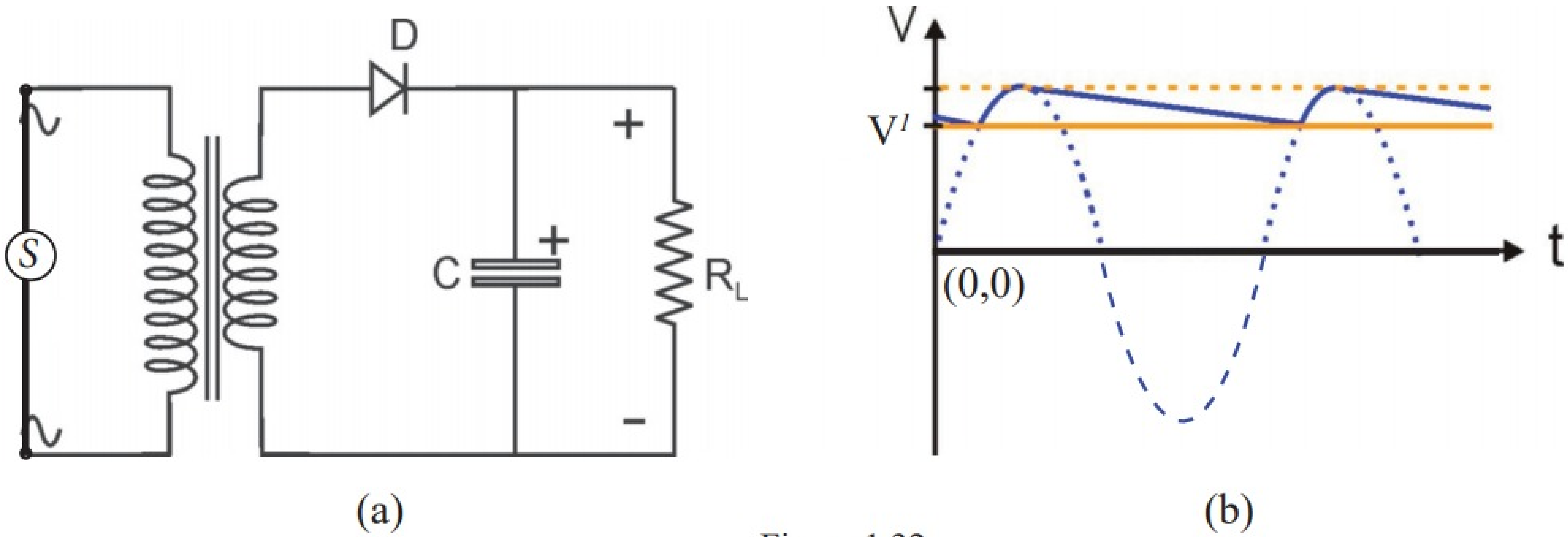
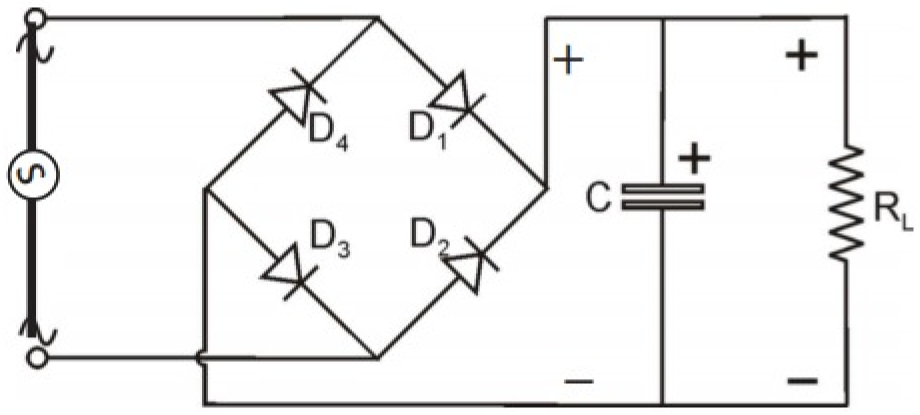


Figure 1.32

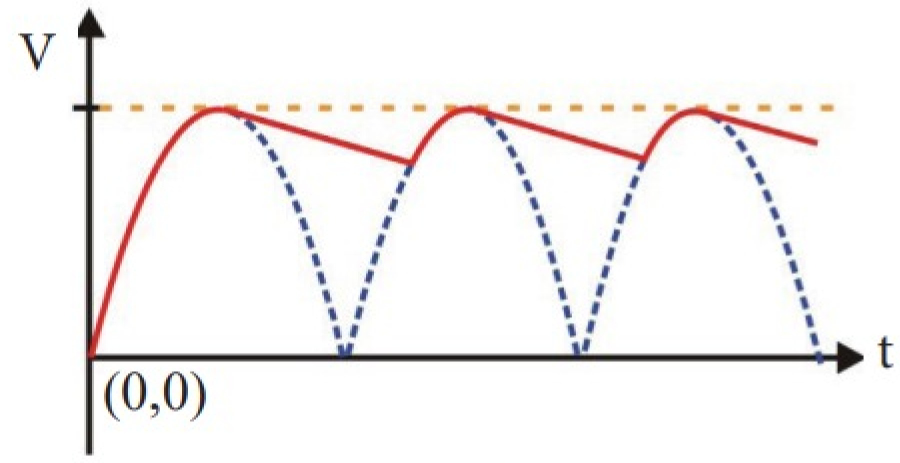
The output voltage of the rectifier circuit to which smoothing has been applied (shown in Figure 1.32 (a)) is a voltage varying between V_m and V^1. In this circuit the capacitor gets charged during the period in which the DC voltage is increasing up to V_m. Then , during the period in which the DC voltage is decreasing, the capacitor gets discharged through R_L providing an output voltage without letting it go down to zero. During the next positive half cycle of the AC voltage the capacitor gets charged again. $(V_m - V')$ is known as the **ripple voltage**.

When the capacitance of the smoothing capacitor is increased the ripple voltage decreases. Since there is only one ripple in each cycle of the alternating voltage, the ripple frequency is the same as the frequency of AC supply voltage. Therefore, the ripple frequency is 50Hz.

In the full wave rectifier circuit, the time between two voltage pulses of the output is less. Therefore when a smoothing capacitor is connected, the ripple voltage is less than that at half - wave rectification.



(a)



(b)

Figure 1.33

The frequency of ripple voltage in full - wave rectification is 100 Hz. The reason is that, there are two ripples in one wave of the alternating voltage.

In general capacitance of the smoothing capacitor should be more than 500 μF (up to about 10000 μF). There may be practical difficulties due to the large size of the capacitors with high capacitance. Capacitors of high capacitance are produced as electrolytic capacitors and there fore their terminals are named as positive and negative, should be connected parallel to the output with the correct polarity.

1.11.6 Identification of terminals and diode data of rectifier diodes.

Figure 1.34 shows the external view of a normal rectifier diode. It has a cylindrical structure made of plastic material. Diodes which can carry a current of 1A, are common in many circuits. They have a cylindrical shape with a diameter of about 3 mm and length of about 5 mm. To identify its cathode a silver ring (band) is marked close to the cathode terminal. When buying a diode it is very important to check for its maximum forward current and peak inverse voltage. In many data sheets, instead of PIV it is given V_{RRM} which is the maximum repetitive inverse voltage.

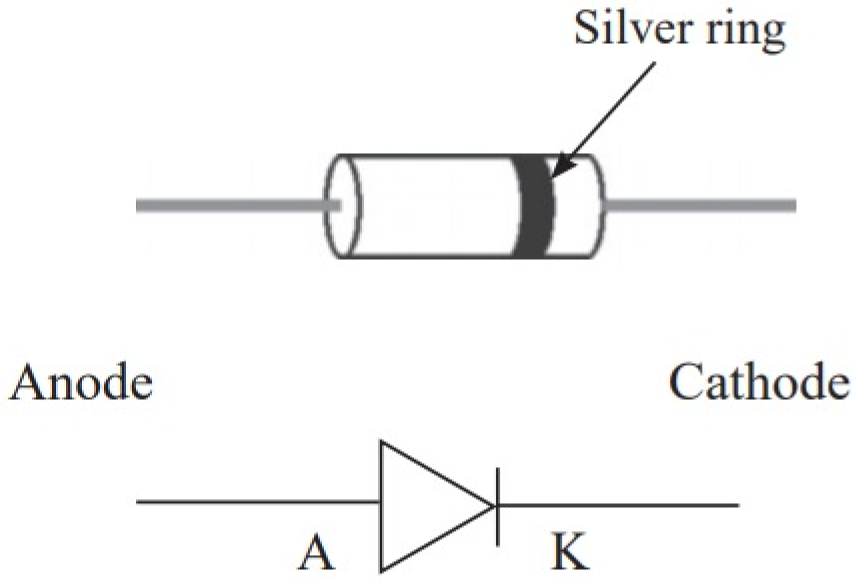


Figure 1.34

Shown below is a data collection of some diodes available in the market.

Table 1.4 Diode data

Identification number of the diode	Maximum forward current	Applicable maximum reverse bias voltage (V_{RRM})
1 N 4001	1 A	50 V
1 N 4002	1 A	100 V
1 N 4003	1 A	200 V
1 N 4004	1 A	400 V
1 N 4005	1 A	600 V
1 N 4006	1 A	800 V
1 N 4007	1 A	1000 V
1 N 5400	3 A	50 V
1 N 5404	3 A	400 V
1 N 5408	3 A	1000 V
BY 127	1 A	1250 V
M R 750	6 A	50 V
M R 754	6 A	400 V
E M 518	1 A	2000 V

In addition to normal diodes, the bridge rectifiers in which there are four diodes arranged in a single pack, are also available in the market. In them the terminals to which the AC voltage should be supplied are marked as ~ and the output terminals are marked as + and -. (Figure 1.35)

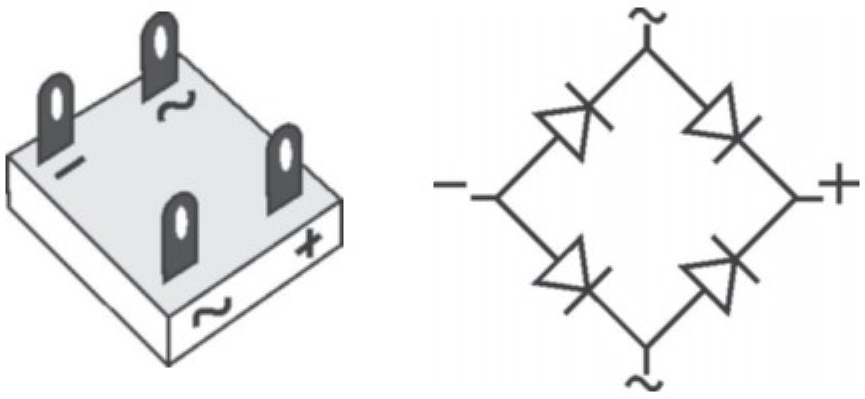


Figure 1.35

1.11.7 Demonstration of the action of a rectifier circuit using cathode ray oscilloscope (CRO)

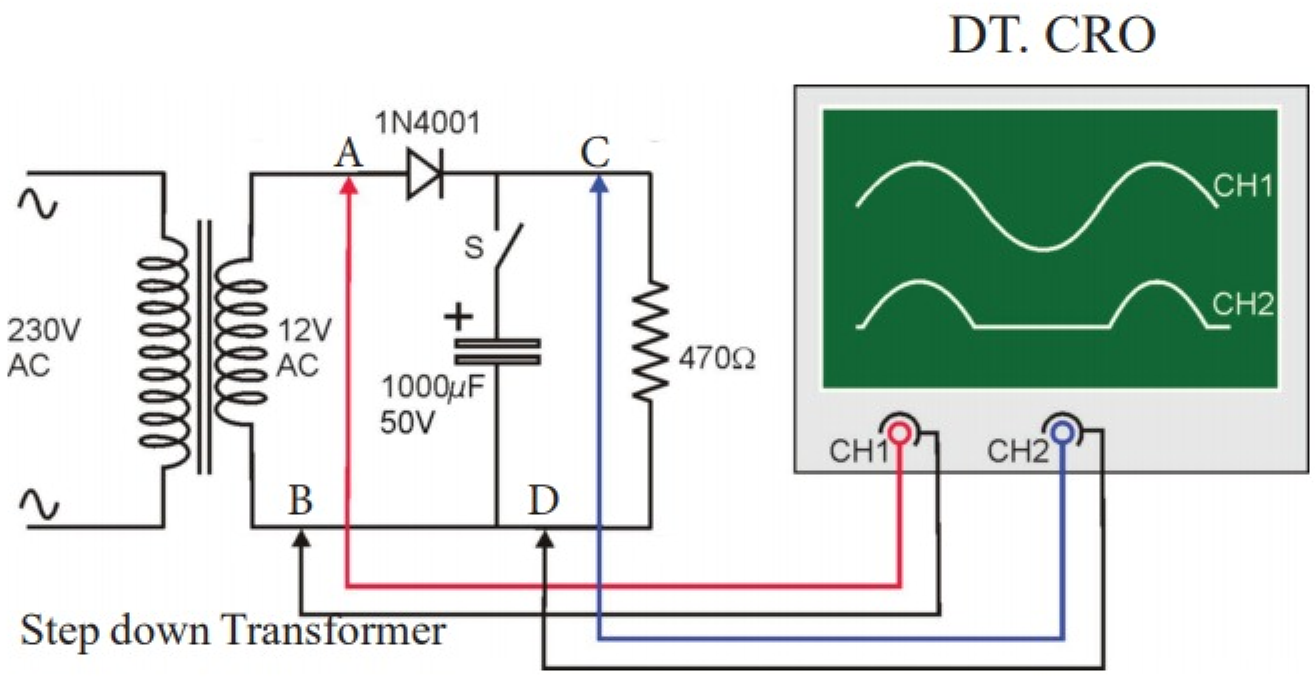


Figure 1.36

Adjust the relevant controls of CRO appropriately. Then the waveforms of the input and the output can be observed on channel one and channel two respectively. Make the switch S closed (ON) and then the output waveform after undergoing the smoothing process can be observed on channel two.

The circuit shown in Figure 1.37 can be used to demonstrate the action of a full - wave rectifier circuit.

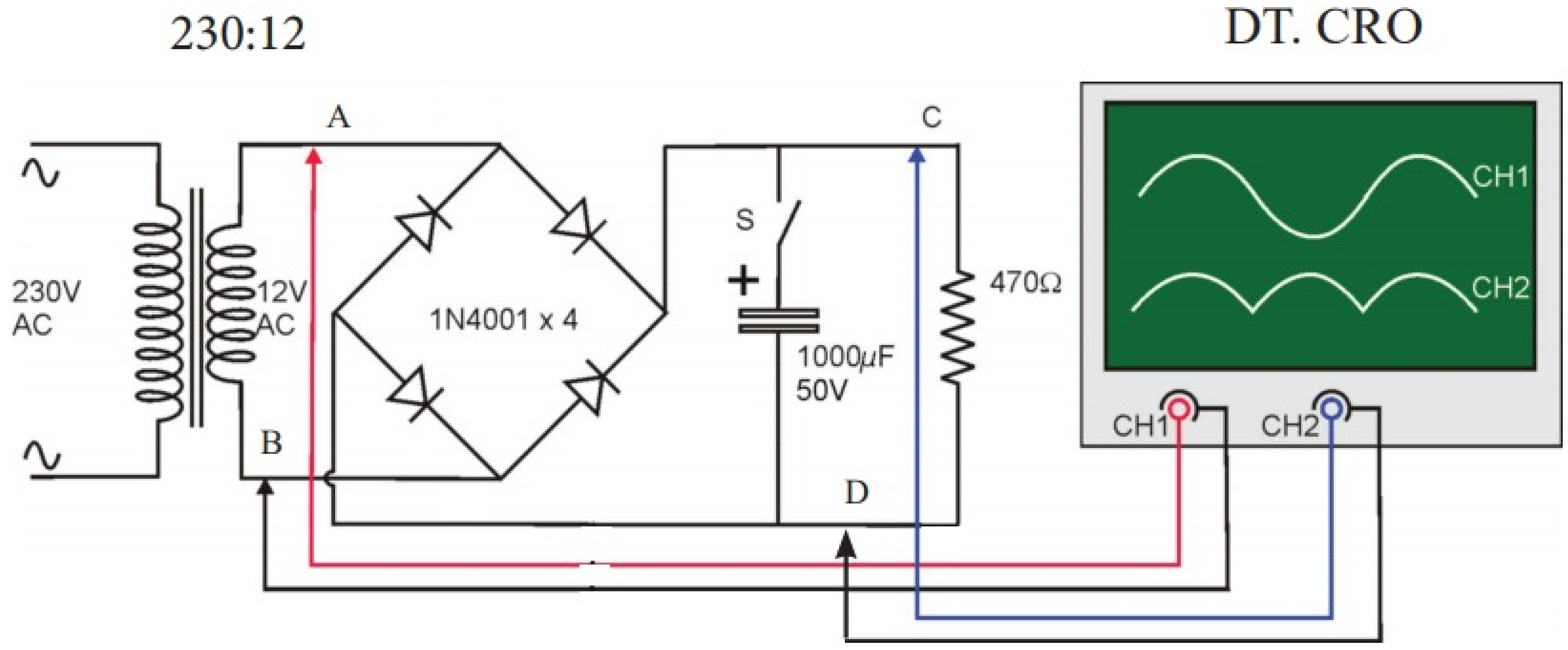


Figure 1.37

Carry out the experiment in the same way as in the case of full-wave rectification and observe the AC waveform. Then close the switch S, so that the rectified voltage undergoes smoothing, and observe that output waveform. In both half wave and full wave circuits, measure relevant ripple voltages using CRO.

Type of diode

So for we have dealt with the rectifier diodes. These diodes are used not only in reciter circuits but also in other circuits like switching circuits, voltage multiplier circuits and wave - clipper circuits. In addition to these rectifier diodes, there are other types as well, used for various applications. Among those types, Zener diode, light emitting diode and photo diode are more familiar. Let us now consider them separately.

1.12 Zener diode

This is a junction diode made of silicon. In rectifier diodes, when the breakdown voltage is exceeded the diode breaks down and conducts a large reverse bias current. In the manufacturing process of zener diodes by controlling the level of doping slightly, they are produced so that they have the property of not getting damaged even when the breakdown voltage is exceeded.

There are two reasons for the flow of a large current in the reverse bias of a diode. One is the **zener breakdown** and the other is **avalanche breakdown**. Let us consider how these breakdowns take place.

© 2020 National Institute of Education. All Rights Reserved.

(i) Zener breakdown

When a p - n junction is made using a semiconductor with a high level of doping the depletion layer becomes very small. (reason is, the required potential barrier can be developed by diffusion of carriers within a short distance from the junction). When this type of a diode is reverse biased the electric field in it is very strong because the potential difference arises within a very short distance. This strong field applies a large force on electrons in the lattice (according to $F = eE$).

When this force becomes enough to break down the bonds, a large number of bonds break down at once producing a large number of minority carriers (free electrons and holes) enabling a large current to flow across the junction. This nature of breakdown of a p - n junction is known as zener breakdown.

(ii) Avalanche Breakdown

In this breakdown the electric field developed due to reverse bias voltage across the junction, is applied a force $(F = Ee)$ on minority carries in the lattice. Because of this force minority carriers gain an acceleration 'a' $(Ee = ma)$. Under this acceleration the velocity of minority carriers increases and they gain a large kinetic energy ($\frac{1}{2} mv^2$). These carriers break bonds in the lattice by colliding with them. The carriers generated by breaking down of bonds also get accelerated and gain a kinetic energy sufficient enough to break down bonds. This process continues as a chain reaction and thereby generates a large number of carriers instantly. The minority carriers thus generated contribute towards the flow of a large reverse bias current at once. This breakdown is known as avalanche breakdown. In rectifier diodes, breakdowns occur at reverse bias voltages due to avalanche breakdown.

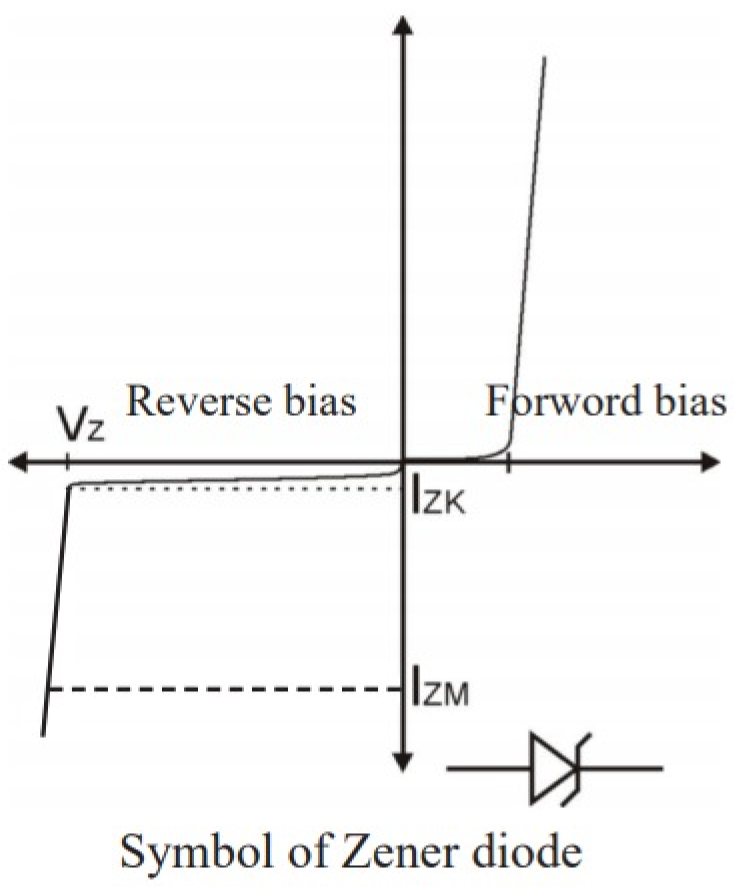


Figure 1.38

Figure 1.38 shows the characteristic curve of a zener diode. Here the breakdown voltage is given as V_z. Zener diodes are used in circuits under reverse bias conditions. At the beginning a very small current flows due to reverse bias voltage and at the breakdown, the relevant voltage is a constant one denoted by V_z, and I_{zk} is known as knee current.

If a large current flows through the zener diode it can get damaged at one stage due to heat generated. Therefore there is a maximum current that the zener diode can withstand. This maximum zener current is given as I_{zM}. The zener diode is used within the limited current range between I_{zk} and I_{zM} . For small zener voltage ($Vz < 6$ V), the zener breakdown takes place and for large zener voltage the breakdown occurs due to avalanche breakdown.

Zener diodes are used in many electronic circuits for various purposes. Among them the two major uses are voltage regulation and making reference voltages. Let us consider them now.

© 2020 National Institute of Education. All Rights Reserved.

1.12.1 Voltage Regulation

If the required voltage is supplied to an electronic device using a rectifier circuit, there may be two reasons for the variation of that supply voltage.

(i) because of the variation of AC voltage by which the power is supplied to the rectifier circuit.

(ii) because of the variation of current drawn by the load connected to the rectifier circuit.

On both of the above occasions, the voltage supplied to the load gets changed and therefore the load may not be functioning correctly. A voltage regulator circuit is connected after the rectification circuit to make the voltage supply to the load not getting changed due to any reason above. A rectified voltage supply without having a regulator circuit is shown in Figure 1.38

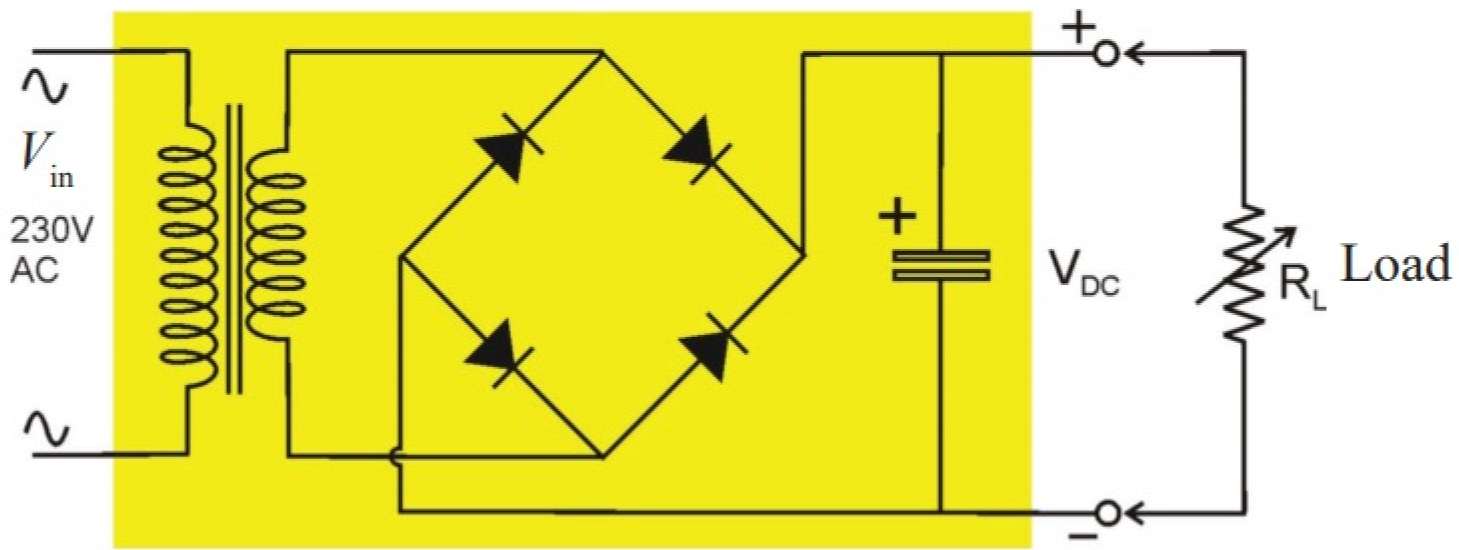


Figure 1.39

The regulation against the above reason (i) is known as 'line regulation' and against the above reason (ii) is known as 'load regulation'. In general it is understood by stating voltage regulation that both these regulations take place. Therefore, the variation of output voltage V_{DC} due to change of AC supply voltage, and due to change of load resistance (R_L) are eliminated. When the load resistance changes, the load current drawn by the power supply also changes when voltage drop across the internal resistance (resistance of transformer windings and diodes) of power supply changes (therefore) V_{DC} changes. ($V_{OUT} = V_{DC}$ – potential drop of power supply)

To get rid of these problems the regulator circuit is introduced between the rectified voltage supply and the load. The simplest regulator circuit is the zener diode with a current controlling resistor connected in series. The property making use of here is, when the zener diode is reverse biased the voltage across it does not exceed the zener voltage (V_Z) of it.

Let us suppose that the output of the rectifier circuit with smoothing is V_S and the current drawn out from it is I_S. The current I_S is flowing through the current controlling resistor R_S. Hence there is a voltage drop of $R_S I_S$ across the resistor R_S. The relevant circuit is shown in Figure 1.40

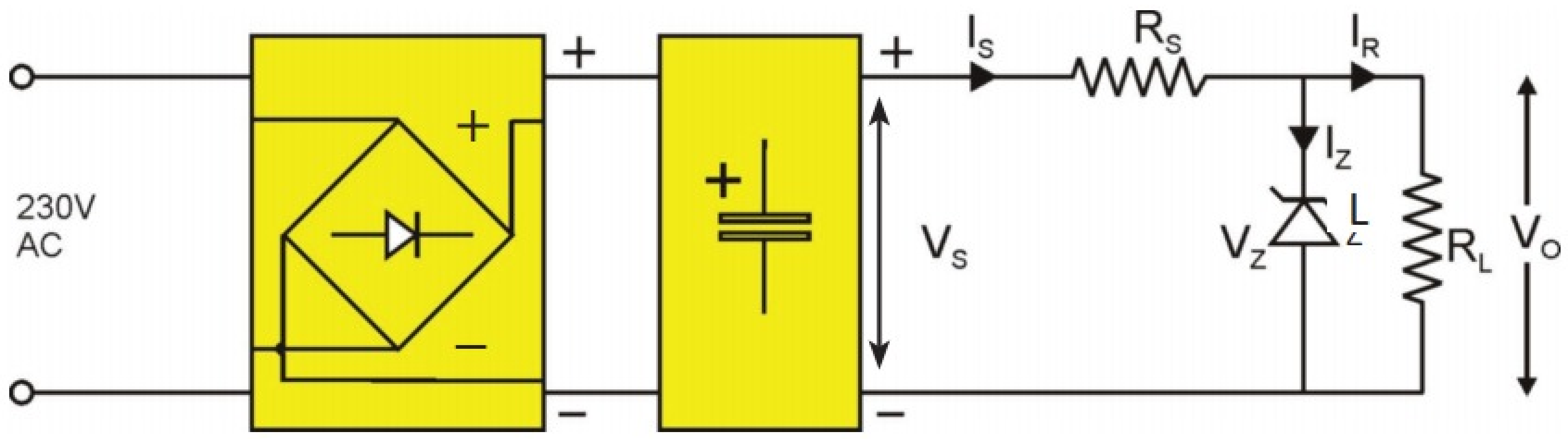


Figure 1.40

The zener diode is connected in parallel with the output and therefore, the output is such that it is equal to V_Z.

$$V_O = V_z$$

The rectifier circuit should be designed so that the value of V_S is always greater than V_z ($V_S > V_z$).

The potential difference $V_S - V_z$ appears across the resistor R_S as its voltage drops ($I_S R_S$), due to current IS flowing through it.

$$\therefore V_S - V_z = R_S I_S$$

$$\therefore \boxed{I_S = \frac{V_S - V_z}{R_S}}$$

If V_S increases due to increase of AC supply voltage, then I_S increases and hence the increased voltage is balanced by the voltage difference $V_S - V_Z$. Hence the output voltage remains constant as $V_0 = V_D$. If R_L is constant then I_L is also constant. Therefore, when I_S increases, I_Z also increases keeping the output voltage constant at V_Z.

In this manner the zener diode regulates the change of output voltage due to change of AC supply voltage. (This regulation is known as 'line regulation').

If the load resistor is changed while the supply voltage is constant, then the load current I_R changes. As I_R changes I_S also changes and therefore the potential drop across R_S changes making V_0 changed.

However, since $I_S = I_L + I_Z$ in this circuit, I_S is kept constant by getting I_Z changed. Then the output voltage can be kept constant at V_Z as earlier. This voltage regulation is called "load regulation".

The maximum current (I_{Max}) which can flow through the zener diode is determined by the power that it can withstand. Let us denote that power as P_M.

$$P_M = V_Z I_{Max} \quad \text{(value of } I_{Max} \text{ should not exceed } I_{ZM})$$

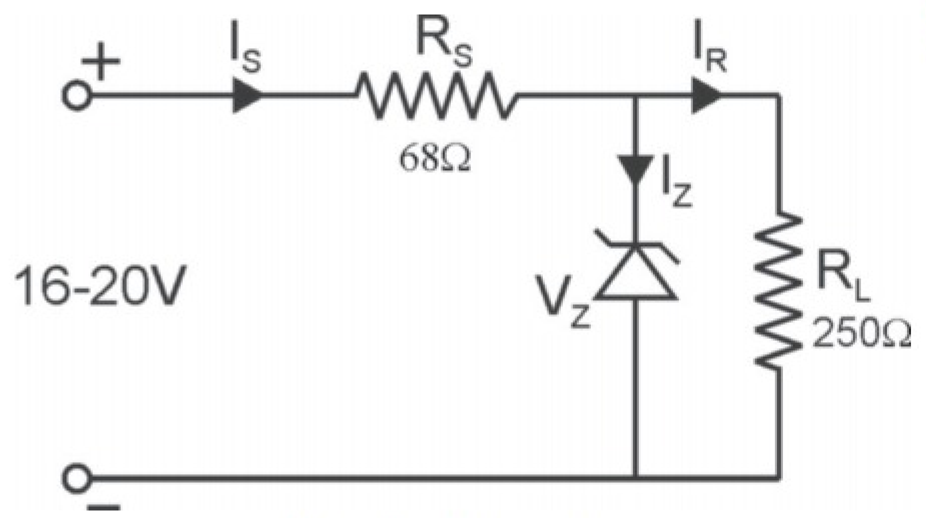
$$\therefore \boxed{P_M = V_Z I_{ZM}}$$

© 2020 National Institute of Education. All Rights Reserved.

When designing a voltage regulator circuit, the values of V_S and R_S should be selected as suitable and should take measures to confine zener current into the range I_{ZK} to I_{ZM}. (Figure 1.40). Since the theories related to designing voltage regulator circuits are beyond the scope of the advanced level physics syllabus, what we do here is only to analyze a voltage regulator circuit which has already been made. It will further be clarified by the following example.

eg. A DC - voltage supply (V_S) having a voltage varying between 16 V and 20 V is used to maintain a constant voltage of 12 V across a load resistor of 250 Ω. A current controlling resistor (R_S) of 68 Ω is used here. Draw the relevant voltage regulator circuit and calculate the following.

(a) Required zener voltage of the zener diode.
(b) The maximum current flowing through R_S.
(c) Current flown through the load.
(d) Under the above load the maximum current flown through R_S.
(e) Power of the resistor R_S to be used.
(f) The maximum current flown through the zener diode when the load is connected.
(g) The maximum current through the diode when the load is connected.
(h) Power of the zener diode to be used in the circuit.



Required circuit
Figure 1.41

(a) since there should be a constant voltage of 12 V across R_L the selected zener diode should be with zener voltage of 12 V.

$\therefore V_Z = 12$ V

(b) Current through R_S is, minimum when the supply voltage is 16 V. On that occasion, R_S

$16 - 12 = I'_S \times 68$

$$\therefore I_S = \frac{4}{68}\text{ A} = \frac{4 \times 1000}{68}\text{ mA} = 58.8\text{ mA}$$

$\therefore$ Therefore, the minimum current flowing through $R_S = 58.8$ mA

(c) Since the voltage across the load is always 12 V

$12 = I_L \times 200$

$$\therefore I_L = \frac{12}{250}\text{ A} = \frac{12 \times 1000}{250}\text{ mA} = 48\text{ mA}$$

(d) When the load is connected current through Rs becomes maximum when the supply voltage is 20 V.

$$\therefore 20 - 12 = I''_S \times 68$$

$$I''_S = \frac{8}{68}\text{ A} = \frac{8 \times 1000}{250}\text{ mA} = 117.6\text{ mA}$$

$\therefore$ Then maximum current flown through $R_S = 117.6$ mA

(e) If the power of the resistor R_S is P_S,

$P_S = I^2_S R_S$

$P_S = (.118)^2 \times 6 = 0.946$ W

Since the closest one available in the market is 68 Ω, 1 W resistor, it should be used here.

(f) Taking I_Z as the maximum current through zener diode when the load is connected, Since the maximum current flowing through the load is 48 mA and the maximum current flowing through R_S (when $V_S = 20$ V) is 117.6 mA,

$$117.6 = 48 + I_S$$
$$\therefore I_S = 69.6\text{ mA}$$

(g) Even when the load is disconnected the output voltage is 12 V. Therefore, the maximum current flowing through R_S is still 117.6 mA. Since the load current (I_L) now, is zero this whole current should flow through the zener diode.
The maximum current flowing through the zener diode = 117.6 mA.

(h) The power which the zener diode should have is, the power that it has when the maximum current is flowing through it.

Since, $P_Z = I_Z \times V_Z$

$P_Z = 0.1176 \times 12 = 1.41$ W

2 W zener diode which is available in the market can be used for this purpose.

© 2020 National Institute of Education. All Rights Reserved.

1.12.2 Making Reference Voltage

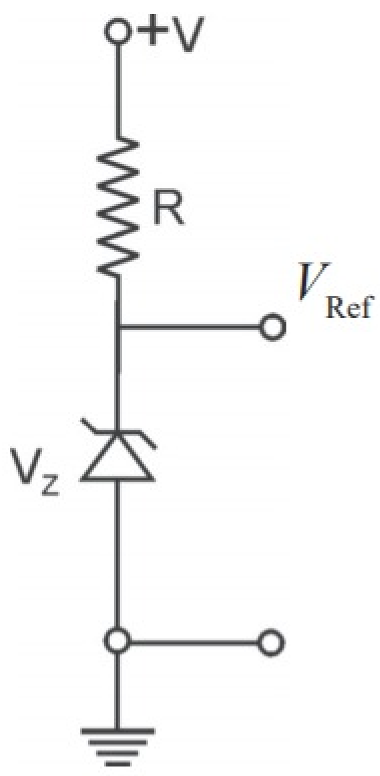


Figure 1.42

Constant reference voltage are needed to operate various electronic circuits. The easiest way to make the voltages is to use zener diodes. There a reverse biased zener diode is connected in series with a current controlling resistor *R*. The resistor *R* should be selected so that a current of about 10 mA is flowing through the zener diode. The reference voltage will be the zener voltage of the diode. (zener diodes produced under E.24 series are available at voltages from 3.3 V to 100V and can be brought under the group 1 N 4728 with power 1W).

eg. Design a circuit to obtain a reference voltage of 6 V from a 12 V, DC power supply.

The zener voltage of the required zener diode should be 6 V. When a current of 10 mA is flowing through the diode, the voltage drop across the resistor R should be (12-6) V.

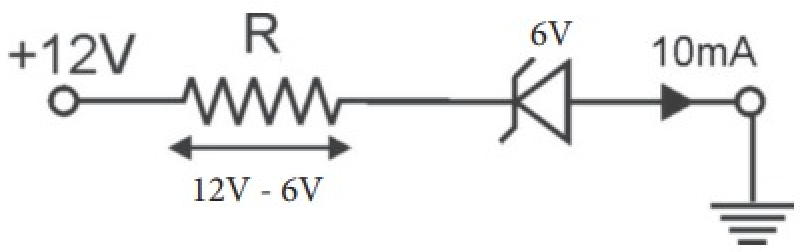


Figure 1.43

$$12 - 6 = \frac{10}{1000} \times R$$

$$\frac{6000}{10} = R$$

$$R = 600\ \Omega$$

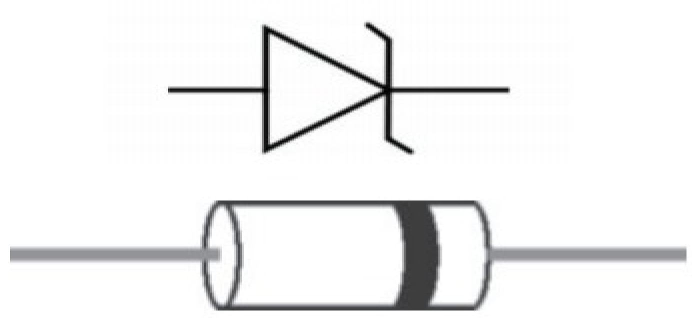
The circuit can be designed as shown in Figure 1.42 by connecting a resistor of 600 Ω in series with the zener diode of 6 V.

(The closet resistance value available in the market is 620 Ω. A zener diode which can be used is 1N 5233, 500 mW)

Zener diode packing and identification of terminals

Several zener diodes are shown in Figure 1.43. Zener diodes having power 1W or less, are packed in glass covers and those having much power are packed in black epoxy covers. Like in normal diodes, in zener diodes also, they are marked with a black or white ring (band) close to the cathode to identify that terminal.

(a)



(b)

Figure 1.44

© 2020 National Institute of Education. All Rights Reserved.

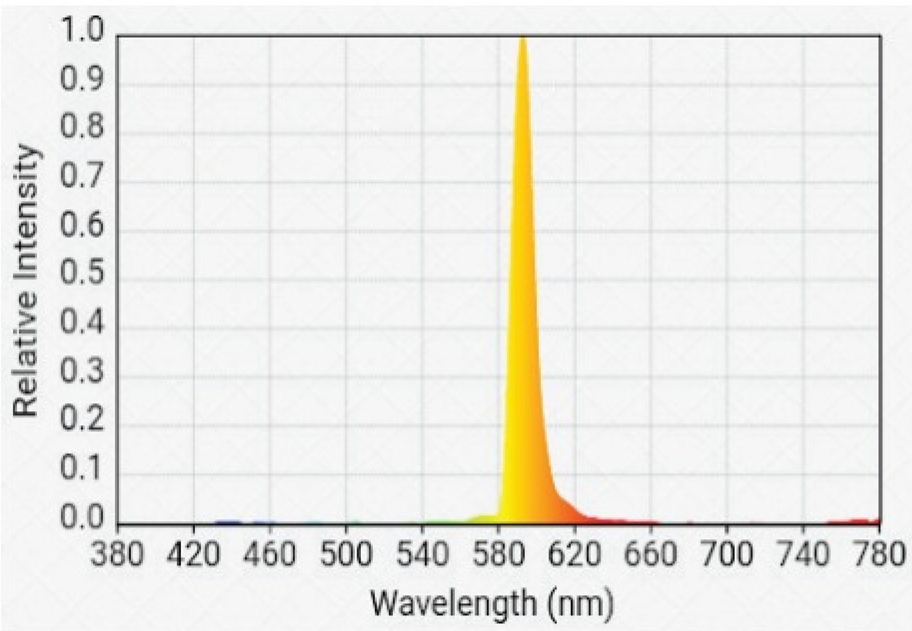
1.13 Light Emitting Diode (LED)

We have learnt earlier that when a p-n junction is forward biased the charge carries flow across the junction and some of them get recombined. In normal diodes made of germanium or silicon, this recombination of carriers emits energy as heat.

The electromagnetic waves emitted in Ge or Si diodes are in the heat radiation region, while in GaAsP diodes those waves are in the read-orange colour region. This occurrence is used in the light emitting diode. In 1962 Nic Holonyak of General Electric Company made the first p-n junction which emits red colour. Afterwards LEDs were produced so as to get any colour in the spectrum. All of those have been produced using p-n junctions made of compound semiconductors of group iii-v elements.

Lights emitted by an LED is confined to a very small range of wavelength. Therefore, it can be considered as a monochromatic source of light. Outer cover of practical LEDs are made of transparent epoxy and have been coloured so as to know its coluor when it is not lit.

Red LEDs made using GaP have a wavelength distribution as shown by Figure 1.44(a) and a data table of p-n junctions which emit various colours, as given in Figure 1.44(b).



(a) Spectrum of Red LED

Colour	Semiconductor
Infra-red (IR)	GaAs
Red	GaP
Orange	GaAsP
Yellow	Al GaIn P
Green	Al GaP
Blue	InGaN
Ultraviolet (UV)	A l N

(b)

Figure 1.45

In the LED there is an n-type semiconductor built on a metal plate and a very thin p-type semiconductor region is formed by means of diffusion, on the n-type semiconductor (Figure 1.45). p-type semiconductor as been connected to the positive electrode using a metal wire while the metal plate has been connected to the cathode. The junction is constructed by the p-type and n-type semiconductors and when the junction is forward biased, charge carriers recombine in the junction and emit electromagnetic radiation. This radiation (light) is emitted from the thin p-type semiconductor which has a large area. The circuit symbol of the light emitting diode is also shown in Figure 1.45.

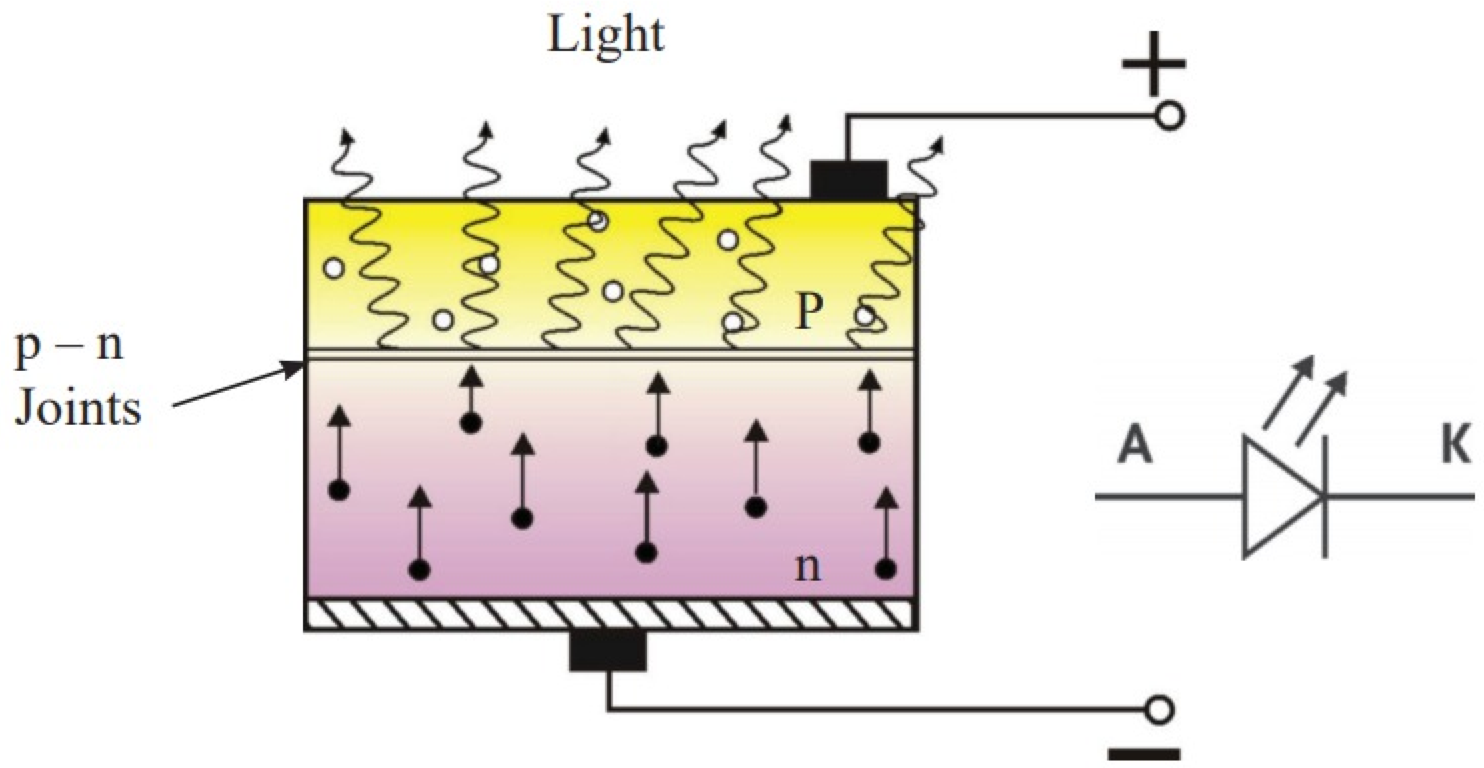
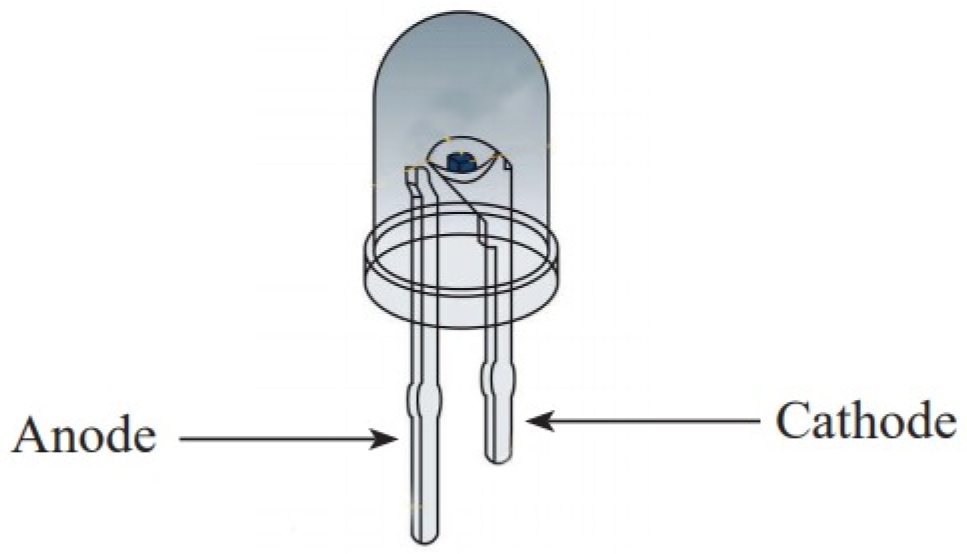


Figure 1.46

Figure 1.45 gives appearance of an LED and the way that the terminals can be identified.



Appearance of LED

Figure 1.47

1.13.1 Application of the light emitting diode

In the early days of LED it came into application as an indicator. In addition they were used in decorative light systems. Now it has found applications as a source of light in domestic electric lamps and searchlights.

1.13.2 Forward bias voltage of LED

We know that for a normal Si-diode the forward bias voltage is about 0.6 V. However depending on the semiconductor, forward bias voltage of an LED made of compound semiconductor lies in the range 1.7 V to 4 V. A normal red-LED emits light well at voltages between 1.8 V to 2.5 V. The intensity of light emitted by the red-LED is directly proportional to the forward current flowing through it. Red-LEDs emit light well at a current flowing though it. Red-LEDs emit light well at a current of 10mA under the forward bias voltage of 2 V. If a normal LED is lit according to the given standards it has a lifetime of more than 50,000 hours. If it is lit under a higher voltage its lifespan gets shortened.

When LEDs are lit using different voltages it should be connected with a current controlling resistor in series. The value that the resistor should have can be calculated easily using Ohm's law.

Eg: An LED should be lit using a 9 V supply. Let us suppose that the rated voltage across the LED is 2 V and the current though it is 10 mA.

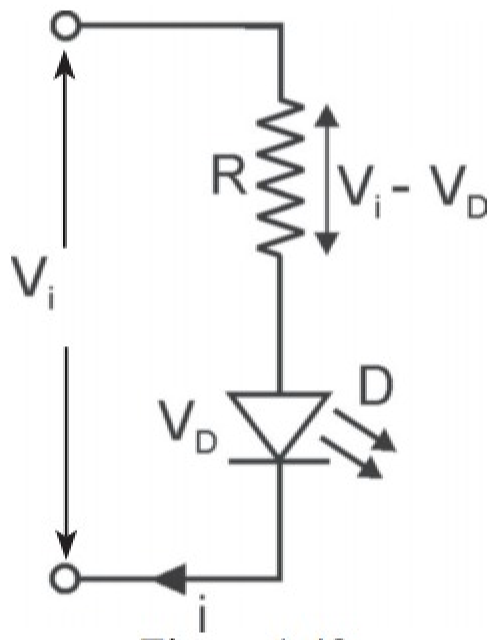


Figure 1.48

$$V_i - V_D = IR$$

$$\therefore 9 - 2 = \frac{10}{1000} \times R$$

$$700 = R$$

$$R = 700\ \Omega$$

Therefore, a circuit with a current controlling resistor (R) of 700 Ω should be used (Figure 1.47)

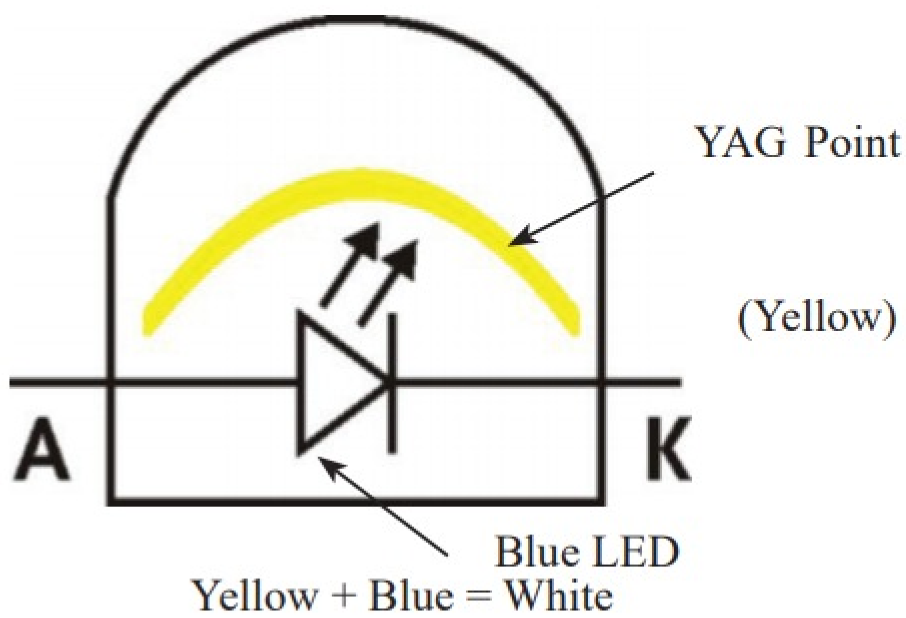
(Resistor which is available in the market having the nearest value is a 750 Ω resistor.)

For extra knowledge

White - LED

There are two methods to get white light from LEDs. The first method is to use three LEDs to get white light. The three LEDs are known as RGB-LEDs. This method is not used to produce normal white light source, but it is used in LCD-LED television screens. In these screens R,G,B three light emitting diodes are used as a backlight which is the source of light behind the LCD screen. RGB-LED is used because it has the three major colours red, green and blue in white light.

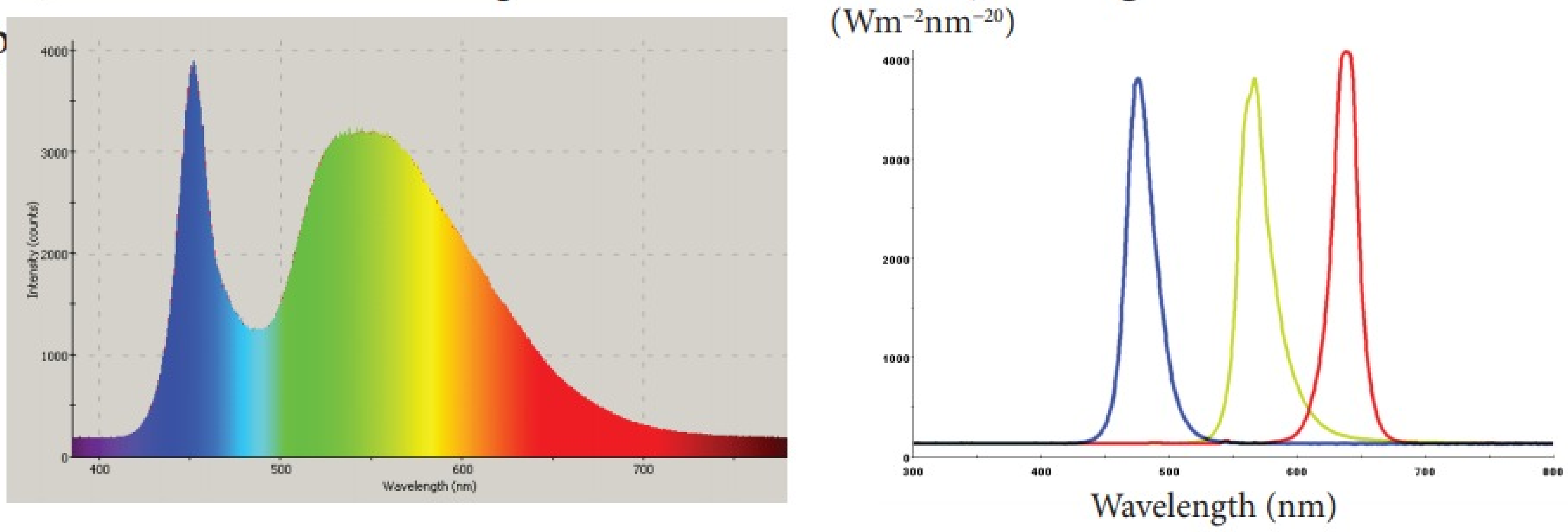
White LEDs now in the marked as sources of white light, were produced based on "YAG fluorescence method" introduced by the Japanese scientist "Nakamura". In this method, one blue-LED is used and above p-n junction there is a layer of fluorescent paint. This paint consists of YAG which is cerium-doped **Y**ttrium **A**luminium **G**arnet.



Structure of YAG LED
Figure 1.49

When the blue LED is lit, as spark of blue light is absorbed by the fluorescent paint and it emits yellow light. The unabsorbed part of the blue light and this yellow light (complimentary colour) mix together and emit white light. This kind of LED pack-ing is shown in Figure 1.48.

The emission spectra of RGB white-LED and YAG white-LED are given below (Figure 1.49). Although it seems that YAG white-LED gives out white light you can observe that, not all the colour components associated with white light are there in the emitted sp



(a) Spectrum of YAG white LED

(b) Spectrum of white RDB LED

Figure 1.50

The peak inverse voltage (PIV) of LEDs is about 5 V. So, if it is used with a reverse bias voltage more than 5 V, the p-n junction in the LED will get damaged. Therefore, care must be taken when LEDs are lit using alternating voltages.

Applications of LED

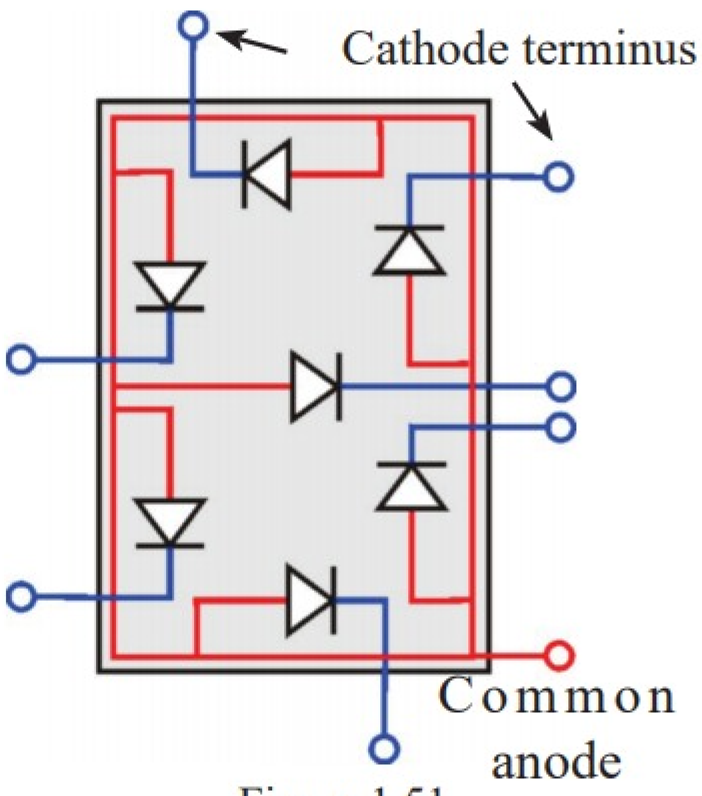


Figure 1.51

- Another application of LEDs is their usage as seven segment LED display to indicate digits. In this display seven LEDs are arranged in a single pack so that the digits can be shown. Either cathodes or anodes of the seven LEDs are connected to a common terminal (common cathode or common anode) and the other terminals are connected individually to another seven terminals coming out. By lighting the appropriate LEDs all the digits from 0 to 9 can be displayed.
- RGB-white LEDs are used as back light source of LCD screens in television receivers. Also in some other television receivers and in displaying panels (advertising boards), pictures or images are displayed using LEDs.
- To detect fluorescent marks in important documents where they are unable to be seen with the naked eye, UV-LEDs are used. It is this UV-LED light with which the banks test whether the banknotes are genuine or not.
- In many remote controlling units the controlling signals are sent using infra-red

radiation produced by IR-LEDs.

1.14 Photo diode

It is known that, there is a very small leakage current through a reverse biased p-n junction due to minority carriers generated by thermal agitation in extrinsic semiconductors. If light is incident on such a junction, some bonds in the attenuated region break down and leakage current increases.

Since junction diodes are produced with an opaque epoxy covering, light can't be incident on it. However, if it is a of glass, covering the reverse leakage (or reverse saturation) current varies depending on the intensity of light incident on the junction. This can be used to obtain a voltage which varies according to the intensity of light.

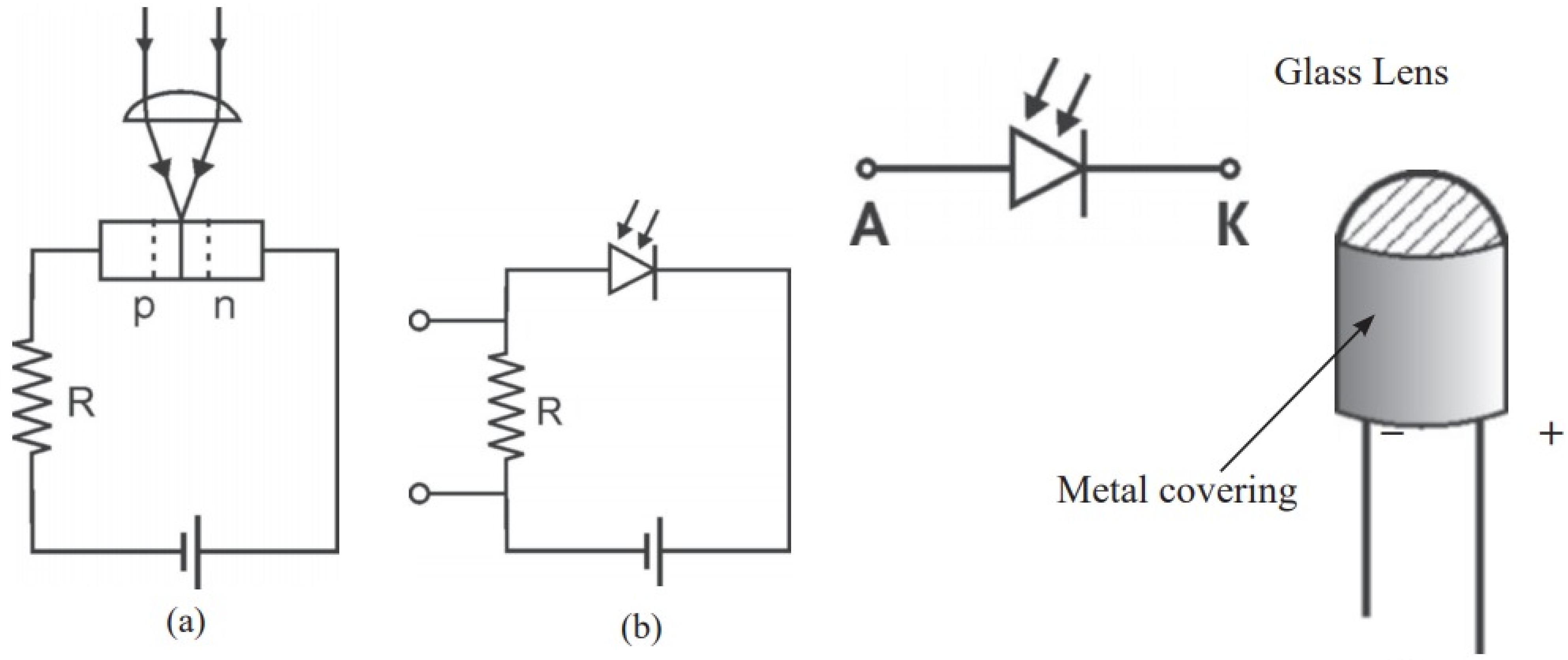


Figure 1.52

To achieve this, the diode is reverse biased (using a battery) and the light is focused onto the junction by means of a small convex lens. There will be a voltage developed across resistor R in the external battery circuit due to the reverse leakage current through the diode. This voltage is used to measure the intensity of light.

Physical appearance of packaging and circuit symbol of the photodiode are shown in Figure 1.51.

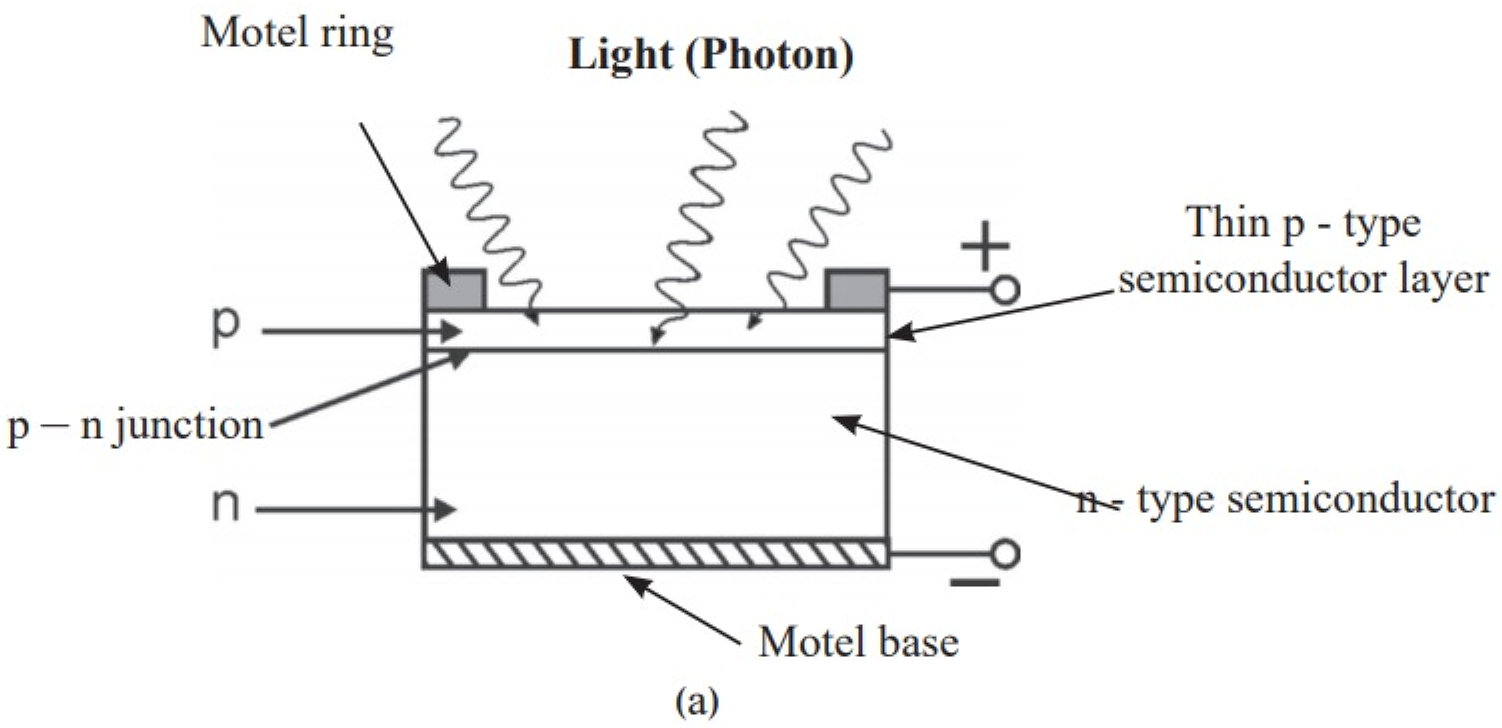
Applications of photo diode

In the way given above, photo diode is used as a photosensitive switch widely in intruder alarms, light meters of cameras and computer data readers.

1.15 Solar cells

The p-n junction can be used to convert solar energy into electrical energy. This kind of set-up is known as a solar cell.

Generally, selenium or silicon is used to make solar cells. Figure 1.52 shows the structure of a silicon solar cell. On the metal base the n-type semiconductor is formed and, on it there is a very thin layer of p-type semiconductor (silicon). Sunlight which is incident on this layer, passes through it and reaches the p-n junction has been built up so that then the negative potential in p-type semiconductor and the positive potential in n-type semiconductor are built.



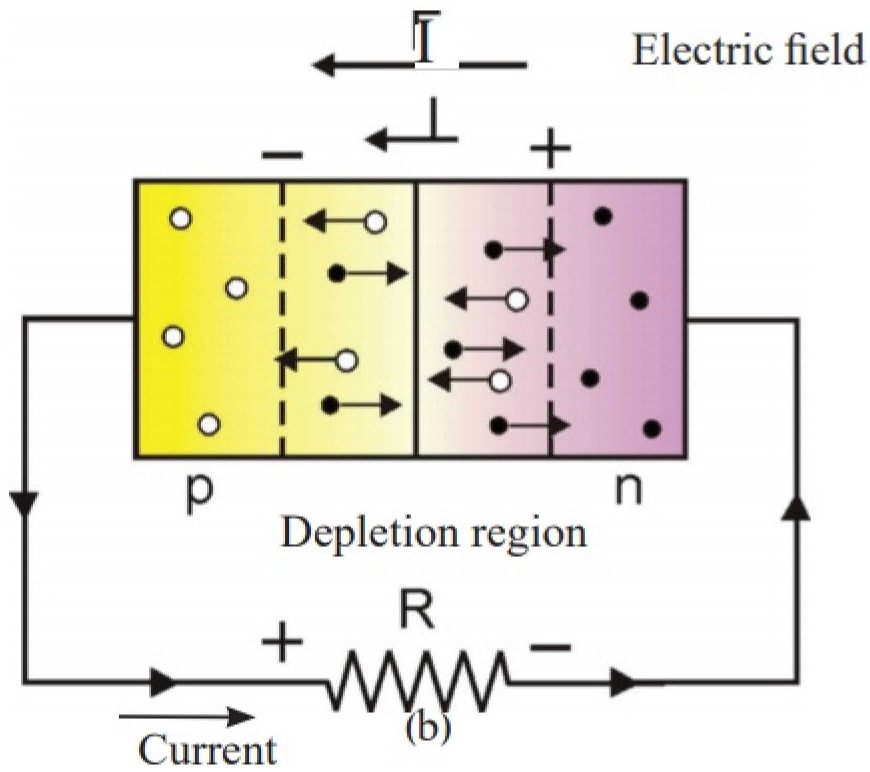
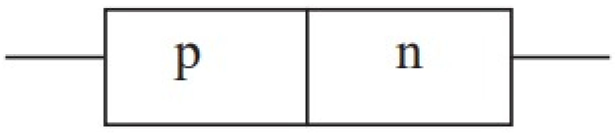
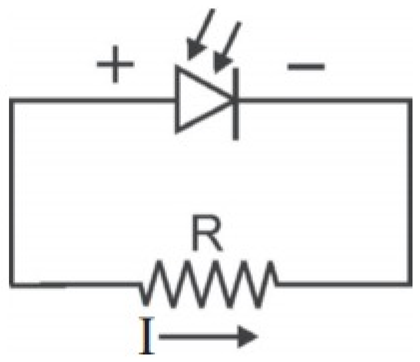


Figure 1.53

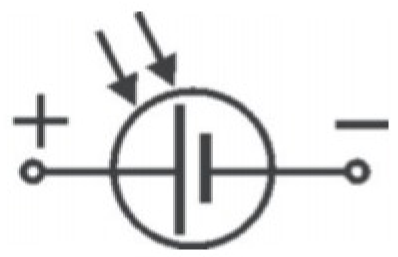
When light is incident on the attenuated region (depletion layer) some bonds in the semiconductor break down and electron-hole pair generation occurs.

The holes and free electrons thus generated drift forwards the n-terminal and p-terminal respectively due to electric field created by the internal potential barrier. When p and n terminals are connected externally a current flows through the external circuit. Therefore when sunlight is incident on the p-n junction it acts as a source of electromotive force, with p-terminal as positive and n-terminals as negative (Figure 1.52 (b)).





(a) Equivalent circuit



(b) Corresponding stardust symbol

Figure 1.54

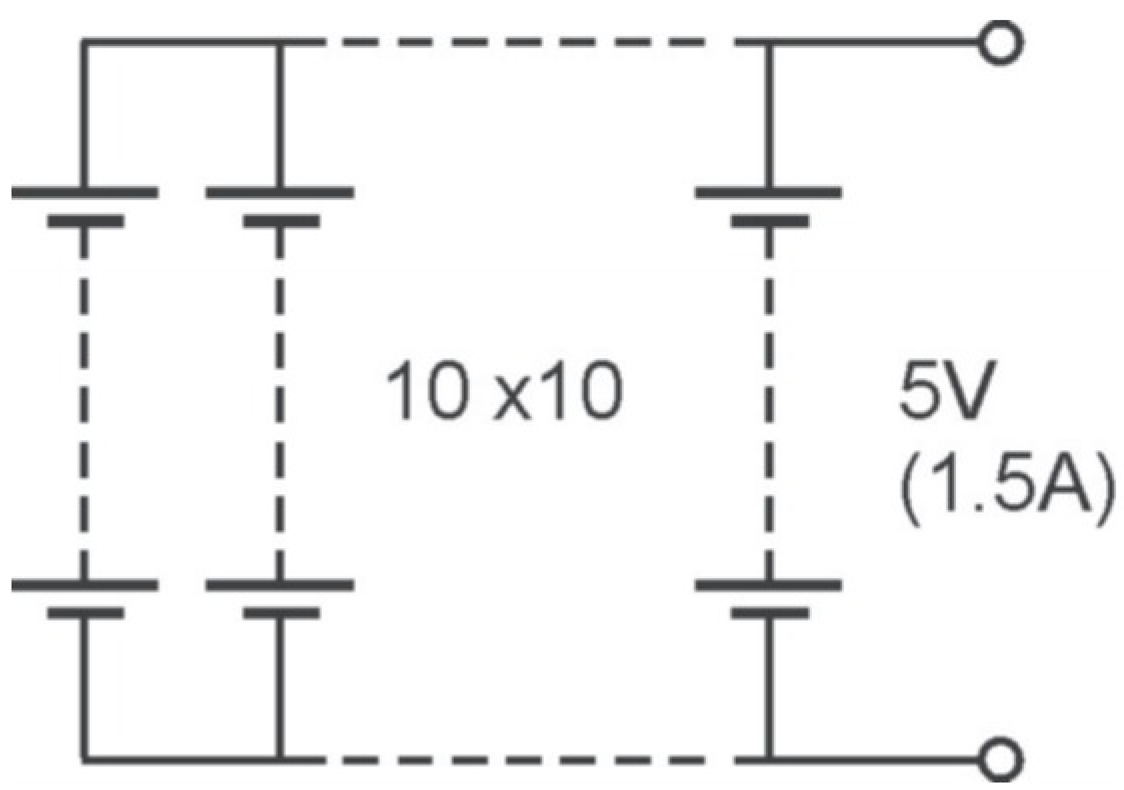
In Figure 1.53, (a) shows the relevant equivalent circuit (b) shows the standard circuit symbol of solar cell. In bright light the emf produced in it is about 0.6 V. The current drawn through the external circuit increases with the increase of surface area of the p-n junction.

1.15.1 Applications of solar cell

Solar cells are used to convert solar energy into electrical energy. In normal daylight a single solar cell produces a voltage of about 0.5 V and can give out a current of about 150 mA. As this is not sufficient for practical purposes, electrical sources capable of supplying a higher voltage and current are made by connecting solar cells in series and parallel combinations.

An example can be given as follows.

When solar cells of 0.5 V, 150 mA are connected in series to form sets containing 10 cells each, and 10 such series cell sets are connected in parallel (Figure 1.54), an electrical source of 5 V (= 0.5 ×10) and 1.5 A (= 0.15 ×10) can be made. Such a combination of solar cell is called a solar panel.



Structure of a solar panel

Figure 1.55

Solar cells are considered to be the future energy source as the Earth receives solar energy at a power of about 1000 W m^{-2} .

The major disadvantages of solar cells are their high cost due to high production cost of pure silicon, and the low efficiency which is about 15%. On the other hand they have remarkable advantages such as not contributing to environmental pollution, availability of solar energy free of charge, not decaying and long lasting.

Solar panels are already in use, to supply electricity to houses and factories. Providing solar power through the electricity network by putting up solar power stations has commenced already. Solar power is used to provide power to motor vehicles and space satellites, through solar panels.

© 2020 National Institute of Education. All Rights Reserved.

© 2020 National Institute of Education. All Rights Reserved.

© 2020 National Institute of Education. All Rights Reserved.

© 2020 National Institute of Education. All Rights Reserved.

Chapter Two

Transistors

2.1 Bipolar Transistors

It has been mentioned earlier that the first transistor was made by John Bardeen, William Shockley and Walter Brattain in the Bel Laboratory in December 1947. The transistor had the same ability to amplify a signal, like the triode valve which was widely used at that time. The word **"TRANSISTOR"** was derived from the two words **TRANS FER-RESI STOR**.

The transistor had been made by forming three regions of p and n semiconductors on a single chip of intrinsic semiconductor. There are two p-n junctions in it, and there are only two ways that p and n type semiconductors can be formed in this manner. Figure 2.1 (a) and (b) show the two formations.

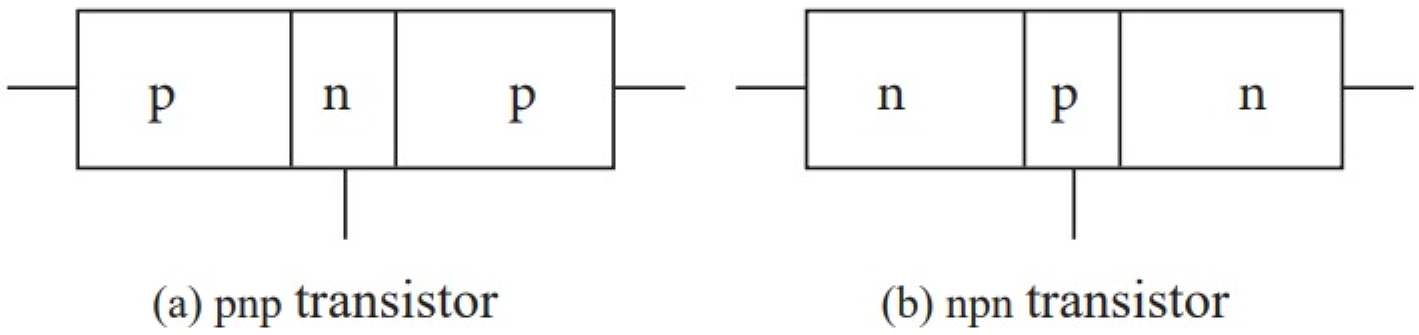
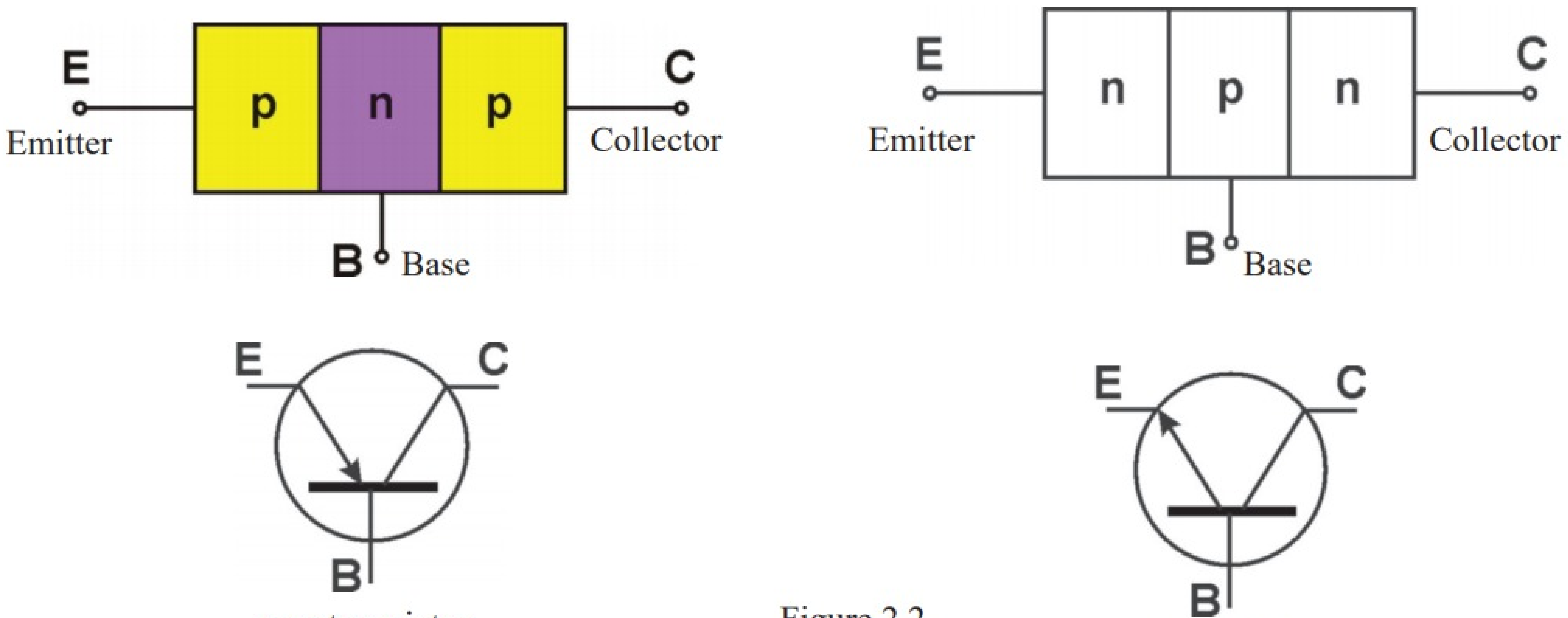


Figure 2.1

The two formations (a) and (b) are known as pnp transistor and npn transistor respectively. To connect a transistor to a circuit there are three terminals which are connected to p , n and p regions (in pnp type) or n , p and n regions (in npn type). Concerning action of the transistor, the region which emits charge carriers is known as the **emitter** and the region where charge carriers are collected is known as the **collector**. The thin region in between them controls the flow of charge carriers and is known as the **base**. They are denoted by E, C and B respectively.

The semiconductor region emits carriers in it. The base region in the middle is a very thin layer of semiconductor and it has been doped with a very low doping level. Since the base region is very thin, charge carriers can easily flow across it but as its doping level is very low, it weakens that flow. This will lead to control the flow of current (will to be discussed later). The collector region is doped with a moderate level and its area has to be made large. Labeling of terminals and the relevant circuit symbols are shown in Figure 2.2.

© 2020 National Institute of Education. All Rights Reserved.



pnp transistor

Figure 2.2

npn transistor

To distinguish between emitter and collector an arrowhead is marked on the emitter and that arrowhead indicates the current flowing direction through the transistor due to carriers (electrons and holes) contributing to action of the transistor. This will be explained later. For the functioning of the transistor both of the charge carriers, electrons and holes are contributing. Therefore, these transistors are called **bipolar transistors**.

2.2 Action and biasing of the transistor.

For simplicity of understanding, let us take a pnp transistor first. In pnp transistors, the majority carriers are holes and hence the direction of current is the same as the direction of the flow of holes. This will make the explanation easy. Let us consider here, that the emitter, collector and base are doped with equal doping levels.

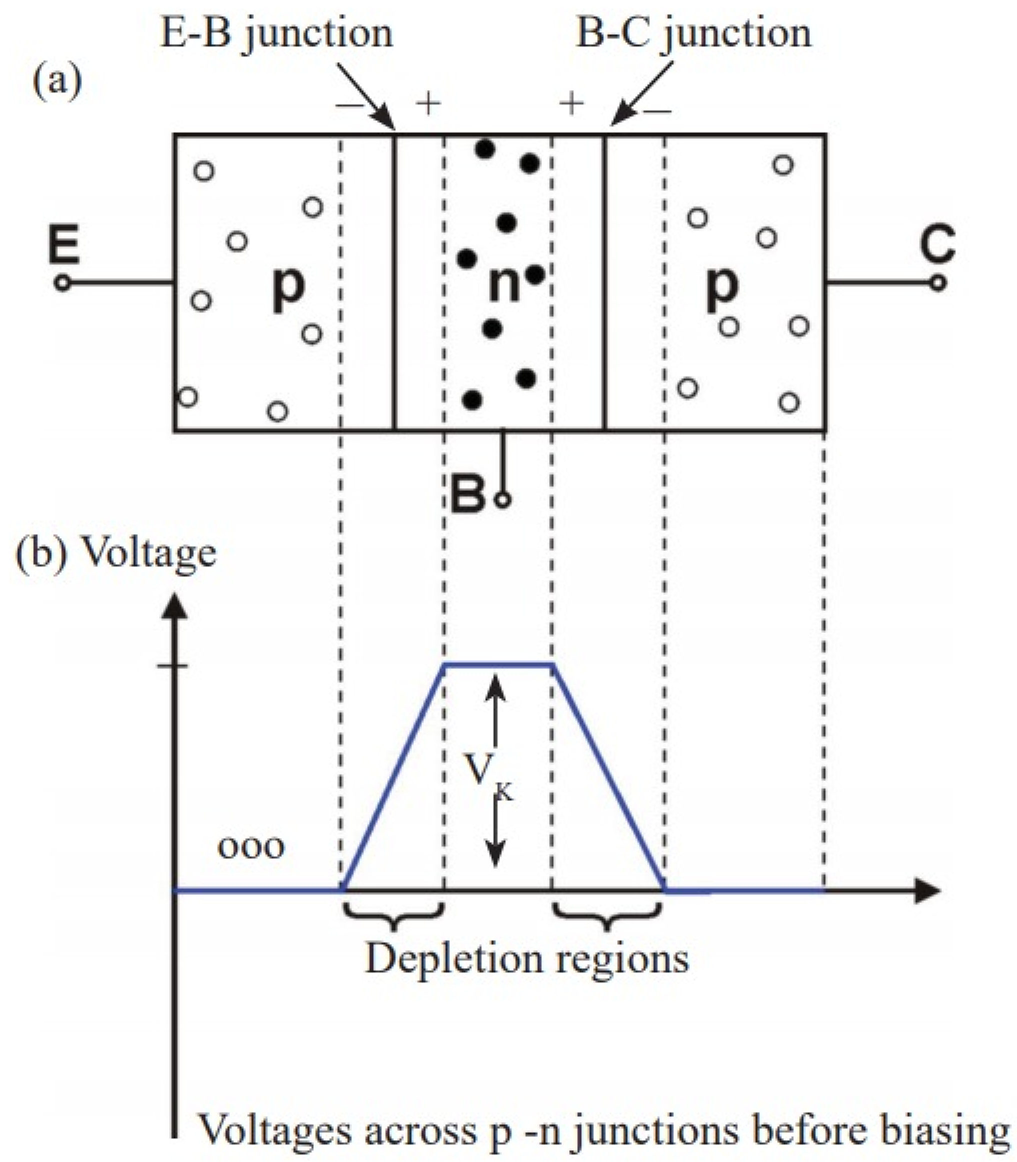
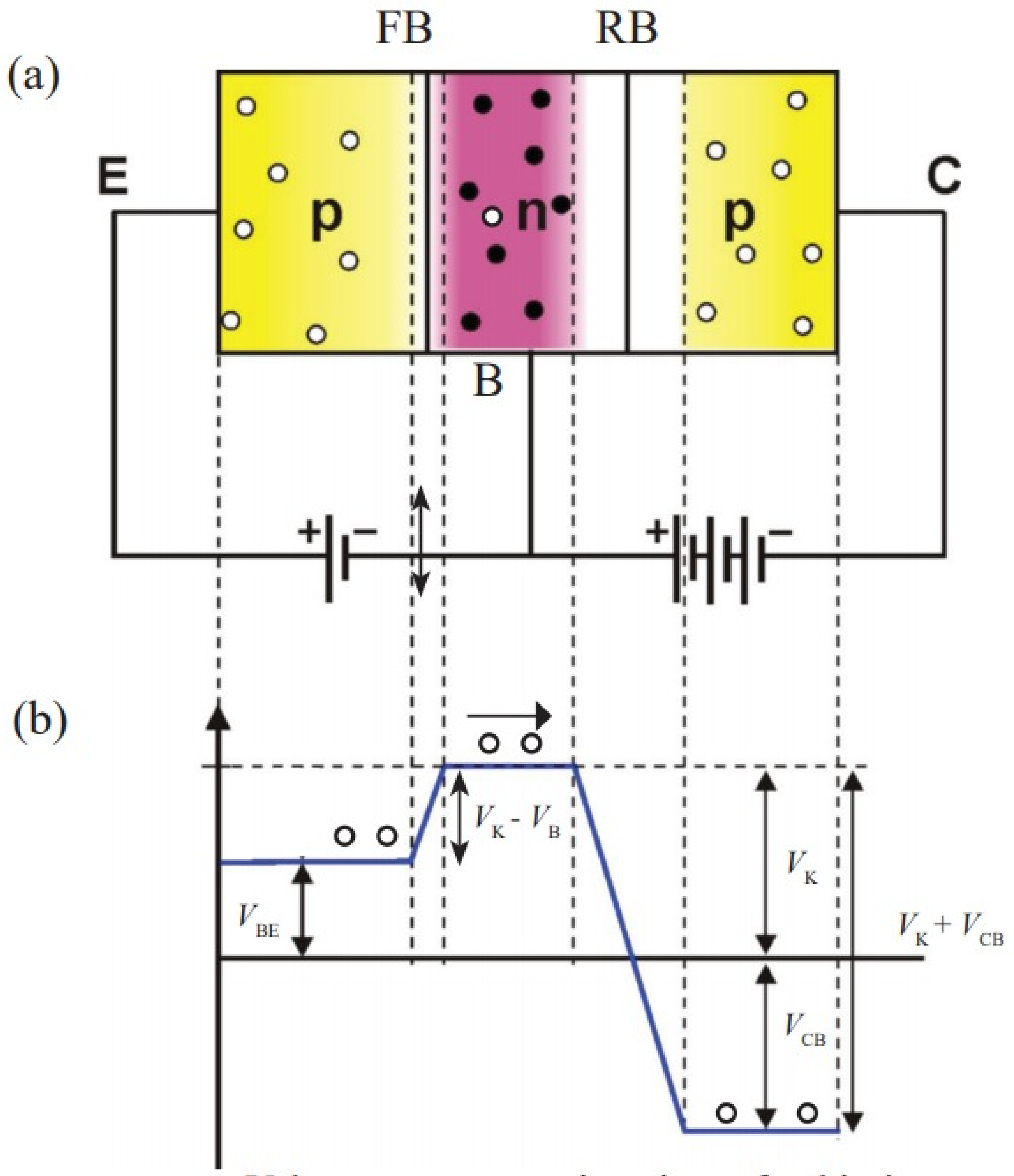


Figure 2.3

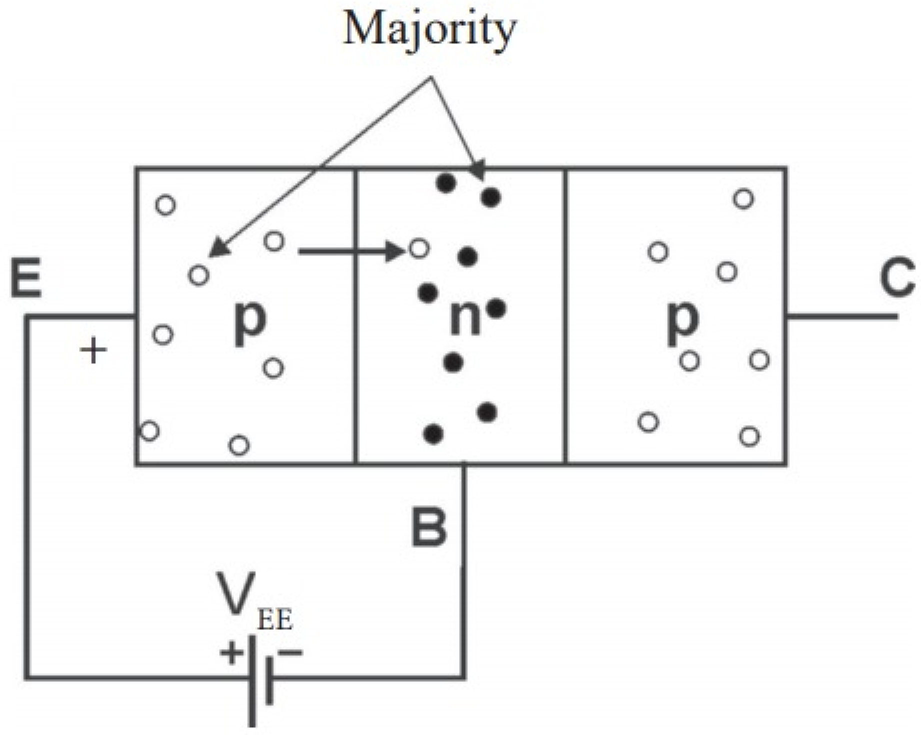
When the terminals of a pnp transistor are free (not connected) there are potential hills (barriers) on holes (major carriers in p-region) across the junctions E-B and B-C. Therefore, holes cannot flow from emitter to collector. By applying appropriate voltages across the junctions the potential hills (barriers) can be removed.



Voltages across p -n junctions after biasing

Figure 2.4

When the transistor is biased using external voltages so that the E-B junction is forward biased and B-C junction is reverse biased, Figure 2.4 shows how the barrier potentials get arranged. Since the E-B junction is forward biased, the barrier potential hill becomes very small and then holes (major carriers) can enter the base region easily. These carriers can easily reach the collector as it is at a higher negative potential. Therefore, to be able to conduct a current the transistor should be biased so that E-B junction is forward biased and B-C junction is reverse biased with a higher reverse voltage. For npn transistors also, p-n junctions should be biased in a similar way.



(a)

To have a clear understanding let us consider the actions at the two junctions separately. When E-B junction is forward biased a large number of holes which are majority carriers in the E-region can enter the B-region as the base in an n-type semiconductor its majority carriers are free electrons. (Figure 2.5 (a)).

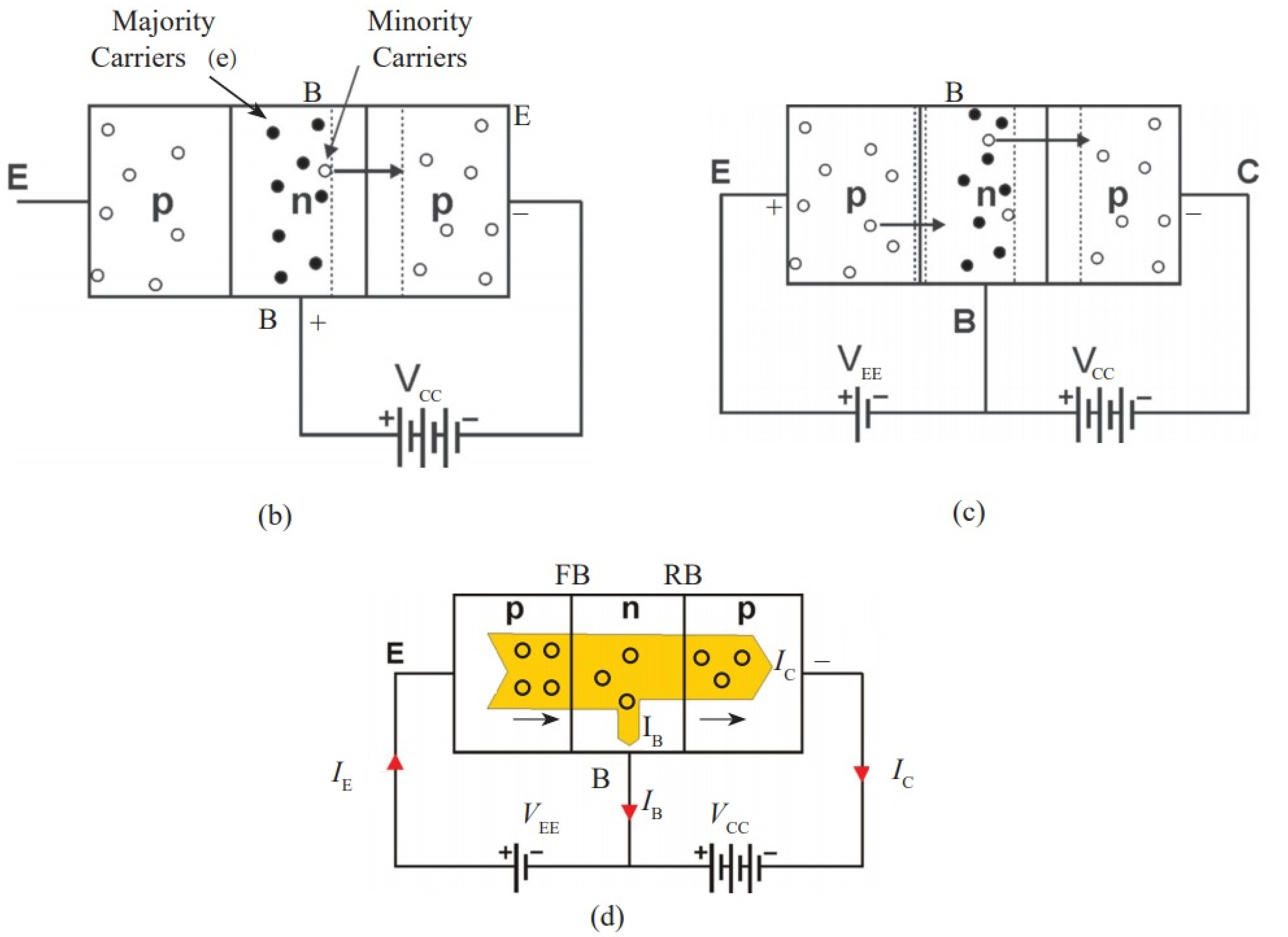


Figure 2.5

Holes which arrive at the base region will be minority carriers there. Although electrons (majority carriers) in the base region cannot flow across the reverse biased B-C junction, the holes arrived at the base region (from E-region) can flow across the B-C junction and reach the collector region easily. Since the base region is in the form of a thin layer the holes (majority carriers) can move across the junction and reach the collector. This is shown by Figure 2.5 (c) and the whole process is shown (in brief) by Figure 2.5 (d).

The total current flowing through emitter is denoted by I_E and a small amount of it flows through the voltage supply V_{EE}. This small amount is denoted by I_B. A large amount of the current I_E reaches collector as I_C. When the forward bias of E-B junction is changed, I_E changes by much and I_B changes a little. When I_B changes a little I_C changes largely. The reason is that the sum of currents I_C and I_B equals I_C.

$$I_E = I_C + I_B$$

In general, I_E is of the order mA and I_B is of the order μA. Accordingly, when I_B is changed even by a little amount I_C is subject to change by a large amount. This is how the transistor achieves its amplification process.

With an npn transistor too, for its operation it should be supplied with voltage so that the B-E junction is forward biased and B-C junction is reverse biased. The biasing voltage should be with the opposite polarity to those of the pnp transistor. Also, in the case of npn transistor, electrons are emitted instead of holes from the emitter towards the base.

The majority carriers of the emitter are electrons and they reach the base easily as E-B junction is forward biased and the base is at a positive potential. Since the base is a p-type semiconductor its majority carriers are holes. Its minority carriers are electrons. The B-C junction is reverse biased and therefore holes which are majority carriers in the base, cannot flow across the B-C junction but the electrons coming from emitter can easily reach the collector which is at a higher positive potential.

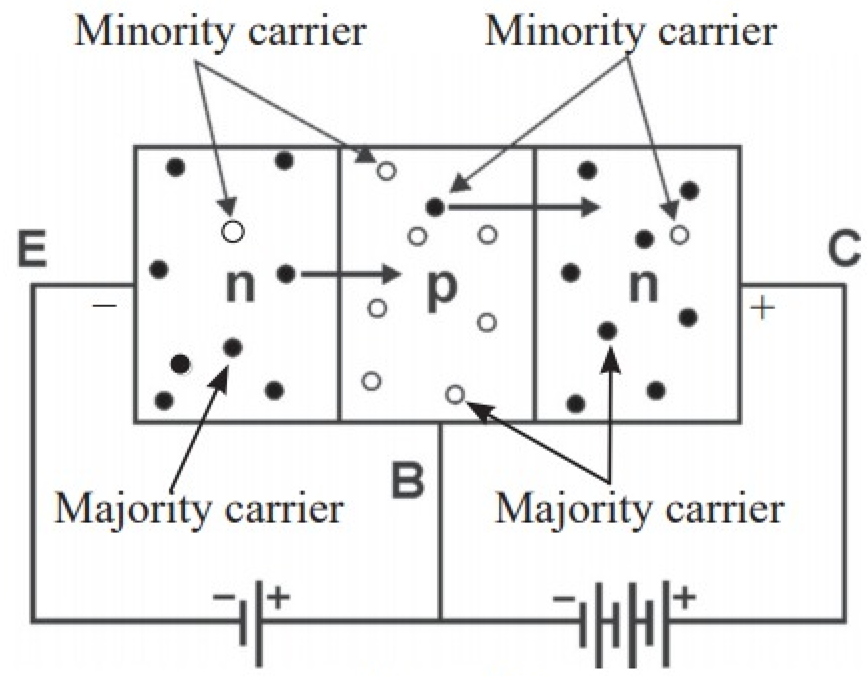


Figure 2.6

For an npn transistor, when the external voltages are supplied so that E-B junction is forward biased and B-C junction is reverse biased the electrons flow easily from emitter to collector. Only a very few electrons are flowing out from the base. Since the current due to flow of electrons is in the direction which is opposite to that of the flow of negative electrons, the current through the transistor flows from collector to emitter.

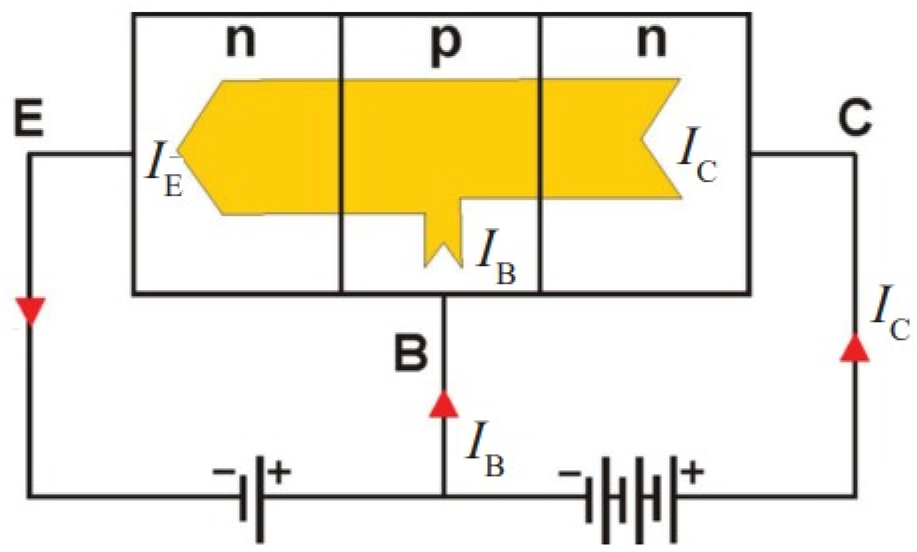


Figure 2.7

Here also the emitter current is equal to the sum of collector current and base current.

$$\boxed{I_E = I_C + I_B}$$

In both of pnp and npn transistors, E-B junction should be forward biased and B-C junction should be reverse biased. The only difference is with their polarity of biasing voltages. In both of the transistors the majority carriers (holes in case of pnp transistors and electrons in case of npn transistors) flow from emitter to collector across the base. According to polarity of their majority carriers, in the pnp transistor current flows from emitter to collector and in the npn transistor current flows from collector to emitter. So, it will be clear to you that the direction of arrowhead, used to identify emitter in the circuit symbol of the transistor, is the direction in which the current flows through emitter.

© 2020 National Institute of Education. All Rights Reserved.

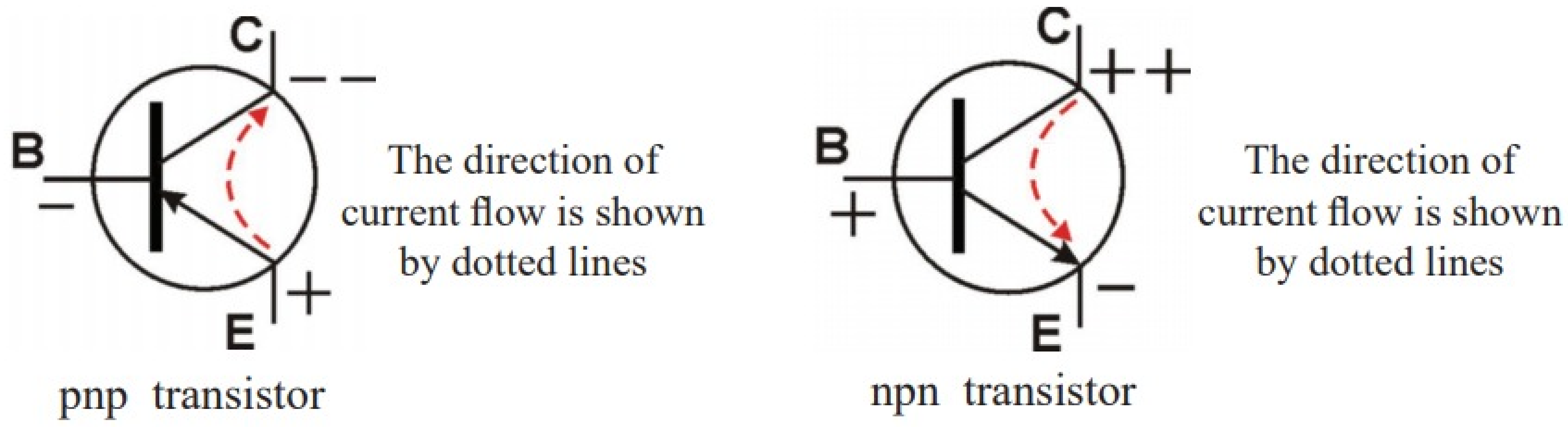


Figure 2.8

Together with the transistor circuit symbols, the current flowing directions and polarities of the potentials which should be at collector and base relative to emitter, are shown in Figure 2.8. By making two positive or two negative signs at the collector, it has been indicated that the collector should be biased with a higher potential than the base.

The current always flows from positive (+) potential to a negative (-) potential. Therefore, the polarities of biasing potentials can be kept in mind without difficulty, by considering the direction of arrowhead in the transistor symbol.

2.3 Configurations of transistor

There are three ways or configurations that a transistor can be connected to an electronic circuit. The transistor has only three terminals E, B and C, but when using it as an amplifier it requires two terminals to input the signal and two terminals to get the output. Therefore, one terminal has to be used in common.

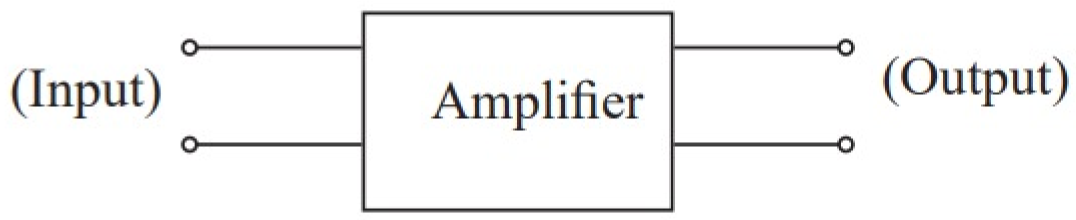


Figure 2.9

Depending on the terminal which is used in common, the transistor can be connected to a circuit in three configurations.

These configurations are,

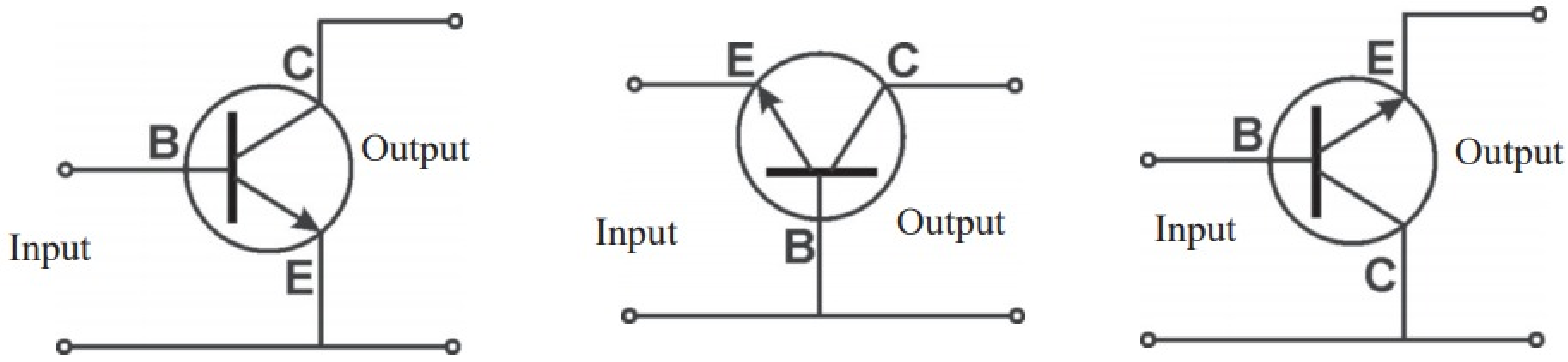
(a) Common Emitter configuration

(b) Common base configuration

(c) Common collector configuration

© 2020 National Institute of Education. All Rights Reserved.

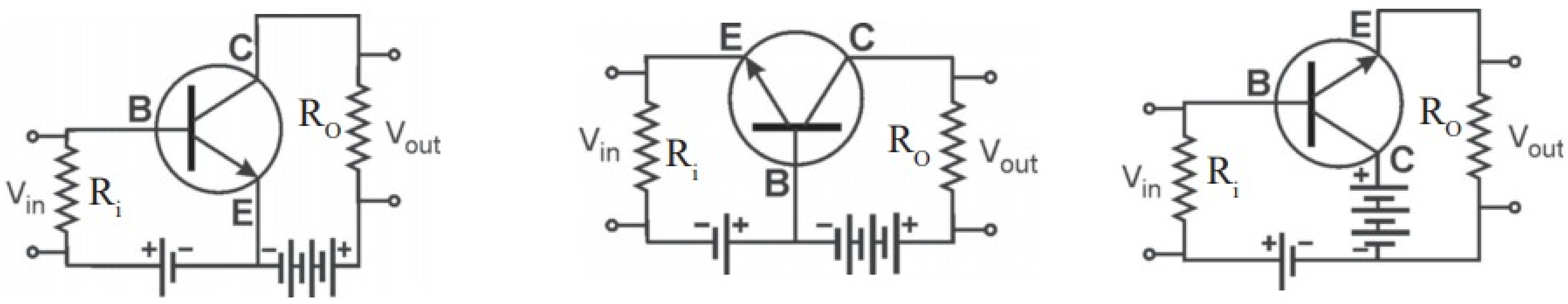
The above three configurations are shown in Figure 2.10.



(a) Common Emitter Configuration | (b) Common Base Configuration | (c) Common Collector Configuration

Figure 2.10

Figure 2.11 gives these configurations with their biasing potentials.

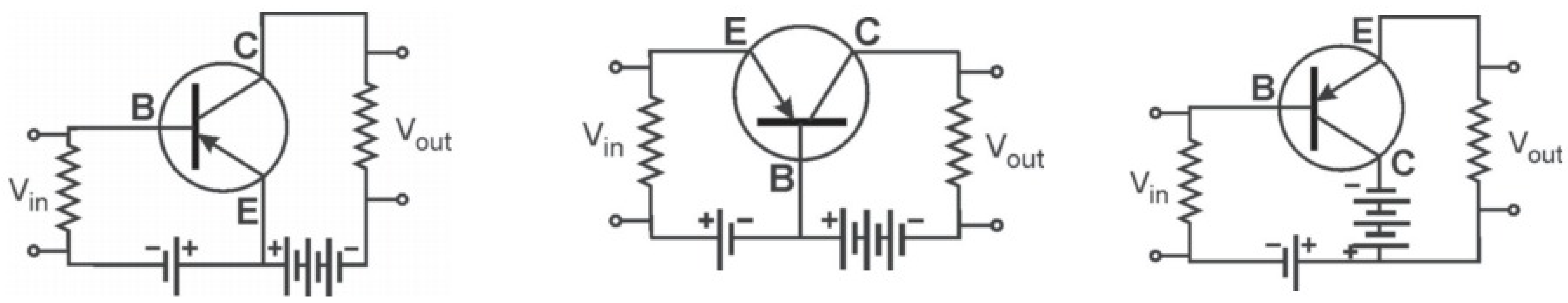


(a) Common Emitter Configuration | (b) Common Base Configuration | (c) Common Collector Configuration

Figure 2.11

When the transistor circuit is used as an amplifier the input signal is applied across R_i and the output is taken out across R_0.

When npn transistors are used in the above configurations, the relevant circuits with their biasing potentials are shown in Figure 2.12.



(a) Common Emitter Configuration | (b) Common Base Configuration | (c) Common Collector Configuration

Figure 2.12

The major difference in these circuits is only the change in polarities of the biasing potentials.

2.4 The notations used with the transistor circuit symbol

To indicate potentials at the terminals and currents flowing through them, several notations are used. English capital letters are used to indicate DC- voltages and DC-currents (Figure 2.13).

(a) npn (b) pnp

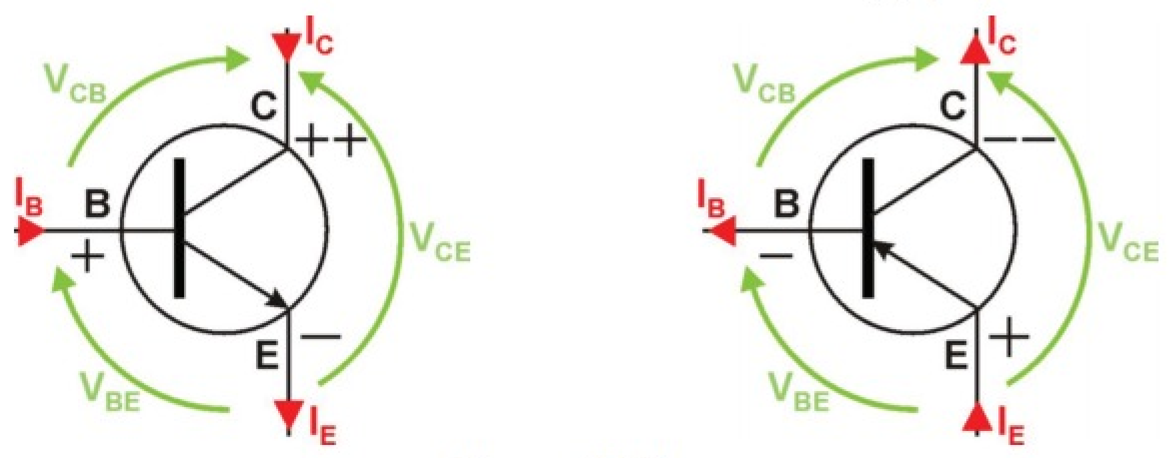


Figure 2.13

As mentioned earlier the letters C, B and E are used to denote Collector, Base and Emitter. To denote currents through them, the symbols I_C, I_B and I_E are used respectively. The potential of a terminal is usually given relative to another terminal. (Absolute potentials can be stated only when a point in the circuit is grounded).

When stating the potential of a terminal, the terminal name should be stated first and the reference terminal is stated next.

eg.

(a) Collector potential relative to emitter $\longrightarrow V_{CE}$

(b) Base potential relative to emitter $\longrightarrow V_{BE}$

(c) Collector potential relative to base $\longrightarrow V_{CB}$

When npn and pnp transistors have been biased correctly, the polarities of potentials can be stated as follows.

	npn transistor	pnp transistor
V_{BE}	positive (+)	negative (-)
V_{CE}	positive (+)	negative (-)
V_{CB}	positive (+)	negative (-)

In addition to this, some potentials are given as V_{BEO}, V_{CEO} and V_{CBO} in data sheet. That means the potentials are given when the remaining terminal, other than the two terminals stated, is kept open-circuited.

eg. V_{BEO} means the potential of base relative to emitter when collector is kept open-circuited.

In the advanced level physics syllabus npn transistors are considered and therefore the circuits coming hereafter are dealt with npn transistors.

2.5 Characteristic curves of common emitter configuration

In amplifier circuits the common emitter configuration of the transistor is widely used. The main reason for this is, it gives both high current gain and high voltage gain compared to other configurations. This will be further discussed later.

To understand the operation of a transistor, the relationships among V_{BE}, V_{CE}, I_B and I_C should be considered. The curves which represent these relationships graphically are called characteristic curves. These curves are presented in three major forms as follows.

(i) Input characteristic (I_B vs V_{BE})

(ii) Output characteristic (I_C vs V_{CE})

(iii) Transfer characteristic (I_C vs I_B)

A circuit which can be used to get readings in plotting these curves for an npn transistor, is shown in Figure 2.14. Let us consider how these curves can be obtained and what properties they have.

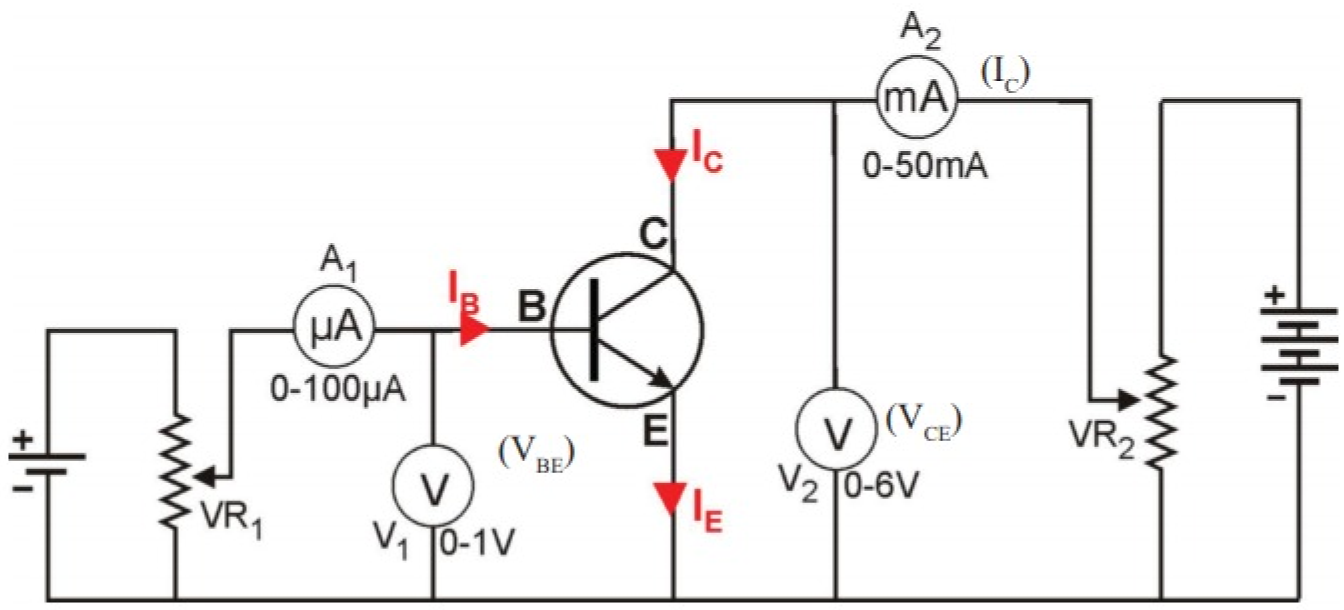


Figure 2.14

2.5.1 (i) Input characteristic

In a common emitter amplifier, input is taken to be the potential at base relative to emitter (V_{BE}). The way the input current I_B varies with V_{BE} is given by the input characteristic. When obtaining this characteristic, the collector potential relative to emitter (V_{CE}) is kept constant. By adjusting the variable resistor VR_2, V_{CE} can be maintained at a constant value in the range 0-6 V. V_{BE} can be varied by varying VR_1. The relevant values of V_{BE} and I_B can be read from the voltmeter V_1 and micro-ammeter A_1 respectively. Using those readings, a graph of I_B versus V_{BE} can be plotted.

Figure 2.15 shows the input characteristic of a silicon transistor.

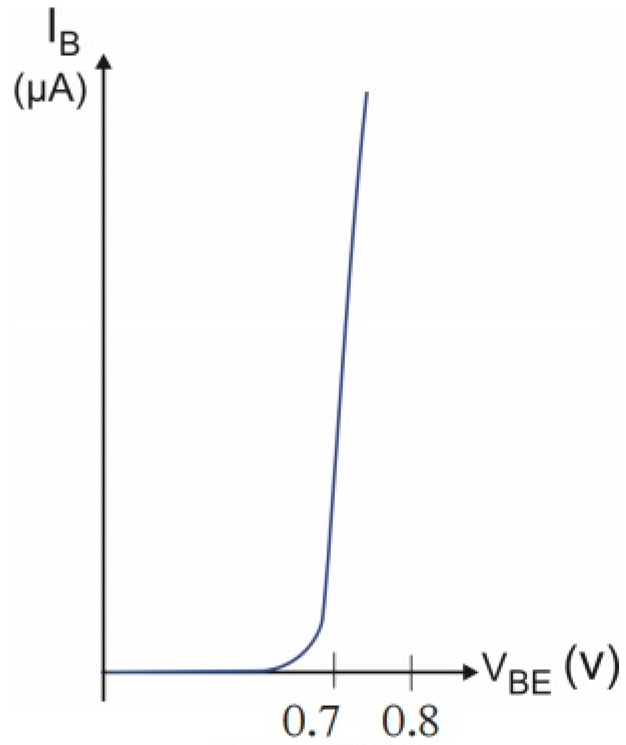


Figure 2.15

It had been stated earlier that a silicon p-n junction becomes forward voltage of 0.6 - 0.7 V. It can be seen from the curve that, up to the value 0.7 V of V_{BE}, I_B is almost zero and beyond that I_B starts to increase rapidly and in linear with (roughly) V_{BE} even for a very small increase of V_{BE}.

If I_B is given as an input to the transistor then it will get amplified, and appear as I_C at the output. Therefore, in the common emitter configuration, transistor is considered to be a current amplifier.

The ratio is known as a DC-current gain of the transistor. (Unless otherwise specified current gain means the DC-current gain)

$$\text{Current gain } (\beta) = \frac{I_C}{I_B}$$

$$\therefore \boxed{I_C = \beta I_B}$$

In the common emitter configuration, typical value of β is about 100 (50-250). However this value changes depending on the transistor and the relevant current gain of a particular transistor can be found from a transistor data book.

According to the above characteristic curve, the cut-off region is where $I_B = 0$. In the linear (active) region $I_C = \beta I_B$ and $I_C \propto I_B$. The current gain (β) is a constant for a given transistor.

In the saturated region, that is where I_C is no longer increasing with I_B, I_C is less than βI_B

$$\therefore I_C < \beta I_B$$

(Further details of linear region and saturated region are given later)

2.5.2 (ii) Output characteristic

The characteristic shows the variation of output current (I_C) with output voltage (V_{CE}) when V_{BE} is kept constant. Keeping V_{BE} constant with the help of VR1 and varying VR_2 (Figure 2.14), the Relevant values of V_{CE} and Ic are measured. Then by plotting I_C against V_{CE}, the output characteristic curve can be obtained.

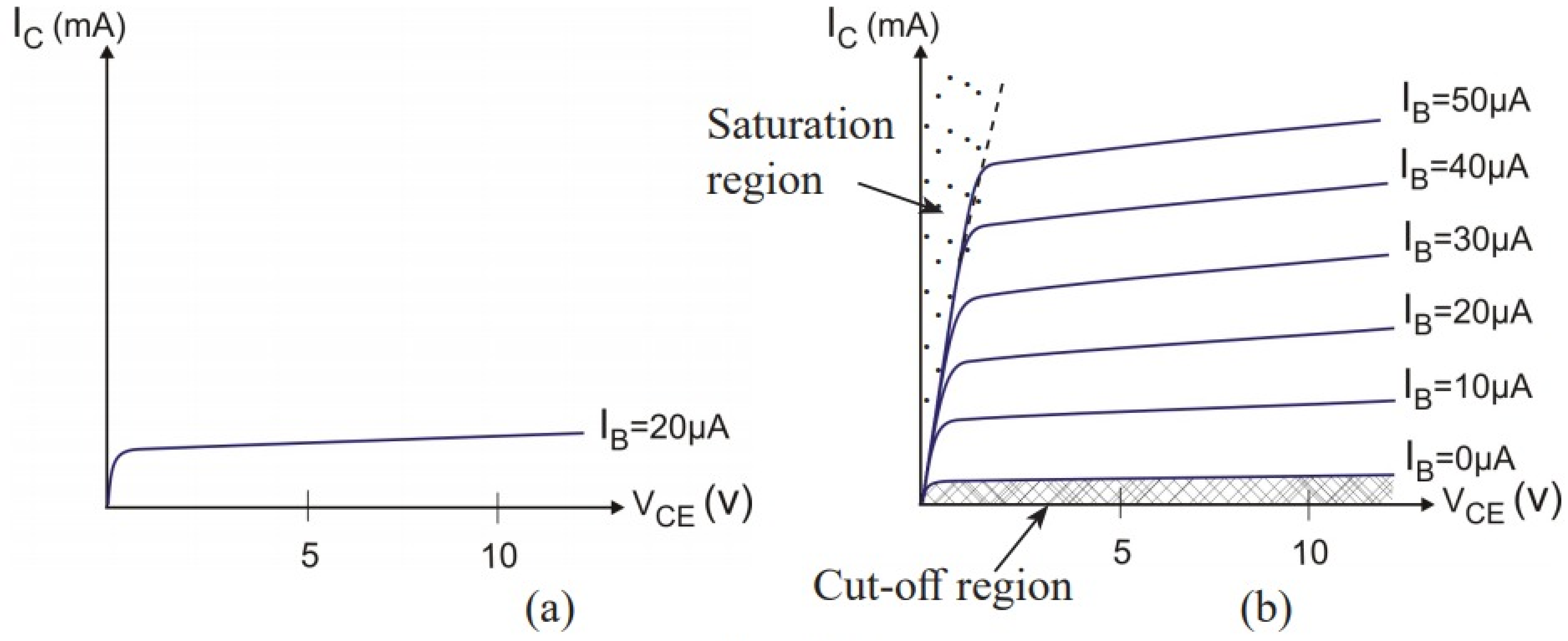


Figure 2.16

When the variation of I_C with V_{CE} is recorded keeping I_B= 20 μA by adjusting VR_1, a curve of the nature shown by Figure 2.16 (a) can be obtained. When the relevant curves are constructed under a constant value of I_B such as 0, 10, 20, 30μA….at a time, a set of curves as shown in Figure 2.16 (b) is obtained. In each I_B value, it is seen from these curves that I_C becomes constant after some small value of V_{CE} is exceeded. Therefore, the horizontal lines of the curves indicate that, after exceeding some small value of V_{CE}, I_C becomes saturated. When I_B= 0, irrespective of the value of V_{CE}, I_C becomes almost zero.

The region in which I_C Ω 0 is known as **cut-off region**. Moreover it can be seen from this set of curves that I_C becomes saturated (this saturation current is determined by I_B) This state is shown by the shaded region close to vertical axis in Figure 2.16 (b). This region is known as the **saturation region**. The region in which I_C varies with I_B is given by the middle part consisting of a set of curves.

This region is known as the **active region** or **linear region**. In this region the transistor operates as an amplifier, making a large variation in I_C due to a small variation in I_B.

2.5.3 (iii) Transfer characteristic

This is the characteristic which shows the variation output (I_C) with the variation of input (I_C). Using the circuit shown in Figure 2.14, I_B is varied by varying V_{BE} with the help of VR_1, and at the same time keeping V_{CE}, constant with the help of VR_2. After getting the relevant values of I_B, and I_C, using A_1 and A_2 respectively, I_C is plotted against I_B. Then the transfer characteristic curve as shown in Figure 2.17 is obtained.

© 2020 National Institute of Education. All Rights Reserved.

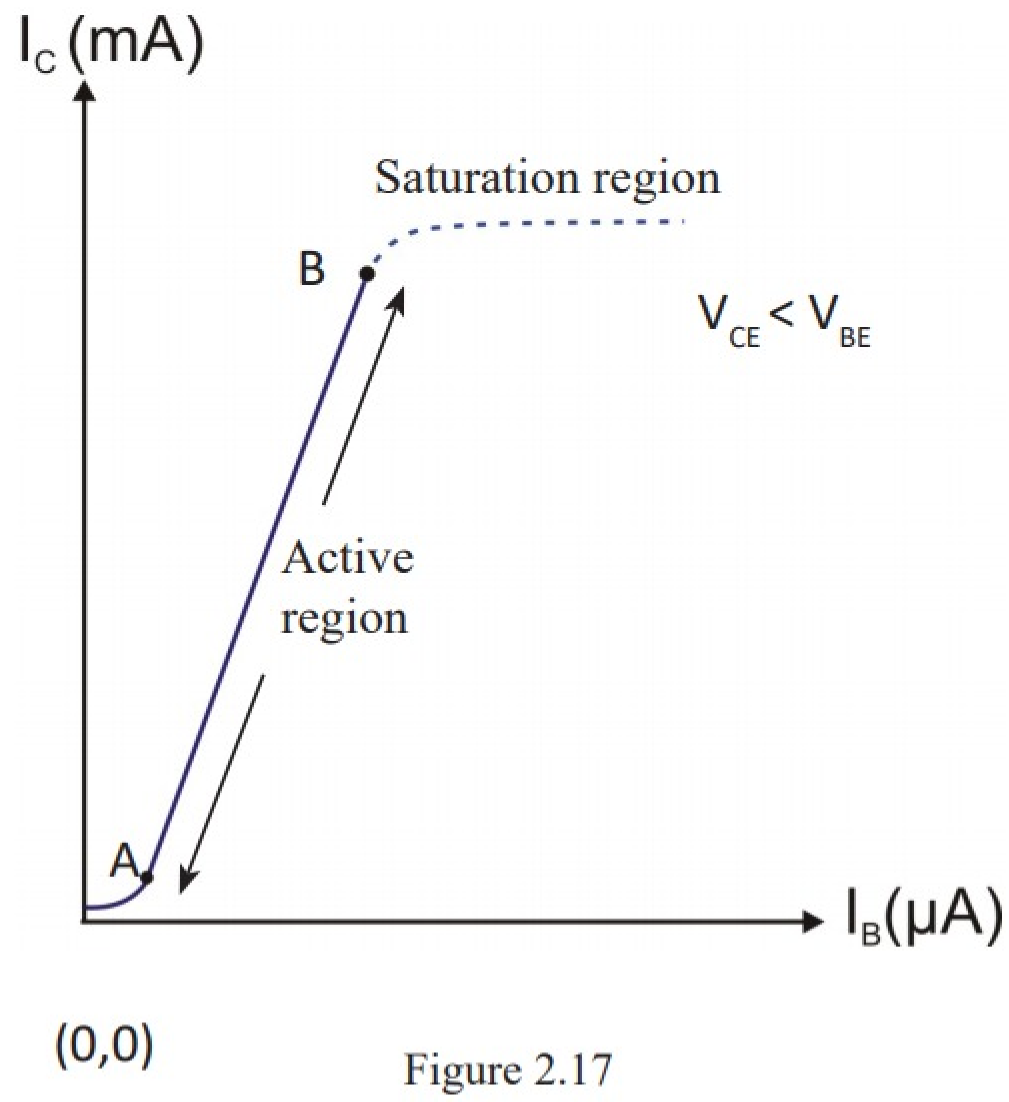


Figure 2.17

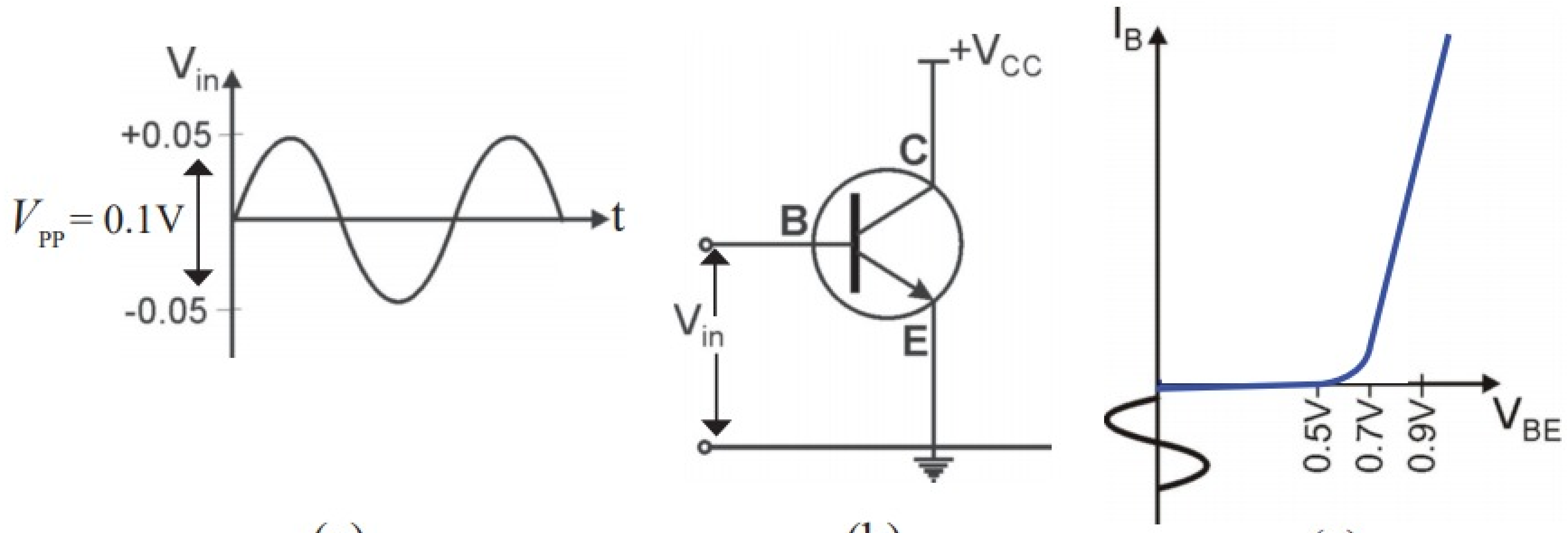
When $I_B = 0$, I_C takes a very small value close to zero. In the linear part of the curve I_C varies in direct proportion with I_B. When I_B is increased beyond the point B on curve V_{CE} comes to a very low value ($V_{CE} < V_{BE}$). Then it is not possible to keep V_{CE} constant at the earlier value by adjusting VR_2. This state, in which $V_{CE} < V_{BE}$ is the saturated state of the transistor. After that state is reached I_C remains constant even though I_B is increased.

The linear region AB is known as the active region of the transistor. The transistor is operated in this region to make it works as an amplifier. There, the change in I_B which is of the order μA is amplified to a change which is of the order mA in I_C.

2.6 Biasing the transistor

2.6.1 Necessity of biasing

Let us consider a situation where a small voltage signal varying with time, is applied to the base of a transistor in the common emitter configuration, so as to get that signal amplified. As an example let us consider an AC-signal of $V_{PP} = 0.1V$ is applied (Figure 2.18 (a)).

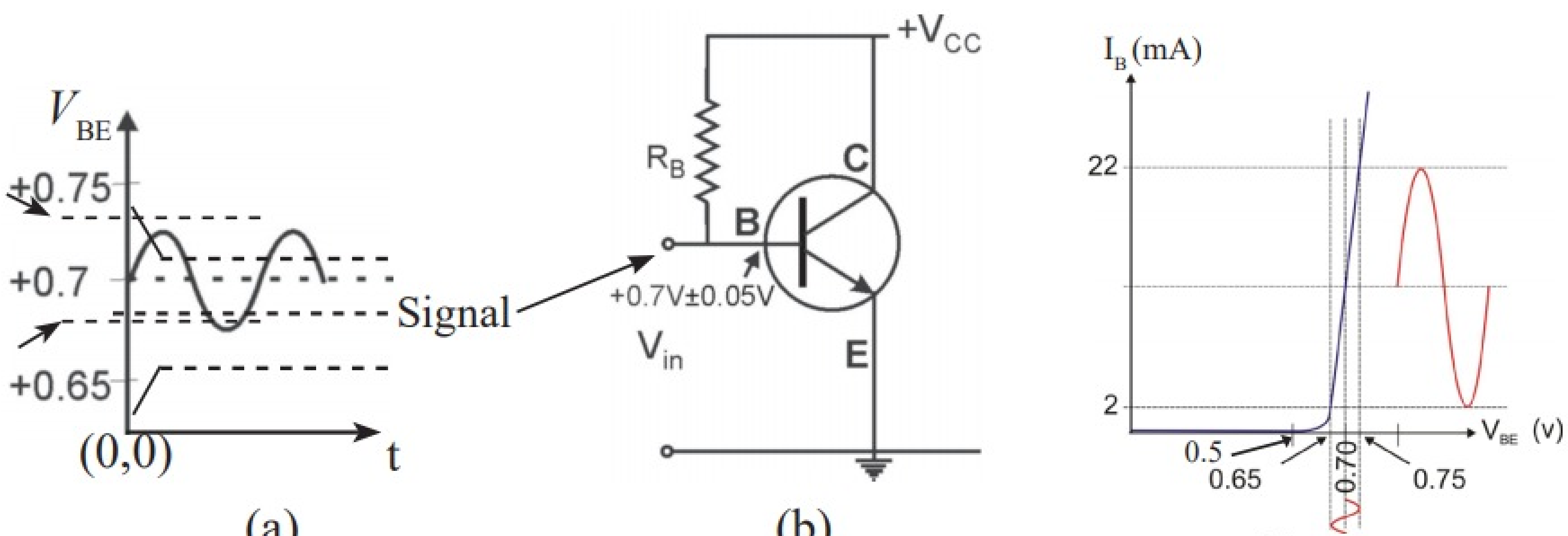


(a) (b) (c)

Figure 2.18

When this signal is applied, the potential at B (V_{BE}) changes from -0.05 V to + 0.05 V and according to input characteristic (Figure 2.18 (c)) I_B remains zero. As $I_B = 0$ the output current I_C is also zero and therefore there is no amplification (no output signal).

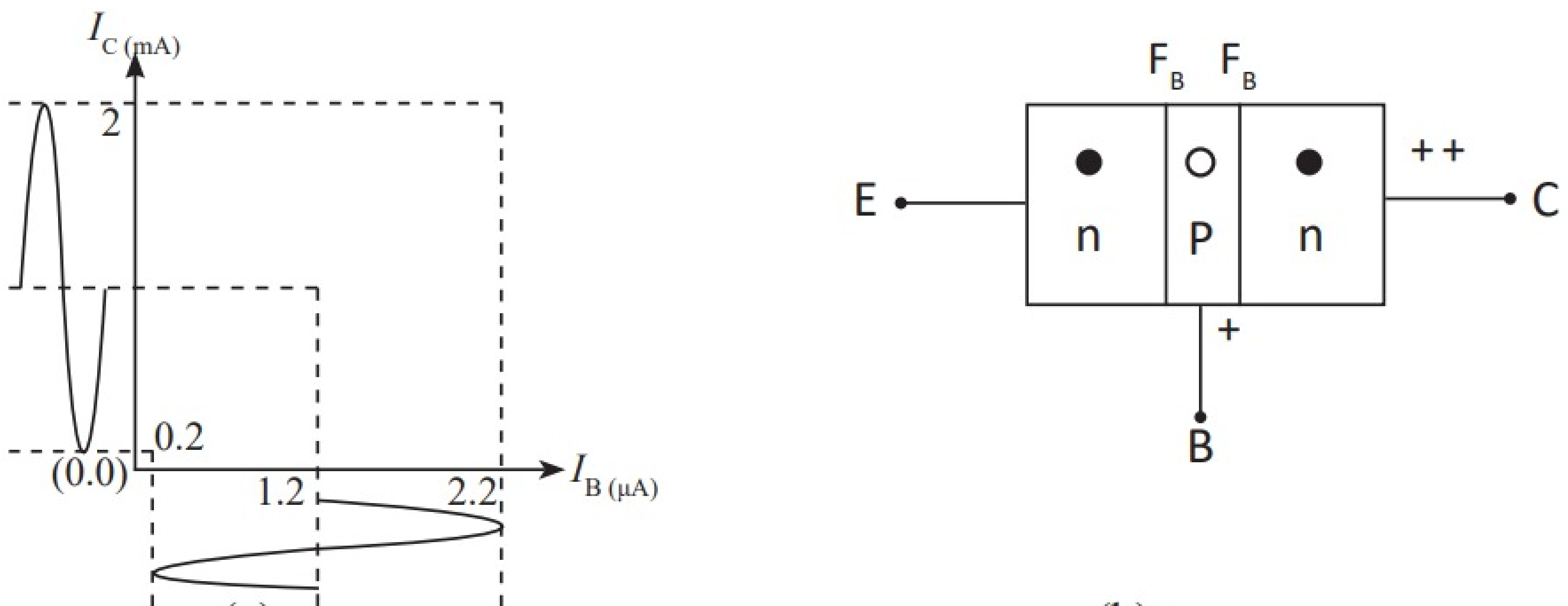
Now consider that the base is kept at a potential +0.7 V using an external voltage supply and then the above signal is applied to the base (Figure 2.19 (b)).



(a) (b) (c)

Figure 2.19

In that case, according to input characteristic I_B varies in the range 2 μA to 22 μA (Figure 2.19(c)).



(a) (b)

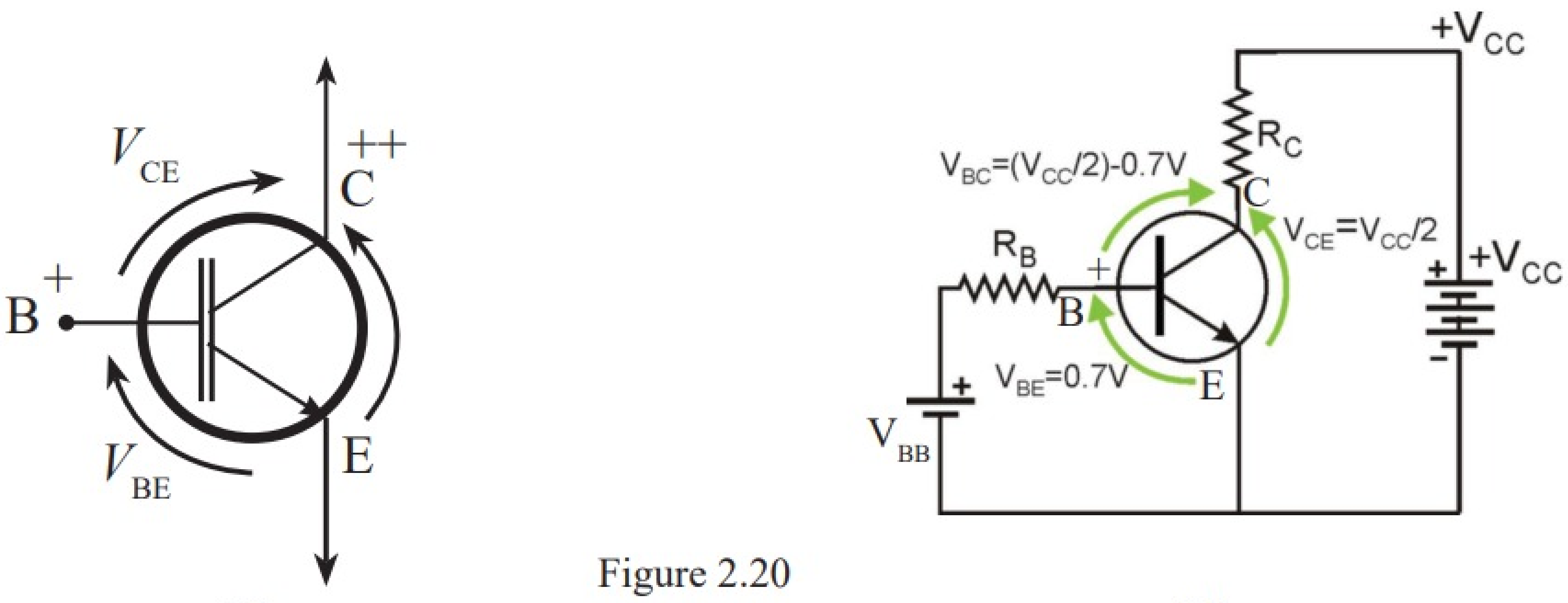


Figure 2.20

(c) (d)

According to this nature of input, the output of the transistor (I_C) changes from 0.2 mA to 2.2 mA as shown by I_C-I_B characteristic (Figure 2.19(C). Therefore, it is clear that a biasing voltage (like +0.7 V above) should be supplied to the base for the transistor to operate as an amplifier. The supply of this kind of a voltage to the base is known as base biasing. In the above biasing we have considered the forward biasing of B-E junction only. In addition to that it is necessary to make B-C junction reverse biased. Since V_{BE} is nearly 0.7 V, V_{CB} could be made sufficiently higher to make B-C junction reverse biased. Using two sets of cells these potentials can be supplied as shown in Figure 2.20. The resistor R_B keeps the base at a potential +0.7 V relative to emitter and the resistor R_C keeps the collector at a potential which is half of the supply voltage V_{CC} ($V_C = \frac{V_{CC}}{2}$). This will be further discussed later.

2.6.2 Biasing the transistor with a single voltage supply

There are several methods of biasing with a single voltage supply and let us consider two such methods which are widely used.

(i) Base-resistor biasing

The simplest method of biasing is the base-resistor method. In this method, the supply voltage terminal (for npn transistors it is the positive potential terminal) is connected to the base through a resistor known as the base-resistors. In the circuit below (Figure 2.21) this base resistor is R_B. To indicate supply voltage the notation V_{CC} is used and for npn transistor amplifiers V_{CC} should be positive(+). To supply required positive potential to the collector, it is connected to the positive terminal f V_{CC} through a resistor R_C . The resistor RB in this circuit maintains a voltage of nearly 0.7V across the base and emitter which should be there as the base biasing voltage (V_{BE}).

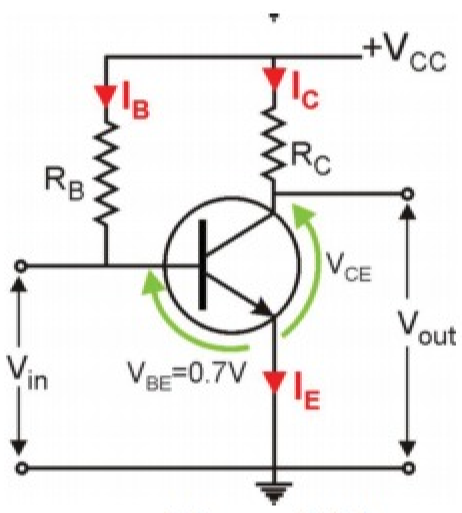


Figure 2.21

When the amplifier is working I_C varies with I_B. As I_C increases the collector potential V_{CE} decreases. When I_C comes to a maximum the potential drop across RC (that is $I_C R_C$) is very nearly equal to V_{CC}. Then $V_{CE} = 0$. When $I_C = 0$, the potential drop across R_C becomes zero (according to $V = I_C R_C$) and $V_{CE} = V_{CC}$. To make use of the entire range of I_C, the value of R_C is chosen so that $V_C = \frac{V_{CC}}{2}$ (It can be made $V_E = 0$ by grounding the emitter. Then $V_C = V_{CE}$ and therefore, V_{CE} can be taken as the collector potential).

Calculating R_B and R_C

According to Kirchhoff's 2nd rule,

for the input circuit loop

$$V_{CC} = I_B R_B + V_{BE} \quad \text{———} \quad (1)$$

for the input circuit loop

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

Since the emitter is grounded $V_E = 0$ and therefore $V_{CE} = V_C$

V_C is kept at $= \frac{V_{CC}}{2}$ as described above.

$$\therefore \; V_{CC} = I_C R_C + \frac{V_{CC}}{2} \quad \text{———} \quad (2)$$

If the current gain of the transistor is β,

$$I_C = \beta I_B \quad \text{———} \quad (3)$$

Considering the current flow through the transistor,

$$I_E = I_C + I_B \quad \text{———} \quad (4)$$

By using the above equations ①, ②, ③ and ④ appropriately and substituting for the supply voltage (V_{CC}), R_B and R_C can be calculated so as to obtain the required value of output current I_C.

Eg. Amplifier circuit has to be designed using a silicon npn transistor of β = 100 and a DC supply voltage of 6 V. Use the base-resistor biasing and draw the relevant circuit diagram. If the value of I_C should be 5 mA, calculate the values of R_B and R_C.

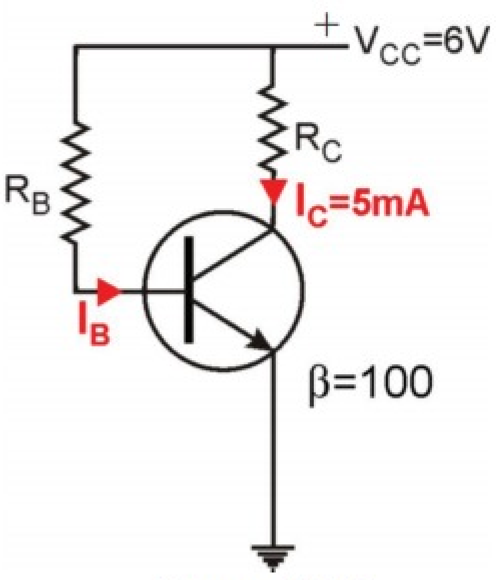


Figure 2.22

Relevant circuit diagram is shown in Figure 2.22.

Applying Kirchhoff's 2nd rule,

$V_{CC} = I_C R_C + V_{CE}$ (for the output circuit loop)

taking $V_{CE} = \frac{V_{CC}}{2}$

$$6 = 5 \times 10^{-3} R_C + \frac{V_{CC}}{2}$$

$$\therefore \; R_C = \frac{6 - 3}{5 \times 10^{-3}}$$

$$\therefore \quad R_C = 600\ \Omega$$

Since $I_C = \beta I_B$

$$5 \times 10^{-3} = 100 \times I_B$$

$$I_B = \frac{5 \times 10^{-3}}{100} = 5 \times 10^{-5}\,\text{A} = 50\ \mu\text{A}$$

Applying kirchhoff 's 2nd rule for the input circuit loop,

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

For a silicon transistor $V_{BE} = 0.7$ V

$$\therefore\ 6 = 50 \times 10^{-6} R_B + 0.7$$

$$\therefore\ R_B = \frac{5.3}{50 \times 10^{-6}} = 1.06 \times 10^{5} = 106\ \text{k}\Omega$$

Therefore R_C should be 600 Ω and R_B should be 106 kΩ. (According to the production standard "E_{24} - series", the resistors available in the market for this purpose are 620 kΩ and 100 kΩ).

(ii) Potential divider biasing

In the base-resistor biasing circuit, it can be seen that some variation occurs in the base current (I_B) leading to a change in the base potential (due to change in potential drop across R_B) and hence a change in base biasing potential V_{BE}. To overcome this drawback a potential dividing resistor arrangement consisting of R_B and R' as shown in Figure 2.23 is used to supply base potential. Because of this potential divider V_B can be made constant. R_B and R' are selected so that $V_B - V_E$ is close to 0.7 V and I is about ten times I_B. Then it can be considered $I \simeq I'$. Therefore, variation in I due to variation in I_B is fairly small. So, V_B can be kept nearly constant all the time with the help of the potential divider.

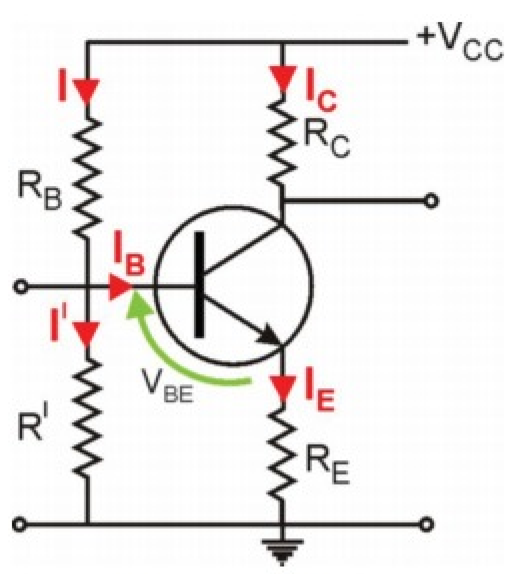


Figure 2.23

Because of some reason such a rising of temperature of the transistor I_C may increase. If I_C increases then I_E also increases (because $I_E = I_C + I_B$) and hence V_E increases (as $V_E = I_E R_E$). Since V_B is constant, this will lead to decrease V_{BE}. This decrease in biasing voltage (V_{BE}) make I_B decreased and this process will bring I_C back to its original value. Therefore, the DC-conditions of the transistor can be kept unchanged. (In this process the emitter resistor R_E plays an important role).

Eg. (i) An amplifier circuit with potential divider biasing is shown in Figure 2.24. Assuming that a silicon transistor has been used, calculate the following.

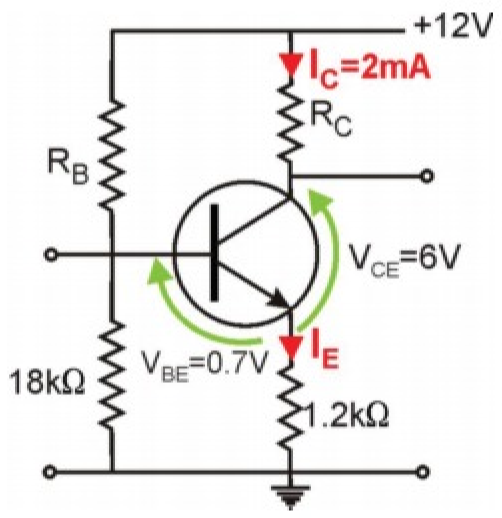


Figure 2.24

i. V_E $(I_C \simeq I_E)$

ii. V_B

iii. V_C

iv. R_C

v. R_B

vi. V_{BC}

vii. What can you conclude by referring to the sign (+ or -) of V_{BC}?

(i) Considering the potential drop across R_E,

$$\begin{aligned} V_E &= I_E R_E \\ &= I_C R_E \quad (\text{as } I_E \simeq I_C) \\ &= 2 \times 10^{-3} \times 1.2 \times 10^3 \\ &= 2.4 \text{ V} \end{aligned}$$

(ii) $V_B = V_E + V_{BE} = 2.4 + 0.7$
$= 3.1$ V

(iii) $V_C = V_E + V_{CE} = 2.4 + 6$
$= 8.4$ V (taking $V_{CE} = \dfrac{V_{CC}}{2}$)

(iv) $$R_C = \frac{V_{RC}}{I_C} = \frac{V_{CC} - V_{CE} - V_E}{I_C} = \frac{12 - 6 - 2.4}{2 \times 10^{-3}} = \frac{3.6}{2 \times 10^{-3}}$$
$$= 1.3 \text{ k}\Omega$$

(v) Considering the potential divider consisting of R_B and 18 kΩ,

$$V_B = \frac{V_{CC} \times 18 \times 10^3}{R_B + 18 \times 10^3} \times V_{CC}$$

Since $V_B = 3.1$ V (From answer (ii) above)

$$3.1 = \frac{18 \times 10^3}{R_B + 18 \times 10^3} \times 12$$

$$\therefore \; R_B + 18 \times 10^3 = \frac{12 \times 18 \times 10^3}{3.1}$$

$$\therefore \; R_B = \frac{12 \times 18 \times 10^3}{3.1} - 18 \times 10^3 = 51.7 \times 10^3 \, \Omega$$

$$R_B \simeq 52 \text{ k}\Omega$$

(vi) $V_{BC} = V_B - V_C = 3.1 - 8.4$

$V_{BC} = -5.3$ V

(vii) As V_{BC} has a negative sign it can be concluded that B - C junction is reverse biased.

Eg :- (ii)

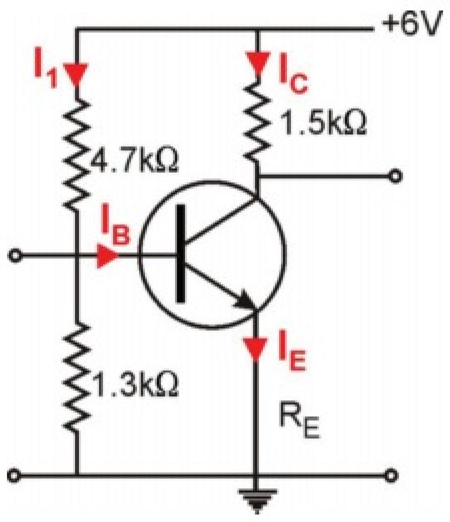


Figure 2.25

An amplifier circuit with potential divider biasing is shown in Figure 2.25. Current I_1 through the potential divider is 25 times the base current I_B. Calculate the values of I_B, I_C, V_C and the current gain of the transistor.

Applying Kirchhoff's 2nd rule to the loop including potential divider,

$I_1(4.7 \times 10^3 + 1.3 \times 10^3) = 6$ (Since $I_B << I_1$)

$$\therefore \; I_1 = \frac{6}{6 \times 10^3} = 1 \times 10^{-3}\text{ A} = 1\text{ mA}$$

$$I_B = \frac{I_1}{25} \quad \text{(Given)}$$

$$\therefore \; I_B = \frac{1}{25}\text{ mA} = 0.04\text{ mA} = 40\ \mu\text{A}$$

Considering V_{CE} and the potential drop across R_C,

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

$$\therefore \; 6 = I_C \times 1.5 \times 10^3 + 3 \qquad \left(\text{taking } V_{CE} \text{ as } \frac{V_{CC}}{2} = 3\text{ V}\right)$$

$$\therefore \; I_C = \frac{6-3}{1.5 \times 10^3}\text{ A}$$

$$= 2\text{ mA}$$

Since $I_E = I_C + I_B$

$$I_E = (2 + 0.04)\text{ mA} = 2.04\text{ mA}$$

$$V_{CE} = V_C - V_E = \frac{V_{CC}}{2} = 3\text{ V}$$

Since $V_E = 0$,

$$V_C = 3\text{ V}$$

$$\text{Current gain } (\beta) = \frac{I_C}{I_B} = \frac{2}{0.04} = 50$$

2.7 Transistor as a switch

A transistor in the common emitter configuration can be used as a switch. Only the cut-off and saturation states are made use of for the switching purpose. When V_{BE} becomes less than 0.5 V and greater than 0.8 V the transistor is in the cut-off and saturation states respectively. (Figure 2.26). When the biasing potential is less than 0.5 V, I_B is close to zero and therefore no collector current (I_C) flows and the transistor behaves as an OFF switch. When the biasing potential is greater than 0.8 V, I_B comes to a maximum bringing the transistor to its saturation. At saturation, V_{CE} is close to zero and hence there flows the maximum current of I_C. This corresponds to the ON state of the transistor switch. (The biasing potentials stated above are for a silicon transistor, and for a germanium transistor the relevant potentials are lower than the stated values.)

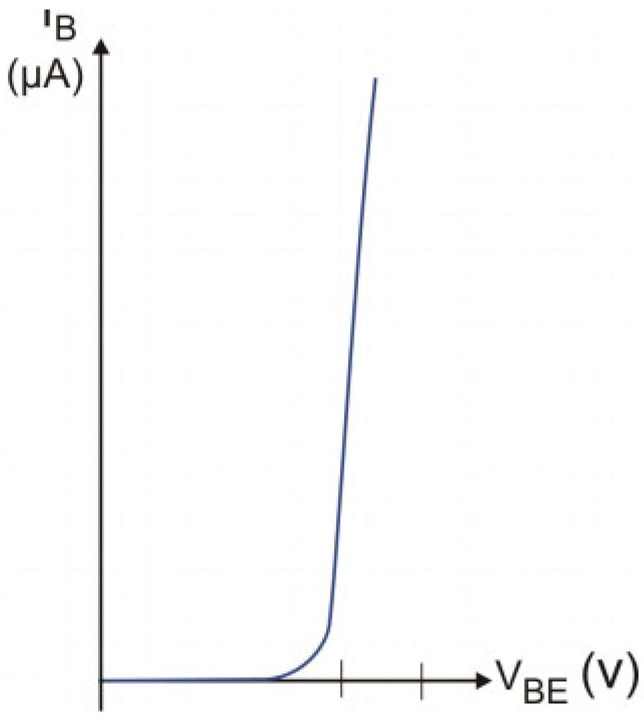
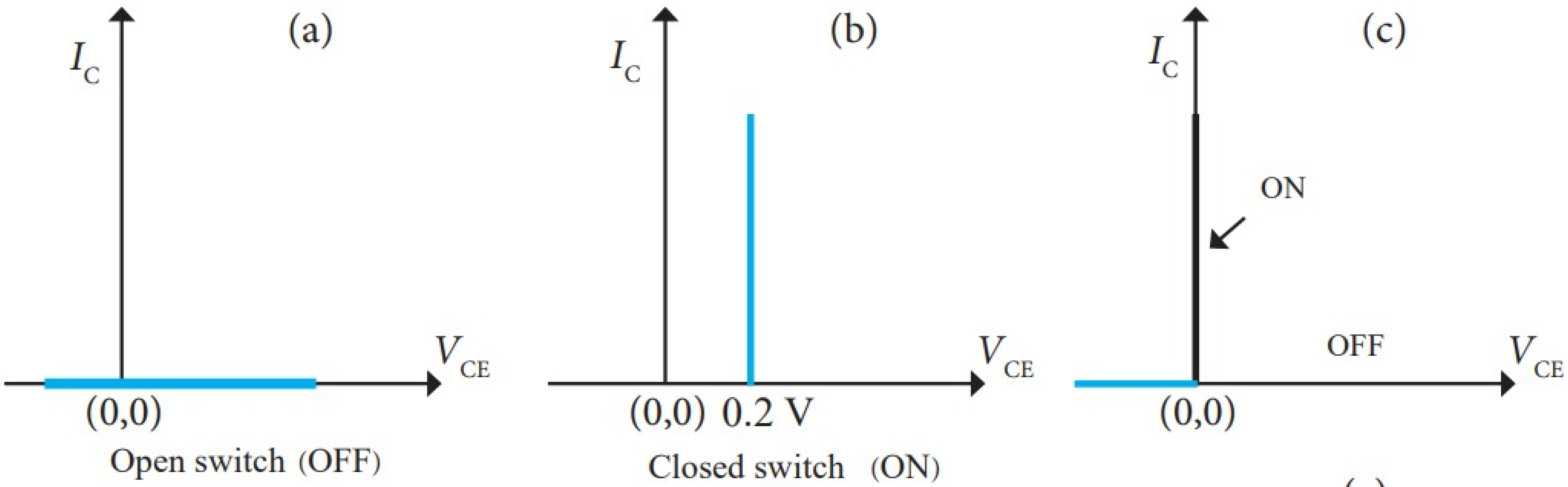


Figure 2.26



(a) (b) (c)

Figure 2.27

When the transistor is in its OFF-switch state, V_{CE} is high ($V_{CE} = V_{CC}$) and $I_C = 0$ (Figure 2.27(a)). When the transistor is in its ON-switch state, V_{CE} is as small as 0.2 V (Figure 2.27 (b)). By considering this small value as close to zero the characteristic shown in Figure 2.27 (c) can be obtained. In the previous chapter the *I-V* characteristic of a mechanical switch has been shown in Figure 1.25 (b). Comparing the characteristics in Figures 1.25 (b) and 2.27 (c) it can be made clear that the transistor too, acts as a switch.

A circuit in which a transistor is operated as a switch by small electric pulses, is shown in Figure 2.28.

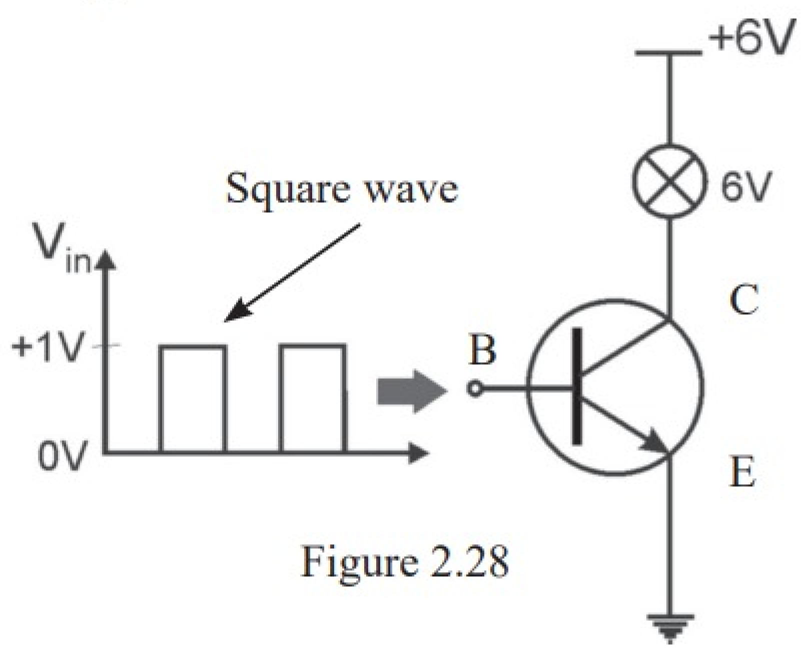


Figure 2.28

A square wave signal varying between 0 V and 1 V is applied to the base. The bulb connected to the collector can be switched ON or OFF by that signal. Since the response time of a transistor is much smaller than that of a mechanical switch, the bulb can be lit or not, much quickly using the transistor switch.

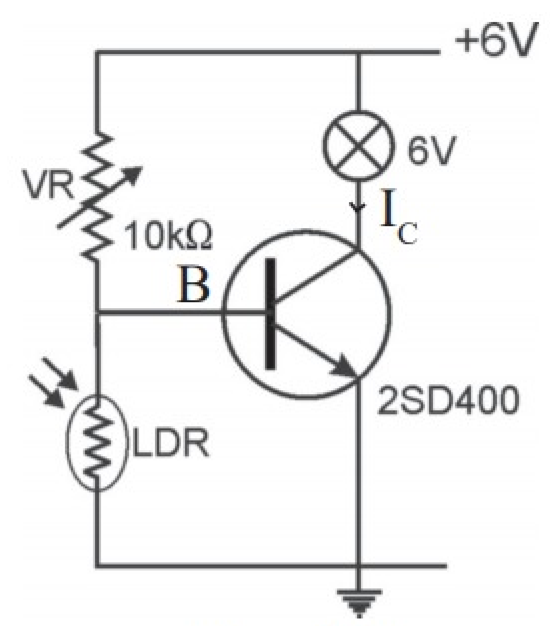


Figure 2.29

The transistor switching circuit shown in Figure 2.29 can be used to light a bulb automatically when there is darkness. The LDR has been used as a light detecting sensor. Resistance of the LDR in the darkness is about 100 kΩ. When light is incident on, its resistance reduces to about 100 Ω depending on the intensity of light. In this circuit the potential divider consisting of the LDR and variable resistor VR makes the potential at the base (VB) increasing gradually when the environment begins to get darker, because of the gradual increase of resistance of the LDR. When VB becomes greater than 0.8 V the transistor acts as a closed switch and allows I_C to flow lighting the bulb. By adjusting VR, the level of darkness required to light the bulb can be changed. If the potential divider is arranged so that the LDR and VR are interchanged with each other then there is light instead of darkness.

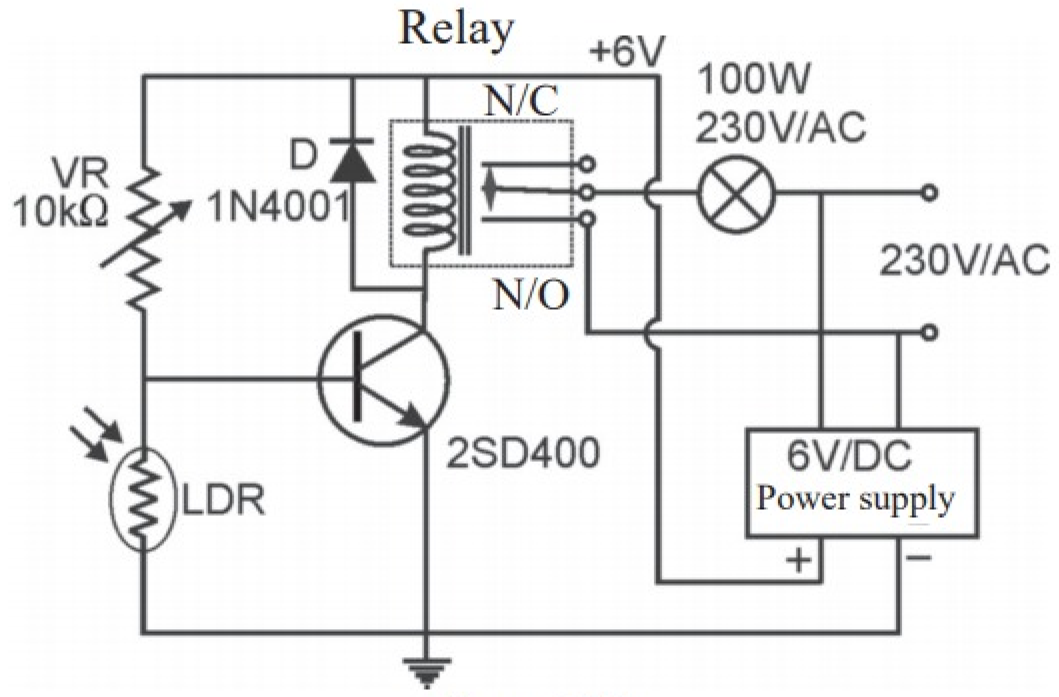


Figure 2.30

The circuit diagram in Figure 2.30 shows how a relay can be used for switching. The circuit is powered by a 6V-DC supply with which the relay can be operated. Using the relay a 230 V, 100 W bulb can be operated safely with the AC-mains supply. A 6 V-relay which is easily found in the market should be used for this circuit. When the coil of the relay is not supplied with electricity the N/C terminal (normally closed terminal) is in closed (ON) position and the N/O terminal is in open (OFF) position. When electricity is supplied terminal becomes closed (ON). The diode D protects transistor against electric surge which occurs due to induction in the relay coil when it is switched ON or OFF. When the coil of the relay is not supplied with electricity the N/C terminal is in closed (ON) position and the N/O terminal is in open (OFF) position. When electricity is supplied the terminal becomes closed (ON). The diode D protects transistor against electric surge which occurs due to induction in the relay coil when it is switched ON or OFF.

2.8 Features of amplifier circuits

In amplifier circuits there are some special features to be considered.

a) Input resistance and output resistance.
b) Current gain, voltage gain and power gain.
c) Frequency response

A detailed study of the above feature is beyond the scope of this book. Therefore, let us discuss some of them in brief.

2.8.1 Current gain of a common emitter amplifier (β)

The DC-current gain β is defined to be,

$$A_i = \beta = \frac{I_0}{I_i}$$

where I_0 = output current
I_i = input current

In this configuration,

$$\therefore \quad \boxed{\beta = \frac{I_C}{I_B}}$$

The common emitter current gain is given as h_{FE} in most of the data sheets according to some other system of symbols. In h_{FE}, the subscripts in capital letters indicate that it is a DC-value. F is to indicate forward bias and E is to indicate common emitter configuration. The typical value of h_{FE} is about 200.

2.8.2 Voltage gain of a transistor amplifier (A_V)

The voltage gain of a transistor amplifier is defined as,

$$A_V = \frac{\text{Change of output voltage}}{\text{Change of input voltage}} = \frac{\Delta V_O}{\Delta V_i}$$

In the emitter configuration as per Figure 2.31,

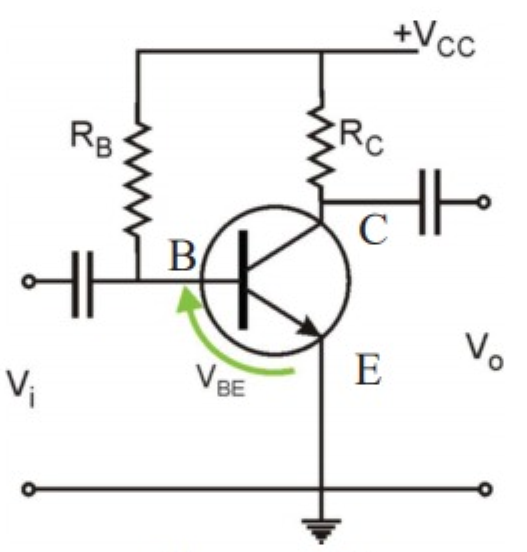


Figure 2.31

$$A_V = \frac{\Delta V_O}{\Delta V_i} = \frac{\Delta V_{CE}}{\Delta V_{BE}}$$

Since $V_{CE} = V_C - V_E$ and $V_E = 0$

$$\therefore \quad \Delta V_{CE} = \Delta V_C$$

Since $V_{BE} = V_B - V_E$ but $V_E = 0$

$$\therefore \quad \Delta V_{BE} = \Delta V_B$$

$$\therefore \quad \boxed{A_V = \frac{\Delta V_C}{\Delta V_B}}$$

The typical value of A_V is about 40.

2.9 Practical applications of the transistor

Using as a voltage and current amplifier

To get a loud sound from a weak electrical signal produced when speaking to a microphone, the signal should be amplified before giving it to a loudspeaker. This amplification can be done by a transistor amplifier. Amplifier circuits are also used to amplify weak signals received by radio receivers. The common emitter amplifier circuit is widely used for amplification.

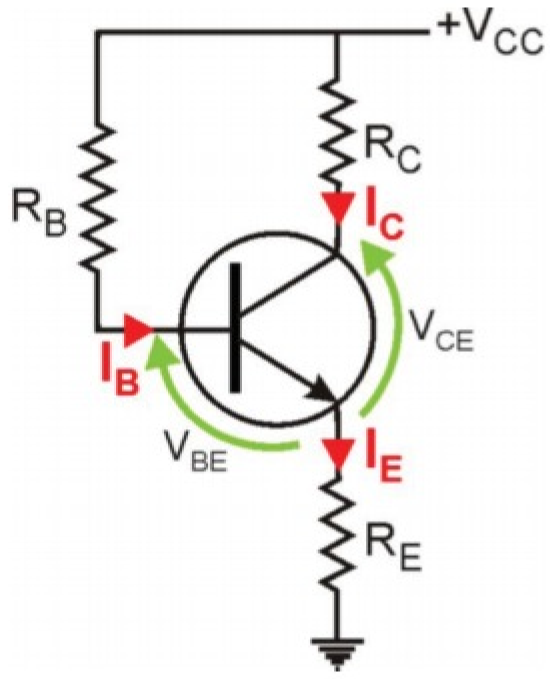


Figure 2.32

The common emitter transistor circuit shown in Figure 2.32 has been biased using a single voltage source V_{CC} and two resistors R_B and R_C. The values of R_B and R_C are chosen so that the potentials V_B and V_C get the required values to get the transistor in its active state. Then DC-currents, I_B of the order of μA and I_C of the order of mA, flow through the transistor.

Consider a situation where a small sinusoidal voltage of the order of mV produced by a signal generator is fed to the base through a capacitor C_1. Then that signal makes V_{BE} vary and hence I_B too, varies accordingly. This is shown using the input characteristic curve in Figure 2.33.

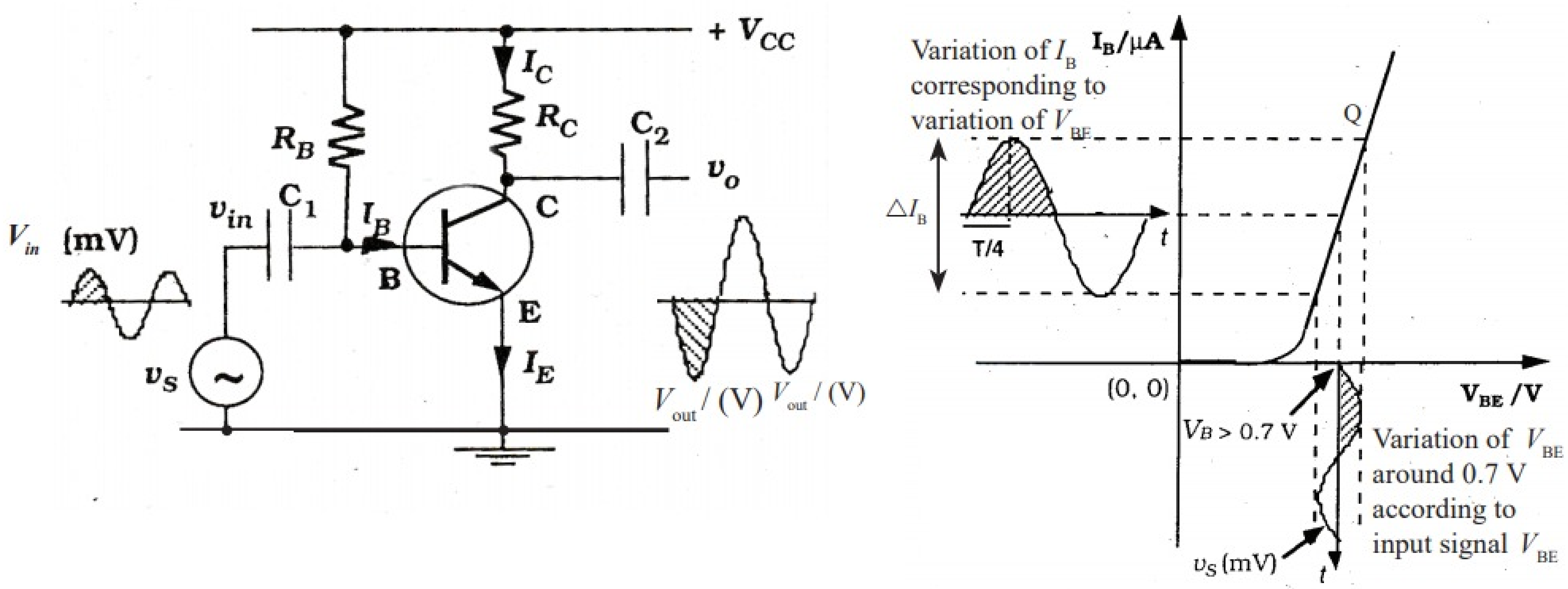


Figure 2.33

Since the transistor is biased so that it is in the active region I_C is proportional to I_B. Therefore, when I_B varies I_C also varies accordingly. This is shown using the transfer characteristic in Figure 2.34.

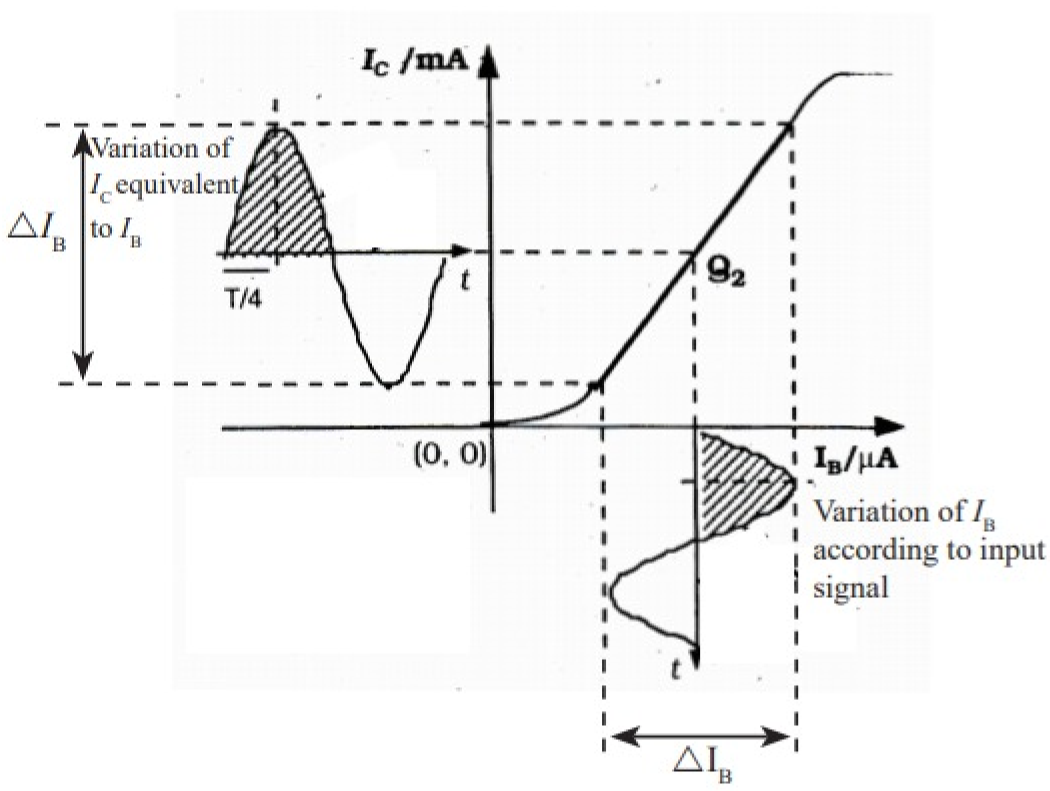


Figure 2.34

Both I_B and I_C vary sinusoidally with time, with the same frequency and phase.

Having a variation of I_C in mA range, corresponding to the variation of I_B in μA range can be considered as a current amplification.

To convert the current amplification in I_C into a voltage amplification, it is necessary to use a collector resistor R_C. The value of R_C should be chosen so that the value of V_C is half the supply of voltage V_{CC}, in order to have an amplified output without a distortion (without a change in the input waveform).

V_{CE} is considered to be the output voltage of a common emitter amplifier. Since the emitter has been grounded, (referring to the circuit diagram shown in Figure 2.32).

$$V_{CE} = V_C$$

This is, $V_{out} = V_C$

$$\therefore \quad V_{out} = V_{CC} - I_C R_C$$

Above equation explains how V_{out} varies with I_C. The output voltage V_{CE} makes a phase difference of π radian with the output current I_C. The input voltage is of few millivolts and the output is of few volts. Therefore, it can be considered as a voltage amplification. This voltage amplification also makes a phase difference of π radian with the input signal.

2.9.1 Action of coupling capacitors

For the amplifier described above to operate successfully, the transistor should be biased so that it is in the active region. To do so, the potentials V_B and V_C should be with their proper values 0.7 V and $Vcc/2$ respectively. The DC-value of V_B can be made constant, as

© 2020 National Institute of Education. All Rights Reserved.

C_1 prevents some part of DC-current I_B going towards the input source. C_1 does not block the AC-signal which comes from the source and goes towards the base.

The capacitor C_2 prevents a part of DC- current I_B going towards the output load and hence the DC-voltage V_C can be kept constant. C_2 does not block the amplified AC- voltage signal going towards the output load.

The main steps of a common emitter transistor amplifier are represented by the following graphs (Figure2.35).

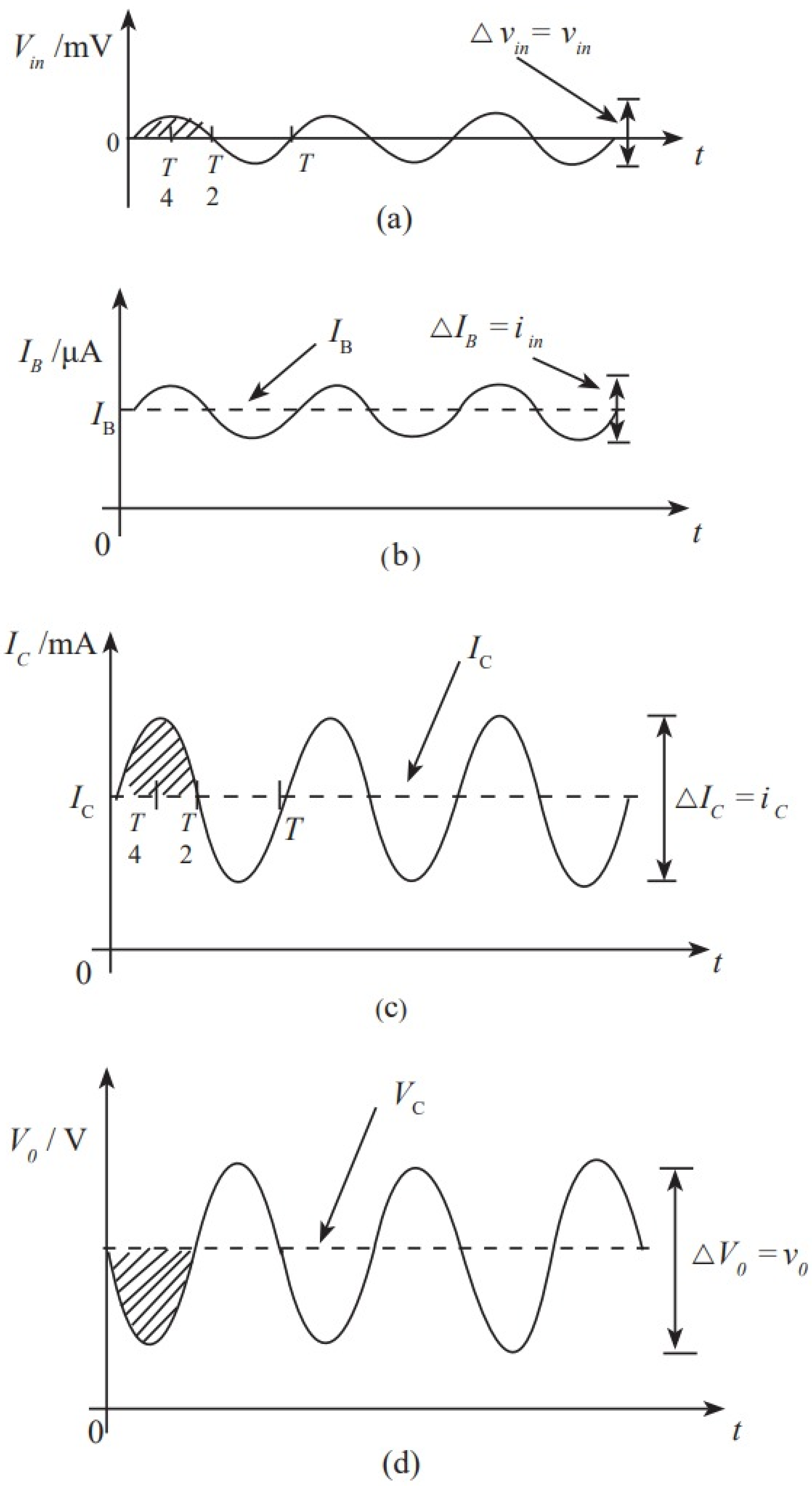


Figure 2.35

The variation of I_B due to input signal V_{in} has been denoted by *im* in Figure 2.35 (a) and (b). Corresponding to the variation *im*, I_C also varies and that variation has been denoted by I_C in Figure 2.35(c). Since I_C varies V_{CE} also varies and that variation has been denoted by V_0 in Figure 2.35(d).

Accordingly, within time *T*/4 (*T* is the period of input sinusoidal signal) during which the current I_C gets increased to its maximum value from its DC-current level, the product $I_C R_C$ also gets increased to its maximum value. Then V_{CE} decreases from its maximum value to a minimum. Accordingly, when I_C decreases V_{CE} increases. Therefore, it is clear that during positive half cycles of each of V_{in}, I_B and I_C, the value of the output signal ($V_0 = V_C = V_{CE}$) gets a lower value than the DC- value of V_0. During negative half cycles it takes a higher value than the DC- value of V_0. These variations have been compared in Figure 2.35. By comparing Figure 2.35 (a) and (b), it can be seen that the input voltage signal (sinusoidal) given to the common emitter voltage amplifier has undergone an inversion (phase difference of 1800) when outputting it. According to matters discussed above it is clear that the common emitter transistor amplifier can be used not only as a current amplifier but also as a voltage amplifier.

$$\text{Voltage gain of the amplifier} = \frac{v_0}{v_{in}} = \frac{\triangle V_0}{\triangle V_{in}} = \frac{\triangle V_C}{\triangle V_B}$$

Voltage gain, current gain and power gain of the common emitter transistor amplifier.

2.9.2 Voltage gain, current gain and power gain for a common emitter transistor

Relating to alternating signals, voltage gain, current gain and power gain are defined as follows.

$$\text{Voltage gain } A_V = \frac{V_{out}}{V_{in}} = \frac{\Delta V_C}{\Delta V_B}$$

$$= \frac{\text{Change in output voltage}}{\text{Change in input voltage}}$$

$$\text{Current gain } A_i = \frac{I_{out}}{I_{in}} = \frac{\Delta I_C}{\Delta I_B}$$

$$= \frac{\text{Change in output voltage}}{\text{Change in input voltage}}$$

$$\text{Power gain } A_p = \frac{P_{out}}{P_{in}} = \frac{V_{out} I_{out}}{V_{in} I_{in}}$$

$$= \frac{\text{output power}}{\text{input power}}$$

2.10 Unipolar Transistors

2.10.1 Field Effect Transistors - FET

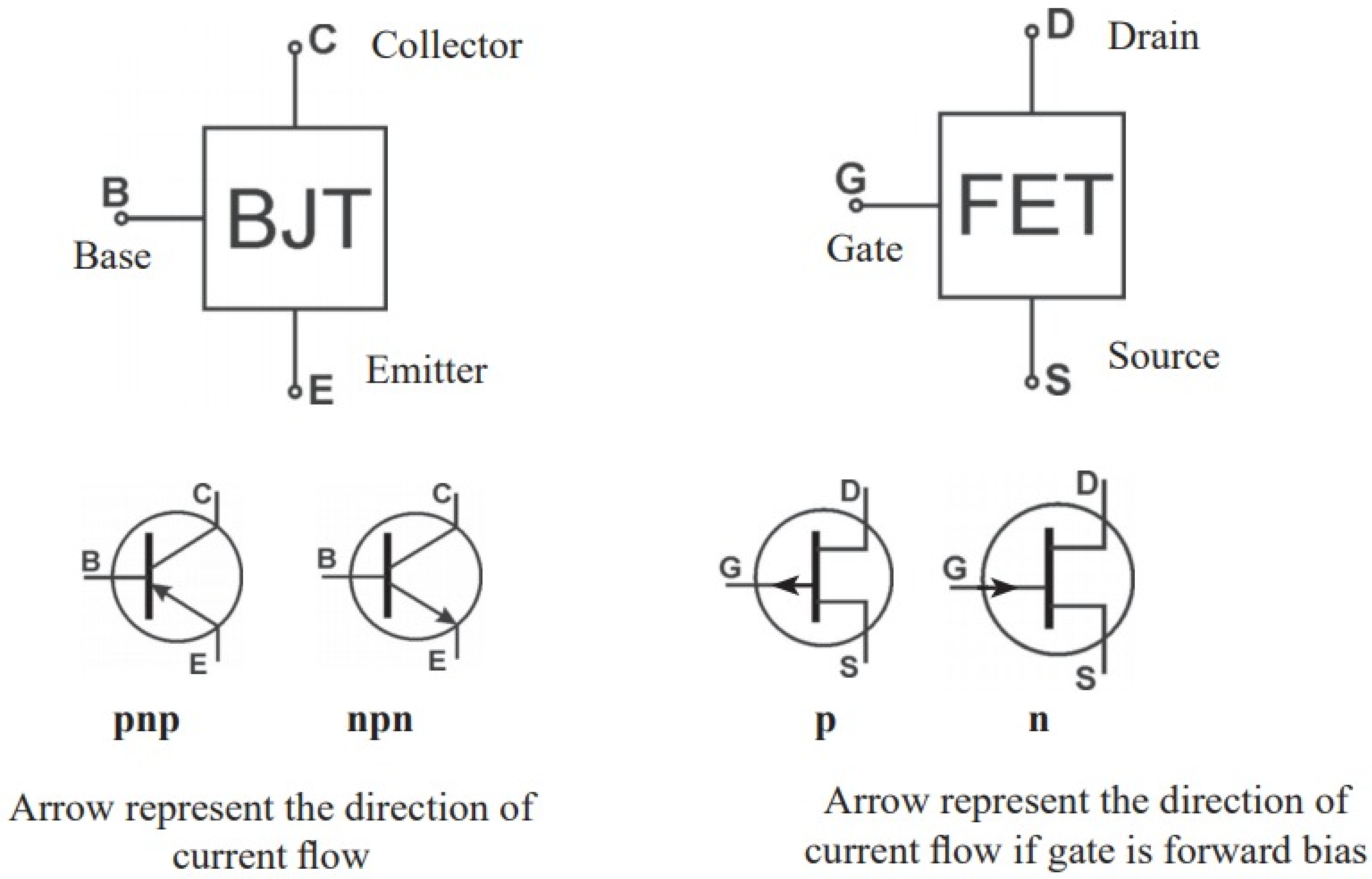
The transistors we have considered so far belong to the category of Bipolar Junction Transistors (BJJ). For their operation both free electrons and holes contributed as charge carries. The other major category of transistors is the unipolar transistors in which only one type of charge carrier is contributing towards its operation. The field effect transistor (FET) belongs to the category of unipolar transistors. Field effect transistors are of two types.

(a) Junction Field Effect Transistor (JFET)
(b) Metal Oxide Semiconductor Field Effect Transistor (MOSFET)

In the advanced level physics syllabus only the JFET is considered.

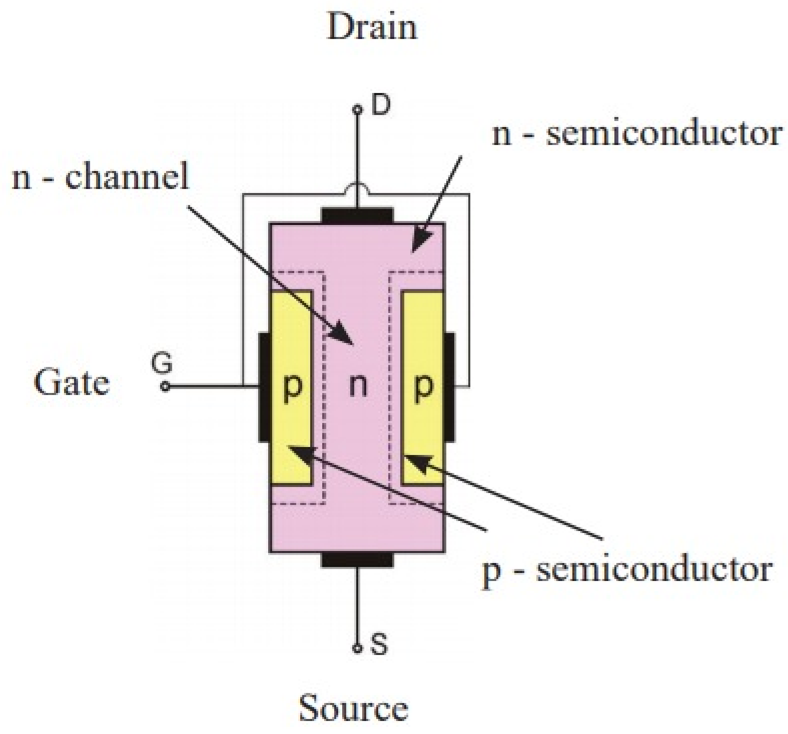
2.10.2 JFET

Like in a normal BJT, in a JFET also there are terminals. Figure 2.36 shows how the terminals are labeled.

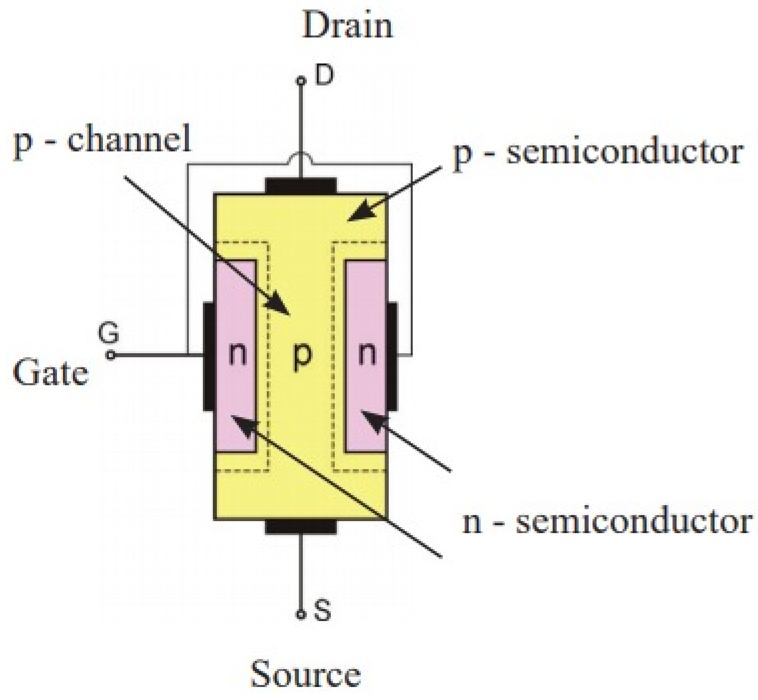


BJT and JFET Symbols
Figure 2.36

Like the two types npn and npn of the BJT there are also two types of JFET as " n - channel " and "p – channel ". Let us consider the basic structure of each type.



(a) n - channel FET



(b) p - channel FET

Figure 2.37

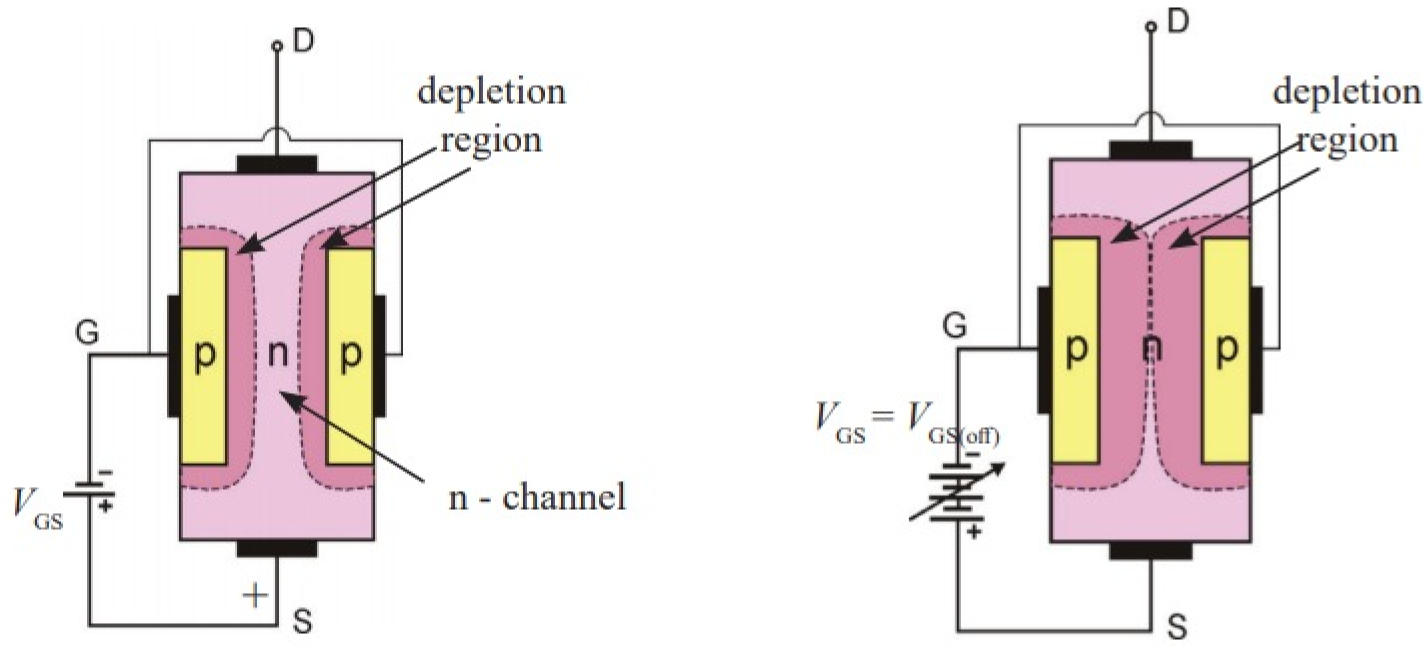
To understand the structure of JFET let us consider that there is a cylindrical n- type semiconductor and around it a p- type semiconductor ring (as for n- channel FET). The cross section of this arrangement is shown in Figure 2.37 (a). The two cross sectional parts of the p - type semiconductor ring are shown on the two sides of the channel (n – type semiconductor). Those two sections are connected by a wire so as to indicate that they belong to the same semiconductor region. There is small depletion layer formed between n- type semiconductor cylinder and p- type semiconductor ring (shown by the dotted lines in Figure 2.37 (a) and (b). The two ends of the n - type semiconductor (channel) have been connected to two terminals coming out of these two terminals, which emits carriers into the channel called the **source** and the terminal which collects those carriers and drives them out from the channel called the **drain**. The p- type semiconductor ring around the n- type semiconductor cylinder is connected to another terminal coming out, and that terminal is called the **gate**. The ring like depletion layer is there between the source (s) and the drain (D).

2.10.3 Action of n- channel FET

(a) When the voltage between gate and source is varied while the drain is left open.

Suppose terminal D is left open and a small voltage (V_{GS}) has been applied so that the p-n junction between gate and source is reverse biased. Then the depletion layer between the gate and n- type channel becomes larger and hence the n- channel becomes narrower (Figure 2.38 (a)).

© 2020 National Institute of Education. All Rights Reserved.



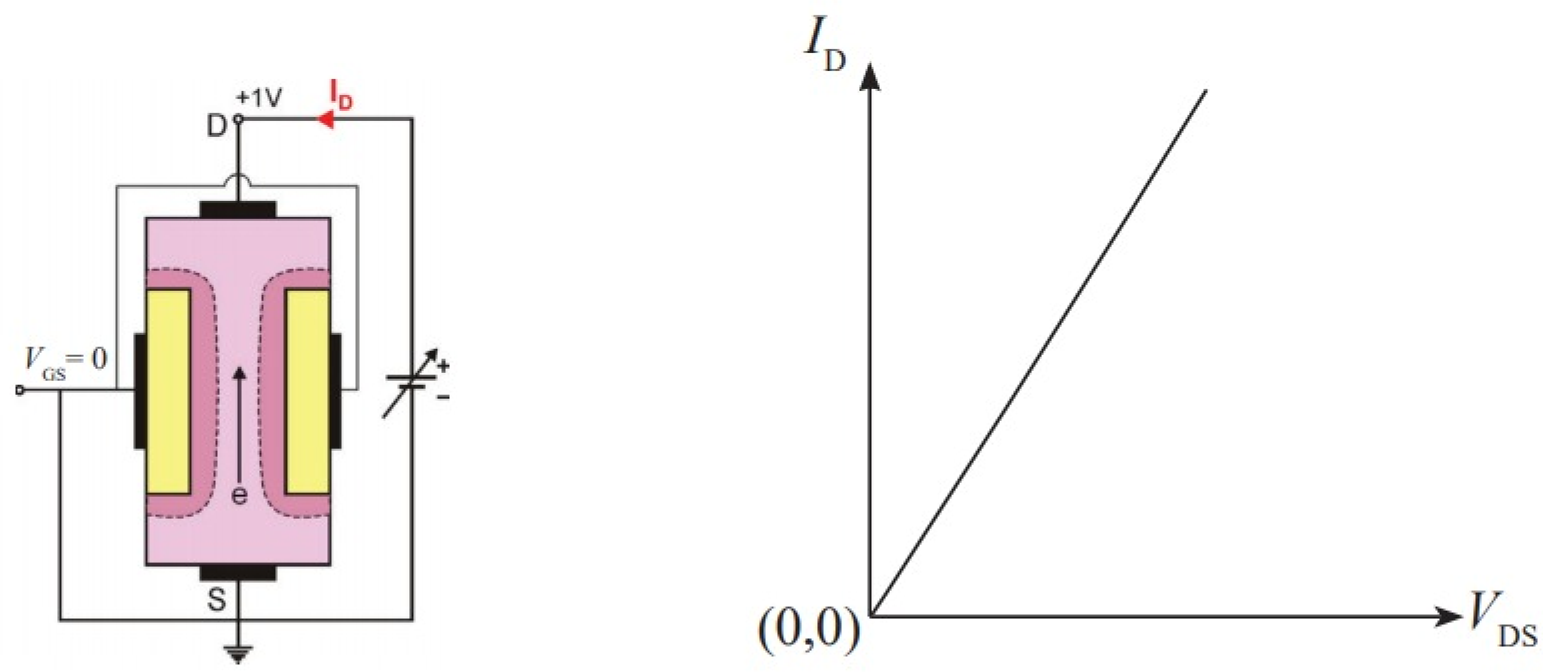
n - channel FET (When $V_{GS} < 0$) n - channel FET (When $V_{GS} = V_{GS(off)}$)

(a) Figure 2.38 (b)

If V_{GS} is increased gradually then the depletion layer becomes larger and finally at some value of V_{GS} the depletion layer closes the channel totally as shown in Figure 2.38 (b). On this occasion the voltage V_{GS} between the gate (G) and source (S) is known as the cut - off voltage and it is denoted by V_{GS} (off). At this cut - off stage, even though a potential difference is applied across the source (S) and drain (D) a current (I_D) does not flow through it since the channel is totally closed. Generally V_{GS} (off) of an FET is a low voltage and depending on the FET used it will be a value in the range 4 V - 8 V. For a given FET V_{GS} (off) is a constant value and it can be found from data sheets. V_{GS}(off) will further be discussed later.

(b) Variation of I_D with V_{DS} when $V_{GS} = 0$

If the gate terminal is grounded and a small voltage is applied as potential, then there will be a flow of electrons from S to D. That is, a current I_D flows from D to S. When V_{DS} is increased gradually, I_D also increases accordingly. With that, the biasing voltage increases as described below, and therefore the depletion layer becomes larger (Figure 3.39 (a)).

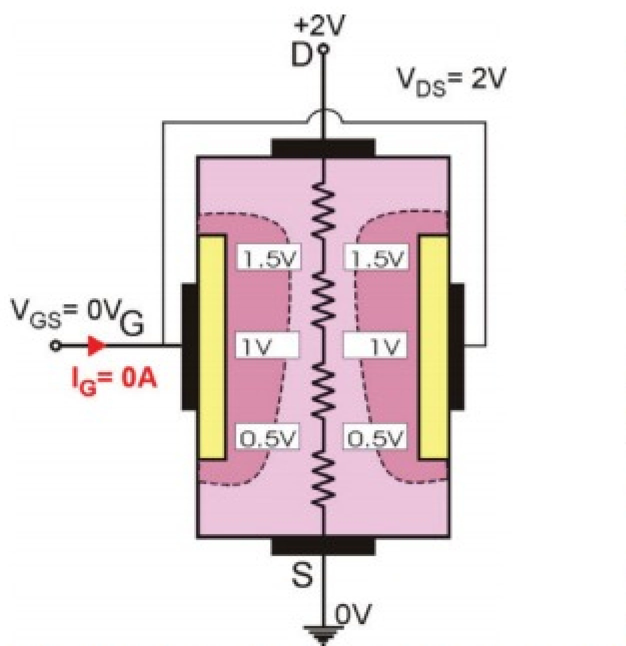


When $V_{GS} = 0$, variation of I_D with V_{DS}

(a) (b)

Figure 2.39

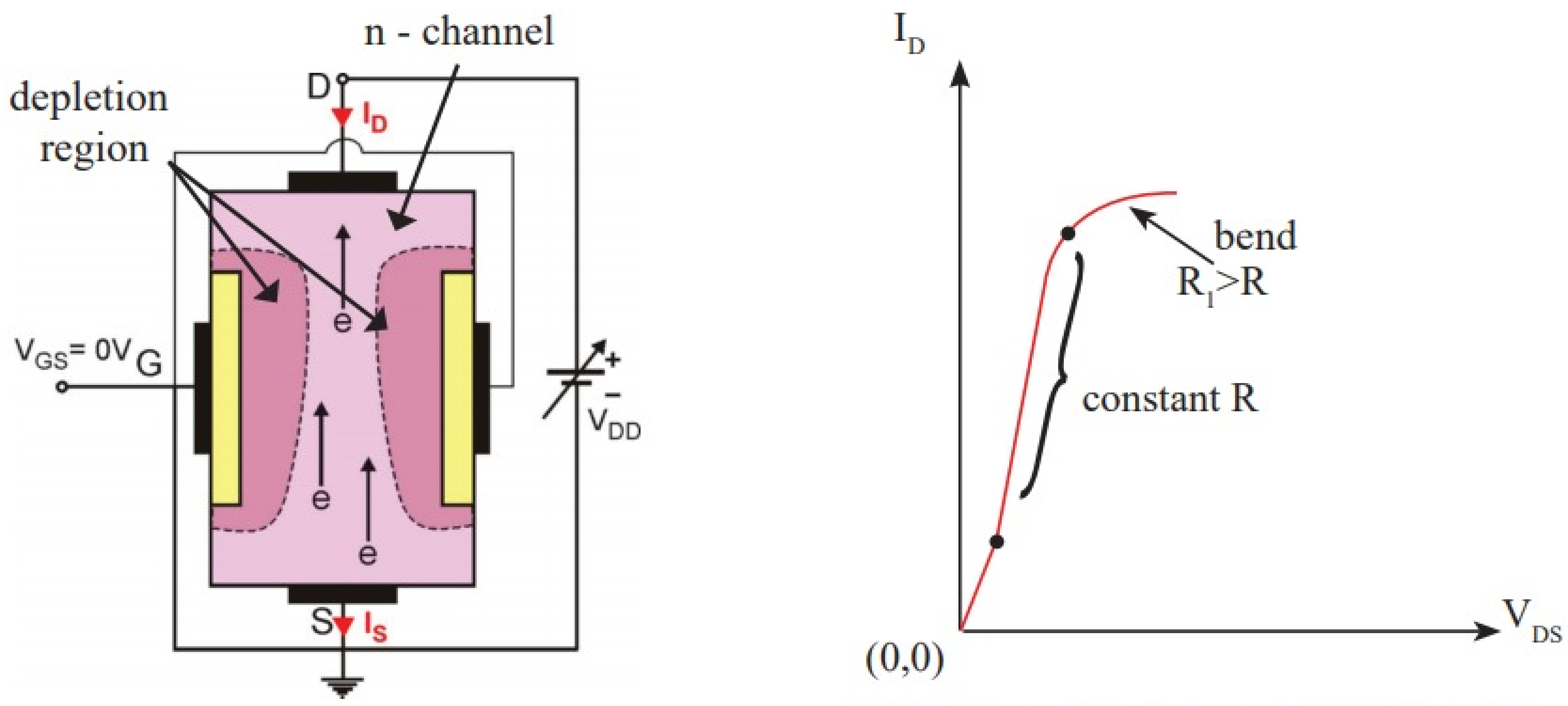
As an example suppose that a voltage of 1 V has been applied between S and D keeping $V_G = 0$. Then the potential at the middle of the channel is $\frac{1 - 0}{2} = 0.5$ V. Therefore, p - n junction behaves as if there were a reverse bias voltage of 0.5 V between the gate and channel. Because of this reverse bias, the depletion layer becomes larger and the channel becomes narrower. In this process the channel behaves as an ohmic resistance and therefore, I_D increases linearly with V_{DS} as shown in Figure 2.40 (b) (gradient of the straight line part of the graph is $\frac{\Delta I_{DS}}{\Delta V_{DS}} = \frac{1}{R}$, where R is the resistance of the channel)



Distribution of reverse bias potential in n - channel through out the p - n junction
Figure 2.40

When V_{DS} is increased further the depletion layer in the upper part of the channel becomes much larger. The reason is that, the positive potential at the upper part of the channel is greater than that at the lower part.

If V_D is + 2 V the distribution of potential in the channel is shown by Figure 2.41. Because of this change in the depletion layer, the channel is not behaving as a uniform resistor any more. Since the resistance of the channel becomes higher now, the increase of I_C with V_{DS} becomes less than earlier. Accordingly the gradient of the graph $\left(\frac{1}{R}\right)$ becomes less than earlier. Therefore, the graph shows a bend (knee shape) corresponding to this state as shown in Figure 2.41 (a).



$V_{GS} = 0$ and $V_{DS} > 0$ V

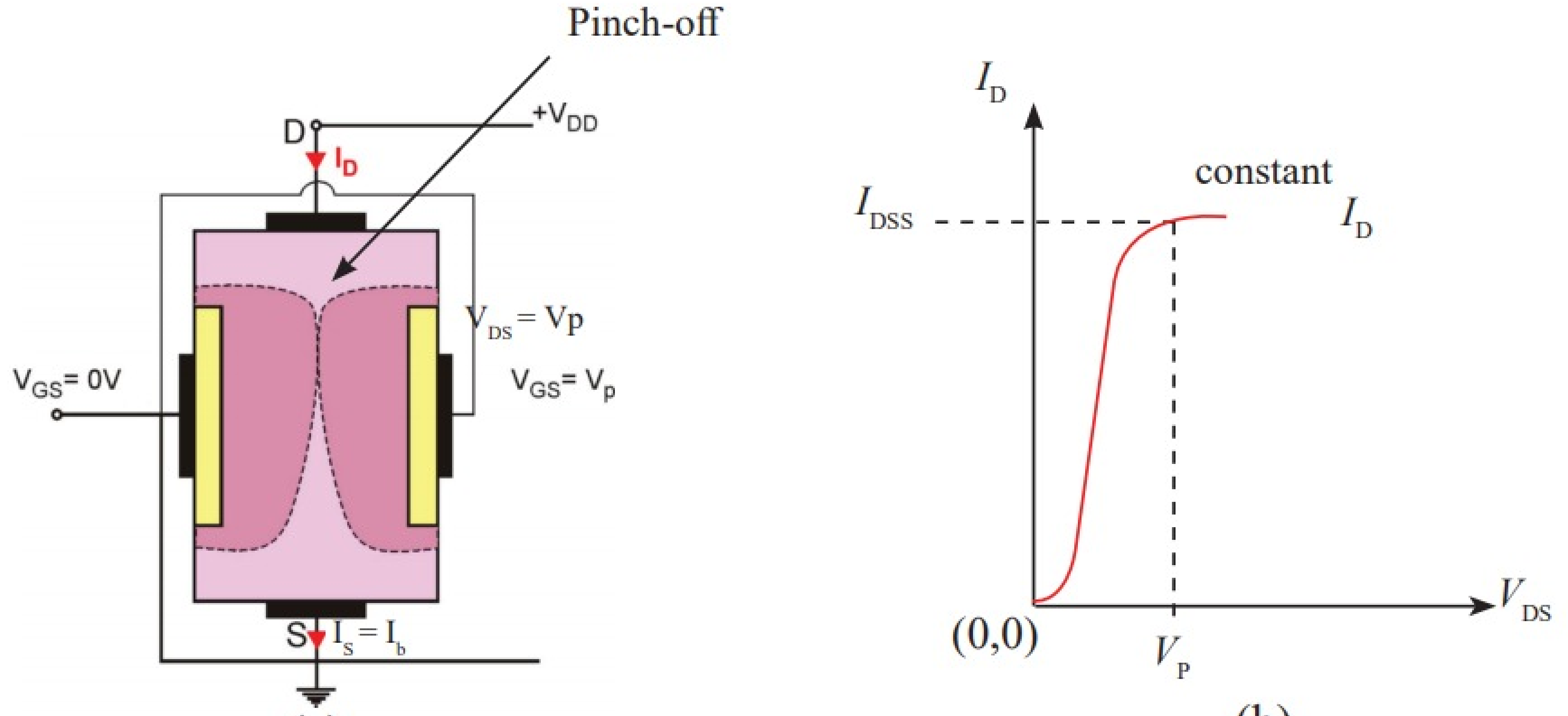
When $V_{GS} = 0$ and $V_{DS} < V_P$ Variation of I_P

(a) Figure 2.41 (b)

According to Figure 2.41 the biasing potentials at the lower part, middle part and upper part of the channel are 0.5 V, 1 V and 1.5 V respectively. For this reason, depletion layer of upper part of the channel gets widened and the channel gets the shape of a cone (Figure 2.41 (a).

© 2020 National Institute of Education. All Rights Reserved.

When V_{DS} is increased further, the depletion layer also gets widened further and the top of it will close the channel (Figure 2.42(a)).



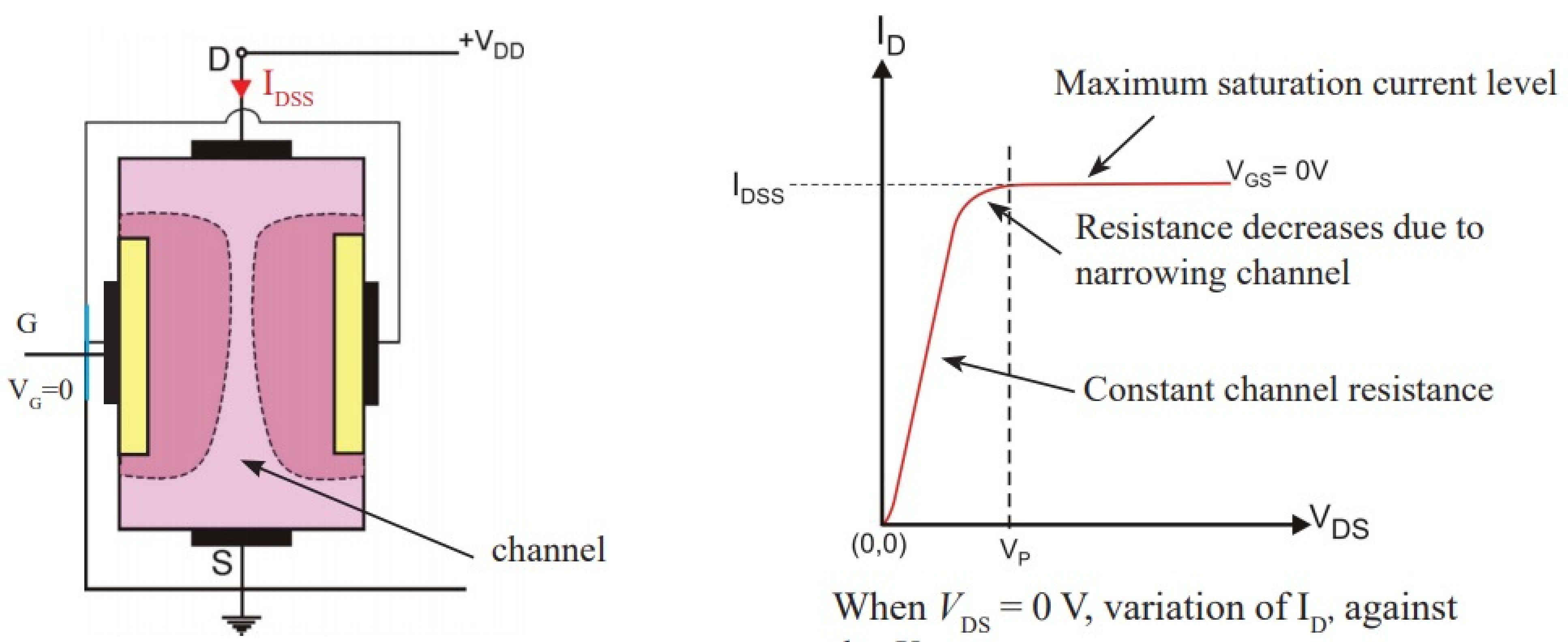
(a)
$(V_{GS} = 0V, V_{DS} = V_P)$

(b)

Figure 2.42

In this manner, if the channel gets closed totally the current I_D should be zero. However, at this stage the current I_D, flowing through the channel has already come to a large value. Therefore, if I_D gets zero at once then the potential drop along the channel will also get zero and hence the depletion layer gets smaller making the channel open again. Then I_D starts to flow again. Therefore this process will come to an equilibrium keeping the value of I_D constant at its previous value, by means of opening the channel a little. This equilibrium state is known as the "pinch off". The value of V_{DS} at the pinch off when $V_{GS} = 0$, is known as the "pinch off voltage" and it is denoted by V_P (Figureb 2.42 (b)).

When V_{DS} is increased beyond V_P, the channel remains open as above and theonly change occurring is that the depletion layer begins to spread downwards (Figure 2.43(a)). Therefore, the current (I_D) flowing in the channel remains constant. I_D has come to its maximum saturation on this occasion (when $V_{GS} = 0$) and the particular I_D is called the maximum saturation current which is denoted by I_{DSS} (Figure 2.43 (b)).



(a)

When $V_{DS} = 0$ V, variation of I_D, against the V_{GS}

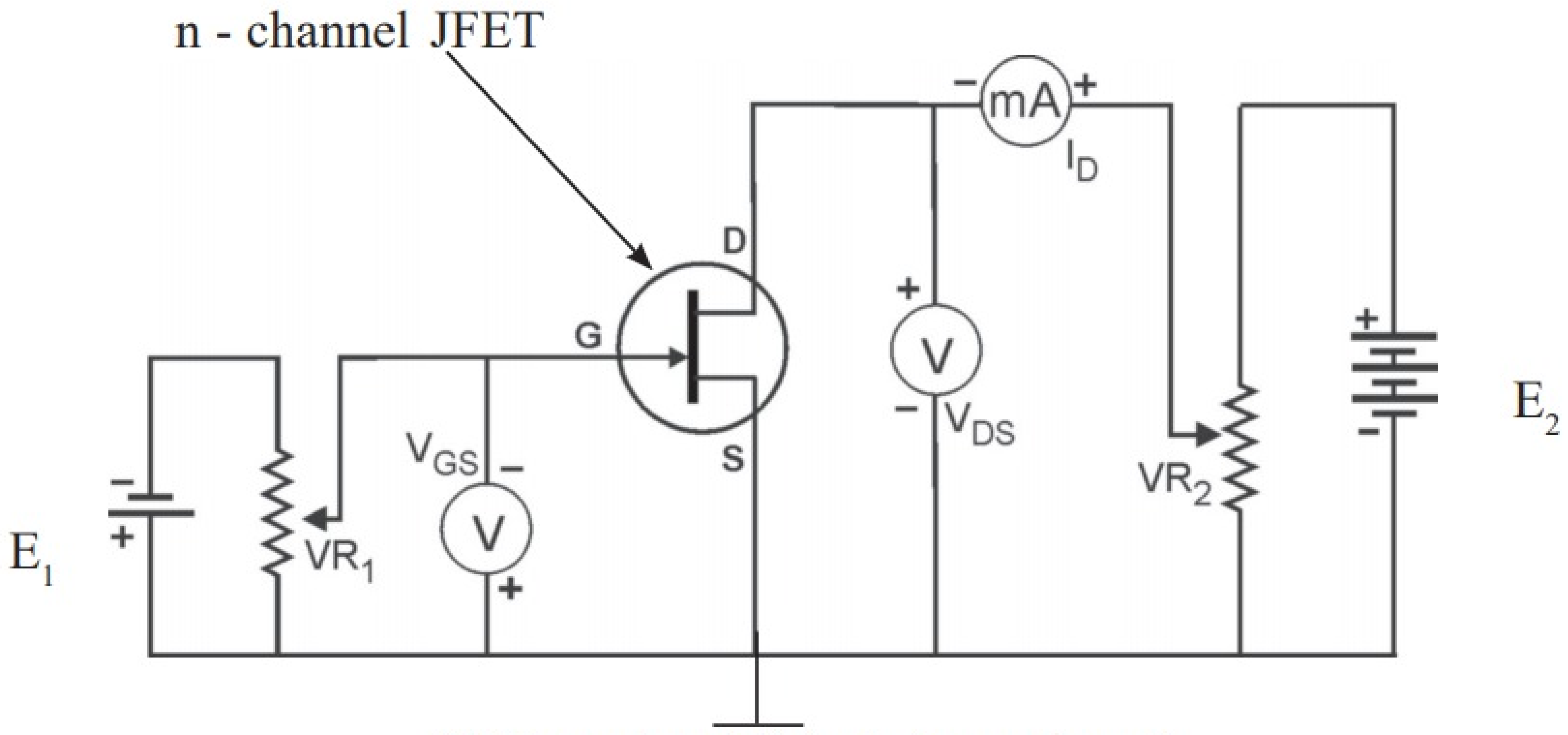
(b)

Figure 2.43

The value of the V_{DS} at which I_D becomes saturated when $V_{GS} = 0$ is called the pinch - off voltage (V_P). The pinch - off voltage (V_P) and the maximum saturation current at pinch - off state (I_{DSS}) are constant for a given FET.

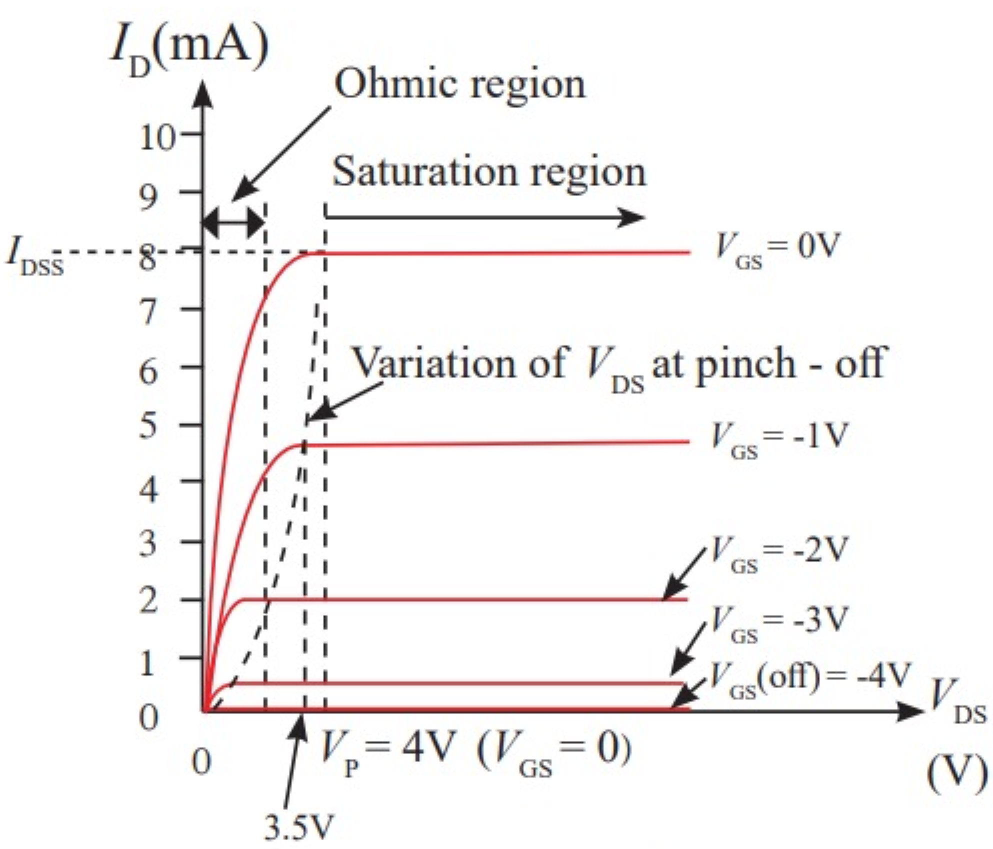
2.10.4 Behaviour of FET when $V_{GS} < 0$

So far we have considered the variation of I_D with V_{DS} when $V_{GS} = 0$. Let us now consider the variation of I_D with V_{DS} when $V_{DS} < 0$ (ie when reverse biased).



JFET experimental circuit for n - channel
Figure 2.44

To measure the value V_{DS} and I_D the circuit shown in Figure 2.44 can be used. By setting VR_1 the value of V_{GS} is kept constant at -1V, -2V, -3V,... etc. at a time. Under each value of V_{GS} adjust VR_2 to kept V_{DS} at various values so as to measure relevant I_D values for a set of V_{DS} values. By Plotting I_D values against relevant V_{DS} values for each biasing voltage of V_{GS}, a set of curves can be obtained as shown in Figure 2.45.



Variation of I_D against V_{DS} for different V_{GS} valuas

Figure 2.45

© 2020 National Institute of Education. All Rights Reserved.

According to these curves it seems that an FET having $V_P = +4$ V has been used. It can be seen that the saturation state has arrived when $V_{GS} = 0$ with a current of $I_{DSS} = 8$ mA. When $V_{GS} = -1$ V the saturation state arrives at a lower value of V_{DS} (3.5 V) with a lower saturation current (4.5 mA) than I_{DSS} (8 mA). On the occasion where $V_{GS} = -2$ V and -3 V, the relevant saturation state arrives with $I_D = 2$ mA respectively.

When V_{GS} is decreased (ie. when reverse biasing is increased) the voltage V_{DS} at pinch - off decreased. The variation of V_{DS} at pinch – off, with V_{GS} is given by the dotted curve in Figure 2.45. When $V_{GS} = -4$ V, I_D becomes zero since the depletion layer closes the channel totally as shown in Figure 2.38 (b). The V_{GS} voltage at which I_D becomes zero, is known as "cut – off voltage" and it is denoted by V_{GS} (off).

For the above FET, $V_p = +4$ V and V_{GS}(off) = -4 V. It should be noticed that the pinch – off voltage (V_p) is a value associated with V_{DS} ($V_{GS} = 0$) and the cut – off voltage is a value associated with V_{GS}. Both these values are constant for a given FET. Also notice that the numerical value of V_p pinch - off voltage and V_{GS}(off) are always equal and therefore, in FET data sheet only one value of the two is given. In most cases only V_{GS}(off) is given. It should be noted that V_p is a positive value and V_{GS}(off) is a negative value for an n - channel FET.

At V_{GS}(off) the channel gets really closed and I_D becomes zero. I_D will always be zero irrespective of the value of V_{DS} (provided $V_{DS} \geq 0$) at the biasing voltage of V_{GS} (off). Consider the pinch - off state at $V_{GS} = 0$. At pinch - off the channel does not get closed totally. Therefore, the current I_{DSS} flows through it. This is the maximum possible current which can be flown through the channel. With other possible V_{GS} values (other than $V_{GS} = 0$) the relevant V_{DS} voltage at pinch - off are less than the value of V_P which is the V_{DS} voltage at pinch - off when $V_{GS} = 0$. Also, when $V_{GS} < 0$ the relevant saturation currents are less than I_{DSS}. Figure 2.46 shows various regions with regard to behaviour of I_D, against V_{DS} when $V_{GS} = 0$.

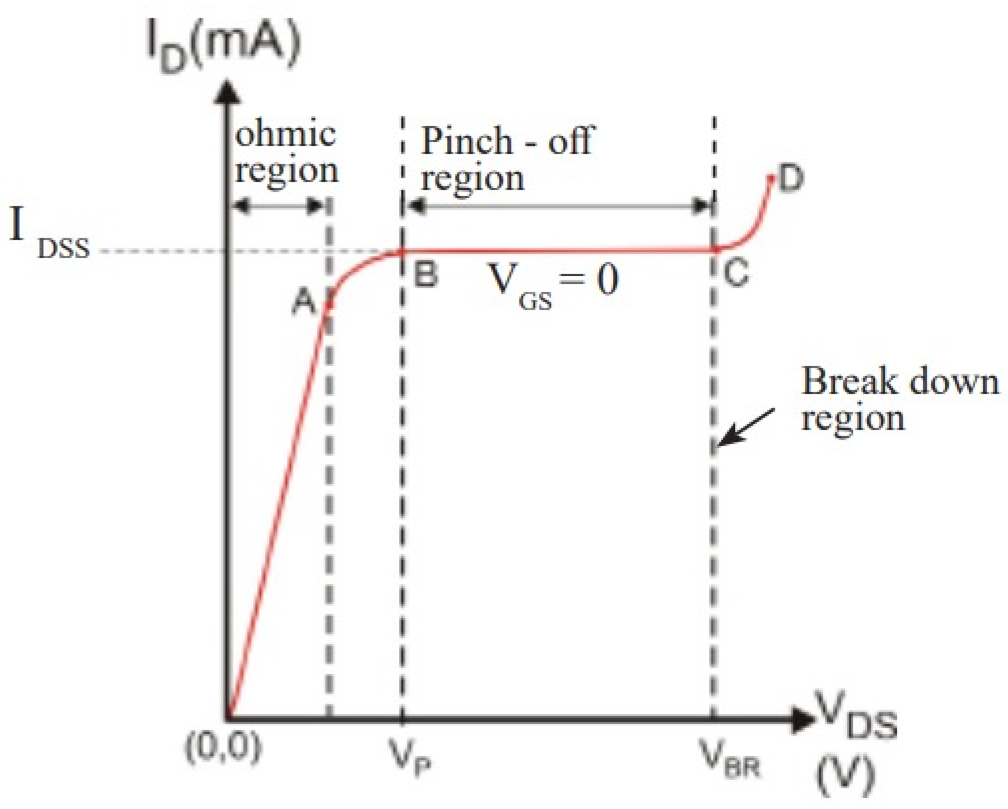


Figure 2.46

© 2020 National Institute of Education. All Rights Reserved.

Region O to A - The region in which I_D varies linearly with V_{DS} is known as the 'ohmic region'

Region A to B - The region in which the rate of increase of I_D reduces due to widening of depletion layer. The drain current tends to gradually reaching a constant value through this region.

Region B to C - At B, the voltage V_{DS} comes to the value V_p which is the pinch – off voltage. The region B to C is called the "pinch - off region" or the "saturation region". It is in this saturation region that the FET can be used as an amplifier. This is contrary to the case of BJT which is not used as an amplifier in its saturation region.

The action of amplification of the FET will be explained later. When the voltage V_{DS} is increased gradually, at a particular voltage the p- n junction in the FET undergoes an avalanche breakdown like in the case of BJT. This particular voltage is known as the breakdown voltage and is denoted by V_{BR}. The breakdown voltage depends on the particular FET and the value of it can be found from data sheets.

When V_{DS} increases beyond V_{BR}, I_D increases rapidly and the FET gets damaged so that it cannot be used again. The relevant region (C to D) on the graph is known as the breakdown region.

$I_D - V_{GS}$ characteristic (transfer characteristic)

In bipolar transistors, the change of output current I_C with the input current I_B can be given as,

$$I_C = \beta I_B$$

where β is a constant and the variation of I_C with I_B is depicted by the transfer characteristic.

In field effect transistors variation of output current I_D with input voltage V_{GS} is given by "Shockley" equation (derivation of this equation is beyond the scope of this book).

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2$$

Where, the saturation current I_{DSS} (at $V_{GS} = 0$) and the pinch – off voltage V_P (at $V_{GS} = 0$) are constant for a given FET while the output I_D changes according to square of the input V_{GS}. Therefore, this relationship is not linear unlike in the case of BJT. This non - linear nature can be seen from the transfer characteristic (I_D against V_{GS}) of the FET. The transfer characteristic can be obtained from the I_D - V_{DS} characteristic (output characteristic) of the FET. Variation of I_D with V_{GS} is depicted by the transfer characteristic shown in Figure 2.47.

© 2020 National Institute of Education. All Rights Reserved.
© 2020 National Institute of Education. All Rights Reserved.

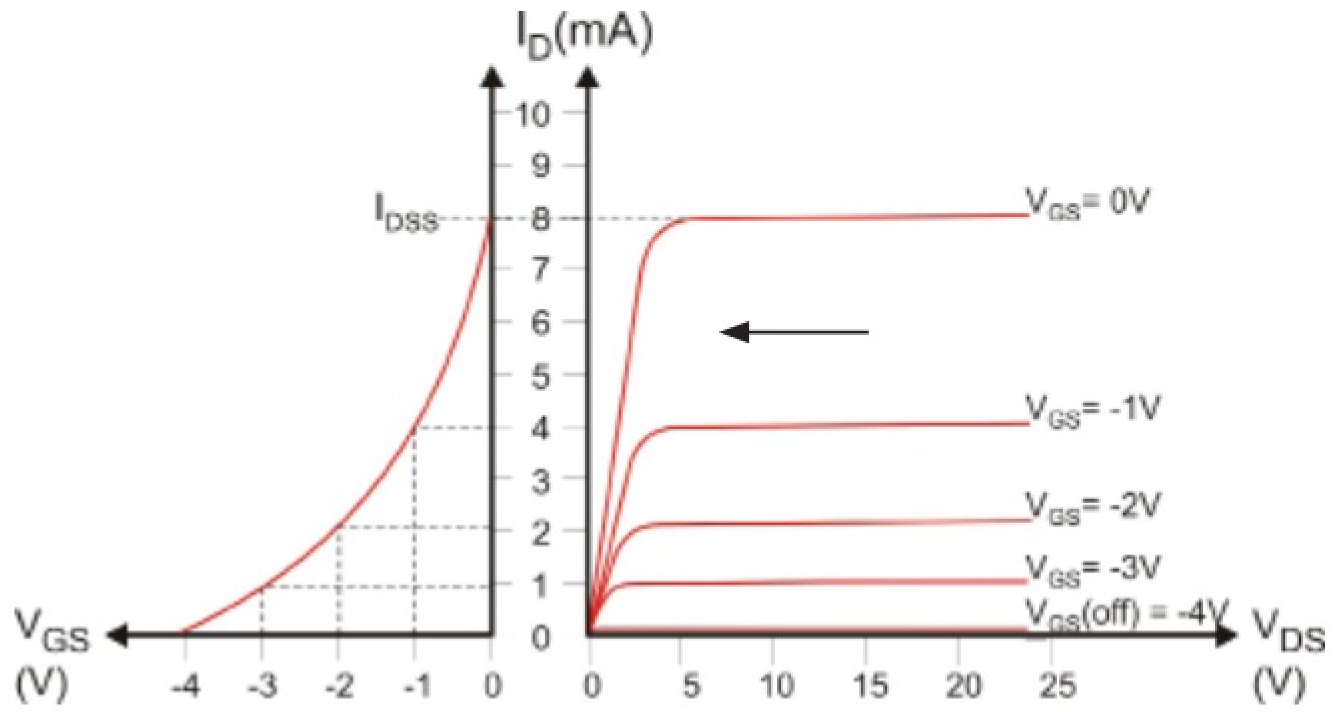


Figure 2.47

2.10.5 FET as an amplifier

Before dealing with FET amplifier, the circuit symbols used for FETs should be considered. There are several symbols used and among them there are two forms which are easier to understand. They are given in Figure 2.48.

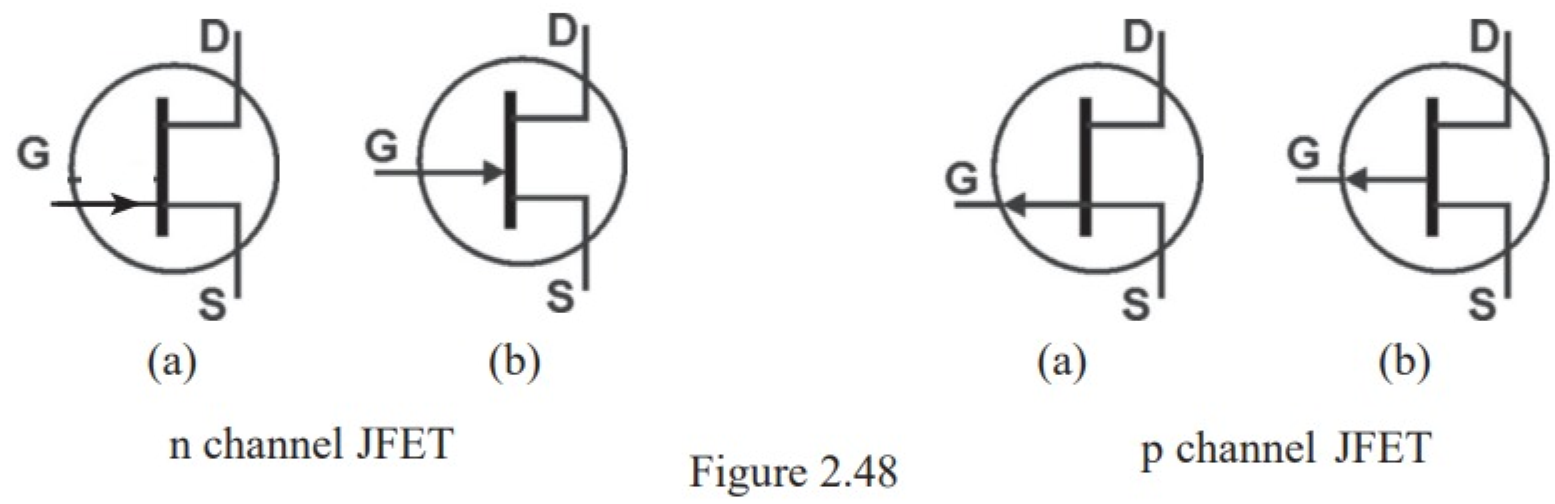


Figure 2.48

In these symbols the line segment with an arrowhead is there to identify the gate terminal (G). From the other two terminals, one which is closer to the gate is the source terminal (S). The remaining terminal is the drain (D). The arrowhead marked on the gate terminal indicates the direction of current which should flow, if the p-n junction is forward biased. However it should be noted thoroughly that the p-n junction of the FET is never made forward biased.

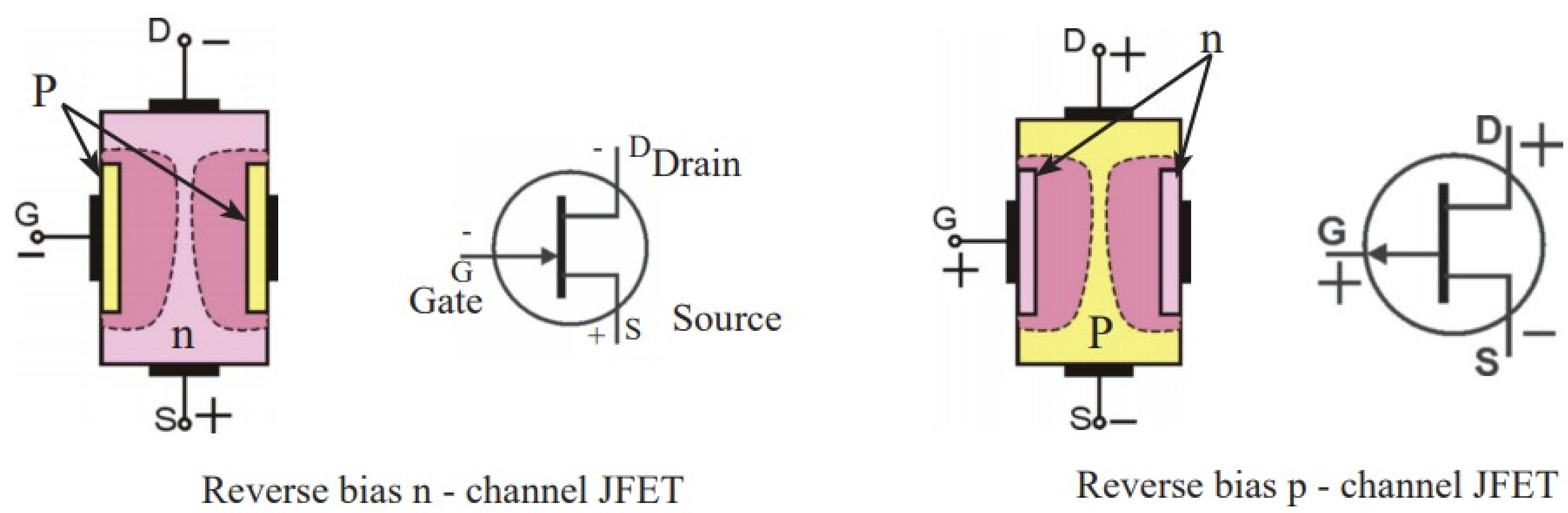


Figure 2.49

If the n- channel FET is made forward biased then the terminal G will be + ve and terminal S will be - ve. Therefore the arrowhead should be towards inside as the current flows from + ve to - ve. In case of p - channel FET, the arrowhead should be in the opposite direction, that is towards outside. This has been shown in Figure 2.49.

Like BJT, the FET also can be used as an amplifier. FET too can be used in three configurations as common – source, common - gate and common – drain.

Let us now consider the common source n – channel JFET amplifier of which the circuit is given in Figure 2.50 and common emitter BJT amplifier circuit.

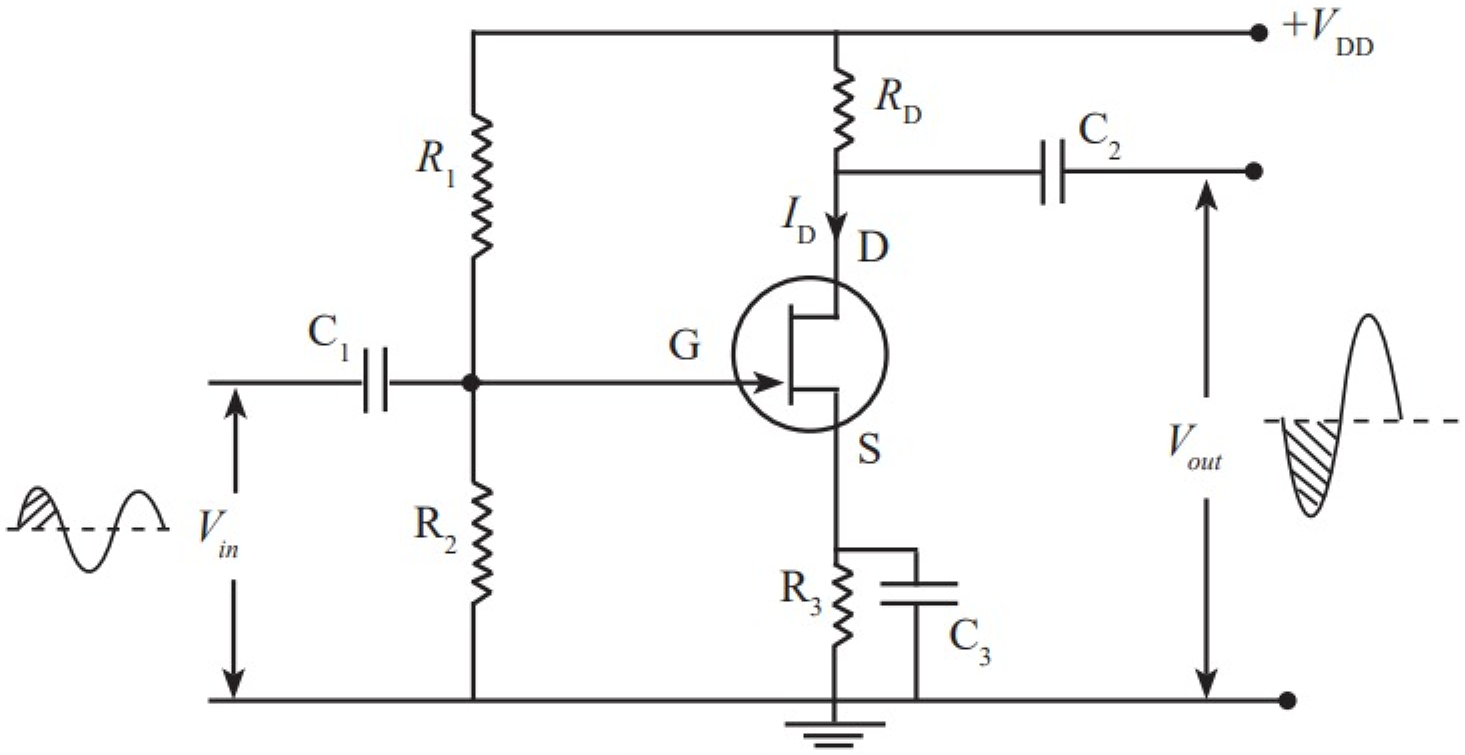


Figure 2.50

The resistance values of the resistors R_1, R_2 and R_3 in this circuit have been chosen so that V_{GS} is maintained at a negative value (i. e $V_G<V_S$). The input capacitor C_1 and the output capacitor C_2 are applied to keep the DC conditions (currents and voltages) of the transistor unchanged while transmitting signals through them. The capacitor C_3 is there to keep the DC conditions across the resistor R_3 unchanged.

When the input signal (V_{in}) is positive and increasing, the effective negative voltage V_{GS} decreases and hence the drain current (I_D) increases. When V_{in} is negative and its magnitude is increasing, the effective negative voltage V_{GS} increases and therefore I_D decreases. So, it is clear that the variations of V_{in} and I_{DS} are in - phase. This will be more clear, if you observe the signal variations shown in the I_D – V_{DS} characteristic curve for the JFET (Figure 2.51). However care must be taken not to make the effective value of V_{GS} positive in using the amplifier. If V_{GS}> 0). the transistor may get damaged due to high current passing through the channel.

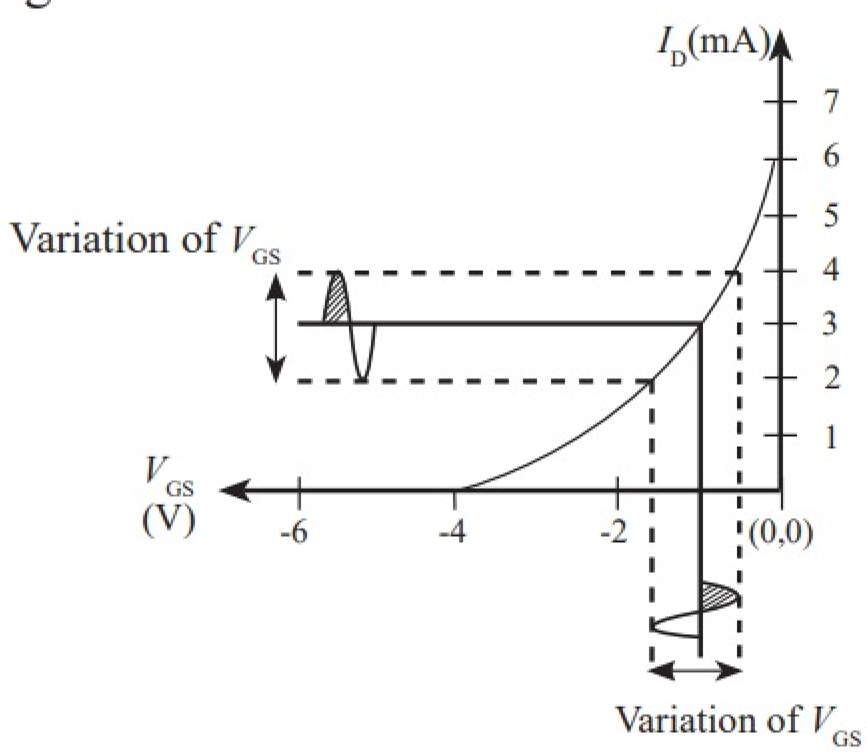


Figure 2.51

When I_D is varying in the circuit the potential drop across R_D (V_{RD}) is also varying according to $V_{RD} = I_D R_D$

Since,

$$V_{out} = V_{DD} - V_{RD}$$
$$= V_{DD} - I_D R_D$$

It seems, when I_D increases, V_{out} decreases, and when I_D decreases V_{out} increases. That is, there is a phase change of π radians between the variations of I_D and V_{out}. So, it is clear that there is also a phase difference of π radians between V_{in} and V_{out}.

As the variation of I_D with the creation of V_{GS} is strong, the amplitude of the output (V_{out}) is high. Accordingly $V_{out} > V_{in}$, and therefore a voltage amplification has taken place.

The p –n junction formed between gate and channel is reverse biased. Therefore, the input resistance of the JFET amplifier is very high (of the order of MΩ) So, the current drawn from the signal source is negligibly small. For this reason the JFET amplifier is more suitable for the amplification of small and weak signals.

2.11 Features of a good amplifier

The features that a good amplifier should have are given below with the reasons in brief, why each feature is important.

(1) Current gain should be large

(This is essential in amplifying the input current signal with a sufficient current amplification)

(2) Voltage again should be large

(This is essential in amplifying the input voltage signal with a sufficient voltage amplification)

(3) Power again should be large

(This is essential for making the output power sufficiently large. Power gain is the product of current gain and voltage gain)

(4) Input resistance should be large

When a signaling device such as a sensor, is connected to the input, the amplifier amplifies the signal received from the device. If the input resistance of the amplifier is high then the current drawn from the signaling device will be small. If a large current is drawn from the device (due to low input resistance) then there will be a considerable voltage drop (reduction) at the input. Therefore, the input voltage may not be sufficient for the correct operation of the amplifier.

(5) Output resistance should be small

It may require a large current from the output to operate a device connected to it (eg. a loudspeaker requires a large current). When the output resistance of the amplifier is low it can deliver a large current through the output.

© 2020 National Institute of Education. All Rights Reserved.

(6) Band width should be large

For any amplifier there is a range of frequencies that it can respond to. The power gain of the amplifier depends on the frequency of the input signal. The range of frequencies which can be amplified with minimum of half the maximum power gain (in the range 0 dB -3 dB) is known as the band width of the amplifier. Eg. If an audio frequency amplifier has a band width 20 Hz -20,000 Hz then it is sensitive to the entire audio range. An audio amplifier having this feature is called a High Fidelity (Hi-Fi) amplifier. For an audio system to be Hi-Fi not only the amplifier but also the input devices (eg. microphone, cassette/CD player) and the loudspeaker (output device) should be having the same band width.

2.12 Comparison of BJT and JFET

BJT	JFET
1. Bipolar device	1. Monopolar device
2. Controlled by current (I_B)	2. Controlled by voltage (V_G)
3. Input resistance is low (few kΩ)	3. Input resistance is very high (few MΩ)
4. Shows a +ve temperature coefficient at high output currents (I_C). Therefore becomes thermally unstable at high temperatures (as current increases)	4. Shows a -ve temperature coefficient at high output currents (I_D). Therefore, prevents becoming thermally unstable at high temperatures (as current decreases)
5. Less suitable for amplifying small signals (as noise is high)	5. Suitable for amplifying small signal (as noise is low)
6. Takes much room when included in ICs	6. Take relatively less room when included in ICs
7. Not used in input circuits of measuring instruments, as the input resistance is low.	7. Used in input circuits of measuring instruments, as the input resistance is very high.

2.12.2 Features of BJT and JFET amplifiers

Feature	BJT	JFET
1. Current gain	High (200)	Very high (20,000)
2. Voltage gain	Medium (40)	Medium(40)
3. Power gain	High (8000)	Very high(800,000)
4. Input resistance	Medium (2500 Ω)	Very high(1 MΩ)
5. Output resistance	Medium (20 kΩ)	Higher medium
6. Phase change of output voltage signal	180°	180°
7. Phase change of output current signal	0°	0°
8. Band width	High	High
9. Application	AF, RF Common amplifiers	AF, RF common amplifiers

Depending on the application, a suitable type of transistor can be chosen with regard to above data. Eg. To make an amplifier with high input resistance the JFET is better. BJTs are widely used today because they are easily available in the market and also they are of low cost. However due to good features of FETs now they are gradually replacing BJTs.

A type of FET which is widely used today is the MOSFET (Metal Oxide Semiconductor Field Effect Transistor). The main difference between the JFET and the MOSFET is that, there is a thin metal oxide (SiO_2) layer in a MOSFET between the gate and the rest of the parts. Metal oxide is an insulator and therefore, the input resistance of the MOSFET is about 10^3 MΩ which is about 10^3 times larger than the input resistance of the JFET. Therefore, practically I_G can be taken as zero for the MOSFET. As a result, the quiescent current (the current which flows in the circuit when the input signal is zero) is very much smaller in the amplifiers made using MOSFETs and hence they are more efficient.

Although the study of FETs other than the JFET is not included in the A/L physics syllabus, this short description is included for your extra knowledge.

Chapter Three

Integrated Circuits and Operational Amplifiers

1. Introduction

An electronic circuit may consist of a large number of electronic components. To make the circuit by assembling individual components, a considerable length of time should have to be spent.

The more the number of components, the more will be the time taken to assemble them. To overcome this difficulty a new technology had been introduced in 1960, enabling almost the whole circuit to be fabricated in a small silicon chip. Such chips are much developed today and they are known as Integrated Circuits (IC).

In an iterated circuit, components such as diodes, transistors, some resistors and some capacitors can be embedded. To connect large components like resistors, capacitors and inductors, and to connect input and output devices externally, the relevant terminals have been provided. The external view of several ICs and the interior of one have been shown in Figure 3.1.

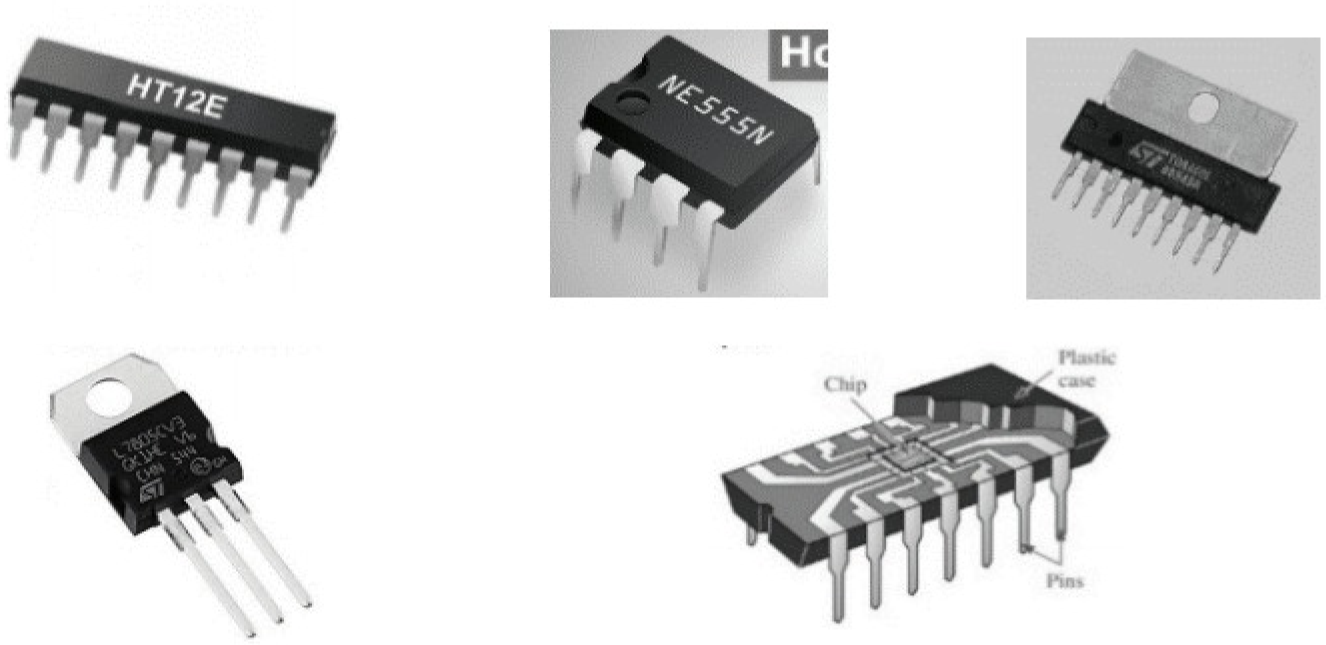


Figure 3.1

Integrated circuits with three terminals as well as more than hundred terminals are in use. The microprocessor used in modern computers is an integrated circuit with more than a million transistors and a large number of terminals. Such an IC is shown in Figure 3.2.

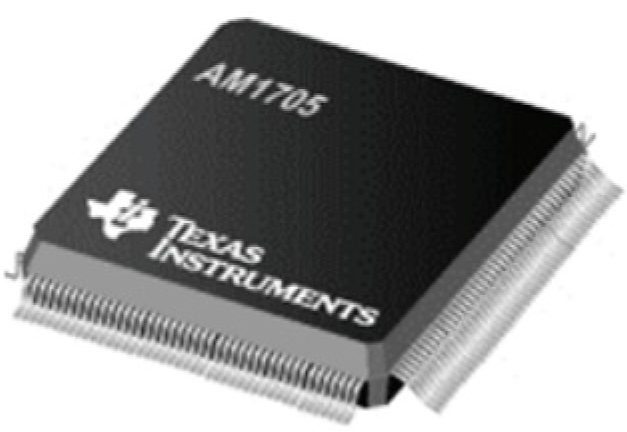


Figure 3.2

Figure 3.3 will be helpful for you to get an idea of how an electronic circuit has been fabricated in a tiny silicon chip.

© 2020 National Institute of Education. All Rights Reserved.
© 2020 National Institute of Education. All Rights Reserved.

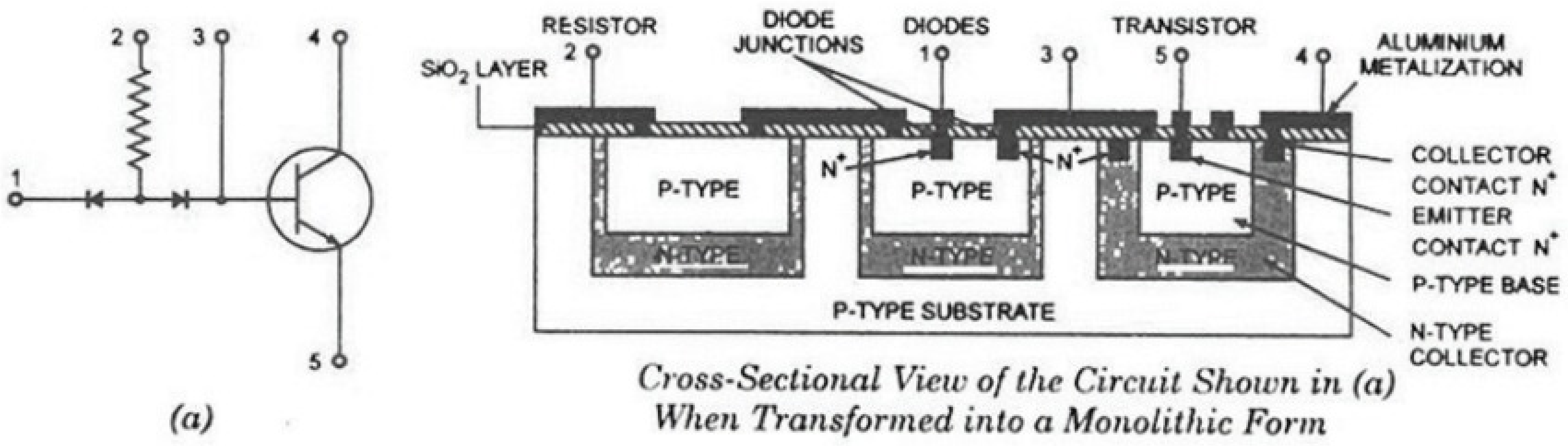


Figure 3.3

3.2 Scales of integration

Depending on the number of components in them, the integrated circuits are scaled as follows.

Table 3.1 Scales of Mtegroton m ICS

Number of components in the chip	Scale of integration
Less than 100	Small Scale Integration (SSI)
Between 100 and 1000	Medium Scale Integration (MSI)
Between 1000 and 10000	Large Scale Integration (LSI)
More than 10000	Very Large Scale Integration (VLSI)

3.3 Advantages and limitations of integrated circuits

When a circuit with a large number of components is in the form of an IC, it has the advantage of occupying less space. There are other advantages too, as ICs are lighter, cheaper and more reliable in operation.

Resistors of high resistance and power, and capacitors of high capacitance, inductors (coils) and transformers cannot be fabricated in a chip. These are the limiting factors regarding the integrated circuits.

3.4 Numbering the terminals of an IC

As mentioned earlier an IC has terminals to connect additional components and to supply power. The terminals are numbered in the manner given below.

As shown in Figure 3.4 , when viewed from above, a notch can be seen on a side of the IC. Mostly, a mark of a dot is also there close to it. When you position the IC so that the notch is on your left, the terminal closer to and below the notch (or the dot) is numbered as terminal 1. The adjacent terminal on the right is terminal 2. The terminals are numbered in this sequence and the terminal at the right end is numbered as the terminal 8 (as with this example). The terminal 9 will be the one which is on the top of the right end. (this can be regarded as numbering the terminals in the anticlockwise direction starting from the terminal closer to and below the notch). This has been shown in Figure 3.4

© 2020 National Institute of Education. All Rights Reserved.

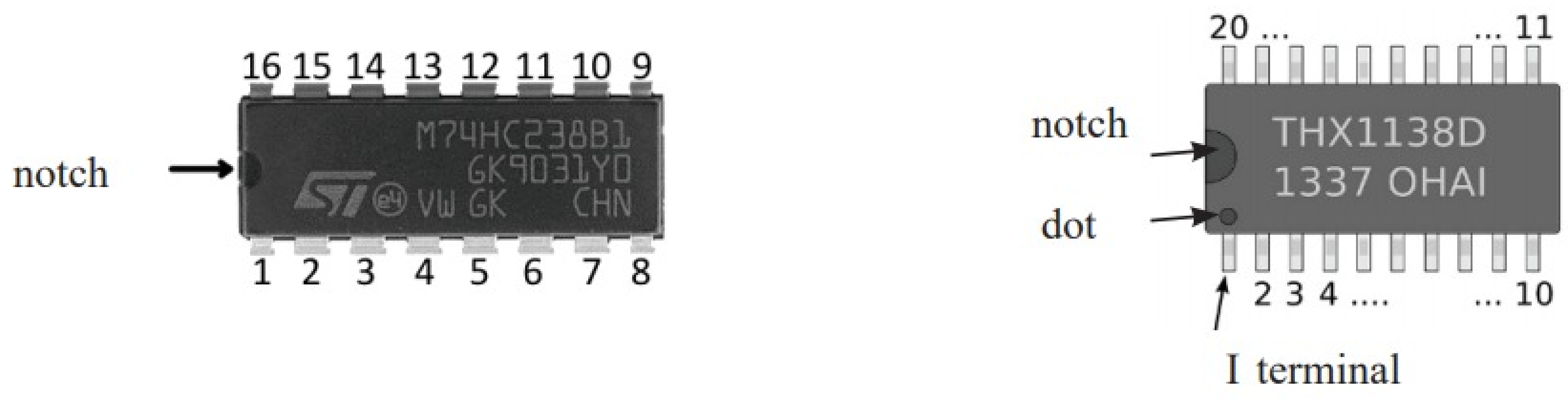


Figure 3.4

3.5 Operational amplifier

Some amplifier circuits have been discussed in the previous chapter. The operational amplifier can be introduced as an amplifier having almost all the good features and in the form of an IC. Because of its specific amplifying features the op-amp can be used to do mathematical operations such as addition, subtraction and multiplication. Hence the name operational amplifier.

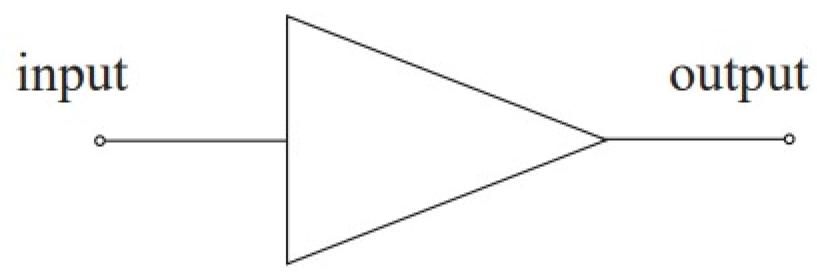
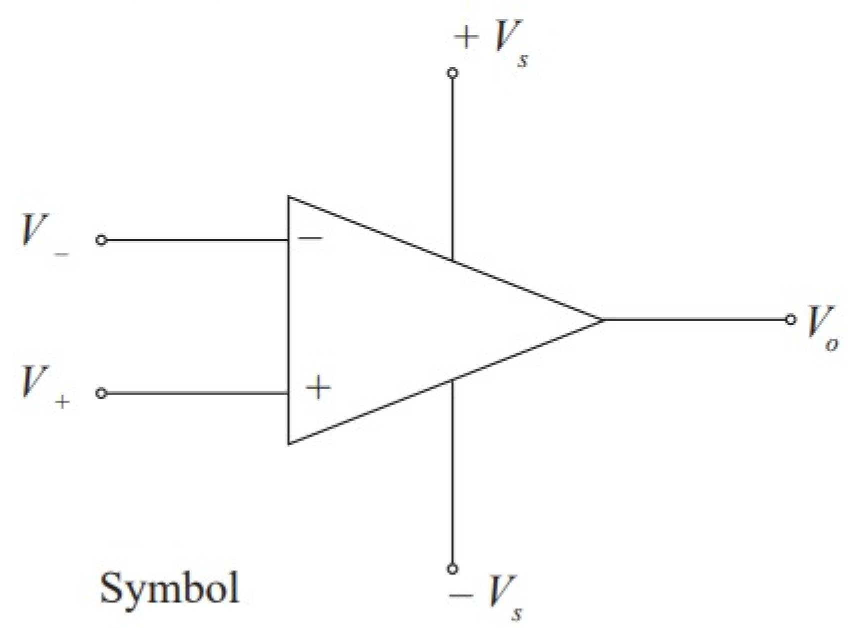


Figure 3.5

An ordinary amplifier has one input only. The signal applied to the input is amplified and then output. The circuit symbol of such amplifiers is shown in Figure 3.5. A special feature of the op-amp is that it has two input terminals. One of which is named as the positive (+) terminal and the other as the negative (-) terminal. The external view and the circuit symbol of the op-amp are given in Figure 3.6.

External view



Symbol

Figure 3.6

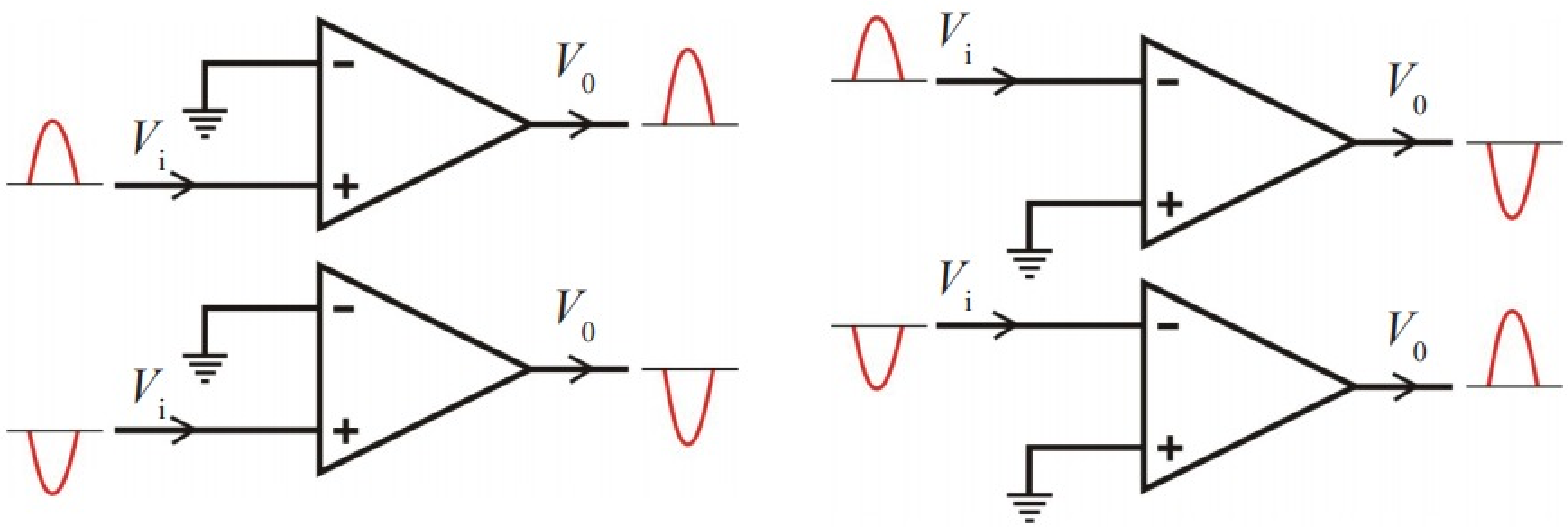
V_+ input is such that, the input voltage signal is output without inversion (i.e the output signal is with the same phase as the input signal).

V_- input is such that, the input voltage signal is output with inversion (i.e the output signal is 180^0 out of phase with the input signal).

By comparing the input and the relevant outputs of the op-amps shown in Figure 3.7 you can realize it.

Special note:

To get the output signals as shown the input voltage (V_i) should be very much smaller (about $V_i < 50\ \mu V$). Although such small voltage signals are not practical, this has been presented only for the mere purpose of explaining the difference between V_+ and V_- inputs. A detailed explanation will be given later.



Applying V_i signal to + the output. Then the output is in - phase. That is, an inversion has taken place

Applying V_i signal to - the output. Then the output is out of - phase. That is, an inversion has taken place

Figure 3.7

Open-loop usage

The use of the op-amp as it is, without any external circuit loop (without feeding any fraction of the output voltage to the input again), is the open-loop state. Needed will be a explained later

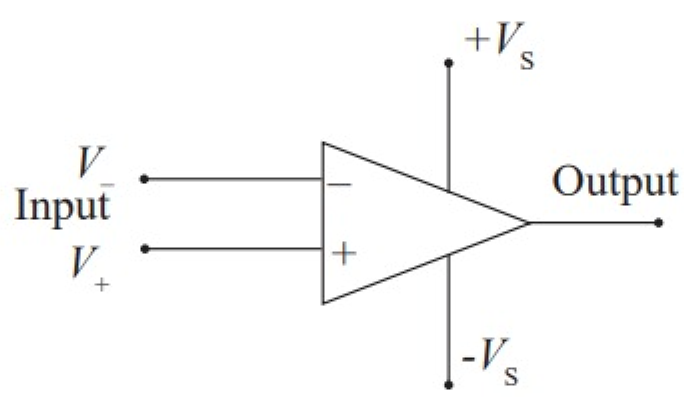


Figure 3.8

When supplying power to the op-amp through the terminals +Vs and -Vs a dual voltage supply with +ve and –ve potential should be used. Ground terminal (with zero potential) should also be there. The +ve and –ve terminals of the dual voltage supply should be connected to the $+V_s$ and $-V_s$ terminals of the op-amp respectively (Figure 3.8). By applying a voltage signal to V_+ or V_- input appropriately, the output can be obtained with or without inversion. To achieve this it is necessary that the output voltage should be symmetrical towards +ve potential with respect to the ground potential. Therefore, the supply voltage should be symmetrical as $+V_S$ and $-V_S$ with respect to ground (zero potential). Figure 3.9 shows how to arrange such a dual voltage supply (symmetrical about ground potential) using cells.

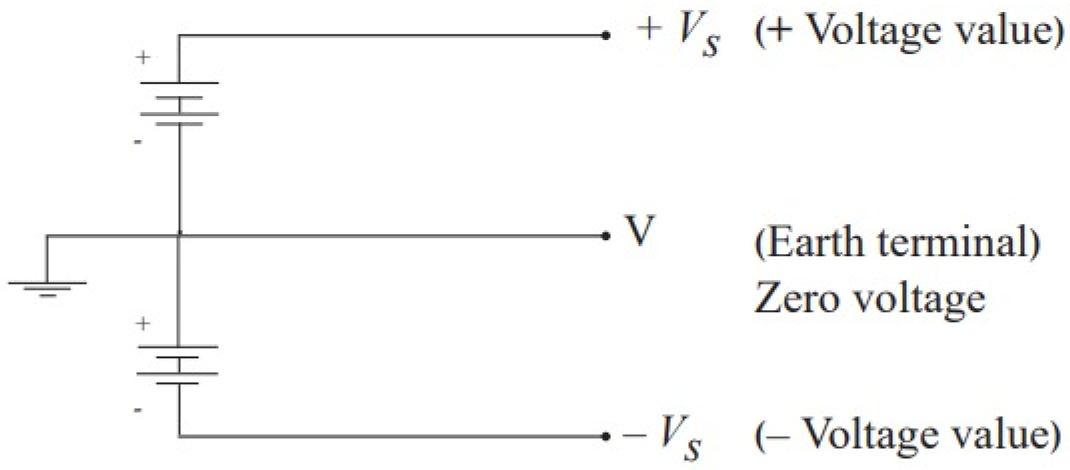


Figure 3.9

3.5.2 Characteristics of open-loop op-amp

Consider the circuit shown in Figure 3.10.

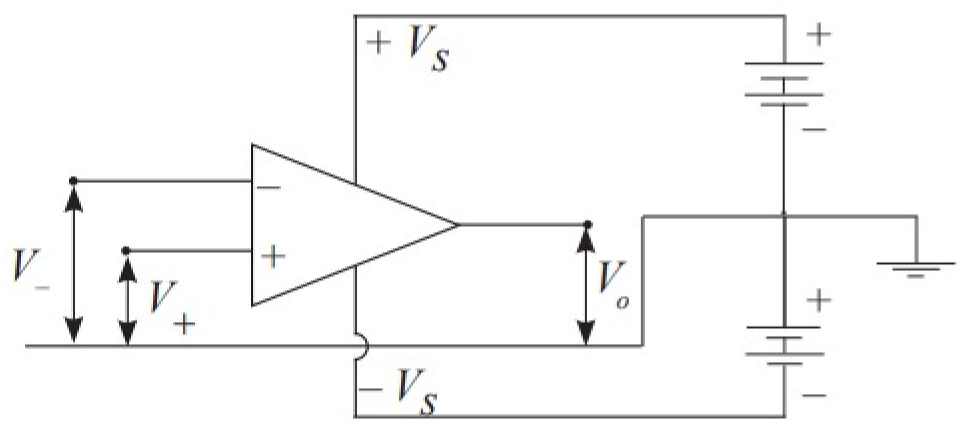


Figure 3.10

The voltage V_- applied to the – input is called the inverting input.

The voltage V_+ applied to the + input is called the non-inverting input.

The difference between the inputs V_+ and V_- is called the difference input.

That is, difference input = $V_+ - V_-$

In fact the op-amp amplifies this difference input.

The output voltage is denoted by V_0.

The differences in input and the output are related as follows.

$$\boxed{V_O = A_O\ (V_+ - V_-)}$$

where A_0 is the open-loop voltage gain of the op-amp.

The characteristic curve, V_0 against $(V_+ - V_-)$ of an op-amp is shown in Figure 3.11.

© 2020 National Institute of Education. All Rights Reserved.

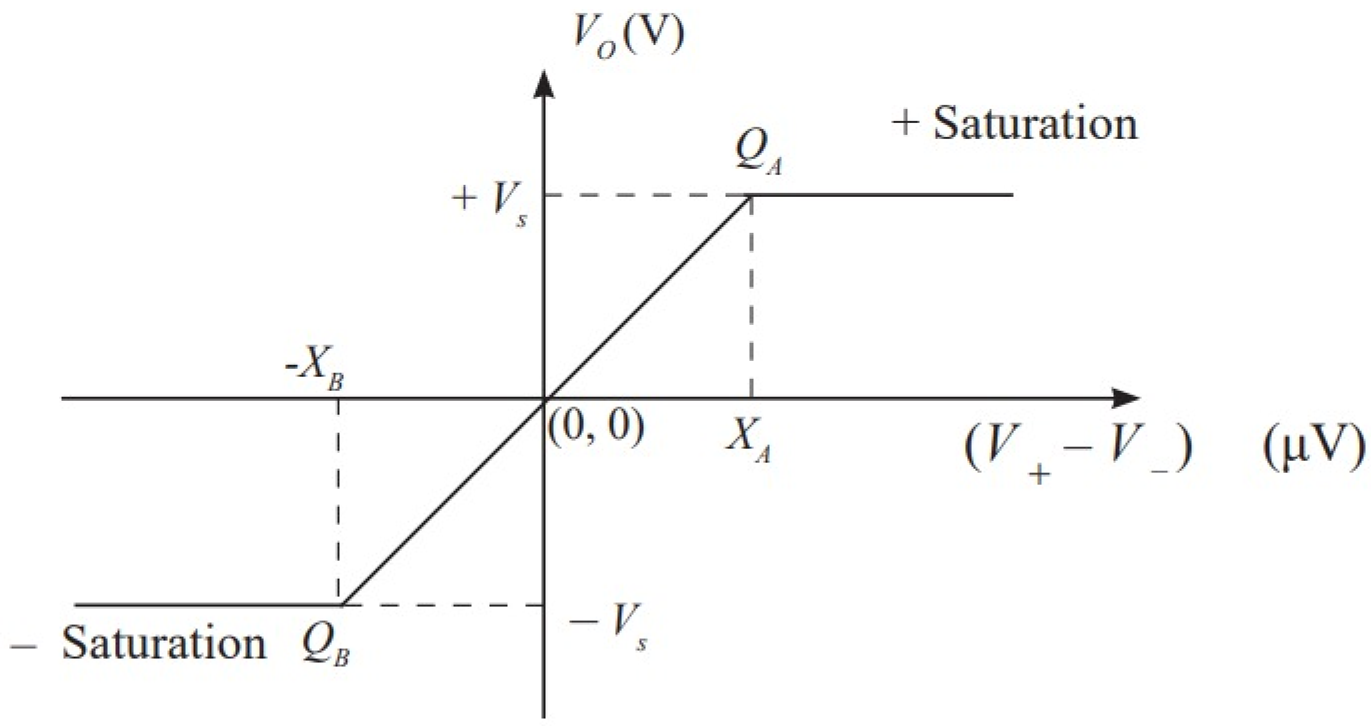


Figure 3.11

Considering the information obtainable from this characteristic curve, the following facts can be stated regarding the op-amp.

1) When the difference input $(V_+ - V_-)$ is positive (i.e. when $V_+ > V_-$), at a particular value X_A (marked on the axis) of the difference input, the output voltage (V_O) reaches a constant positive value. This positive value is nearly equal to the positive voltage ($+V_S$) of the supply voltage . This is known as the saturation at positive. This saturation region is on the right of point Q_A marked on the graph.

2) When the difference input $(V_+ - V_-)$ is negative (i.e. when $V_+ < V_-$), at a particular value X_B (marked on the axis and X_B is such that $X_B = X_A$) of the difference input, the output voltage (V_O) reaches a constant negative value. This negative value is nearly equal to the negative voltage ($-V_S$) of the supply voltage. This is known as the saturation at negative. This saturation region is on the left of the point Q_B marked on the graph.

3) The magnitude of voltage at points X_A and X_B is very much smaller (about 100 μV). This means the op-amp becomes saturated at a very small value of difference input. Therefore, it is clear that the output voltage of the op-amp (V_O), varies linearly with the input voltage $(V_+ - V_-)$ only within a very small range (from $-X_B$ to $+X_A$).

Accordingly, the open loop op-amp behaves as amplifier within a very much smaller range of input voltage, where there is a linear variation between input and output voltages. This is in the region from Q_A to Q_B on the curve. This region is known as the linear region.

Considering, $V_O = A_O (V_+ - V_-)$

$$(V_+ - V_-) = \frac{V_O}{A_O}$$

So, it can be seen that the magnitude of $(V_+ - V_-)$ required to get the saturation value of V_O , is dependent upon A_0 (the open loop gain).For an op-amp with a higher value of A_0, the relevant $(V_+ - V_-)$ value is less. That is, the more the value of A_0 is, the narrower the linear region will be.

© 2020 National Institute of Education. All Rights Reserved.

Consider an op-amp having the open loop voltage gain 10^5 and operating with a $\pm$ 15 V dual voltage supply.

$$V_O = A_O\,(V_+ - V_-)$$

At saturation $V_O = V_S$

$$\therefore\ V_S = A_O(V_+ - V_-)$$

$$\therefore\ (V_+ - V_-) = \frac{V_S}{A_O}$$

$$= \frac{\pm 15}{10^5}$$

$$= \pm 150\ \mu\text{V}$$

$\therefore$ The linear region is confined to a very narrow region, from -150 μV to +150 μV.

If the open-loop voltage gain had been 10^6 ,

$$V_+ - V_- = \frac{V_S}{A_O}$$

$$= \frac{\pm 15}{10^6}$$

$$= \pm 15\ \mu\text{V}$$

Then the linear region would have been confined to a very much narrower region, from -15 μV to +15 μV.

3.5.3 Specific properties of the op-amp

The operational amplifier has the following specific properties.

1) Open loop voltage gain is very large. In the ideal case it should be infinity and for the practical op-amp the voltage gain is about 10^5.

2) Input resistance is very large. In the ideal case, it should be infinity and for the practical op-amp this resistance is about from $10^6\ \Omega$ to $10^{12}\ \Omega$. Because of this property, the current drawn by the circuit from the input is negligible. Therefore, the effect (loss) on the input voltage is minimized.

3) Output resistance is very small. In the ideal case, it should be zero and for the practical op-amp this resistance is 100 Ω - 200 Ω . This makes the process more efficient. (Energy loss at the output is low).

3.5.4 Voltage comparator

Consider the op-amp circuit shown in Figure 3.12.

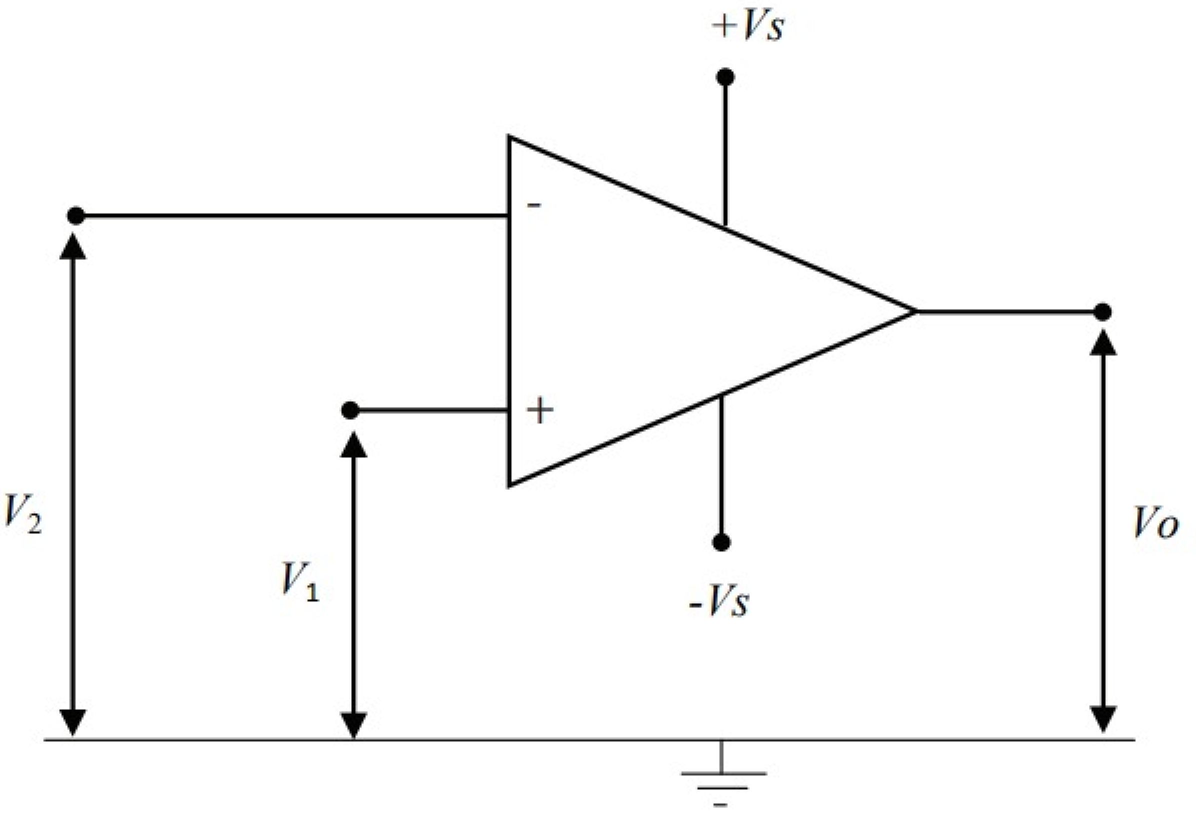


Figure 3.12

In this circuit, when $V_1 > V_2$ the difference input is positive and the output becomes nearly $+V_S$ (saturation at positive). This happens practically as the linear region of the open loop op-amp is very much narrow.

When $V_1 < V_2$, the difference input is negative and the output becomes nearly $-V_S$ (saturation at negative). This also happens practically as the linear region of the open loop op-amp is very much narrow.

Therefore, the op-amp compares the input voltage V_1 and V_2 and gives two output levels as,

$$\text{if } V_1 > V_2 \text{ then } V_0 = +V_S$$
$$\text{if } V_1 < V_2 \text{ then } V_0 = -V_S$$

Therefore, the op-amp has functioned as a voltage comparator.

Eg. A voltage comparator circuit designed using an op-amp is shown in Figure 3.13. The comparator is provided with two input voltages V_1 and V_2.

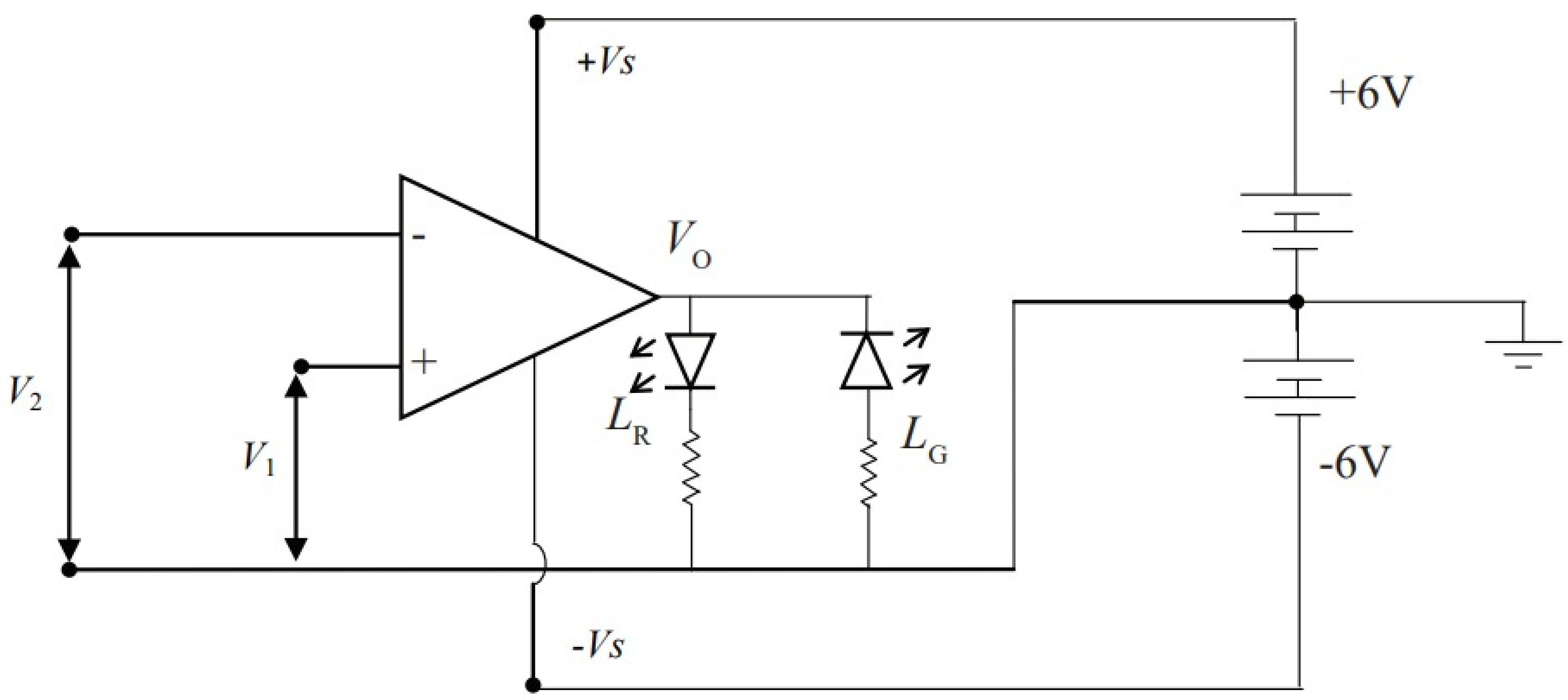


Figure 3.13

The comparator compares V_1 and V_2 and,

if $V_1 > V_2$ it lights the red LED (L_R)
if $V_1 < V_2$ it lights the Green LED (L_G)

When $V_1 > V_2$, $V_0 = +V_S = +6$ V. Then L_R becomes forward biased and it lights. During this period L_G does not light as it is reverse biased.

When $V_1 < V_2$, $V_0 = -V_S = -6$ V. Then L_G becomes forward biased and it lights. During this period L_R does not light as it is reverse biased.

3.5.5 Op-amp as a switch

If one input of the voltage comparator described above is kept at a constant voltage, then depending on the other input the comparator can be made to switch ON/OFF. This is shown in Figure 3.14.

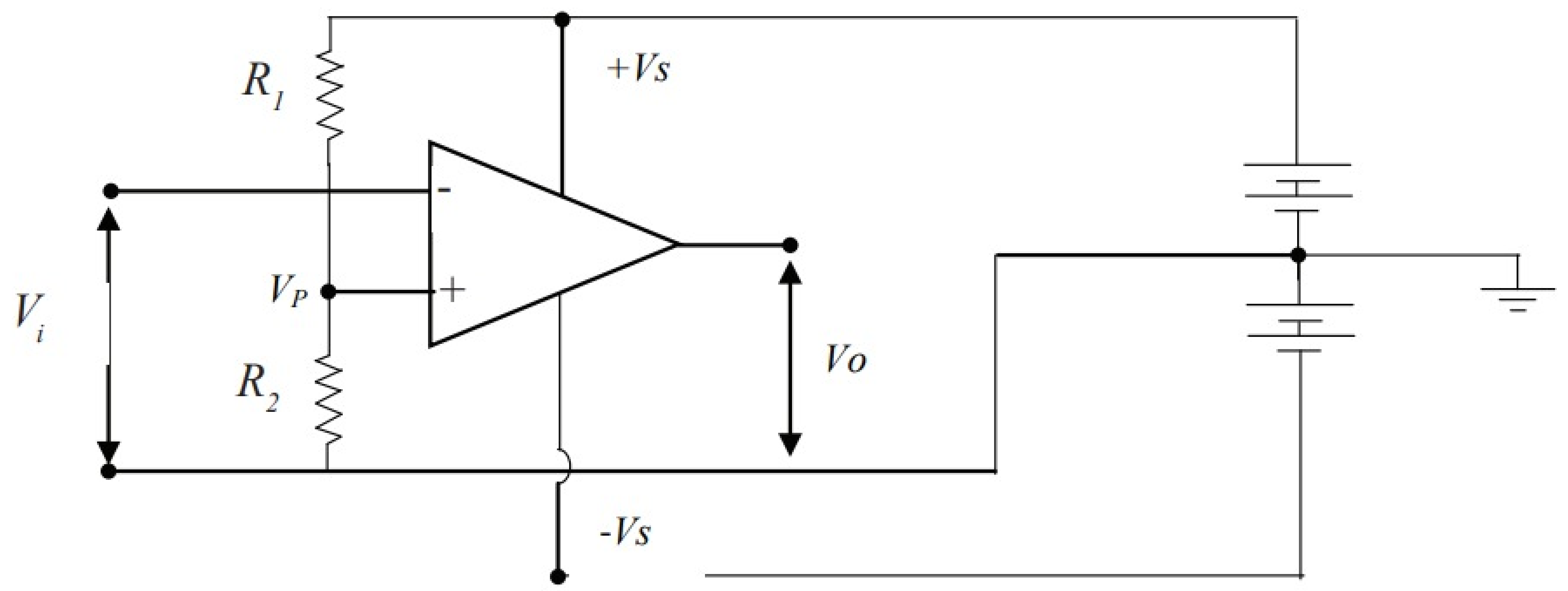


Figure 3.14

Some constant voltage is given to the + input of the op-amp by means of the potential divider consisting of resistors R_1 and R_2. Take this constant voltage as V_P. Now, depending on the voltage V_i applied to the – input of the op - amp the output V_o is decided as,

if $V_i < V_p$ then $V_0 = +V_S$
if $V_i > V_p$ then $V_0 = -V_S$

This is dipcted by the graph shown in Figure 3.15.

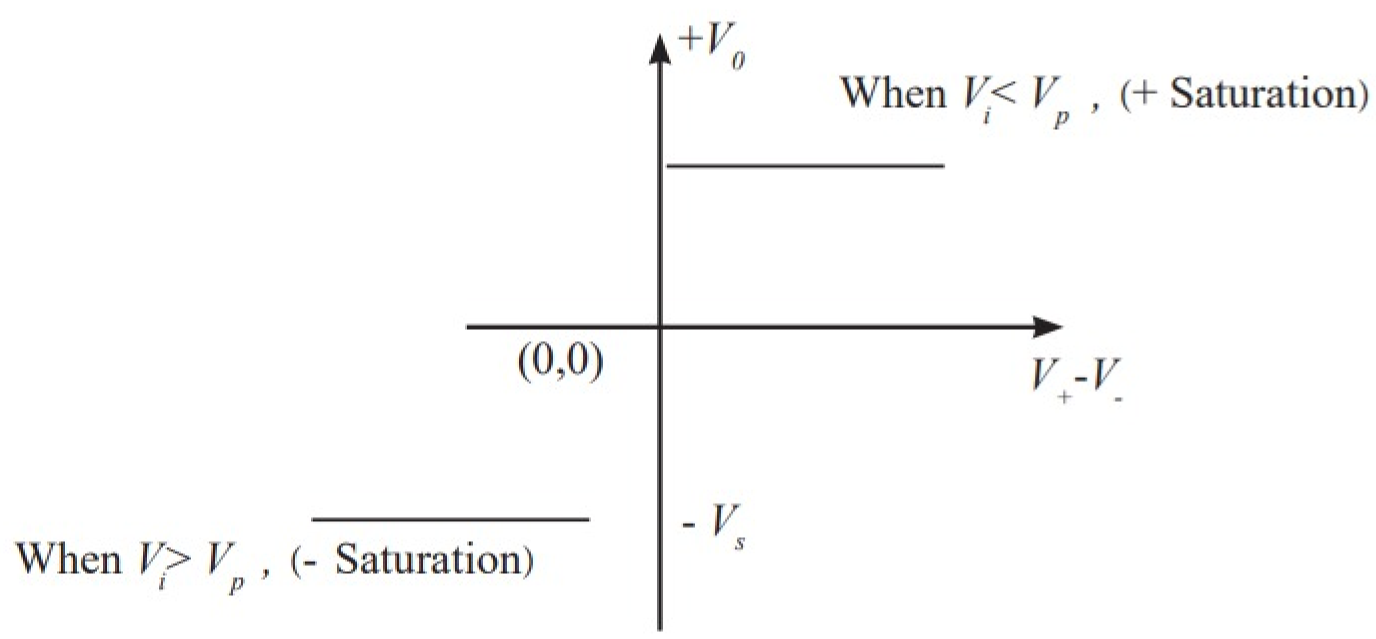


Figure 3.15

According to the above, since the output appears as a high or low voltage level depending on the value of input V_i, this circuit can be used as a switch operated by V_i .

When output becomes $+V_s$ (nearly) it can be taken as switching ON and when it becomes $-V_S$ (nearly) can be taken as switching OFF. Or else becoming output $-V_S$ can be taken as switching ON and it becoming $+V_s$ can be taken as switching OFF. (Although it is taken as the magnitude of the output, voltage is V_s at saturation, practically it may be slightly less than V_s. This is due to the fact that there are tiny voltage drops across some components in the circuit).

We can also arrange the circuits so that the switch can be operated by the voltage applied to the + input. For that, the voltage V_P should be given the – input.

In that case the output voltage will be as follows.

if $V_i > V_p$ then $V_0 = +V_S$

if $V_i < V_p$ then $V_0 = -V_S$

3.5.6 Practical data

The integrated circuit μA 741 is an op-amp which is widely used today. The external view and the internal pin connection diagram are shown in Figures 3.16 (a) and Figures 3.16 (b) respectively.

(a) Figure 3.16 (b)

The 8th terminal of the I_C is not connected to any circuit point. It is there only for symmetry of pin arrangement. The 1st and 5th terminals are used only for a special adjustment and it is not included in the A/L physics syllabus. (When no signal is applied to the op-amp inputs, the output voltage should be zero. However, practically there may be a small output voltage. This may be due to some unbalanced condition of the electronic components. In such a case by applying suitable voltages to the 1st and 5th terminals using a potential divider, the output voltage can be made zero). For the IC µA 741, it is necessary to use a dual voltage supply with V_S in the range $\pm$ 5V to $\pm$ 15V . $+V_S$ and $-V_S$ voltages should be given to the 7th and 4th terminals respectively. The open loop gain of this op-amp is 10^5.

3.5.7 Close-loop usage

When the op-amp is used with open-loop, because of its enormous voltage gain, the linear region is very much narrower. Therefore, even with a tiny input voltage signal (of the order of μV) it becomes saturated. To overcome this, when the op-amp is used as an amplifier its high voltage gain is reduced. Then its linear region gets widened. To achieve this, a suitable fraction of output voltage is fed to the inverting input again, using an external resistor. Since the output voltage is 180^0 out of phase with the input voltage, this arrangement reduces the voltage gain. Applying a suitable fraction of output signal to the input so that he amplification is reduced, is called a **negative feed-back**.

When the op-amp is used as an amplifier in which a negative feed-back has been applied by connecting an extra resistor it is known as the op-amp with close-loop. The external resistor used to apply the appropriate fraction of output voltage to the input is called the **feed-back resistor**. It is denoted by R_f in Figure 3.17.

For simplicity the voltage supply is not shown in the circuit.

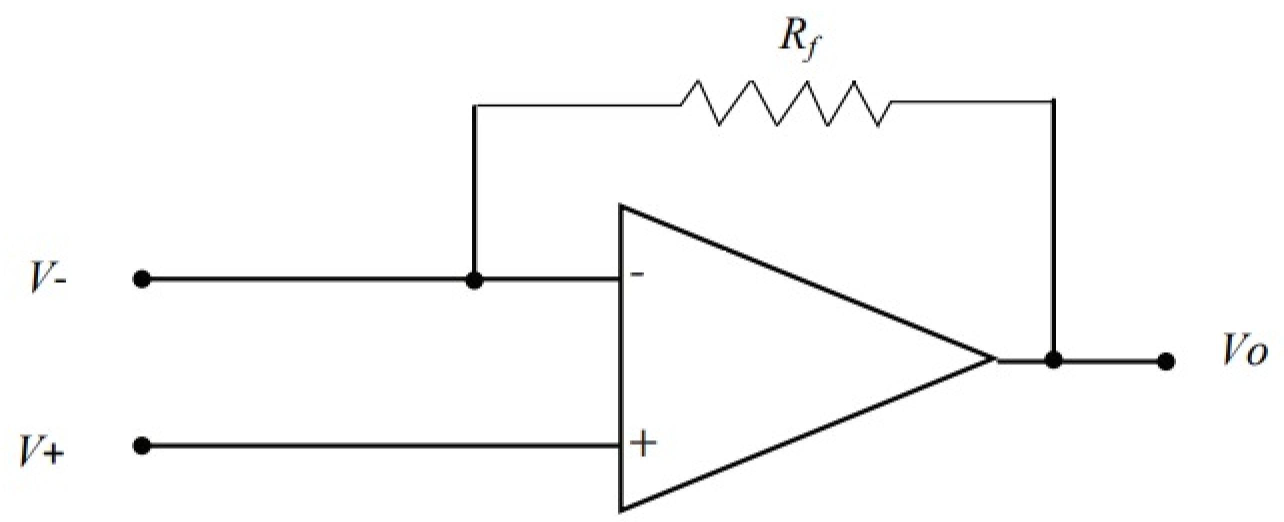


Figure 3.17

3.5.8 Golden Rules

There are two rules known as " Golden Rules " which are extremely useful for circuit analysis of the op-amp. They are as follows.

I. When an op-amp operates in the **linear region**, the voltage difference between the two op-amp input terminals is equal to zero.

   i.e. $(V_+ - V_-) = 0$

   (In a calculation carried out earlier, the possible magnitude of $(V_+ - V_-)$ in the linear region was found to be limited to 150 μV. Such a tiny voltage can be regarded practically, as equal to zero.)

II. The currents flowing into the op-amp input terminals, equal zero.

   (Since the input resistance of the op-amp is very high, under the voltage inputs supplied, the current drawn into the op-amp μA 741, is about 0.08 μA. This can be regarded practically, as zero)

© 2020 National Institute of Education. All Rights Reserved.

3.5.9 Op-amp with close –loop

The open-loop op -amp gives a huge voltage gain. By applying a negative feed-back through an external circuit loop, this huge voltage gain can be brought to a finite practical value. Then the linear region of the op-amp widens and it can be used as a practical amplifier circuit. There are two types of such amplifiers.

1. Inverting amplifier
2. Non-inverting amplifier

3.5.10 Inverting amplifier

The term "inventing " means that the output voltage signal of the amplifier is 180° out of phase with the input voltage signal. Such an amplifier circuit is shown in Figure 3.18.

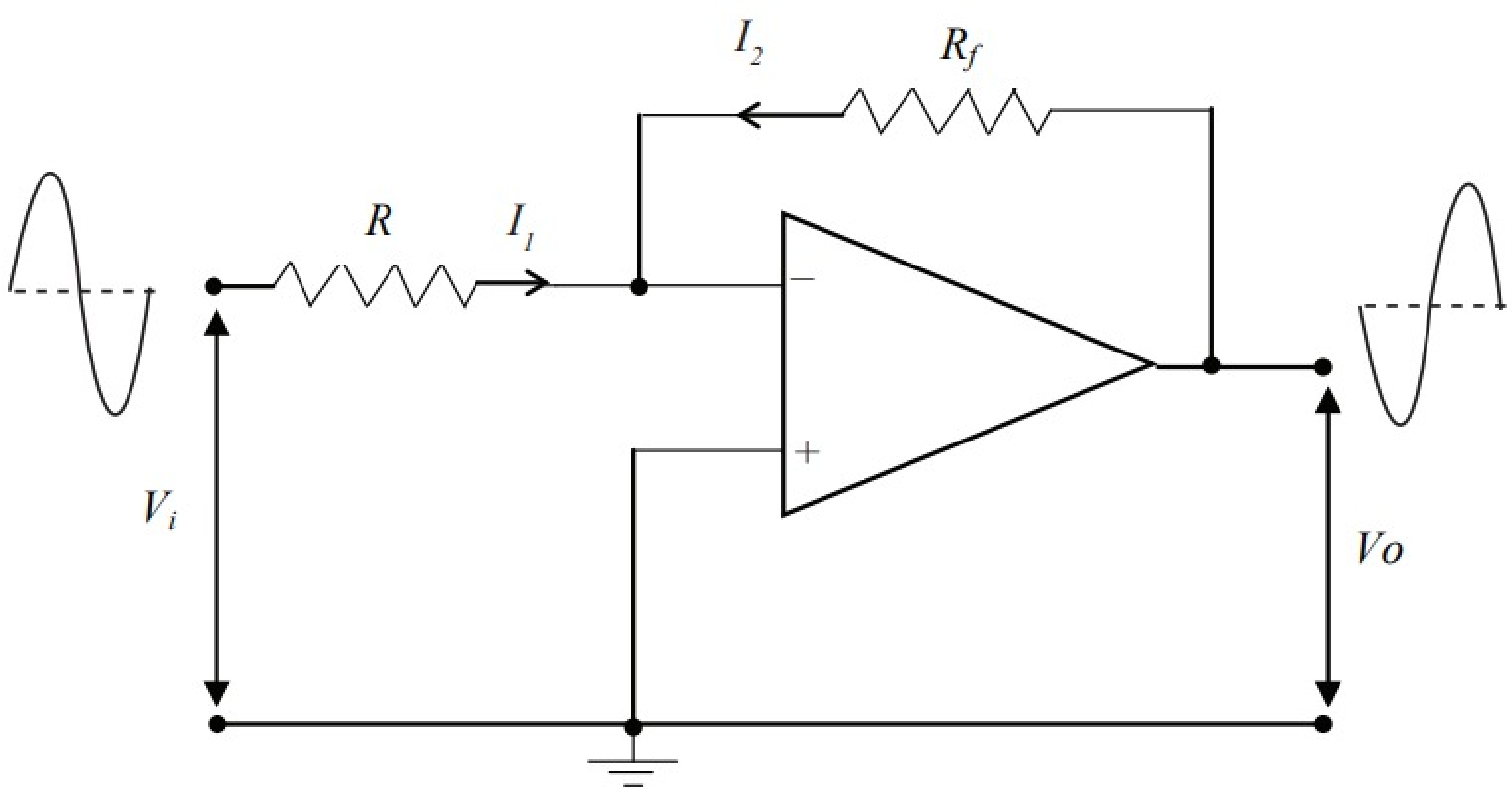


Figure 3.18

Since the + input of the op-amp has been grounded it is zero, when the amplifier is operated in the linear region,

According to " Golden rule I" ,

$$(V_+ - V_-) = 0$$

But due to $V_+ = 0$ (as + input has been grounded)

$$V_- = 0$$

This – input terminal, the potential of which has became zero is considered to be a "Virtual earth".

Then, Potential difference across $R = V_R = V_i - V_- = V_i - 0 = V_i$
Potential difference across $R_f = V_{Rf} = V_o - V_- = V_o - 0 = V_o$

Taking, I_1 as the current flowing to the – input terminal through resistor R, due to input voltage (V_i)
I_2 as the current flowing to the – input terminal tough resistor R_f, due to output voltage (V_o)

© 2020 National Institute of Education. All Rights Reserved.

According to " Golden rule II" ,

The current flowing into – input is zero. Therefore, from Kirchhoff 's 1st rule,

$$I_1 + I_2 = 0$$

$$\therefore \quad \frac{V_R}{R} + \frac{V_{Rf}}{R_f} = 0$$

$$\therefore \quad \frac{V_i}{R} + \frac{V_O}{R_f} = 0$$

$$\frac{V_O}{R_f} = -\frac{V_i}{R}$$

$$\frac{V_O}{V_i} = -\frac{R_f}{R}$$

$\frac{V_O}{V_i}$ is the **voltage gain** of the close-loop inverting amplifier.

if it is denoted by G_V, then,

$$\therefore \quad \boxed{G_V = -\frac{R_f}{R}}$$

This expression which stands for G_V has a negative sign on its right hand side. This – sign indicates that the output voltage makes a phase difference 180^0 with the input voltage. That is , the output is inverted with respect to the input. (Observe the waveforms at the input and output terminals of the amplifier circuit given in Figure 3.18). With the appropriate selection of the resistance values of R and R_f , the voltage gain of the inverting amplifier can be set as required.

The characteristic of V_0 against V_i for an inverting amplifier is shown in Figure 3.19.

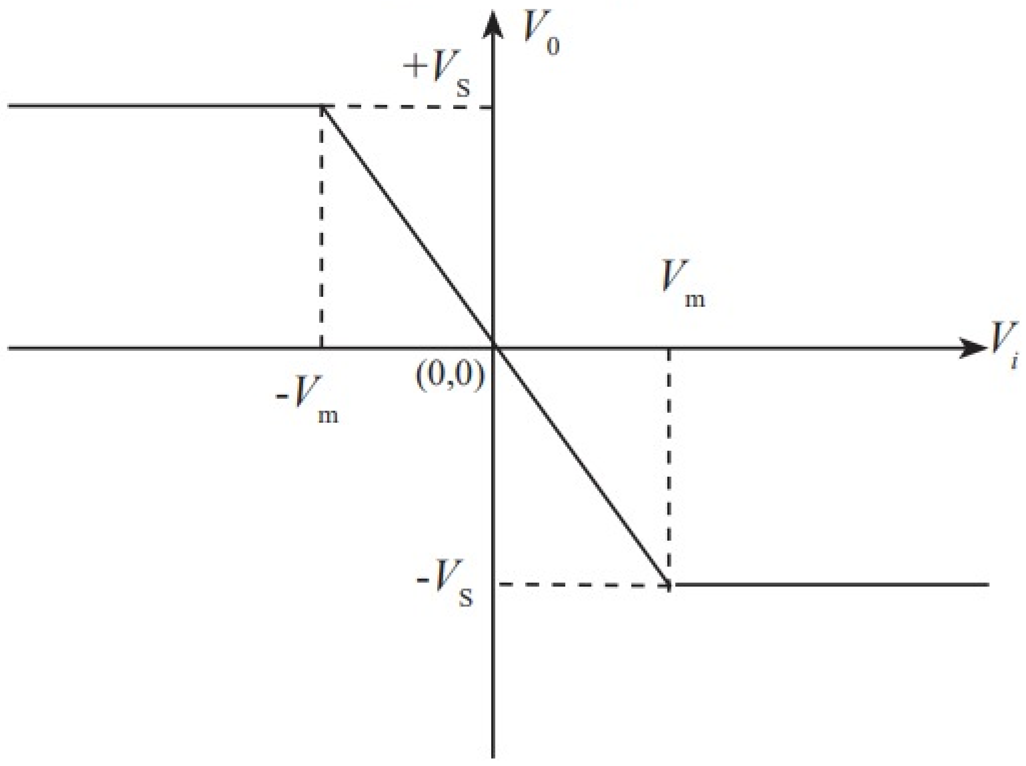


Figure 3.19

In the linear region

$$G_v = \frac{V_0}{V_i}$$

$$\therefore \quad V_0 = G_V . V_i$$

(Since this is an inverting amplifier, the gradient is negative)

The magnitude of the output voltage (V_0) does not exceed the magnitude of the supply voltage. ($\pm V_S$)

To operate the amplifier in the linear region, magnitude of the input voltage V_i should not exceed the value given by V_m then the output will come to a saturation. (At saturation the magnitude of V_0 is nearly equal to V_S. This may be about 0.8Vs practically).

3.5.11 Non-inverting amplifier

In this amplifier " non-inverting" means that the output voltage of the amplifier is in - phase (no phase difference) with the input voltage. A non-inverting amplifier circuit is shown in Figure 3.20.

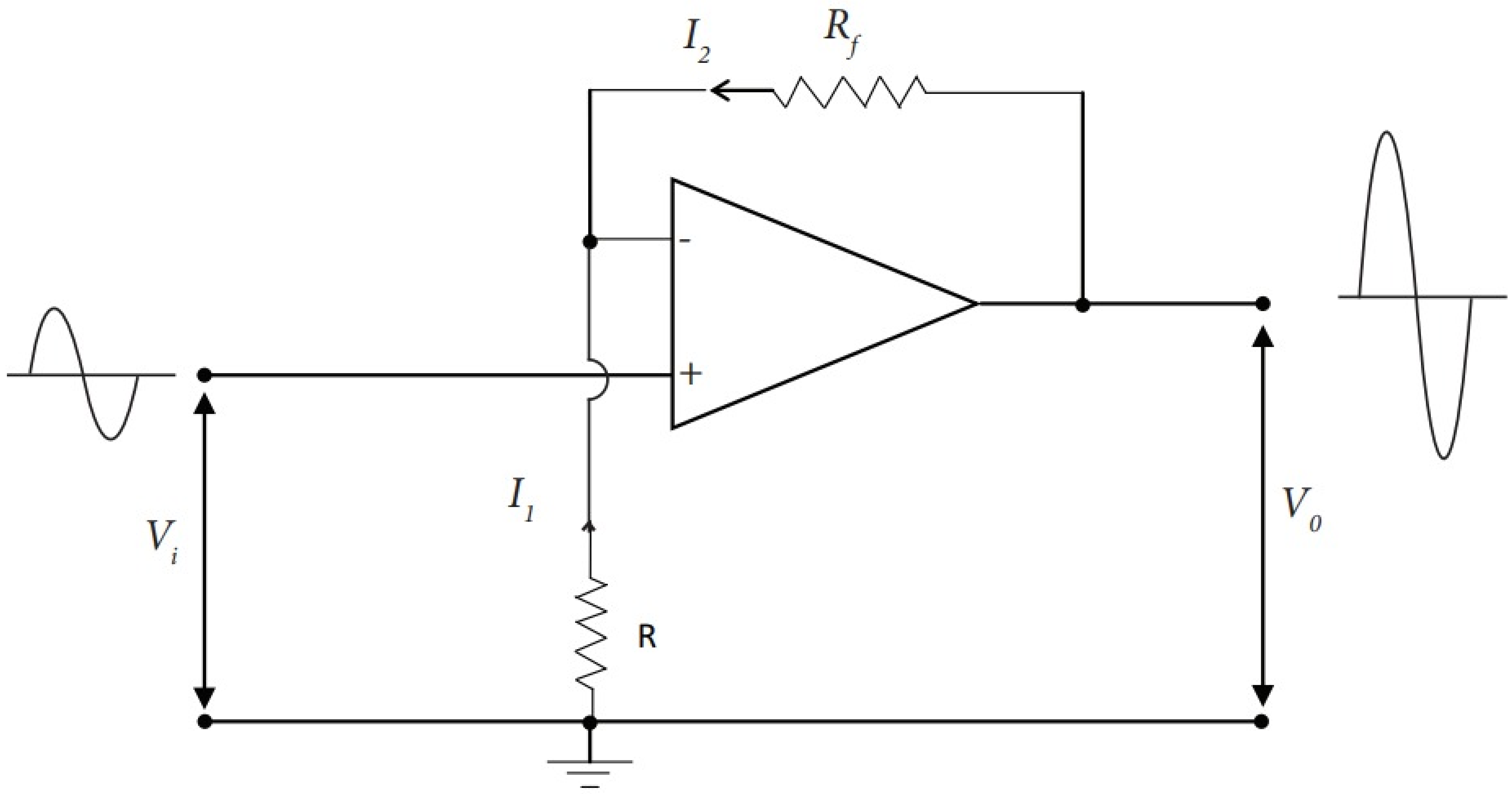


Figure 3.20

In this circuit,

$$V_+ = V_i$$

When operated in the linear region,
According to " Golden rule I" ,

$$V_+ - V_- = 0$$

$$\therefore \quad V_i - V_- = 0$$

$$\therefore \quad V_- = V_i$$

Potential difference across $R_f = V_{Rf} = V_O - V_- = V_O - V_i$

Potential difference across $R = V_R = V_- - 0 = V_i - 0 = V_i$

According to " Golden rule II", no current is flown into the input terminal.

$$I_1 = I_2$$

$$\therefore \quad \frac{V_R}{R} = \frac{V_{Rf}}{R_f}$$

$$\therefore \quad \frac{V_i}{R} = \frac{V_O - V_i}{R_f}$$

$$\frac{V_O}{R_f} = V_i\left(\frac{1}{R} + \frac{1}{R_f}\right)$$

$$\therefore \quad \frac{V_O}{V_i} = 1 + \frac{R_f}{R}$$

$\frac{V_O}{V_i}$ is the voltage gain of the close-loop non-inverting amplifier.

If it is denoted by G_v then,

$$\therefore \quad \boxed{G_v = 1 + \frac{R_f}{R}}$$

The right hand side of this equation is positive. That means the output voltage is in-phase with the input voltage. That is, the output is not inverted with respect to the input. (Observe the relevant waveforms shown with the circuit diagram).

With the appropriate selection of the resistance values of R and R_f, the voltage gain of the non-inverting amplifier can be set as required.

The characteristic of V_O against V_i is shown in Figure 3.21.

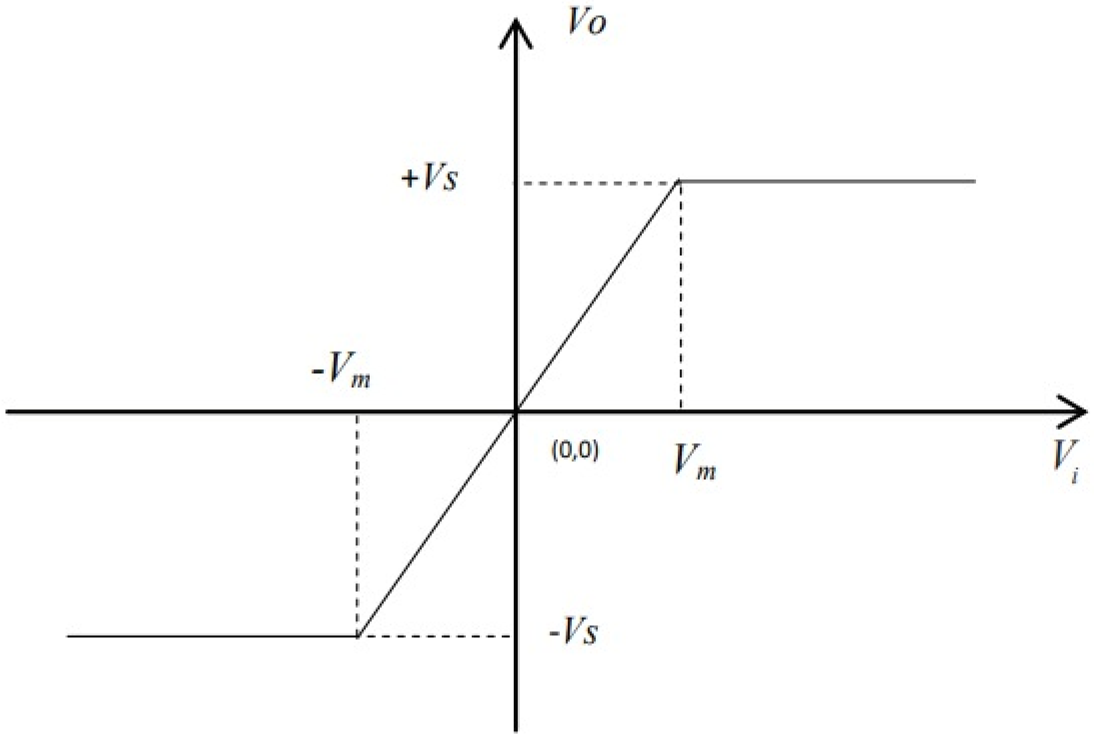


Figure 3.21

In the linear region,

$$G_v = \frac{V_O}{V_i}$$

$$V_O = G_v V_i$$

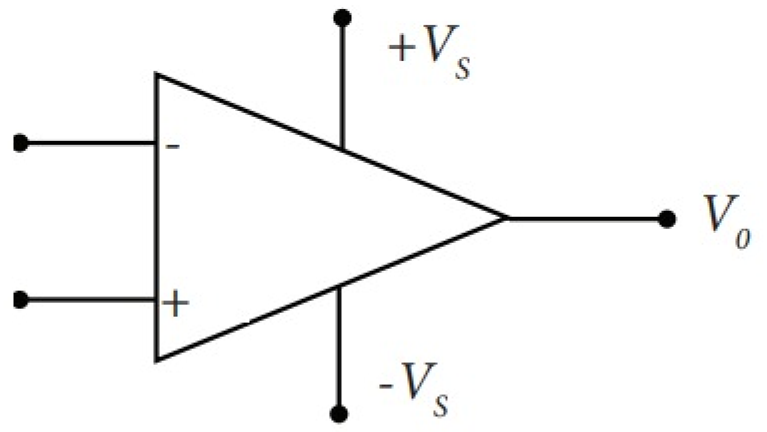
$\therefore$ The gradient of this graph $= G_v$ (since this is a non-inverting amplifier, the gradient is positive)

The magnitude of the output voltage (V_O) does not exceed the magnitude of the supply voltage. ($\pm V_S$)

To operate the amplifier in the linear region, magnitude of the input voltage (V_i) should not exceed the value given by V_m in the graph. If the magnitude of V_i exceed V_m then the output will come to a saturation.(At saturation the magnitude of V_o is nearly equal to V_s. This may be about 0.8 V_s practically).

Worked examples

(1).



The open-loop voltage gain of the op-amp shown in the above Figure is 10^5. It is powered with a $\pm$ 15 V dual voltage supply. Consider that the output voltage of the op-amp at saturation is $\pm$ 15 V.

1) What is the magnitude of the difference in input at the moment the output becomes saturated?
2) If + input is provided with a constant voltage of 2.0 V,

   a) What is the voltage applied to – input when the op-amp becomes saturated at positive voltage?
   b) What is the voltage applied to – input when the op-amp becomes saturated at negative voltage?
   c) What is the range of voltage that can be applied to input so that the op-amp operates in the linear region?

Answer

(i) $V_O = A_O (V_+ - V_-)$

$\therefore \quad 15 = 10^5 \ (V_+ - V_-)$

$$\therefore \quad (V_+ - V_-) = \frac{15}{10^5} \text{ V}$$

$$= 15 \times 10^{-5} \text{ V}$$

$$= \underline{\underline{150 \ \mu\text{V}}}$$

(ii) (a) $V_O = A_O (V_+ - V_-)$

$$\therefore \quad (V_+ - V_-) = \frac{V_O}{A_O}$$

As $V_O = +15$ V at + ve saturation,

$$(V_+ - V_-) = \frac{15}{10^5}$$

$$= 15 \times 10^{-5}$$

$$\therefore \quad (2.0 - V_-) = 15 \times 10^{-5}$$

$$\therefore \quad V_- = 2.0 - 15 \times 10^{-5} = 1.99985 \text{ V}$$

$\therefore$ V_- voltage at + ve saturation, = $\underline{\underline{1.99985 \text{ V}}}$

(b) $$V_O = A_O (V_+ - V_-)$$

$$\therefore \quad (V_+ - V_-) = \frac{V_O}{A_O}$$

As $V_O = -15$ V at - ve saturation,

$$(V_+ - V_-) = \frac{-15}{10^5}$$

$$\therefore \quad (2.0 - V_-) = -15 \times 10^{-5}$$

$$\therefore \quad V_- = 2.0 + 15 \times 10^5$$

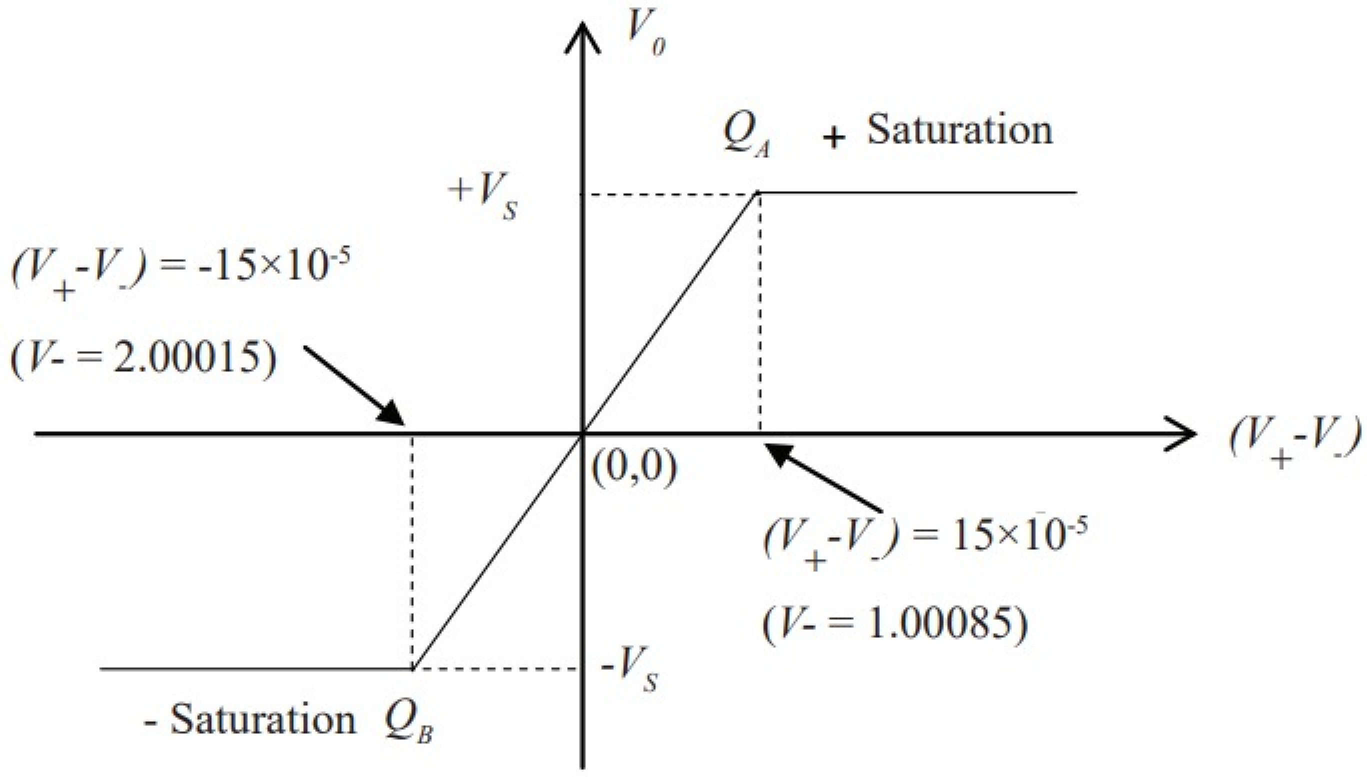
$$= 2.00015 \text{ V}$$

$\therefore$ V_- voltage at - ve saturation, = $\underline{\underline{2.00015 \text{ V}}}$

(c) Since the linear region is between the – saturation and + saturation regions,

The voltage range that can be applied to V- so that the op-amp operates in the linear region is from 1.85 V to 2.00015 V.

Observe the following graph with care. Then you will be able to realize the answer further well.

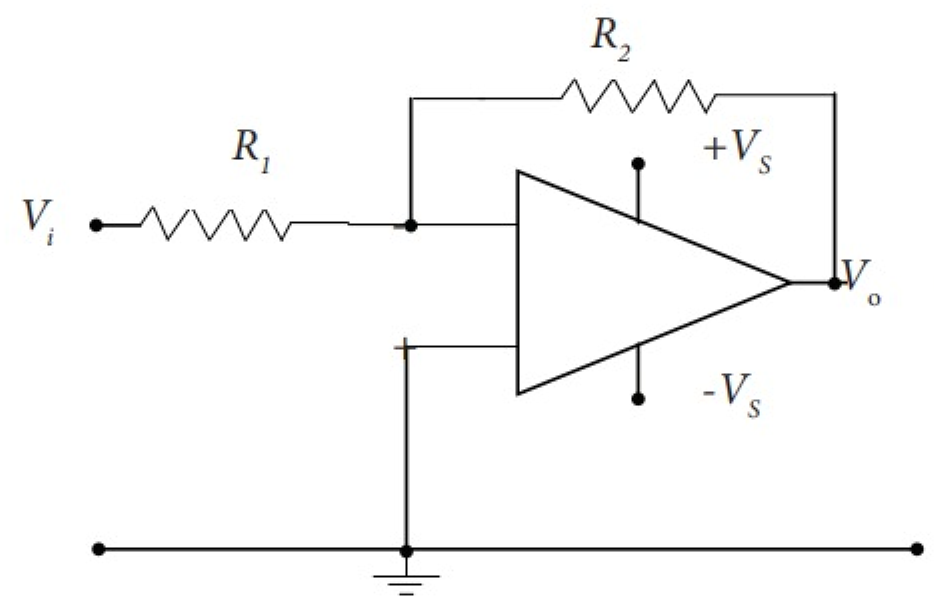


To be operated in the linear region,

The range of $(V_+ - V_-)$ is from $-\ 15 \times 10^{-5}$ V to $+15 \times 10^{-5}$ V

The range of V_- is from 1.99985 V to 2.00015 V.

(2).



The op-amp circuit shown here is provided with $+V_s = +15$ V and $-V_s = -15$ V. Consider that the magnitude of the output voltage at saturation is equal to the magnitude of the supply voltage.

(i) Select the suitable resistance value for R_1 and R_2 from the values given below, so that the voltage gain of this op-amp becomes 12.
10 kΩ, 12 kΩ, 15 kΩ, 68 kΩ, 100 kΩ, 120 kΩ

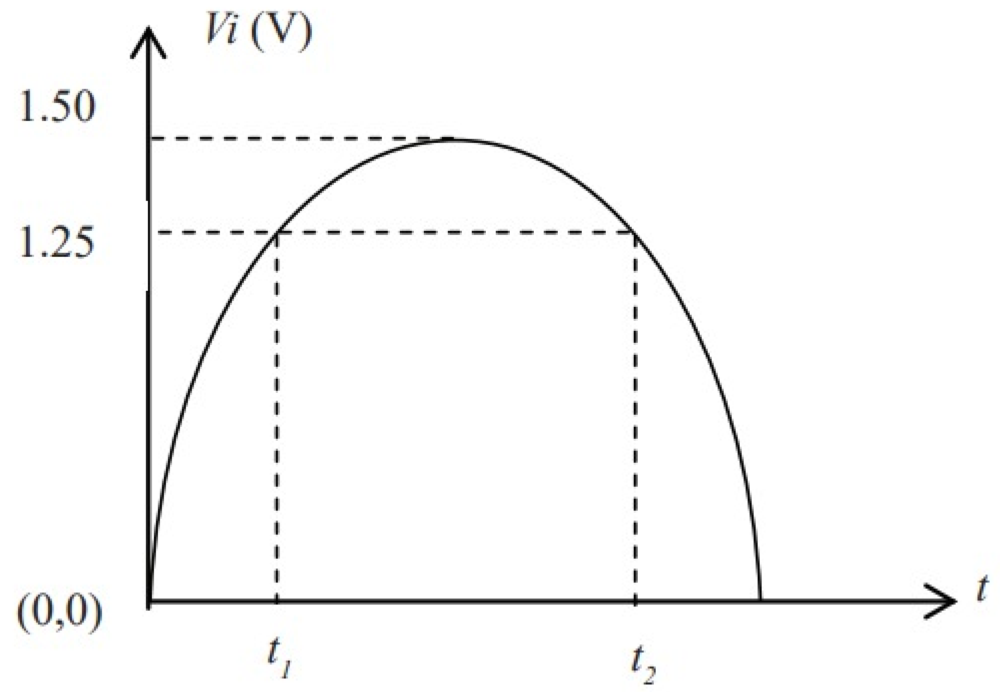
(ii) (a) What is the magnitude of the output voltage V_o, if the voltage gain of the amplifier is 12 and the voltage applied as V_i input is 0.75V ?

(b) Is this output voltage inverted or non-inverted with respect to the voltage?

iii) What is the output voltage when a voltage of 1.5 V is applied for V_i ?

iv) What is the value of V_i when the output of this amplifier just comes to saturation?

v) The variable voltage shown in the following graph is applied for V_i input. Draw a graph to show how the output voltage of the amplifier (V_0) varies with time (t). Mark the relevant voltage values at t_1 and t_2 on the graph.



Answer

(i) $R_1 = 10$ kΩ

$R_2 = 120$ kΩ (since $G_V = \frac{R_2}{R_1}, 12 = \frac{R_2}{R_1}$)

© 2020 National Institute of Education. All Rights Reserved.

(ii) (a) $G_V = -\frac{V_O}{V_i}$ (Since this is an inverting amplifier)

$$12 = -\frac{V_O}{0.75}$$

$\therefore \quad V_O = -0.75 \times 12$

$= -9.0 \text{ V}$

$\therefore$ Magnitude of the out put $= \underline{\underline{9.0 \text{ V}}}$

(b) Inverted

(iii) $G_V = -\frac{V_O}{V_i}$

$$12 = -\frac{V_O}{1.5}$$

$\therefore \quad V_O = -1.5 \times 12$

$= -18.0 \text{ V}$

Since supply voltage is – 15 V, V_0 cannot be – 18 V. Therefore, the output should be saturated at negative.

$\therefore$ Output voltage = $\underline{\underline{-15\text{V}}}$

(iv) $G_V = -\frac{V_O}{V_i}$

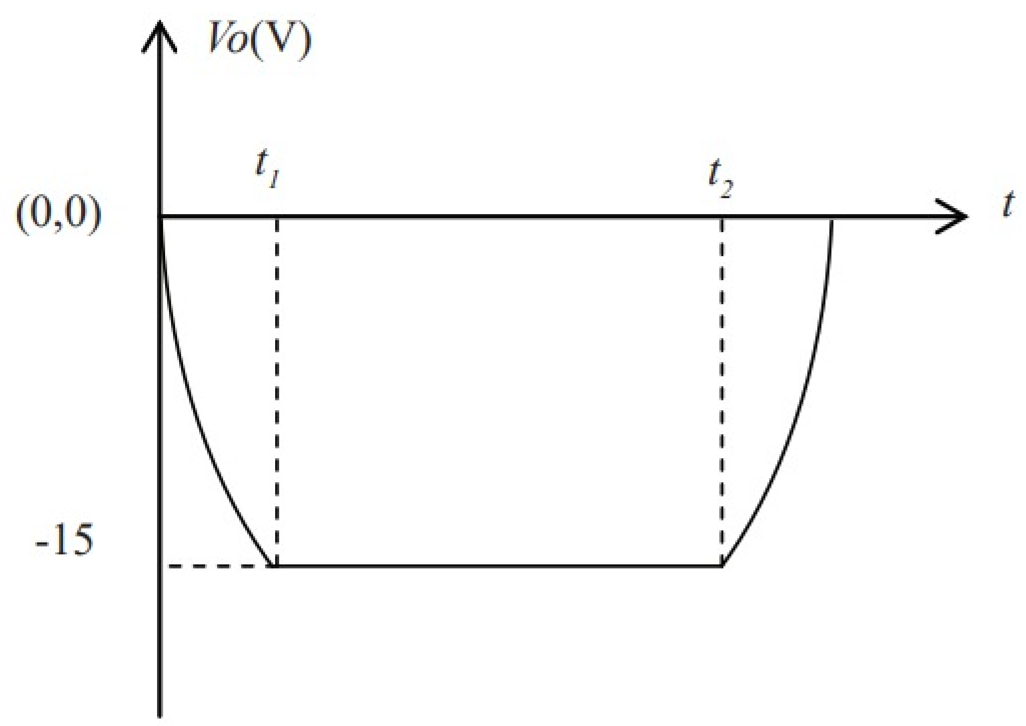
When saturated at negative $V_0 = -15$ V,

$$12 = -\frac{(-15)}{V_i}$$

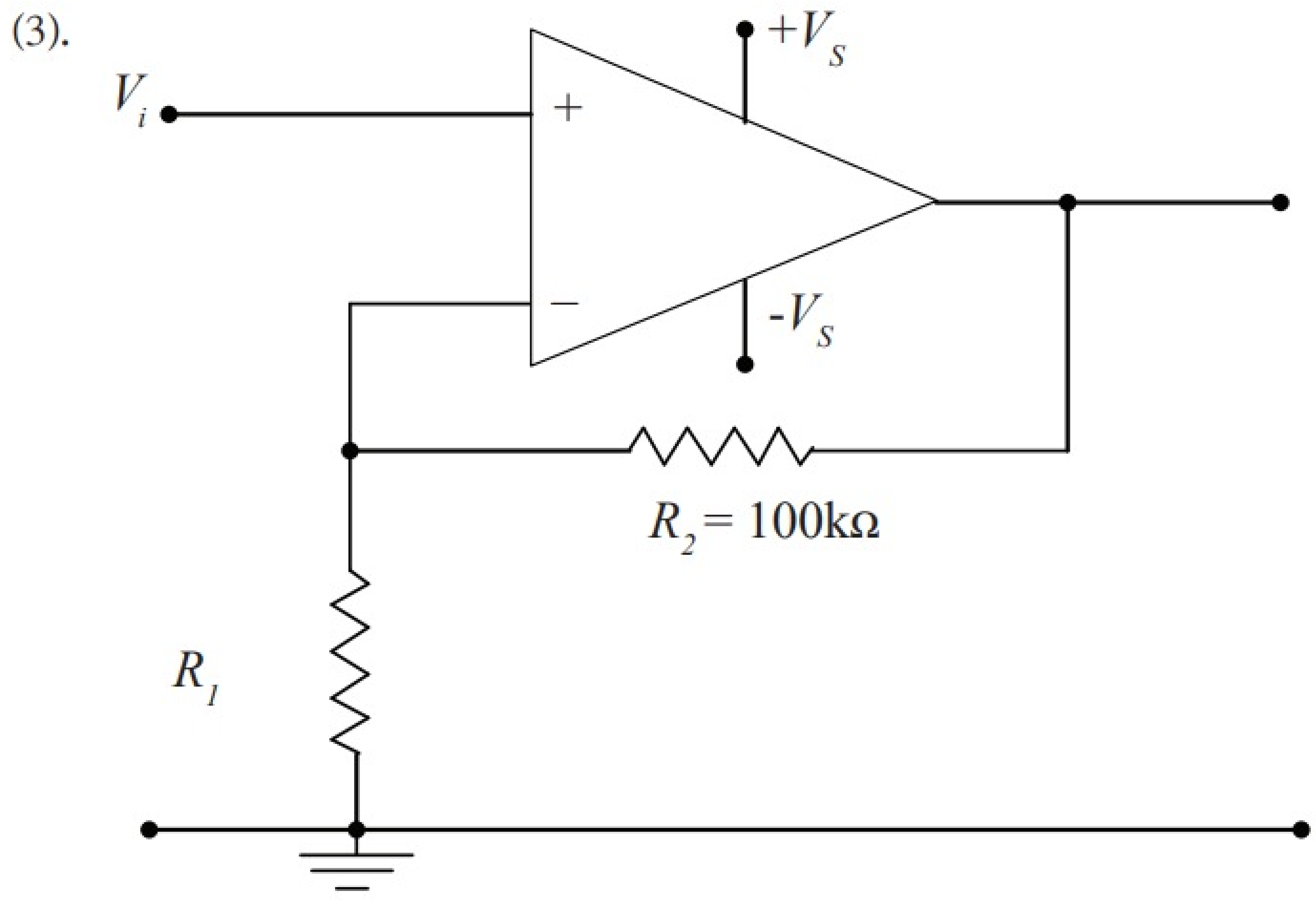
$\therefore \quad V_i = \frac{15}{12}$

$= \underline{\underline{1.25 \text{ V}}}$

(v)



(3).



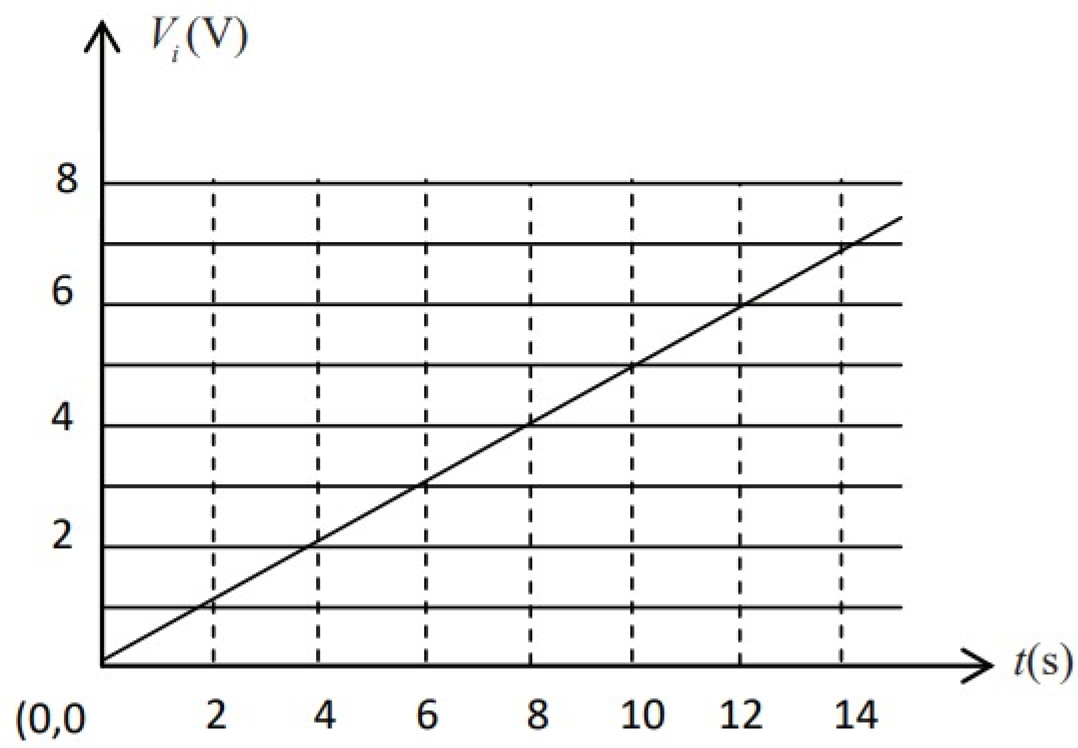
The figure shows a close-loop op-amp circuit. The voltage supply to the circuit is with $+V_S$ as +12V and $-V_S$ as -12 V. Consider that the output voltages at saturation are equal to the relevant values of the supply voltage.

(i) Calculate the value of the resistance R_1 so as to obtain an output voltage of 4.5V from a constant voltage of 1.5 V applied to the input V_i. (The value of R_2 is given as 100 kΩ). When answering the other parts, consider R_1 you obtained as the answer for this part.

(ii) Calculate the output voltage (V_0) when a constant voltage of -2V is applied as the input (V_i).

(iii) What is the minimum value of V_i required to make the output of this amplifier saturated at positive potentials.

(iv) A voltage varying with time as shown in the following graph, is applied as the input V_i. Draw a graph to show how the output of this amplifier (V_0) varies with time (t).



Answer

(i) $G_V = \frac{V_O}{V_i} = \frac{4.5}{1.5} = 3$

Since this is a non - inverting amplifier

$$G_V = 1 + \frac{R_2}{R_1}$$

$$\therefore \quad 3 = 1 + \frac{100\ \text{k}\Omega}{R_1}$$

$$\therefore \quad \frac{100\ \text{k}\Omega}{R_1} = 2$$

$$\therefore \quad R_1 = \frac{100\ \text{k}\Omega}{2}$$

$$\therefore \quad R_1 = 50\ \text{k}\Omega$$

(ii) $G_V = \frac{V_O}{V_i}$

$$\therefore \quad 3 = \frac{V_O}{-2}$$

$$\therefore \quad V_O = -6\ \text{V}$$

(iii) $G_V = \frac{V_O}{V_i}$

When saturated at + ve, $V_O = 12$ V

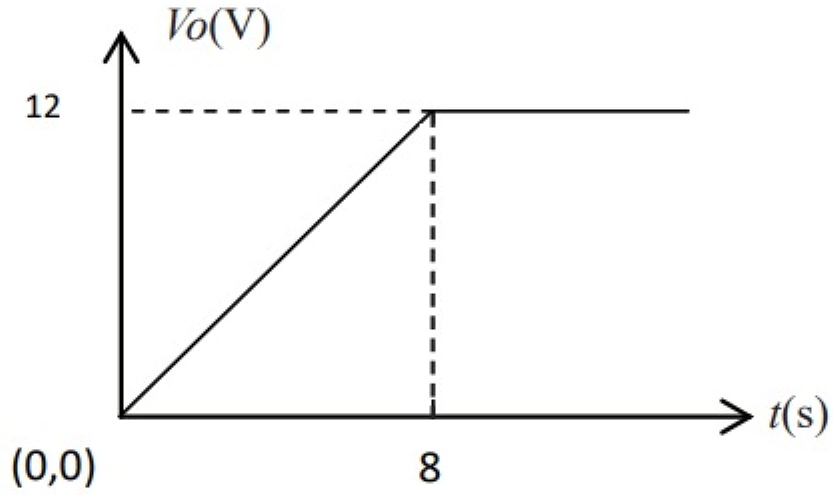
$$\therefore \quad 3 = \frac{12}{V_i}$$

$$\therefore \quad V_i = \frac{12}{3}$$

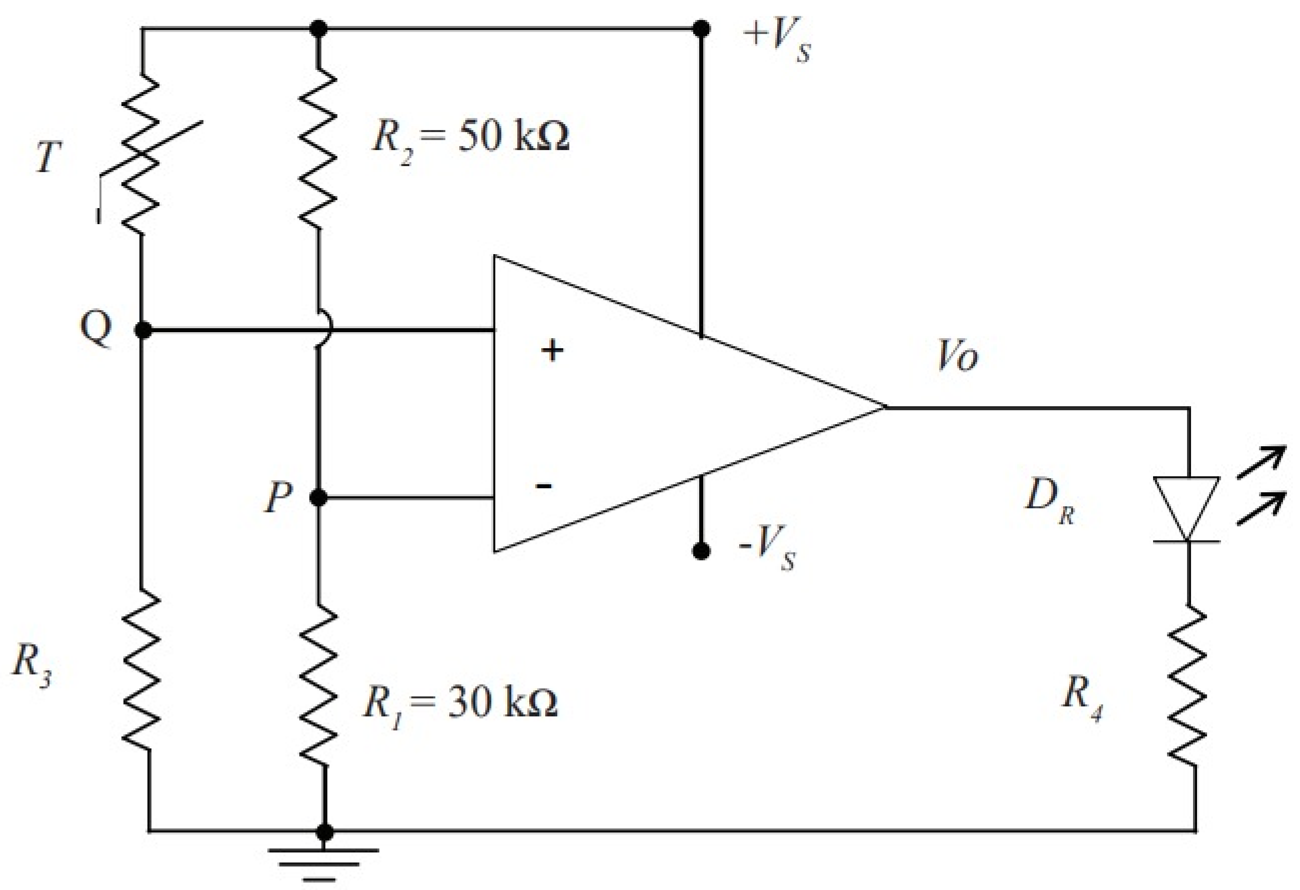
$$= 4\ \text{V}$$

$\therefore$ Required minimum V_i value $= 4$ V

(iv)



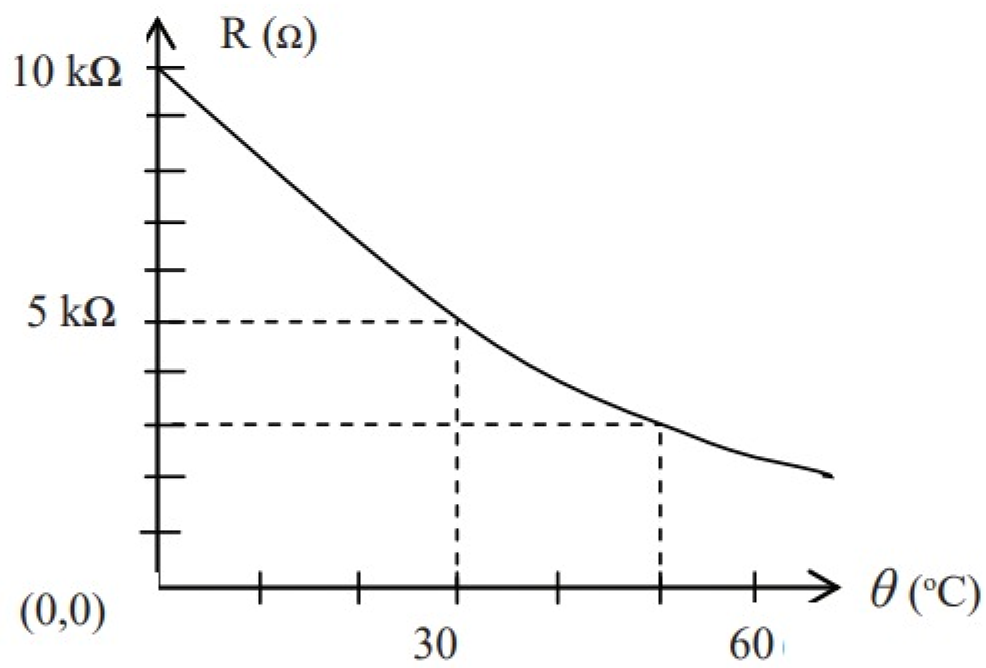
(4)



In the operational amplifier shown, the – input has been supplied with a definite voltage using a potential divider consisting of resistors R_1 and R_2. The power supply is such that $V_S = \pm 6$ V

(i) Calculate the voltage at P (-input).

The + input of the amplifier is supplied with a voltage by a potential divider consisting of a thermistor T and a resistor R_3 . The temperature (θ) - resistance (R) characteristic curve of the thermistor is given below.



(ii) What is the resistance of the thermistor at room temperature which is 30 °C ?

(iii) (a) What is the voltage at Q (+input) at 30 °C? Take the value of R_3 as 1 kΩ for this calculation?

(b) Will the red-LED denoted by D_R light or not? Give reasons.

(iv) Which voltage value, the voltage at Q should be greater than, for the output (V_0) to be saturated at positive potential?

(v) What is the resistance of the thermistor at 500 °C?

(vi) What should be the value of R_3 for the red-LED to light when temperature increases beyond 50 °C? (Consider the resistance value of the resistors other than that of thermistor is unchanged).

Answers

(i) $V_P = \dfrac{R_1}{R_1+R_2} \times V_s$

$= \dfrac{30}{30+50} \times 6$

$= \dfrac{30}{80} \times 6$

$= 2.25$ V

(ii) 5 kΩ

(iii) (a) $V_Q = \dfrac{R_3}{R_3+R_T} \times V_s$ (R_T is the resistance of the thermistor at 30 °C)

$= \dfrac{1}{(1+5)} \times 6$

$= 1$V

(b) D_R does not light.

In this case, as $V_p = 2.25$ V and $V_Q = 1$V. It makes the condition $V_- > V_+$. Therefore, output is saturated at the negative potential. Therefore, D_R does not light as it is reverse biased.

(iv) should be greater than 2.25V

(v) 3 kΩ

(vi) For D_R to light at 50 °C the output should be saturated at positive potential. Therefore, the voltage at V_+ input should reached 2.25 V.

$\therefore\ 2.25 = \dfrac{R_3}{R_3+R'_T} \times 6$ (R'_T is the resistance of the thermistor at 50 °C)

$2.25 = \dfrac{R_3}{R_3+3} \times 6$

$\therefore\ 6R_3 = 2.25R_3 + 6.75$

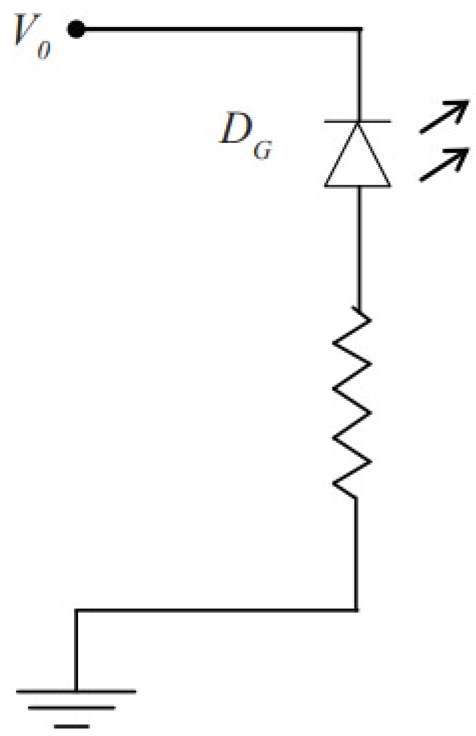
$\therefore\ 3.75R_3 = 6.75$

$\therefore\ R_3 = \dfrac{6.75}{3.75}$

$= 1.8$ kΩ

© 2020 National Institute of Education. All Rights Reserved.
© 2020 National Institute of Education. All Rights Reserved.

(vii)



Chapter four

Digital Electronics

4.1 Analogue signals and Digital signals

Digital electronics is of much importance as it has contributed a lot to the progress of the field of electronics. It is based on the operation of electronic circuits with digital signals. It is more useful and advantageous than the earlier means of operation of electronic circuits with analogue signals. Therefore, let us consider analogue signals and digital signals now.

When heating a bowl of water, its temperature rises from a lower to a higher. It does not get the higher temperature at once. The temperature rises continuously (without gaps) from lower temperature to higher temperature. The variation of temperature of that bowl of water is an example for an analogue signal.

When you use a torch at night you light the bulb on and off. It has no intermediate state other than the two state of lighting and not lighting. Such a signal which has only two well defined states, is an example for a digital signal.

The main difference between an analogue signal and a digital signal can be stated as follows.

When two states (values) of an analogue signal are considered, the signal can take any state (value) between the two states (values) in a continuous manner. Unlike this, a digital signal can have only two definite states (values). There are no in-between states (values) in it. Digital signal is a discrete signal having only two states (values). By referring to figure 4.1 you will get a better understanding.

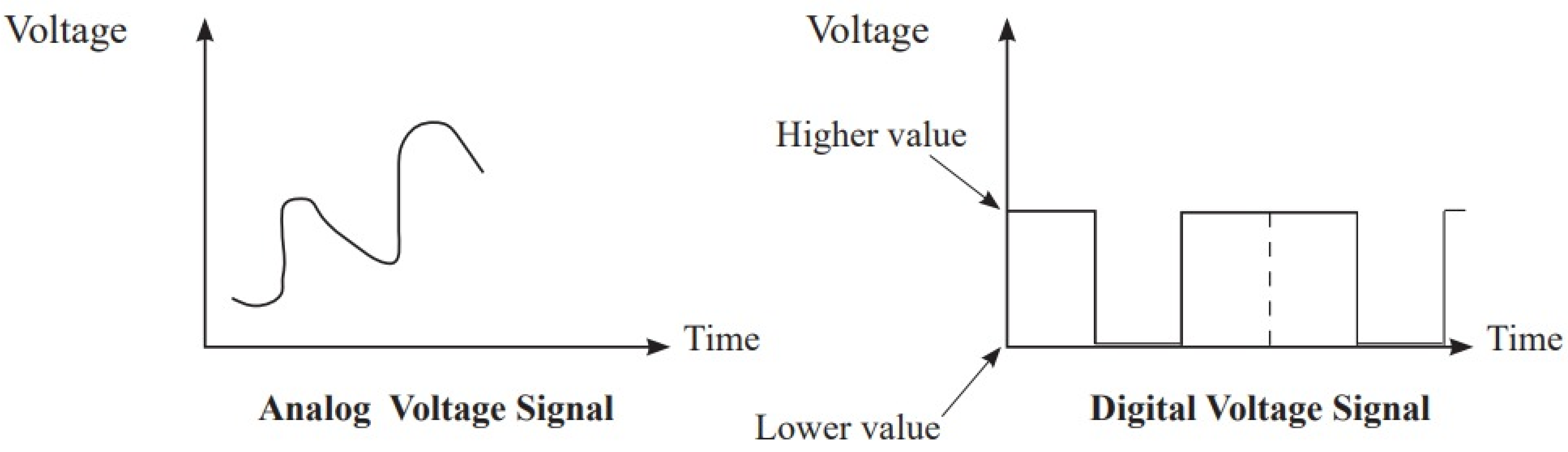


Figure 4.1

Since there is a large number of intermediate states, it is difficult to represent an analogue signal using numbers. However, a digital signal which has only two definite states, can be represented easily using numbers (this will be explained later). Therefore, let us discuss the number systems next.

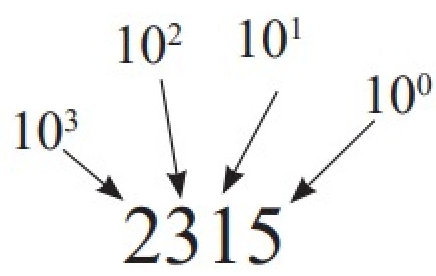
4.2 Decimal numbers and Binary numbers

The number system we use in our day-to-day life is the decimal number system. It consists of ten different digits. 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 are the digits in the decimal number system. The positional value of digits in a decimal number changes in multiples of 10. The positional value of digits in a decimal number changes in multiples of 10. The positional value of digits increases when going from right to left of the number.

Example:

Let us take the decimal number 2315

Mark the positional values of the digits



Positional value of the digit 5 which is at the fight end is 10^0 or 1
Positional value of the digit 1 which is the next on the left is 10^1 of 10
Positional value of the digit 3 which is the even next on the left is 10^2 or 100
Positional value of the digit 2 which is at the left end is 10^3 or 1000

Accordingly, the value of the number 2315 can be analyzed as,

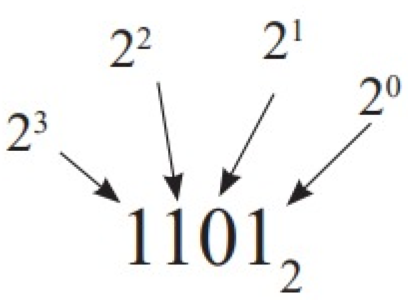
$(1000)\times2+(100)\times3+(10)\times1+(1)\times5$ within brackets are the positional values.
$= 2000 + 300 + 10 + 5$
$= 2315$

In the similar way, the binary number system too, can be understood. There are only two different digits in the binary number system. The two digits are 0 and 1. In the binary number system the positional value of the digits in a binary number changes in multiples of 2. The positional value increases from right to left.

Eg. Consider the binary number 1101

To indicate that a particular number is a binary number, 2 is written as a subscript at the right end of that number. That is, as 1101_2

Let us consider the positional values now



Positional value of digit 1 at the right end is 2^0 or 1.

Positional value of the digit 0 which is the next on the left is 2^1 of 2.
Positional value of the digit 1 which is the even next on the left is 2^2 or 4.
Positional value of the digit 1 at the left end is 2^3 or 8.

Therefore, the decimal value of the binary number 11012 is,

$(8) \times 1 + (4) \times 1 + (2) \times 0 + (1) \times 1$ within brackets are the positional values.
$= 8 + 4 + 0 + 5$
$= 13$

If it is required specifically, to indicate that a particular number is decimal, then 10 is written as a subscript at the right end of that number.

Eg. 13_{10}

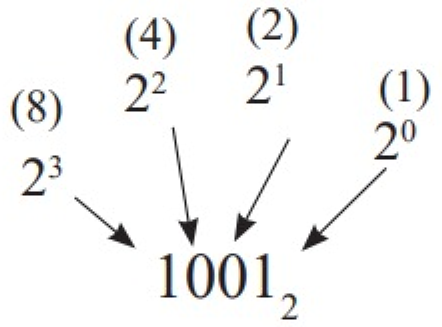
According to the above explanation,

$1101_2 = 13_{10}$

So, now you can try to express a binary number in the decimal from (i.e. to find the decimal equivalent of a binary number).

Worked examples

(1) Express the number 1001_2 in decimal form.

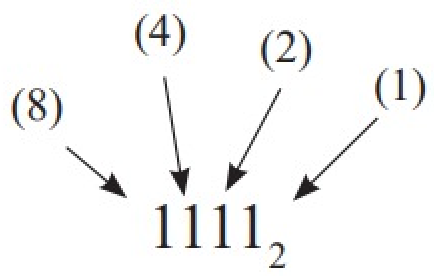


Let us obtain the decimal form, considering the positional values as follows.

$$(8) \times 1 + (4) \times 0 + (2) \times 0 + (1) \times 1$$
$$= 8 + 0 + 0 + 1$$
$$= 9$$

$\therefore\ 1001_2 = 9_{10}$

(2) Write the binary number 1111_2 as a decimal number.



$$1111_2 = (8) \times 1 + (4) \times 1 + (2) \times 1 + (1) \times 1$$
$$= 8 + 4 + 2 + 1$$
$$= 15_{10}$$

Conversion of decimal number into binary numbers

Let us study how to write a decimal number in the binary form, referring to following examples.

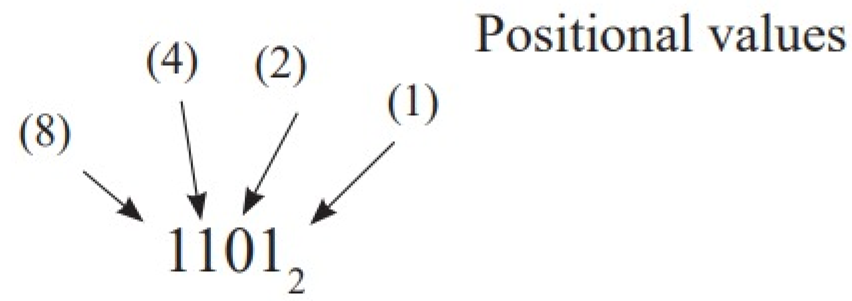
(i) Consider the number 13_{10}

Divide the number by 2 repeatedly until you get the quotient as zero. In each step of division write the remainder on the right and side as shown.

2 | 13
2 | 6 — 1 (the remainder represents the positional value 1)
2 | 3 — 0 (the remainder represents the positional value 2)
2 | 1 — 1 (the remainder represents the positional value 4)
0 — 1 (the remainder represents the positional value 8)

Accordingly, the number 13_{10} has,
1 in the positional value of 8
1 in the positional value of 4
0 in the positional value of 2
1 in the positional value of 1

Let us now write the number in binary form.



$\therefore\ 13_{10} = 1101_2$

(ii) Write the number 11_{10} in the binary form

2 | 11
2 | 5 — 1
2 | 2 — 1
2 | 1 — 0
0 — 1

↑ Right end
Left end

The binary number can be obtained by writing the digits of remainders from left to right in the direction show by the arrow. That is 1011.

$\therefore\ 11_{10} = 1011_2$

4.3 Voltage levels of digital signals

In digital electronics the two states (values) of a digital signal are expressed as 0 and 1. Two voltage levels are used to mark 0 and 1. The two standard voltage levels used are 0 V and 5 V. Generally 0 V is used to mark 0 and 5 V is used to mark 1.

According to the above, the waveform (series of voltage pulses) relevant to some digital voltage signal can be represented by a binary number as follows.

In a digital voltage signal each voltage level is taken in equal lengths of time.

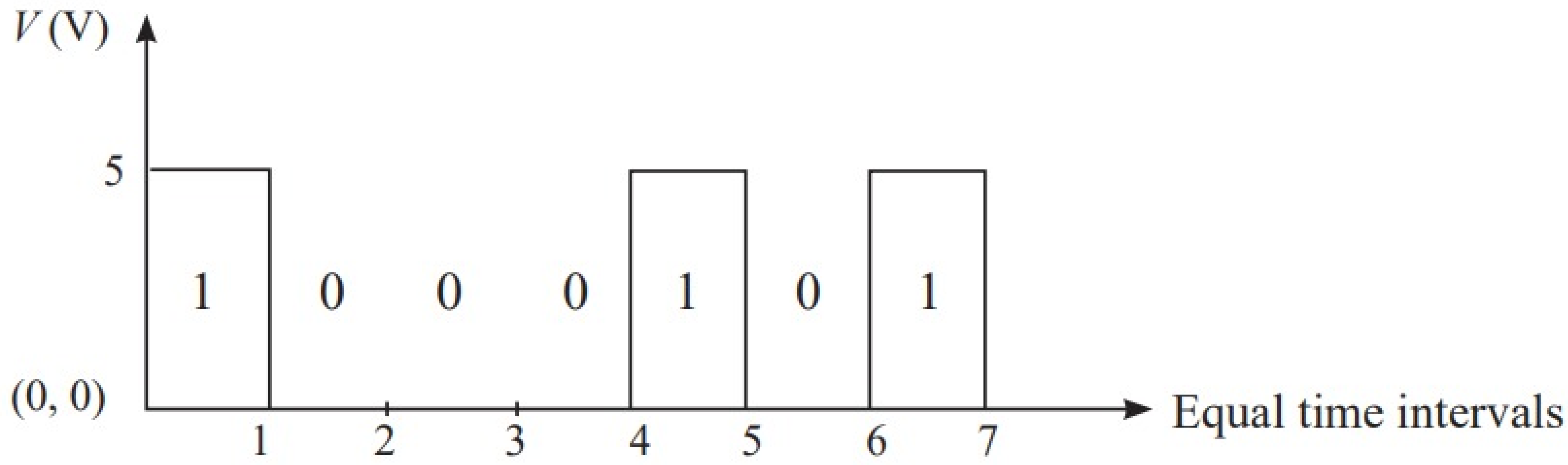


Figure 4.2

Since 5 V level is taken as 1 and 0 V level as 0, the binary number represented by the digital signal given in Figure 4.2 is,

$$1000101_2$$

In the above signal, 0 and 1 are marked at relevant places for you to make it clear.

Eg.

The voltage level representation of a digital signal is given in Figure 4.3. Represent it by a binary number.

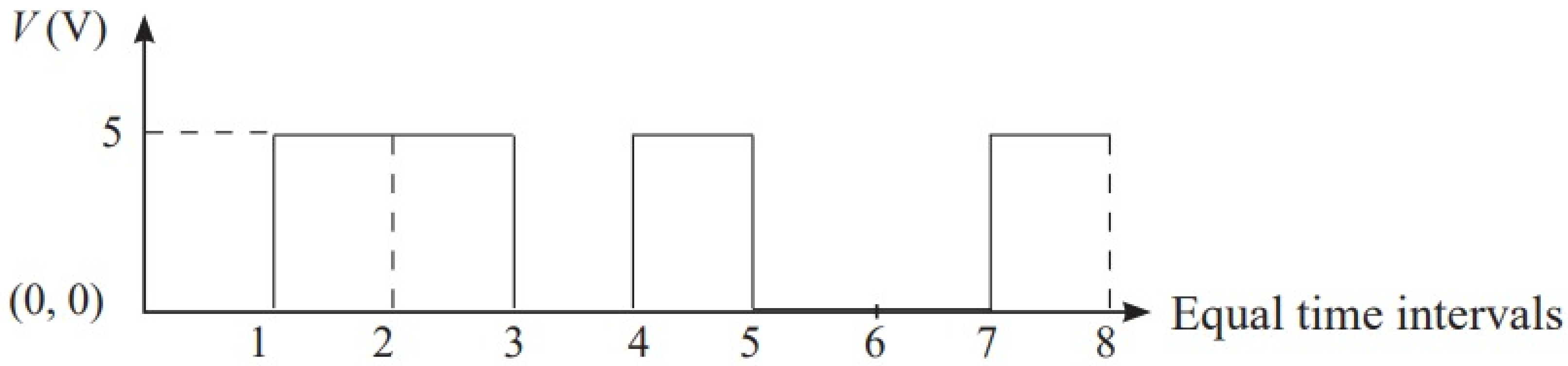


Figure 4.3

Writing the relevant digits,

0 for 0 V and
1 for 5 V

From left to right of the signal, the required binary number can be obtained.

The relevant binary number $= 01101001_2$

4.4 Digital electronic circuits

The electronics circuits which operate relating to digital signals and also control those signals are called digital electronic circuits.

The basic elements used in making digital electronic circuits are the logic gates. There are several types of logic gates. Each logic gate gives out an output after executing the relevant logic operation on the input/inputs applied to it. Let us consider seven such operations.

Those logic gates are named as follows.

NOT, AND, OR, NAND, NOR, XOR, XNOR

The logic gates were first made using mechanical switches.

Next they were made using relays which can be switched by electrical signals. Since the response time of mechanical switches was high, the logic gates made using them were slow in operation. Later, logic gates were made using transistor switching circuits and they could be operated with a high speed. Today the logic gates are produced in the form of IC, using modern technology. This will be discussed later.

Let us now discuss each of these gates in detail.

4.5 NOT gate

The circuit symbol of the NOT gate is given in Figure 4.4.

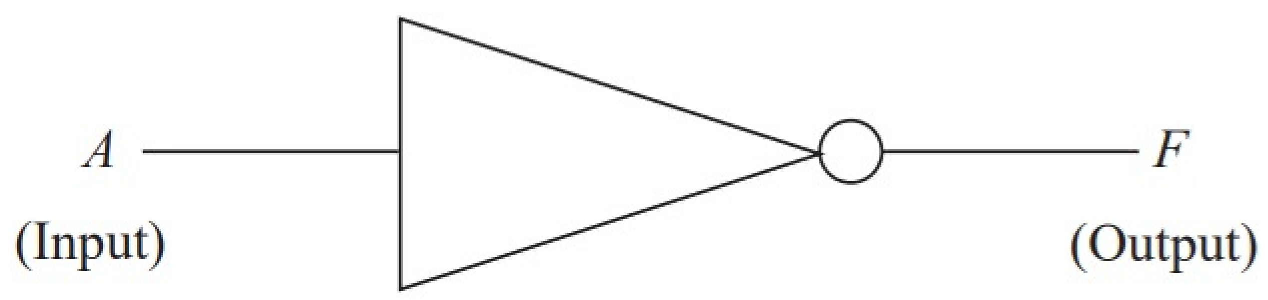


Figure 4.4

To understand the logic of operation of a logic gate a specific table is given. That table which includes the states of input and output, is known as the **truth table.**

The truth table for NOT gates is much simple. It is given below. It is simple because the NOT gate has only one input.

A	*F*
0	1
1	0

The small bubble (or circle) marked at the output of the circuit symbol of NOT gate is to indicate that the output is inverted with respect to input.

According to the above, it is clear that the logic expression for the NOT gate can be given as **"taking the inversion of input".**

Note that,

Inversion of 0 is 1

and Inversion of 1 is 0

The logic operation done by a logic gate can be represented as an expression according to a particular symbolic manner. These expressions are written according to a mathematical procedure introduced by the mathematician **George Boole**. That Procedure of mathematics is called Boolean algebra, and the expressions written using it are known as Boolean expressions.

The Boolean expression for the NOT gate is written as,

$\boxed{F = \bar{A}}$ Where $\bar{A}$ is the inversion of A

If $A = 0$ then $\bar{A} = 1$

If $A = 1$ then $\bar{A} = 0$

The action of a NOT gate can be demonstrated using the simple switching circuit with a mechanical switch shown in figure 4.5.

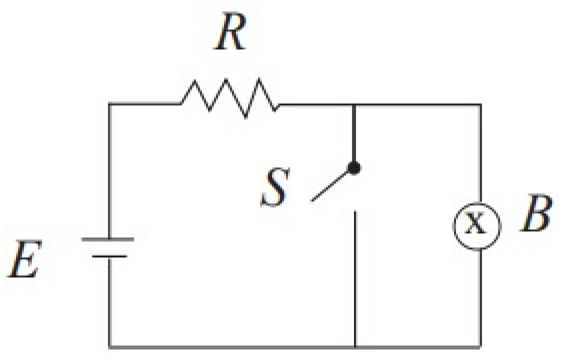


Figure 4.5

The emf of the cell E and the resistance value of R are chosen so that, when the switch S is open (OFF), the bulb B lights well.

When S is closed (ON) the circuit gets short circuited across the bulb B and cuts the current through it. Then the bulb does not light.

Consider, the input S as,

0 when the switch S is open (OFF)

1 when the switch S is open (ON)

The output F as,

0 when the bulb B does not light.

1 when the bulb B lights.

Then the operation of the circuit can be given in a table as follows.

S	B
0	1
1	0

This is the NOT operation.

There are several methods of making logic gates. They are made using diodes, transistors and resistors.

The NOT gate has only one input and one output. It should be noted that non of the logic has more than one output.

4.5.1 Making a NOT gate

The circuit diagram of a NOT gate made using one transistor and two resistors is shown in figure 4.6.

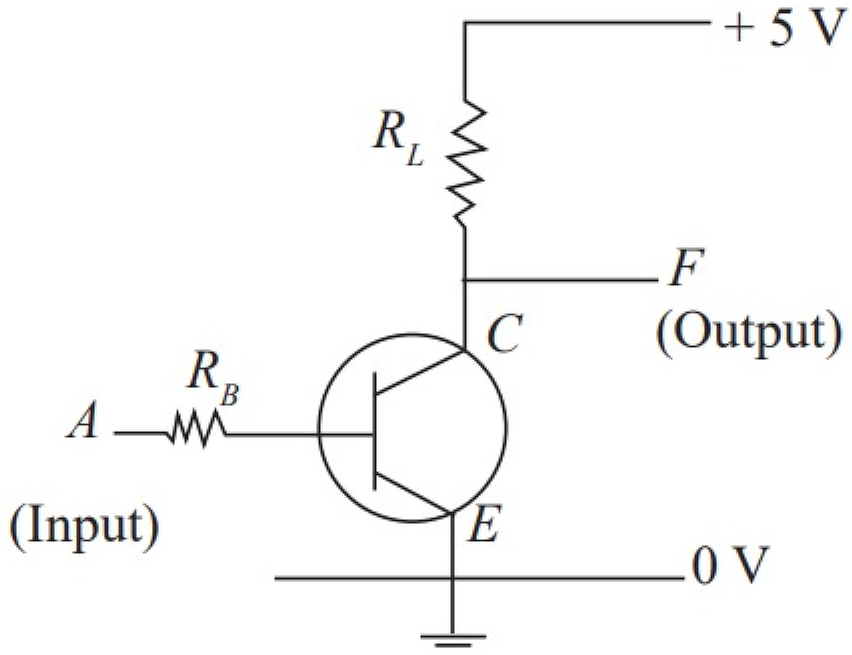


Figure 4.6

When zero voltage (logic 0) is applied to input *A*, the transistor becomes cut-off. Then there is an open circuit condition between *C* and *E* of the transistor and hence the voltage at *F* becomes nearly +5V (logic 1). That is, if… $A = 0$ then $F = 1$.

When the input *A* is supplied with +5V (logic 1) then the transistor becomes saturated and there is nearly a short circuit condition between *C* and *E* of it. Then the voltage at *F* becomes nearly zero (logic O). That is, if $A = 1$ then $F = 0$.

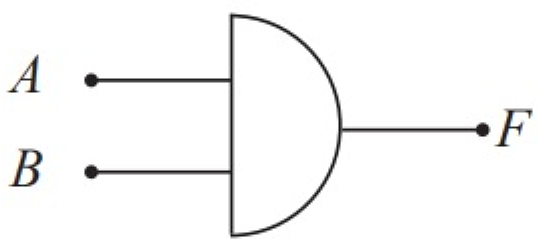
Therefore, this circuit behaves as a NOT gate.

(It should be considered that the resistance value of R_B has been chosen so that the transistor comes to saturation with the input of +5 V)

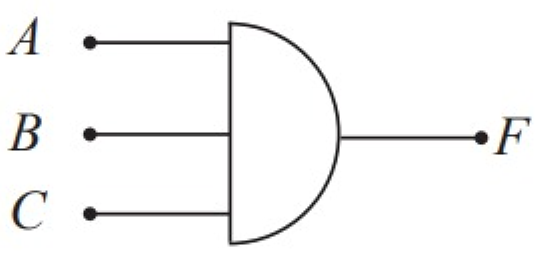
This type of switching circuit made using transistors and resistors are known as RTL (Resistor Transistor Logic).

4.6 AND gate

The AND gate can be with at least two or more inputs. Circuit symbols for the AND gate are given in figure 4.7.



With two input, A and B



With three input, A, B and C

Figure 4.7

The truth table for the AND gate with two inputs is given below.

A	*B*	*F*
0	0	0
0	1	0
1	0	0
1	1	1

The truth table for the AND gate with three inputs is given below.

A	*B*	*C*	*F*
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

According to the above truth tables it is clear that the AND operation is, to give the output as logic 1 only when both inputs are at logic 1 state. Therefore, a logic expression can be given as **"When both *A* and *B* inputs are 1, the output is 1"** for the two input AND gate similarly for the three-input AND gate, the logic expression is "when all three inputs *A*, *B* and *C* are 1, the output is 1".

Boolean expression for the two-input AND gate is written as,

$$\boxed{F = A.B}$$

Accordingly, it can be realized that,

$0.0 = 0$, $0.1 = 0$, $1.1 = 0$ and $1.1 = 1$

Boolean expression for the three-input AND gate is written as,

$0.0 = 0$, $0.1 = 0$, $1.0 = 0$ and$1.1 = 1$

$$\boxed{F = A.B.C}$$

A three input AND gate can be made using two of two-input AND gates as shown in figure 4.8

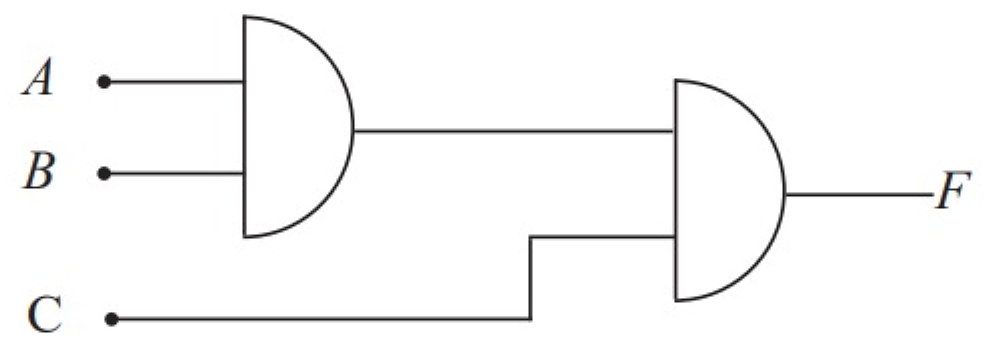


Figure 4.8

It is assigned to you to prepare a truth table for the operation of this logic gate arrangement taking *A, B, C* as inputs and *F* as the output.

The action of two-input AND gate can be demonstrated with the simple switching circuit shown in figure 4.9.

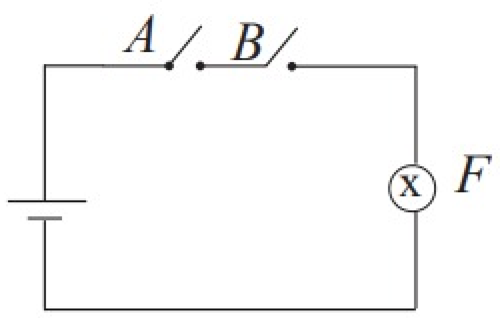


Figure 4.9

Let us take that,

Inputs : Opening (OFF) of a switch *A* or *B* is logic 0
Closing (ON) of a switch *A* or *B* is logic 1

Outputs : not lighting the bulb *F* is logic 0
Lighting the bulb *F* is logic 1

According to inputs *A* and *B*, the output *F* (lighting or no lighting the bulb) is determined as in the following table.

A	*B*	*F*
0	0	0
0	1	0
1	0	0
1	1	1

Only when both *A* and B switches are closed (NO), the bulb lights. That is only when both *A* and *B* inputs are 1, the output *F* is 1.

This is the AND operation.

4.6.1 Making an AND gate

The circuit given in Figure 4.10 shows how a two-input AND gate can be made using two diodes and a resistor.

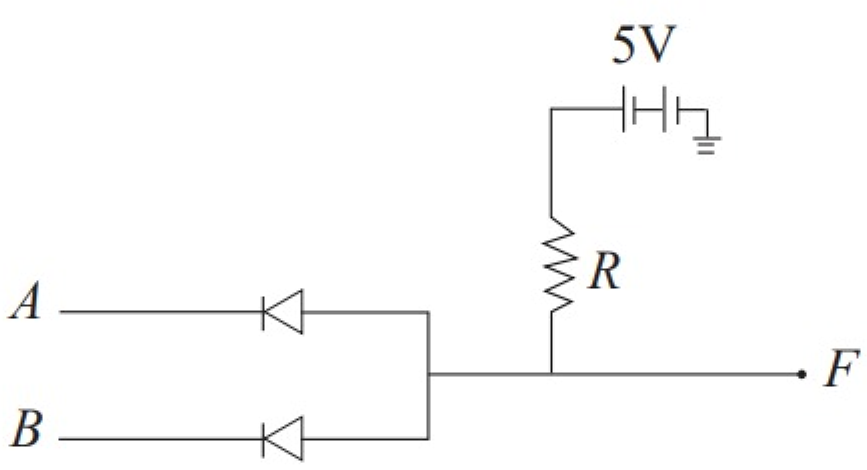


Figure 4.10

If only *A* or only *B* both A and *B* are at the zero voltage then the diode/diodes will become forward biased due to voltage supplied by the 5 V battery via *R*. Then there will be a forward voltage drop of 0.7 V across a silicon diode. Therefore the voltage at *F* will be 0.7. This can nearly be taken as logic 0.

If both *A* and *B* are supplied with 5V then both diodes become reverse biased. Then a current does not flow through *R* and there is no voltage drop across *R*. Therefore the battery voltage which is 5V appears at *F*. That is, *F* is at logic 1. The relevant voltage table and the truth table are given below.

$A_{(V)}$	$B_{(V)}$	$F_{(V)}$
0	0	0.7
0	5	0.7
5	0	0.7
5	5	5

Voltage table

A	B	F
0	0	0
0	1	0
1	0	0
1	1	1

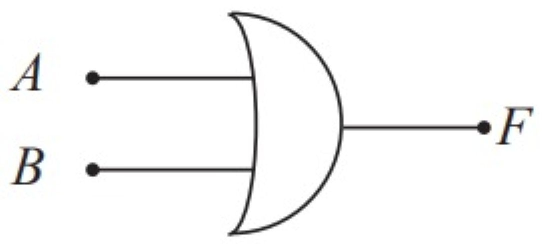
Truth table

Figure 4.10

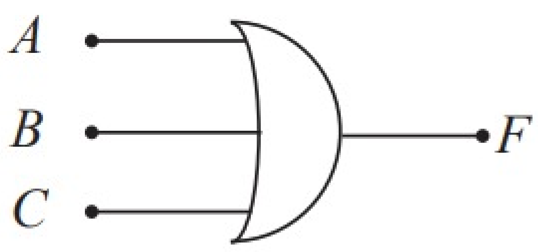
This type of circuit made using diodes and resistors are known as DRL (Diode Resistor Logic).

4.7 OR gate

OR gate can be with at least two or more inputs. The circuit symbols for the OR gate are shown in figure 4.11.



with two inputs *A and B*



with three inputs *A*, *B*, *C*

Figure 4.11

The truth table for two-input OR gate is given below.

A	B	F
0	0	0
0	1	1
1	0	1
1	1	1

It is clear from the truth table that in the OR operation the, given output is as logic 1 when at least one of the inputs A and B is at logic 1.

Boolean expression for the OR gate is,

$$\boxed{F = A + B}$$

The + sign in this expression has a different meaning than the meaning of + sign in mathematics. The + sign used here is in accordance with the Boolean algebra. It should be realized referring to the truth table that,

$$0 + 0 = 0$$
$$0 + 1 = 1$$
$$1 + 0 = 1$$
and $$1 + 1 = 1$$

© 2020 National Institute of Education. All Rights Reserved.

The truth for the two-input OR gate can be given as "When the input *A* or *B* is 1, the output is 1.

A three-input OR gate can be made using two-input OR gates as shown in figure 4.12.

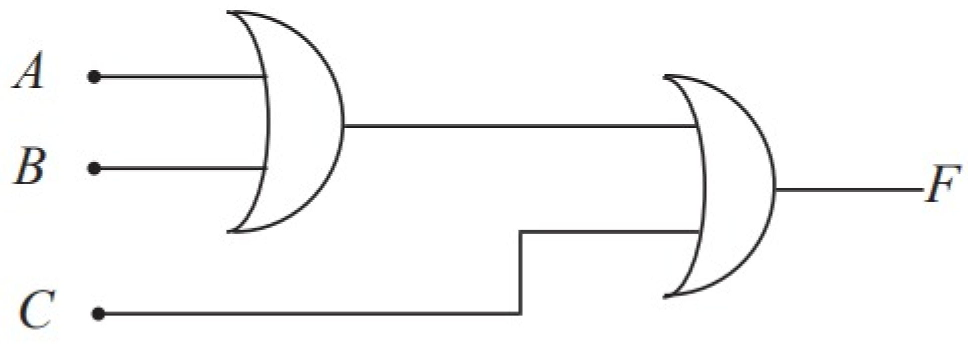


Figure 4.12

Try whether you can prepare a truth table for this logic gate arrangement.

The action of an OR gate can be demonstrated using a simple switching circuit with mechanical switches as shown in Figure 4.13.

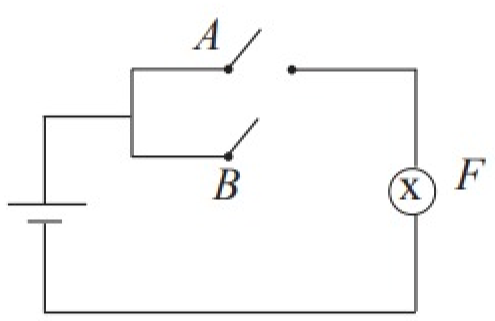


Figure 4.13

Let us take that,
Input : opening (OFF) of a switch *A* or *B* is logic 0
closing (ON) of a switch A of *B* is logic 1
Output : not lighting the bulb *F* is logic 0
lightning the bulb *F* is logic 1

The operation of this circuit can be represented by a table as follows. It has been taken A and B as the inputs and F as the output.

A	B	F
0	0	0
0	1	1
1	0	1
1	1	1

When the switch *A* or the switch *B* is closed (NO), the bulb lights.
That is, when the input *A* or input *B* is 1 the output *F* is 1.
It is clear that this is the OR operation.

4.7.1 Making an OR gate

The circuit given is figure 4.14 shows how a two-input OR gate can be made using two diodes and a resistor. (in the from of DRL).

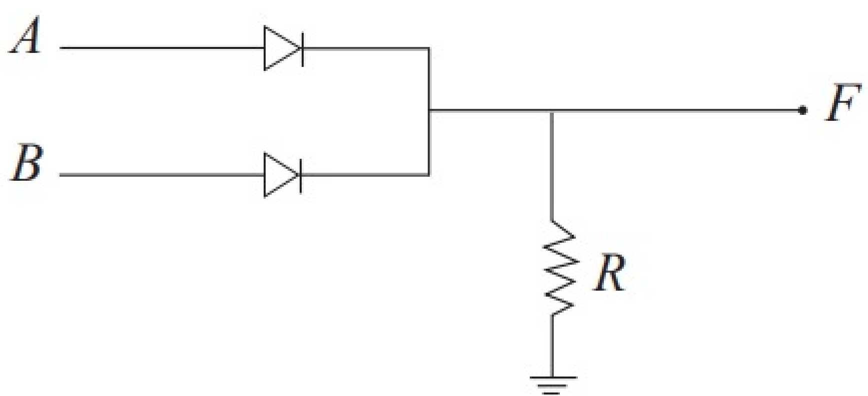


Figure 4.14

When, at least one of *A* and *B* inputs is supplied with 5V (logic 1), that diode becomes forward biased and therefore irrespective of the other input, the output is equal to (5-0.7) V = 4.3V. This can be taken as logic 1. (The forward bias voltage drop has been taken as 0.7 V considering a silicon diode).

When both *A* and *B* inputs are supplied with zero voltage (logic 0), none of the diodes is forward biased. Therefore, the output *F* remains at zero voltage (logic 0)

The relevant voltage table and the truth are given below.

$A_{(V)}$	$B_{(V)}$	$F_{(V)}$
0	0	0
0	5	4.3
5	0	4.3
5	5	4.3

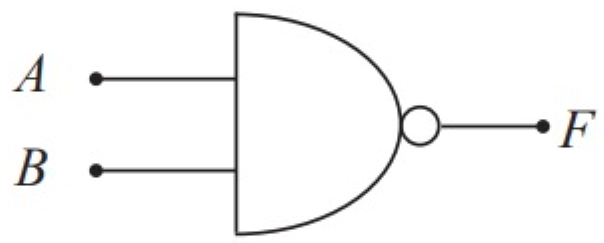
Voltage table

A	*B*	*F*
0	0	0
0	1	1
1	0	1
1	1	1

Truth table

4.8 NAND gate

The circuit symbol and the truth table of a two-input NAND gate are given in Figure 4.15.



Voltage table

A	*B*	*F*
0	0	1
0	1	1
1	0	1
1	1	0

Truth table

Figure 4.15

The logic operation of the NAND gate is the inverse of that of the AND gate. It will be clear to you if you compare the relevant truth tables from the two gates.

A NAND gate can be made by passing the output of an AND gate through a NOT gate as shown in Figure 4.16.

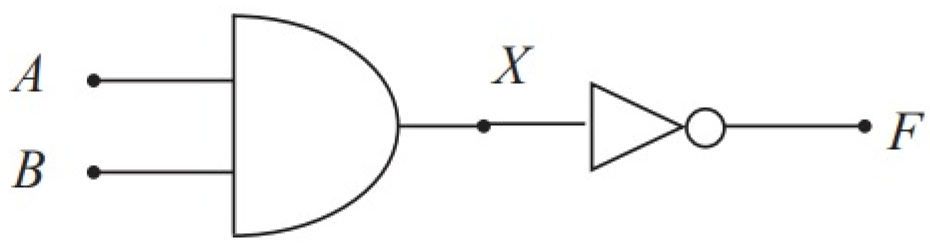


Figure 4.16

A and B are the inputs.
X is the output of the AND gate
F is the final output

© 2020 National Institute of Education. All Rights Reserved.

Let us make it clear with the help of following truth table.

A	B	X	$F\,(=\overline{X})$
0	0	0	1
0	1	0	1
1	0	0	1
1	1	1	0

The output F is same as the output of a NAND gate.

According to Boolean expression,

$X = A.B$ (expression of the AND gate)

and $F = \overline{X}$ (expression of the NOT gate)

$\therefore \quad F = \overline{A.B}$

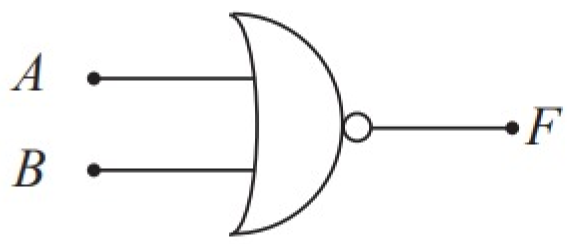
Therefore, the Boolean expression for the NAND gate is written as,

$$F = \overline{A.B}$$

The logic expression for the NAND gate can be given as **"when A and B inputs are 1, the output is 0".**

4.9 NOR gate

The circuit symbol and the truth table for a two-input NOR gate are given in figure 4.17.



Symbol

A	B	F
0	0	1
0	1	0
1	0	0
1	1	0

Truth table

Figure 4.17

By comparing the two truth tables for OR gate and NOR gate it can be realized that the logic operation of the NOR gate is the inversion of that of the OR gate.

Figure 4.18 shows how a NOR gate can be obtained by passing the output of an OR gate through a NOT gate.

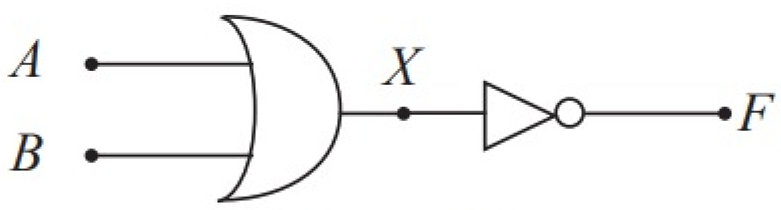


Figure 4.18

Let us prepare a truth table for the above.

A	B	X	$F\ (=\overline{X})$
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0

This output F is the same as the output of a NOR gate.
According to Boolean expression,

$X = A + B$ (expression for the PR gate)
$F = \overline{X}$ (expression for the NOT gate)
$\therefore\ F = \overline{A+B}$

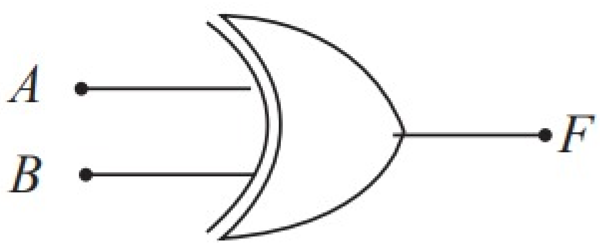
Therefore, the Boolean expression for the NOR gate is given as,

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

The logic expression for the NOR gate can be given as **"when A or B input is 1, the output is 0".**

4.10 XOR gate

The circuit symbol and the truth table for a two-input XOR gate are given in Figure 4.19. XOR is also written as Ex – OR. (X or Ex is used as the shorten form of "exclusive).



Symbol

A	B	F
0	0	0
0	1	1
1	0	1
1	1	0

Truth table

Figure 4.19

It is clear from this truth table that the logic operation of the XOR gate is to output 1 when only one of the inputs A or B is 1. So, if both A and B inputs are 1 then the output will be 0.

The Boolean expression for the XOR is written as,

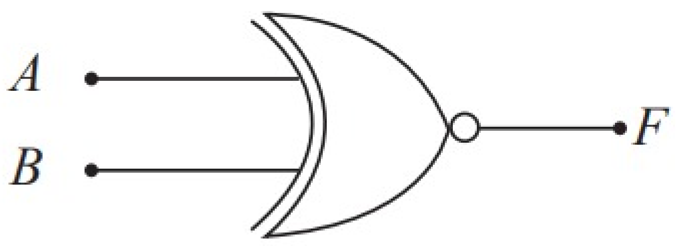
$$F = A \oplus B$$

The logic expression for the XOR gate can be given as, **" When only one of the inputs A or B is 1, then output is 1".**

© 2020 National Institute of Education. All Rights Reserved.

4.11 XNOR gate

The circuit symbols and the truth table for a two input XNOR gate are in figure 4.20. XNOR is also written as Ex-NOR.



Symbol

A	B	F
0	0	1
0	1	0
1	0	0
1	1	1

Truth table

Figure 4.20

By comparing the truth tables for XOR and XNOR gates, it can be seen that the output of XNOR gate is the inversion of the output of XOR gate.

An XNOR gate can be obtained by passing the output of an XOR gate through a NOT gate as shown in figure 4.21.

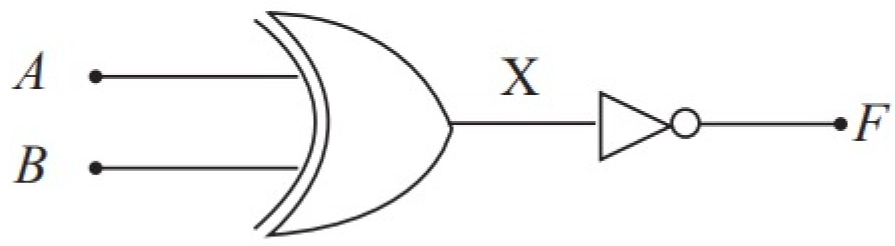


Figure 4.21

Let us prepare a truth table for this arrangement.

A	B	X	$F\ (=\overline{X})$
0	0	0	1
0	1	1	0
1	0	1	0
1	1	0	1

This output F is the same as the output of an XNOR gate.
Considering Boolean expressions,

$X = A \oplus B$ (expression for the XOR gate)

$F = \overline{X}$ (expression for the NOT gate)

$\therefore\ F = \overline{A \oplus B}$

Therefore, the Boolean expression for the XNOR gate is written as,

$$\boxed{F = \overline{A \oplus B}}$$

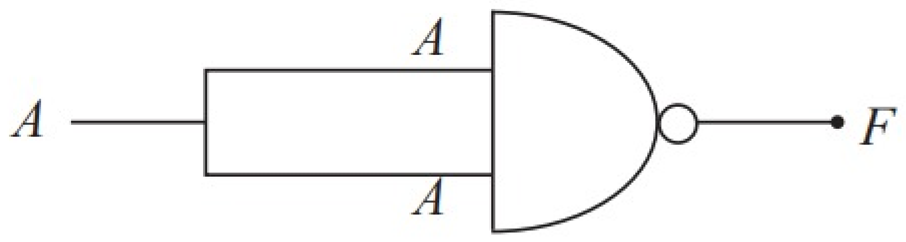
© 2020 National Institute of Education. All Rights Reserved.

The logic expression for the XNOR gate can be given as **"When only one of the inputs A or B is 1, the output is 0"**

Among the logic gates that we have dealt with so far, there is an additional feature in the NAND gate and the NOR gate. That is, by using NAND gates only or NOR gates only, the other gates can be made. Such examples are given below.

4.12 Making other logic gates using NAND gates only

(i) Making a NOT gate



A NOT can be made by connecting the two input terminals of a NAND gate together and taking it as one input.

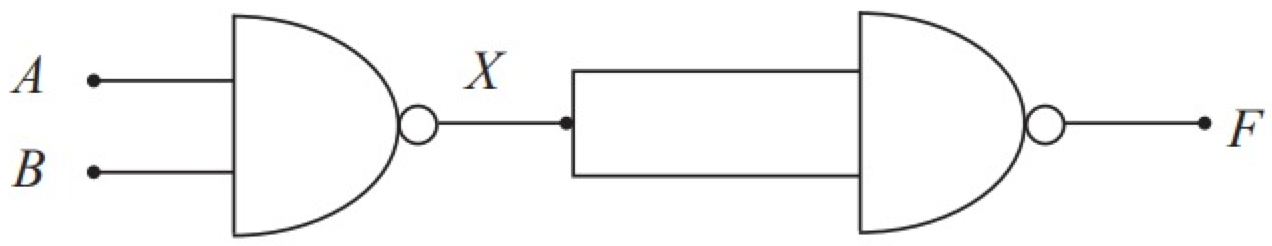
If $A = 1$ then $F = 0$

If $A = 0$ then $F = 1$

$\therefore$ $F = \bar{A}$ (behaves as a NOT gate)

(ii) Making as AND gate

This can be done as follows



According to Boolean expressions,

$X = \overline{A.B}$

And $F = \overline{X}$

$\therefore$ $F = \overline{\overline{A.B}}$

$= A.B$

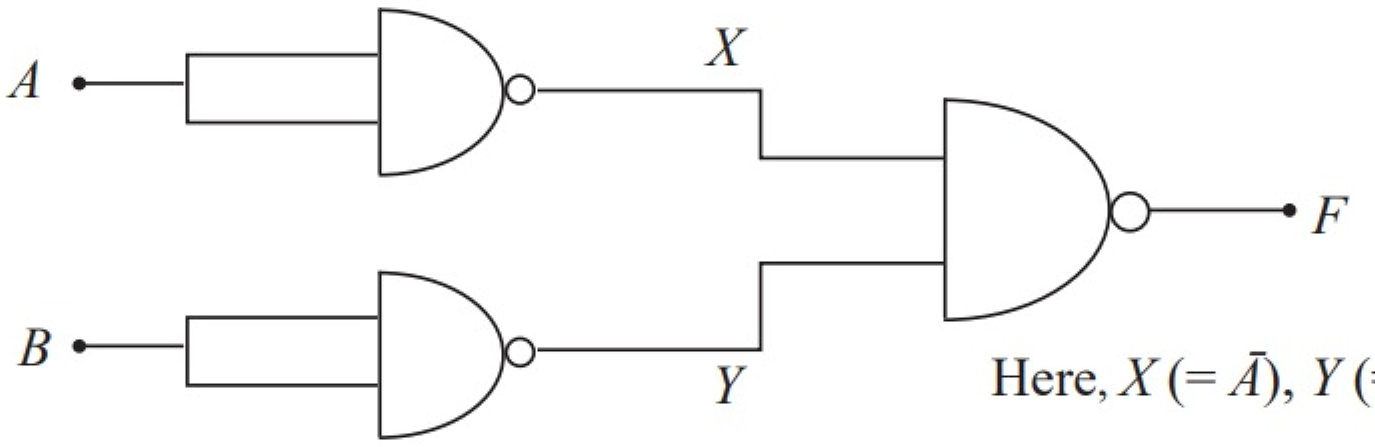
Consider the truth table,

A	B	X	$F (= \overline{X})$
0	0	1	0
0	1	1	0
1	0	1	0
1	1	0	1

This output F is the same as the output of the AND gate. Therefore the required AND operation has been obtained.

© 2020 National Institute of Education. All Rights Reserved.

(iii) Making an OR gate



Here, $X\,(=\bar{A})$, $Y\,(=\bar{B})$ and $F(=\overline{X.Y})$

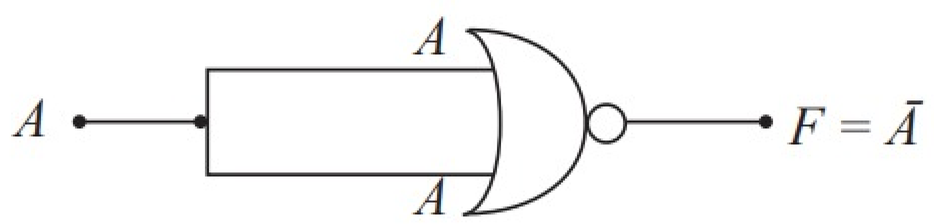
Consider the truth table relevant to this circuit.

A	B	X $(=\bar{A})$	Y $(=\bar{B})$	F $(=\overline{X.Y})$
0	0	1	1	0
0	1	1	0	1
1	0	0	1	1
1	1	0	0	1

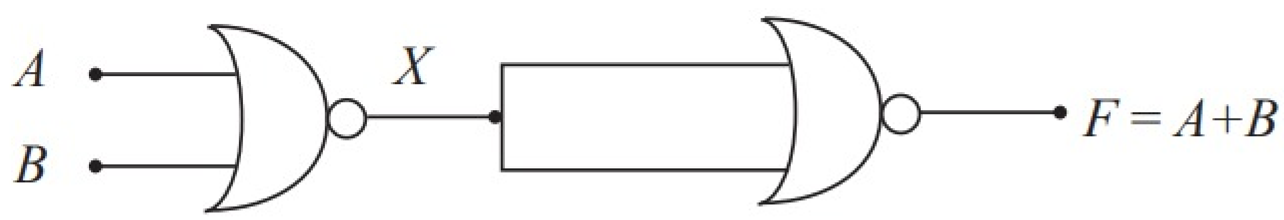
The output F is the same as the output of the OR gate.

4.13 Making other logic gates using NOR gates only

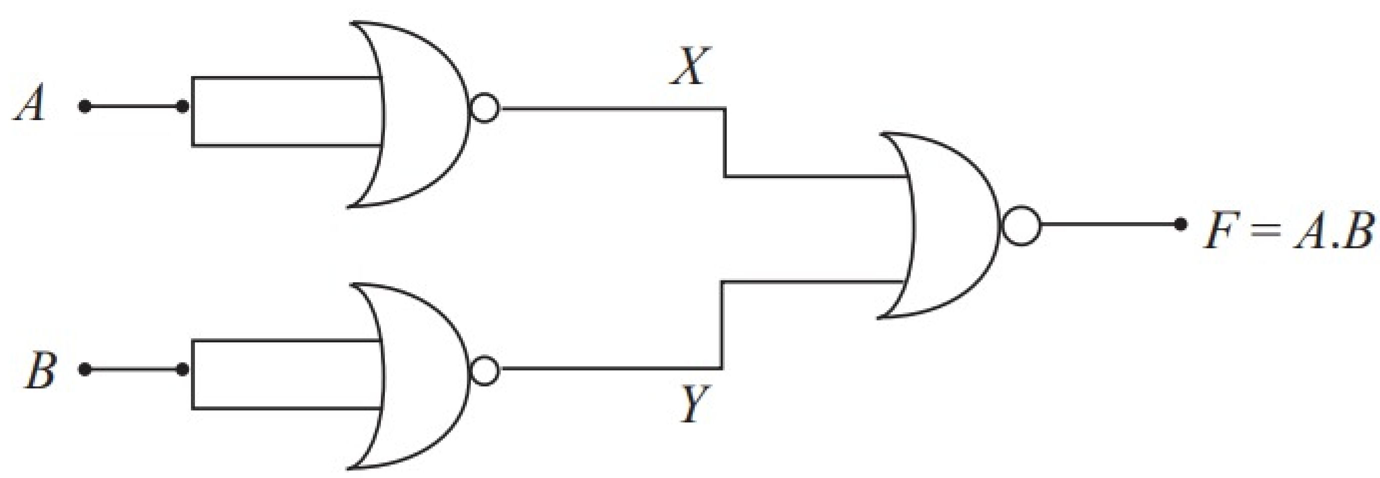
(i) NOT gate



(ii) OR gate



(iii) AND gate



It is left to you as an assignment, to check the correctness of these by preparing the relevant truth tables.

Because of the reason that the other logic gates can be made using NAND gates only or NOR gates only, the two gates NAND and NOR are known as **universal gates.**

In practical usage the inputs of the logic gates should always be at either 0 (0 V) or 1 (5 V) level. The inputs should not be left open (floating). If they are left open then the inputs must not be definite and hence outputs may not be correct.

To apply inputs to logic gates without making inputs open, an appropriate circuit arrangement can be applied as given by the following examples in Figure 4.22 and Figure 4.23.

i)

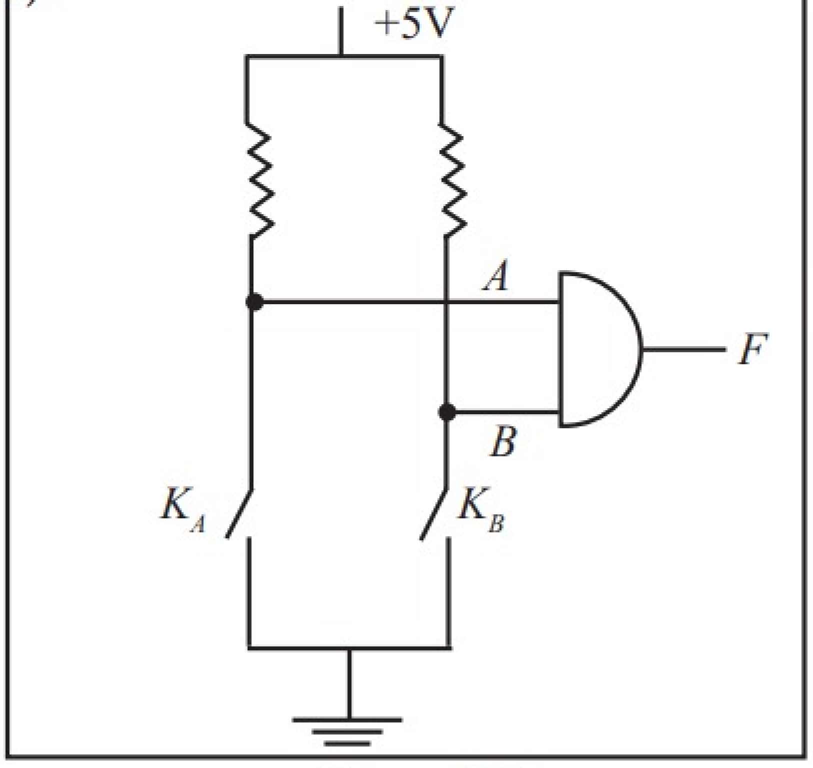


Figure 4.22

When the key K_A is open (OFF),
Input A gets +5 V voltage via 10 kΩ resistor.
Therefore, A = logic 1

When K_A is closed (ON),
Input A gets the ground voltage (zero).
Therefore, A = logic 0

Similarly by opening or closing the key K_B, the input B can be supplied with logic 1 or logic 0 respectively.

ii)

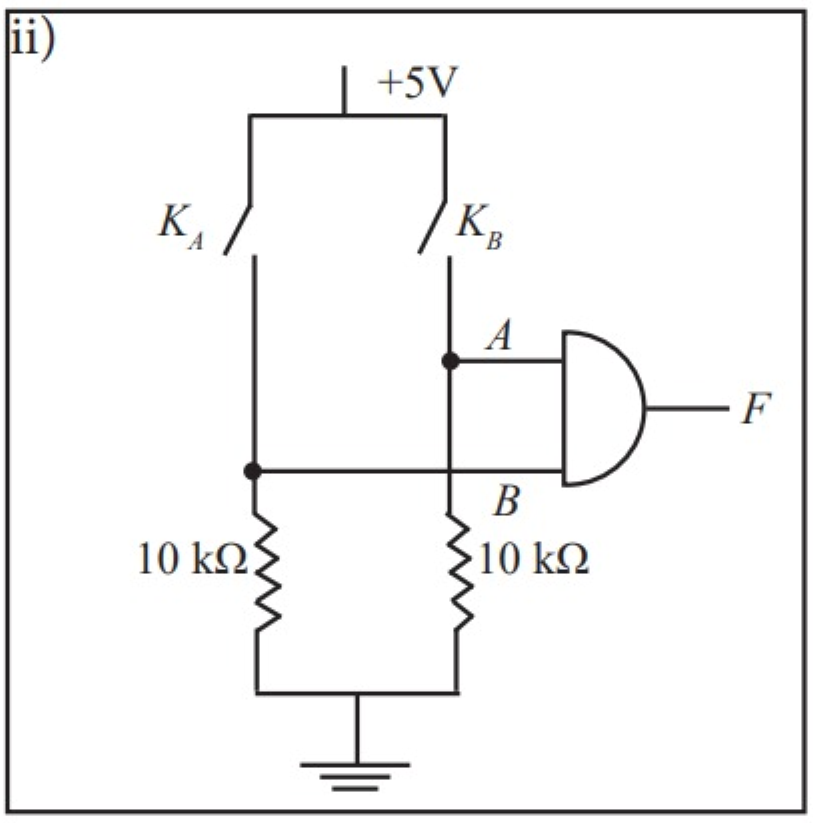


Figure 4.23

When K_A is open (OFF),
Input A gets the ground voltage zero via 10 kΩ resistor.
Therefore, A = logic 0

When K_A is closed (ON),
Input A gets the +5 V voltage
Therefore, A = logic 1

Similarly by opening or closing the key K_B, the input B can be supplied with logic 0 or logic 1, respectively.

In the above two examples each input is always supplied with a logic level either 0 or 1, and never left open.

Earlier it has been described that DRL (Diode Resistor Logic) and RTL (Resistor Transistor Logic) are two forms of circuit used in the production of logic gates. In addition to them there is another method of producing logic gates using combinations of transistors. That is called TTL (Transistor Transistor Logic). In the TTL form junction transistors are used to produce ICs. Further more, there is another way of producing logic gates using field effect transistors (FETs). There the FETs made with semiconductors and a thin metal oxide layer are used. They are known as MOSFET (Metal Oxide Semiconductor FET). The MOSFETs are used as complimentary pairs (pairs having combinational action with each other) in making logic gates. That form of circuit is known as CMOS (Complimentary Metal Oxide Semiconductors).

The power supply for TTL circuits should exactly be with 5 V. As for CMOS circuits they can be supplied with any constant voltage in the range 3 V to 15 V. TTL circuits are speedier in operation than CMOS circuits (reason is the use of junction transistors in TTL circuits). Therefore, for the circuits operated with high frequency signals, TTL type is better.

Logic gates are produced in the form of ICs as TTL type and CMOS type. You can buy them referring to their IC numbers. The photographs of such integrated circuits are shown in Figure 4.24.

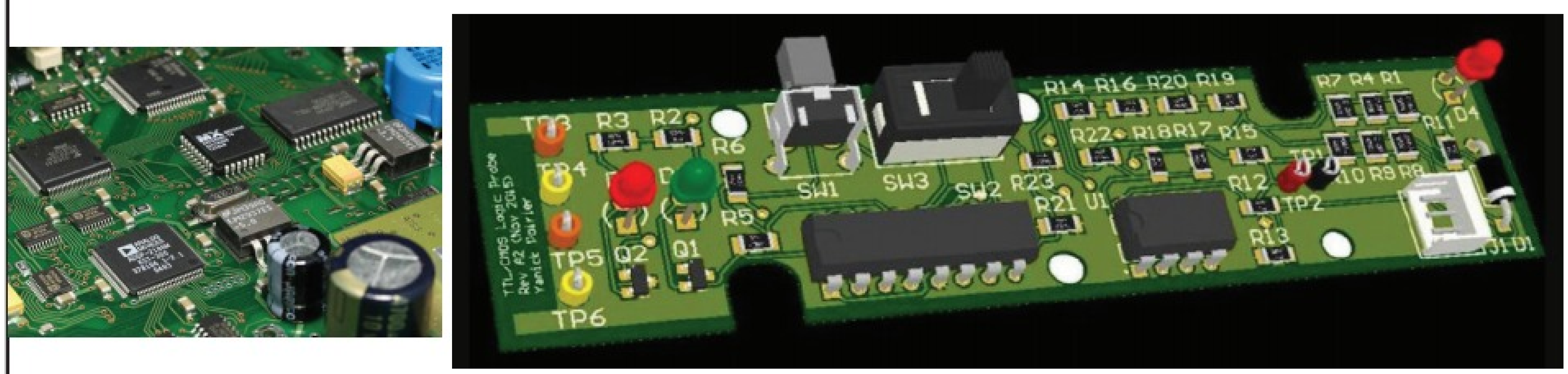
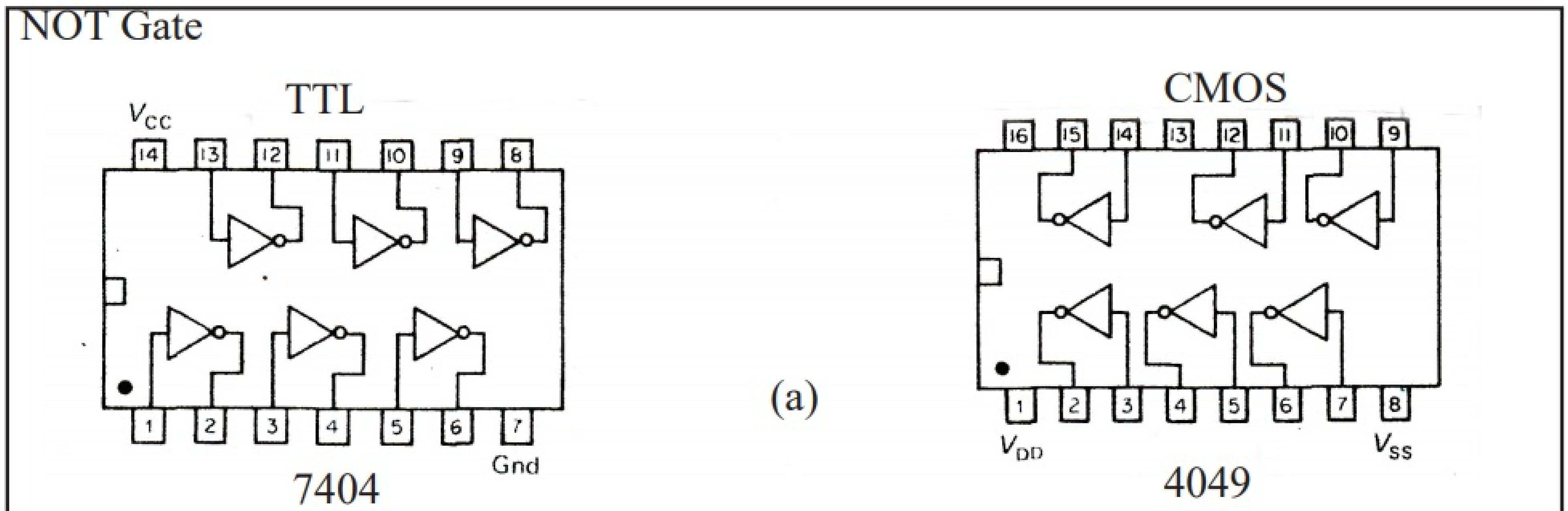


Figure 4.24

The IC numbers of TTL and CMOS ICs with NOT gates or AND gates or OR gates, and the diagrams showing their gate arrangement are given below in Figures 2.25 (a), (b) and (c).



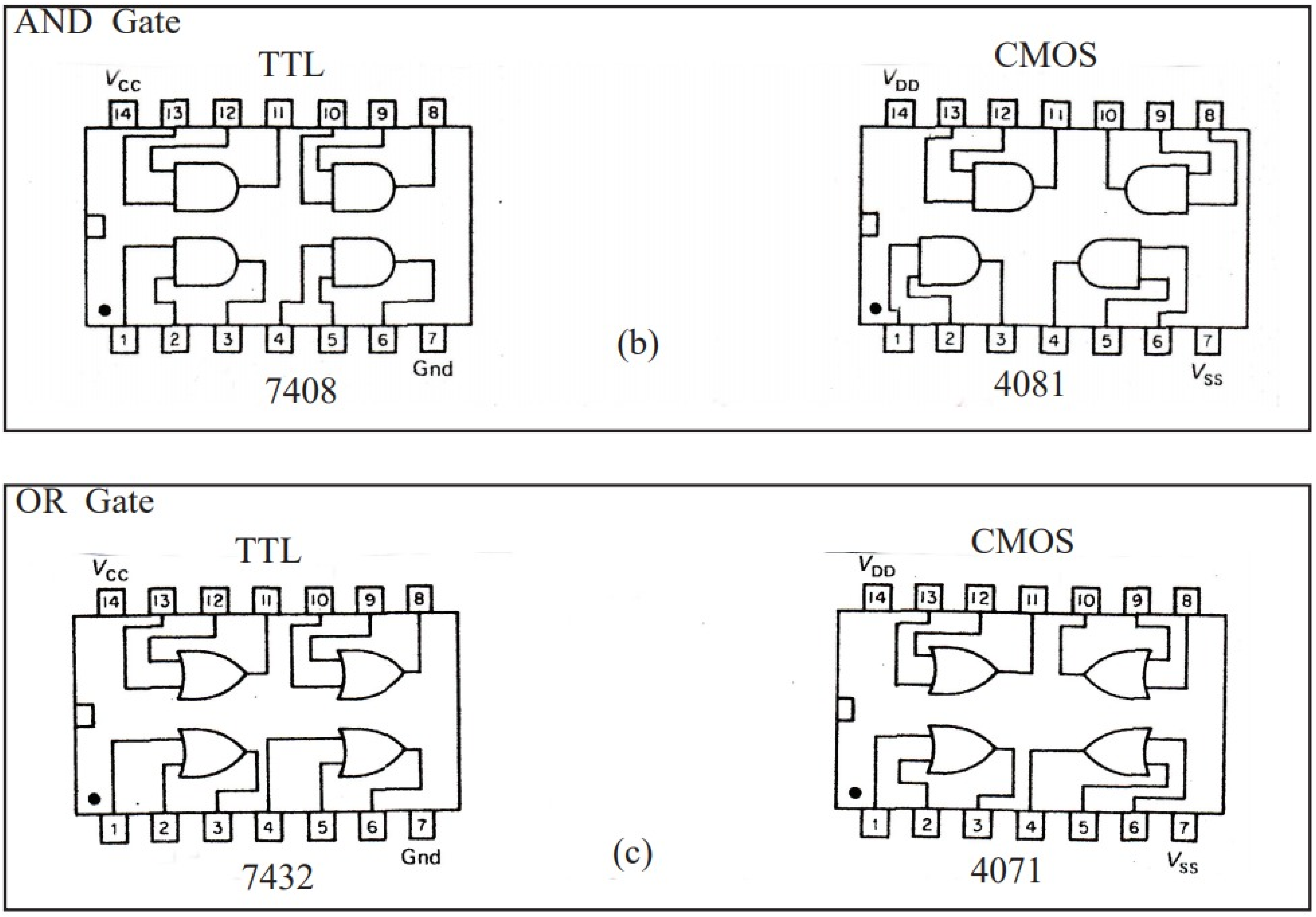


Figure 4.25

4.14 Designing logic gate circuits

The circuits which are capable of giving the relevant outputs, after operating the input digital signals according to a particular logic, can be introduced as logic circuits. When making logic circuits, basically logic gates are combined appropriately to design **logic gate circuits.** Such a designing of a logic gate circuit requires several major steps to be followed as given below.

1. Obtaining the expression which represents the logic process for the relevant problem.
2. Preparing a truth table representing that logic process.
3. Writing the Boolean expression relative to that truth table.
4. Designing the circuit using logic gates so that it operates according to that Boolean expression.

To have a better understanding of the process of designing logic gate circuits, let us consider it in steps given below.

1. Obtaining the truth table relevant to a given Boolean expression.
2. Writing the Boolean expression relevant to a given truth table.
3. Designing a logic gate circuit relevant to a given truth table.
4. Obtaining the truth table relevant to a given logic gate circuits.
5. Writing the Boolean expression relevant to a given logic gate circuit.
6. Designing a logic gate circuit relevant to a given Boolean expression.
7. Designing a logic gate circuit behaving according to given conditions, to fulfill a given task.

© 2020 National Institute of Education. All Rights Reserved.

1. Obtaining the truth table relevant to a given Boolean expression.

Consider the Boolean expression $F = A.B + \bar{A}.\bar{B}$

It has two inputs A and B

Relating to them $A.B$ and $\bar{A}.\bar{B}$ are obtained.

To obtain the truth table let us prepare a table considering all the above inputs, outputs and operations as follows.

A	B	$\bar{A}$	$\bar{B}$	$A.B$	$\bar{A}.\bar{B}$	$F = (A.B + \bar{A}.\bar{B})$
0	0	1	1	0	1	1
0	1	1	0	0	0	0
1	0	0	1	0	0	0
1	1	0	0	1	0	1

Writing $\bar{A}$ considering A

Writing $\bar{B}$ considering B

Writing $A.B$ considering A and B

Writing $\bar{A}.\bar{B}$ considering $\bar{A}$ and $\bar{B}$

Writing F considering $A.B$ and $\bar{A}.\bar{B}$

Using the above table the basic (simple) truth table can be given as follows.

A	B	F
0	0	1
0	1	0
1	0	0
1	1	1

2. Writing the Boolean expression relevant to a given truth table.

As an example let us obtain the Boolean expression for the truth table given below.

A	B	F
0	0	1
0	1	1
1	0	0
1	1	0

At first let us select only those states where $F = 1$, and write those states in terms of A and B inputs.

A	B	F
0	0	1
0	1	1

This can be written as $F = \bar{A}.\bar{B}$
($\bar{A} = 1$ and $\bar{B} = 1$ as $A = 0$ and $B = 0$.
That is $\bar{A}.\bar{B} = 1.1 = 1$)

This can be written as $F = \bar{A}.B$
($\bar{A} = 1$ as $A = 0$. Then as $B = 1$, $\bar{A}.B = 1.1 = 1$)

© 2020 National Institute of Education. All Rights Reserved.

Since $F = 1$ in any of the above two states $\bar{A}.\bar{B}$ and $\bar{A}.B$, the relevant Boolean expression can be written as,

$$F = \bar{A}.\bar{B} + \bar{A}.B$$

3. Designing a logic gate circuit relevant to a given truth table

As an example let us design the circuit relevant to the truth table given below.

A	B	F	
0	0	1	← $F = \bar{A}.\bar{B}$
0	1	1	← $F = \bar{A}.B$
1	0	0	
1	1	1	← $F = A.B$

At first the relevant Boolean expression should be written. For that let us write the Boolean expression for the states where $F = 1$ (These are written on the right of the table).

Accordingly the relevant Boolean expression is,

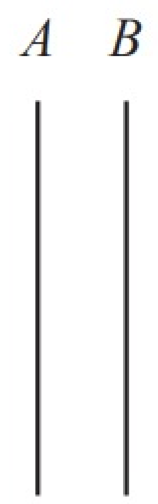
$$F = \bar{A}.\bar{B} + \bar{A}.B + A.B$$

The inputs are A and B.

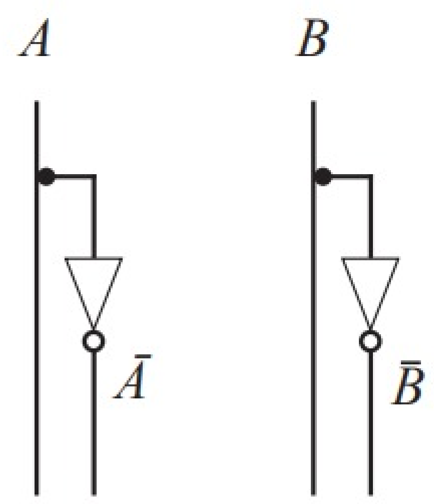
With them $\bar{A}.\bar{B}$, $\bar{A}.B$ and $A.B$ should be obtained. To do that NOT gates and AND gates are required. Finally a three-input OR gate is required to get $\bar{A}.\bar{B} + \bar{A}.B + A.B$.

Let us build the relevant circuit as follows.

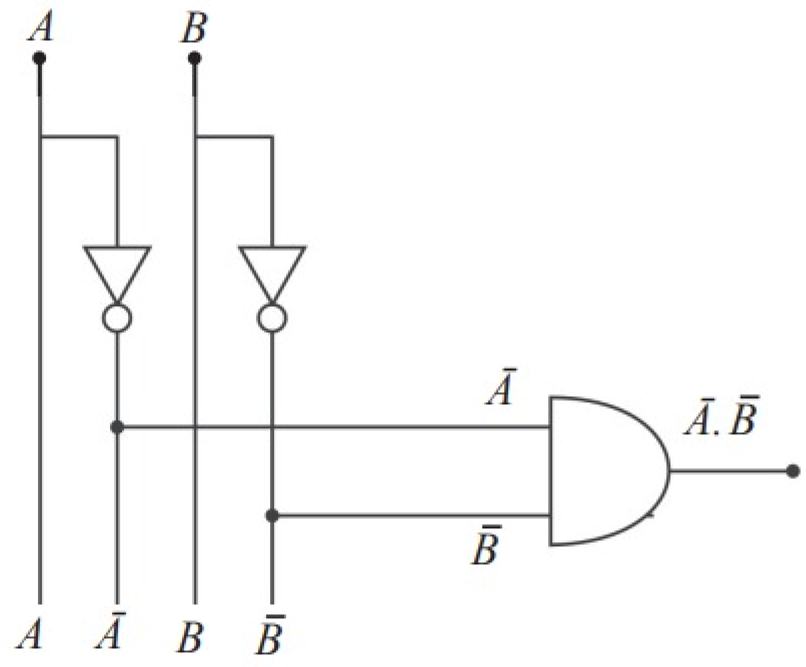
(i) Show input lines A and B



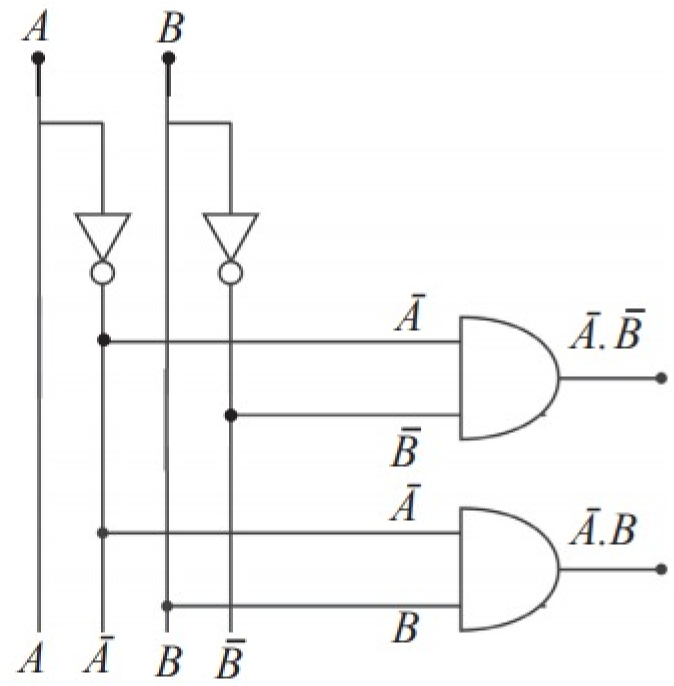
(ii) Obtain $\bar{A}$ and $\bar{B}$ using NOT gates



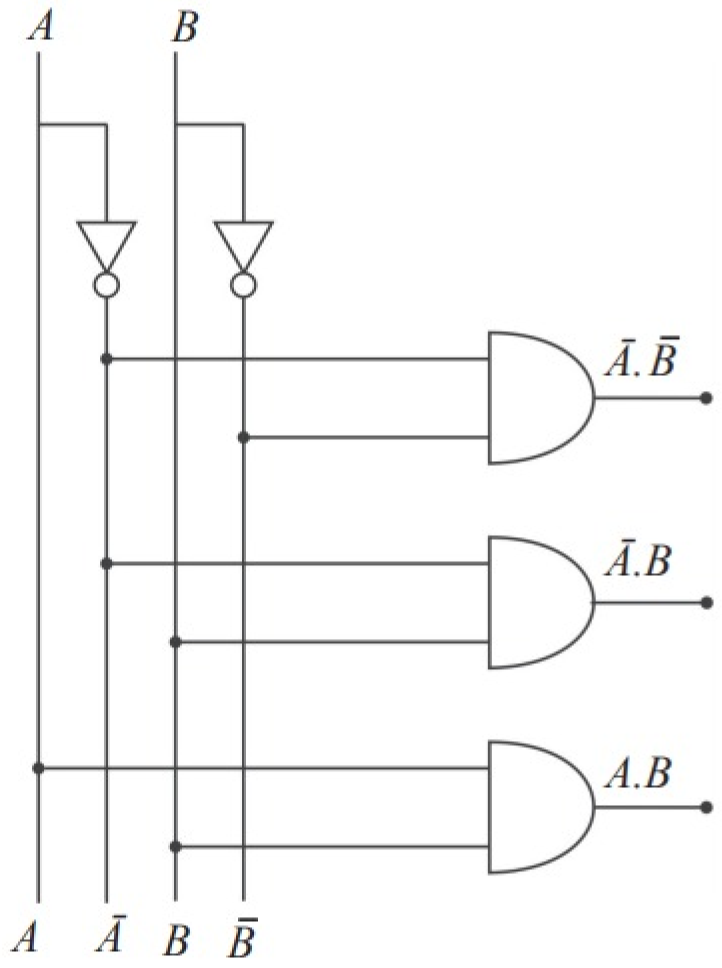
(iii) Obtain $\bar{A}.\bar{B}$ using an AND gate



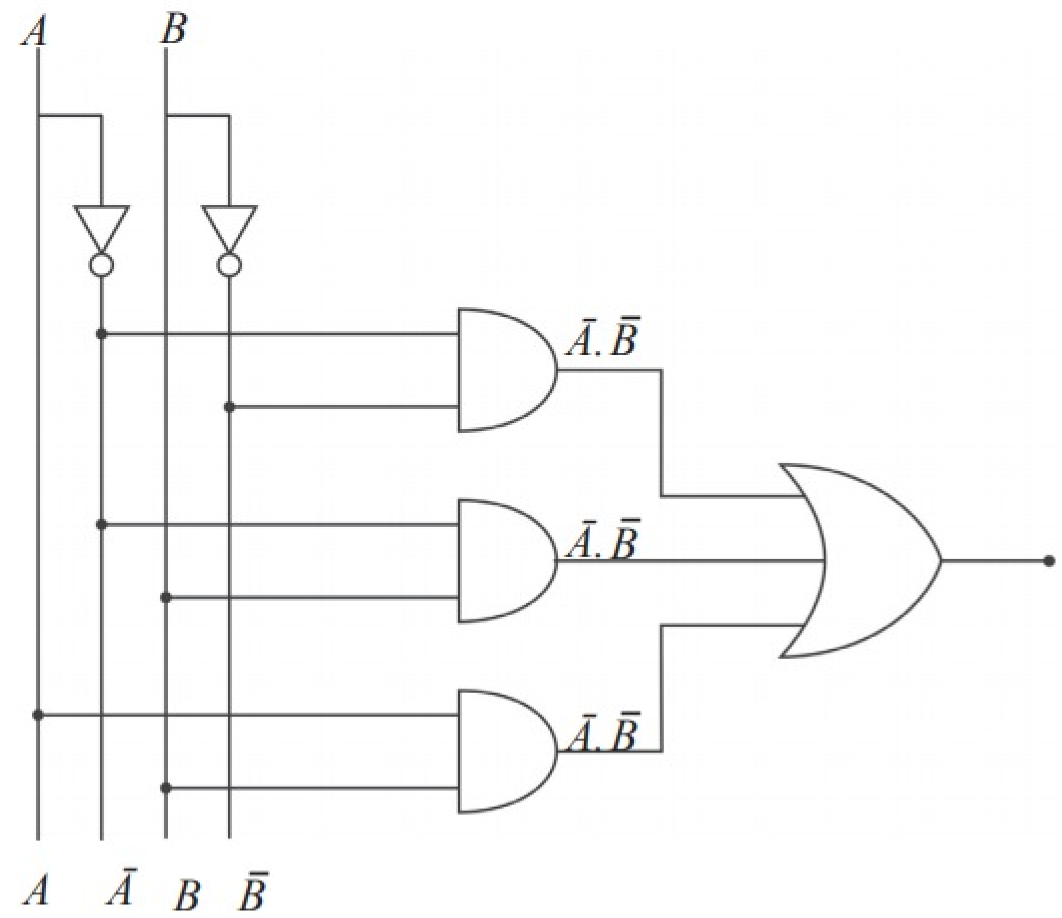
(iv) Obtain $\bar{A}.B$ using another AND gate



(v) Obtain $A.B$ using one more AND gate .



(vi) Finally using an OR gate with 3 inputs, let us draw the complete logic gate circuit so as to obtain the output $\bar{A}.\bar{B} + \bar{A}.B$ as follows.

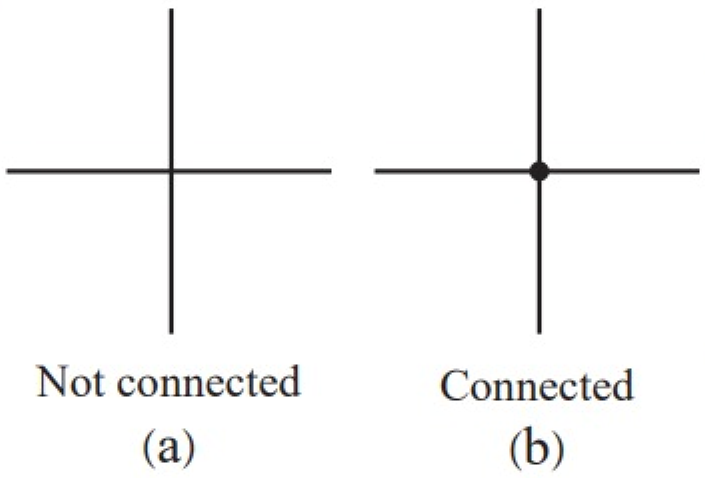


$F = \bar{A}.\bar{B} + \bar{A}.B + A.B$

Note:

When drawing these circuit diagrams there are two methods to indicate whether two wires are connected or not when they cross each other.

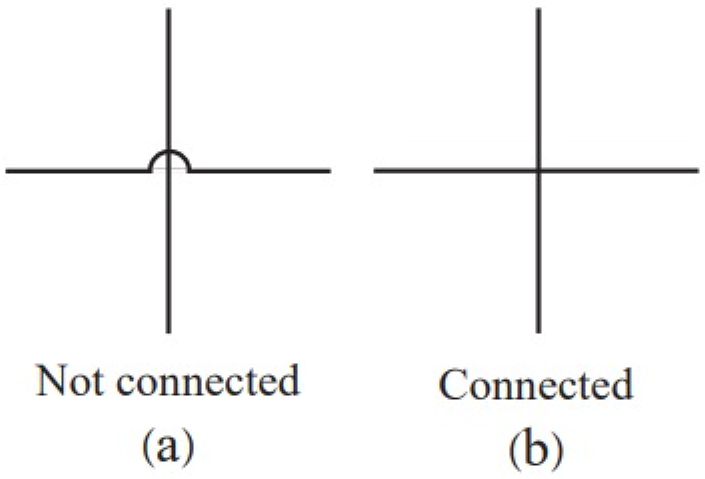
(i)



In this method, when two wires are crossing each other without getting connected, the wires are drawn as shown in Figure (a).

If the wires are not connected they are drawn as shown in Figure (b).

(ii)



In this method, when two wires are crossing each other without getting connected, the wires are drawn as shown in figure (a).

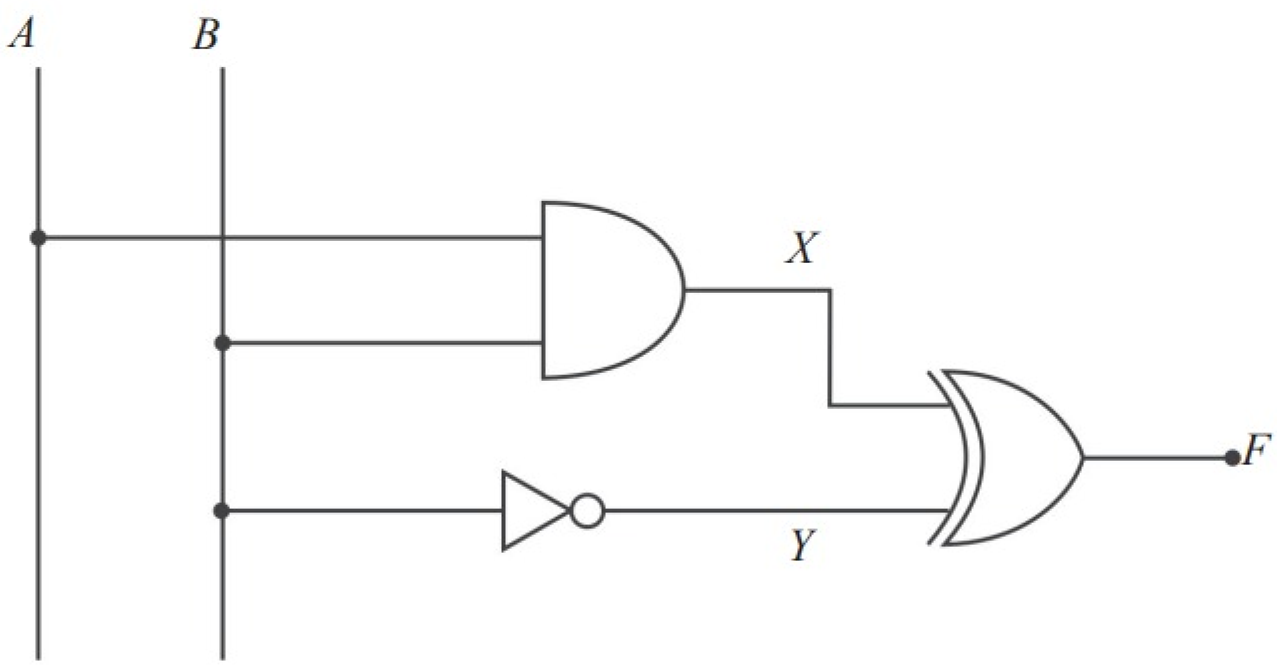
If the wires are not connected they are drawn as shown in Figure (b).

It should be noted that the method (i) above is used in this book hereafter.

In the circuit designed above, input *B* is fed to a NOT gate and the output of that NOT gate is used to feed an AND gate as the input. Similarly input *A* is fed to a NOT gate and the output of it has been used to feed two other AND gates as the input. When feeding the output of a logic gate to some other gates as inputs, there can be a reduction in output voltage level of the initial logic gate, and that may lead to malfunction of the circuit. Therefore there is a limit to the number of gate inputs that can be connected to the output of a logic gate. That limit depends on the type of the logic gate.

4. Obtain the truth table relevant to a given logic gate circuit.

You can learn it with the help of following logic gate circuit.



Using Boolean expressions

$X = A.B$ (expression for the AND gate)

$Y = \bar{B}$ (expression for the NOT gate)

and $F = X \oplus Y$ (expression for the XOR gate)

$\therefore$ $F = A.B \oplus \bar{B}$

As the next step prepare a descriptive table including $A.B$ and $\bar{B}$ in addition to the inputs A,B and the output F.

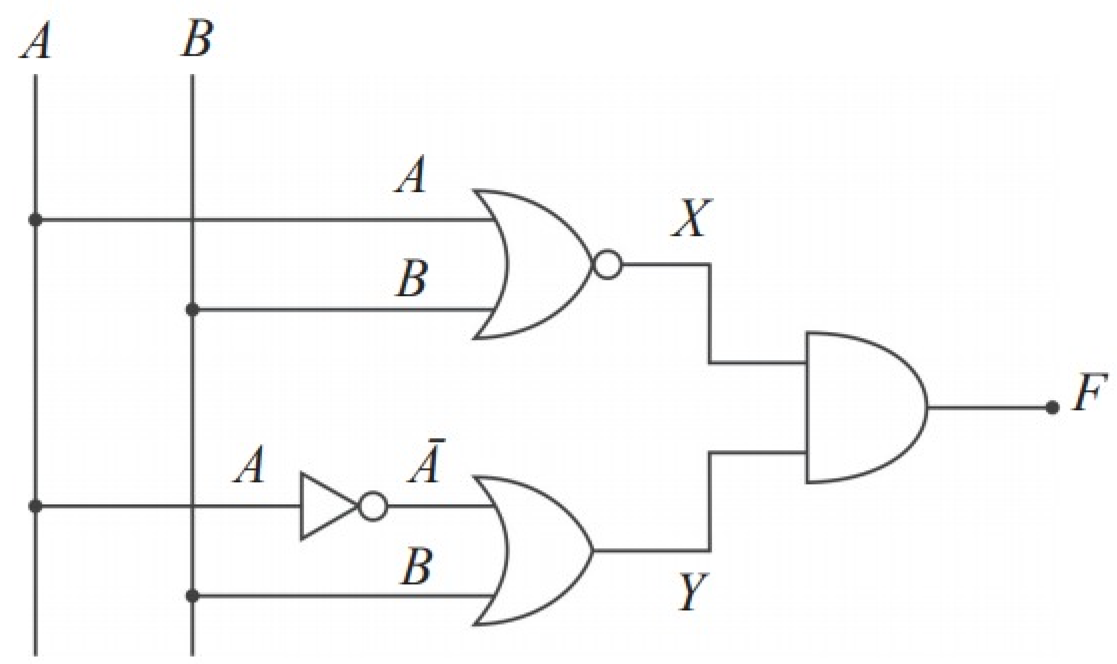
A	B	$A.B$	$\bar{B}$	$F = A.B \oplus \bar{B}$
0	0	0	1	1
0	1	0	0	0
1	0	0	1	1
1	1	1	0	1

Now you can prepare the relevant truth table given below.

A	B	F
0	0	1
0	1	0
1	0	1
1	1	1

5. Writing the Boolean expression relevant to a given logic gate circuit

Consider the following logic gate circuit.



Considering NOR gate,

$X = \overline{A+B}$

Considering OR gate,

$Y = \bar{A} + B$

Considering AND gate,

$F = X.Y$

$= (\overline{A+B}) . (\bar{A} + B)$

∴ The relevant Boolean expression is,

$$F = (\overline{A+B}) \,.\, (\bar{A} + B)$$

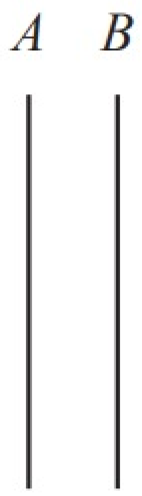
6. Designing a logic gate circuit relevant to a given Boolean expression.

Consider the Boolean expression given below.

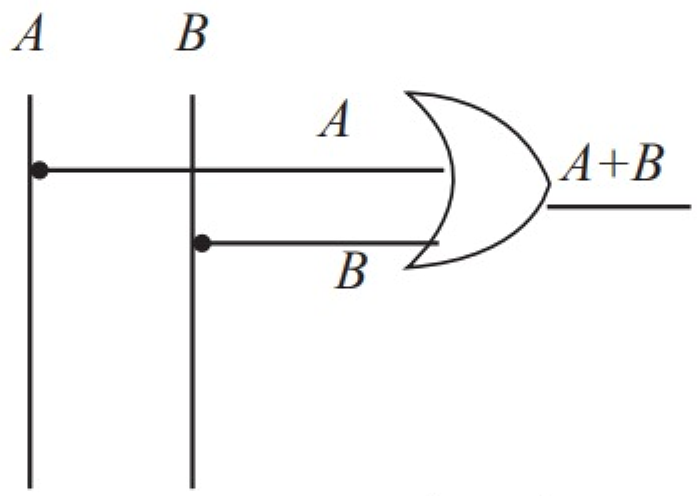
$$\text{F} = (\bar{A}+B) \,.\, (\overline{\bar{A}+B})$$

Designing the circuit (a similar one has been described earlier also in step 3).

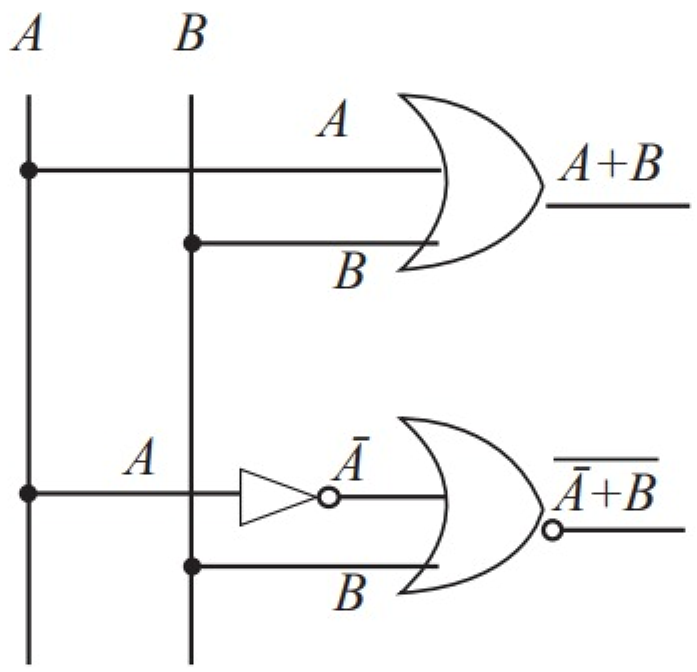
(i) showing input lines



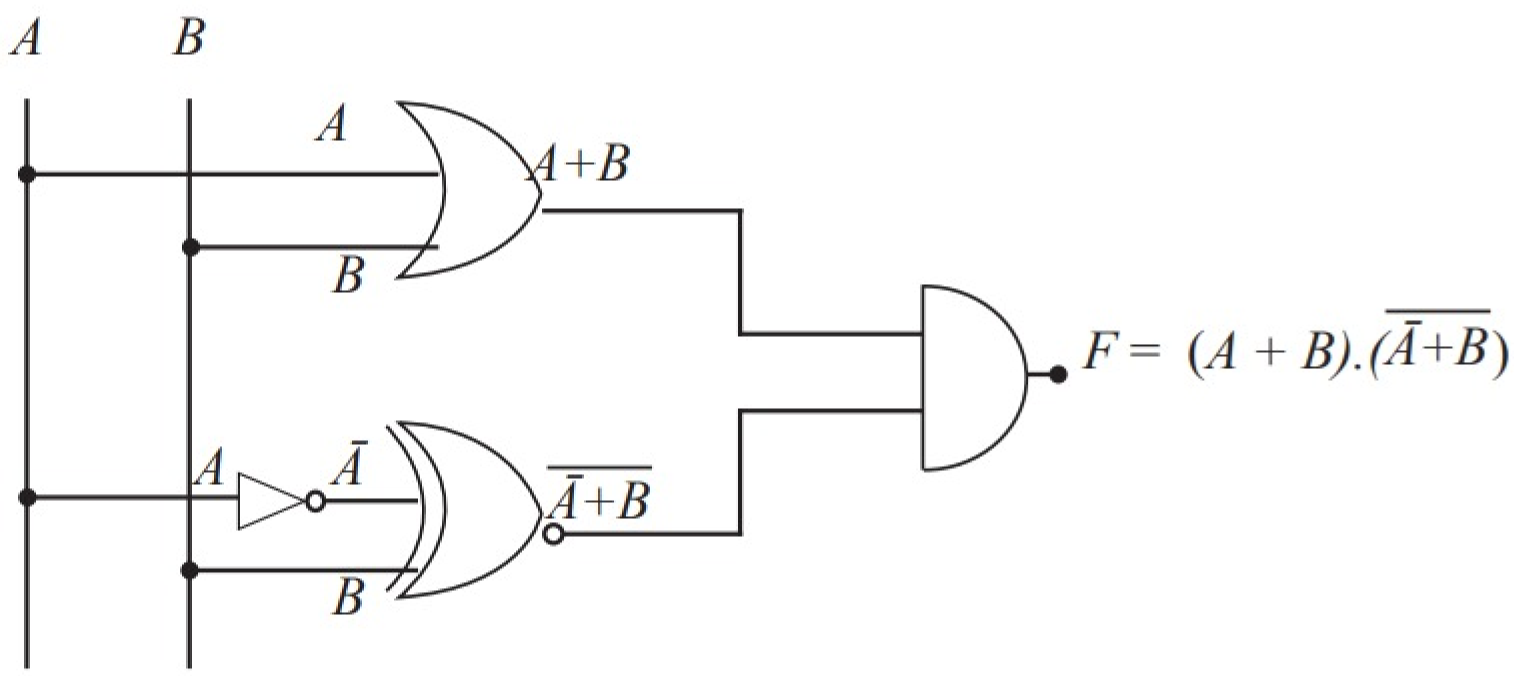
(ii) Obtaining A+ B using an OR gate



(iii) Obtaining $\bar{A}$ using a NOT gate and then obtaining $\overline{\bar{A}+B}$ using a NOR gate.



(iv) Completing the circuit diagram using AND gate to get the output F.



7. Designing a logic gate circuit behaving according to given conditions, to fulfill a given task.

To understand this whole process the example given below will be helpful to you.

An emergency lamp has to fulfill the conditions given below.

(i) The lamp lights at night (in the darkness) only when there is no mains electricity.

(ii) The lamp does not light during daytime (when light is there) irrespective of whether mains electricity is or is not there.

For the circuit two sensors are provided. The sensors output logic signals A and B such that,

$A = 1$ when mains electricity is there.

$A = 0$ when mains electricity is not there.

$B = 1$ at night (in the darkness)

$B = 0$ during daytime (when light is there)

To get the lamp lighted as required, an output F should be obtained so that in (i) above output is logic 1 (F=1) and in (ii) above output is logic 0 (F=0). (That is it $F = 1$ then the lamp lights and it $F = 0$ then the lamp does not light)

Taking A and B as inputs and F as the output, let us design a logic gate circuit so that the given conditions (i) and (ii) above are satisfied.

To achieve this, the relevant truth table should be prepared first. According to the conditions (i) and (ii) above the relevant logic expression is **"The output F should be 1 only when the input A is zero and the input B is 1."**

The relevant truth table is as follows.

A	B	F	
0	0	0	← $A=0$ (no electricity), $B=0$ (daytime) ∴ $F=0$ (lamp doesn't light)
0	1	1	← $A=1$ (no electricity), $B=1$ (nighttime) ∴ $F=1$ (lamp lights)
1	0	0	← $A=1$ (electricity available), $B=0$ (daytime) ∴ $F=0$ (lamp doesn't light)
1	1	0	← $A=1$ (electricity available), $B=1$ (nighttime) ∴ $F=0$ (lamp doesn't light)

Let us now obtain the relevant Boolean expression.

To get it we should consider only the states where $F = 1$.

We have only one such state here as follows.

© 2020 National Institute of Education. All Rights Reserved.

A	B	F
0	1	1

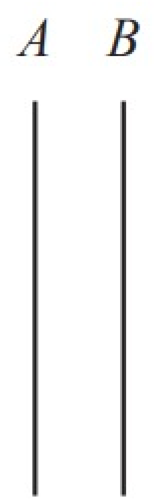
← This state can be represented by $\bar{A}$,B ($\bar{A}$=1 when A=0. Also B=1, Therefore, $\bar{A}.B = 1.1 = 1$)

Therefore, the relevant Boolean expression is,

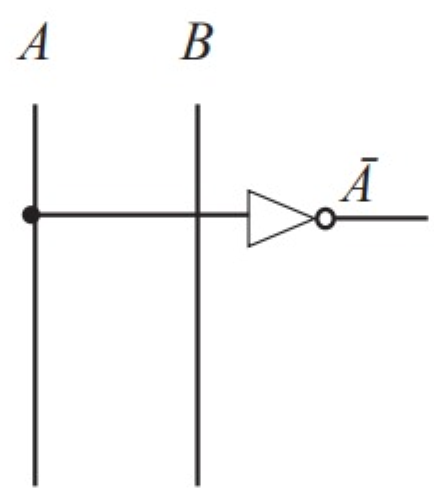
$$F = \bar{A}.B$$

As the final step let us design the logic gate circuit relevant to this Boolean expression.

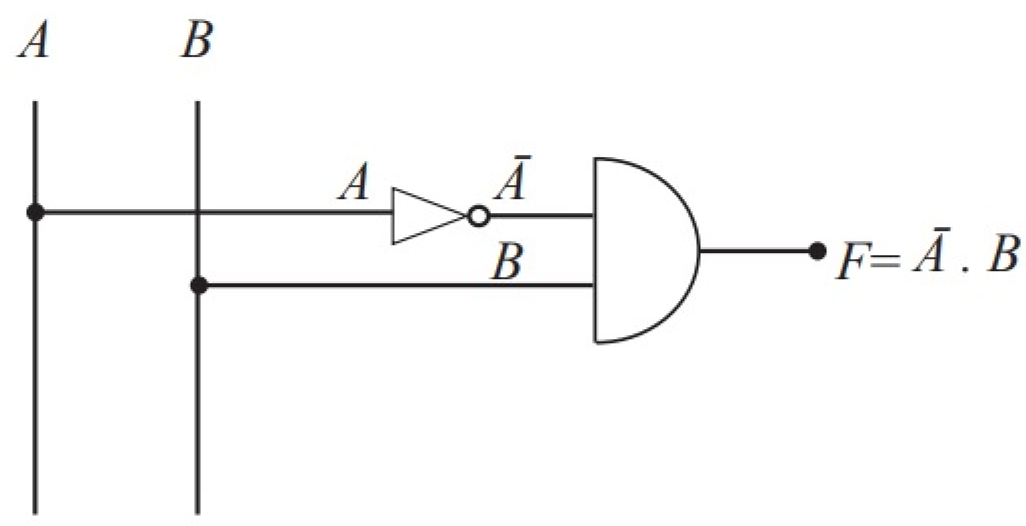
First let us show input lines A and B.



Next obtain $\bar{A}$ using a NOT gate.



To get the required output F, $\bar{A}.B$ should be obtained using an AND gate.



The switching circuit used to light the lamp is switched by this output F, and then the emergency lamp lights according to the conditions given.

4.15 Sequential Logic Circuits

In the logic gate circuits we have discussed so far, the output at an instance is decided by the inputs present at that particular instance. The logic circuits in which the output is decided in the above manner are called **combinational logic circuits**. However there is a special kind of logic circuits in which the output is decided not only on the present inputs but also on the previous output. This kind of a circuit has a memory of the output that prevailed at the previous state. This kind of circuits are known as **sequential logic circuits.**

4.16 Flip-Flops

The flip- flop can be introduced as a basic circuit of the kind of sequential circuits. It can also be considered as a memory element because it has a memory of its previous output. Although a memory cannot be there in a single logic gate, a particular combination of logic gates known as a flip-flop can store a binary digit 0 or 1, also stored digit can be read when required. There are several types of flop-flops. Let us discuss a simple type known as Set-Reset flip flop (SR flop –flop).

The S-R flip flop has two input terminals named as *S* input and *R* input. *S* input is known as the **Set** input and the *R* input is known as the **Reset** input. This flip-flop has two output terminals known as Q and $\bar{Q}$. $\bar{Q}$ is the inversion of Q (that is when $\bar{Q} = 0, \bar{Q} = 1$ and when $\bar{Q} = 1, \bar{Q} = 0$). In addition to Q, $\bar{Q}$ is also there for the reason that it is useful when the flip-flop is connected to another circuit component. Depending on the requirement Q or $\bar{Q}$ or both can be used.

The circuit symbol of the SR flip-flop is shown in Figure 4.26.

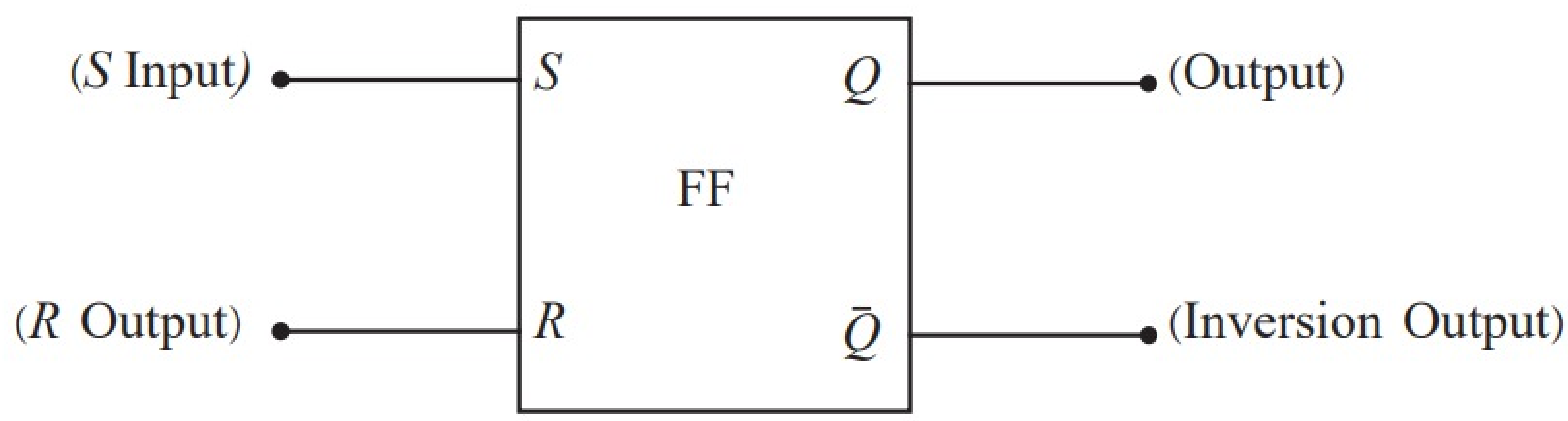


Figure 4.26

4.16.1 Preparing *SR* flip - flops

A set-reset flip-flop can be made by the appropriate combination of two NOR gates. To get the influence of the previous output, its output is fed to its input as a feed-back. The logic gate circuit of such a *SR* flip-flop made using two NOR gates is shown in Figure 4.27.

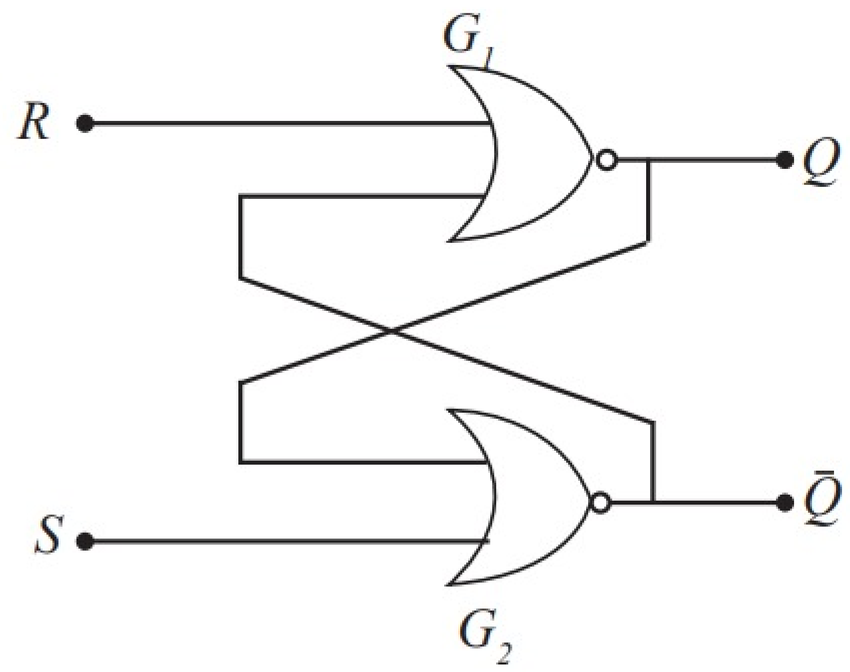


Figure 4.27

To understand the operation of this flip-flop let us consider the different ways that the inputs can be applied to S and R inputs.

1. Applying inputs as $S = 0$ and $R = 0$ (unchanged state)

Let us use the circuit given in figure 4.28 to study this. In it the previous (old) outputs of Q and $\bar{Q}$ are given in cages.

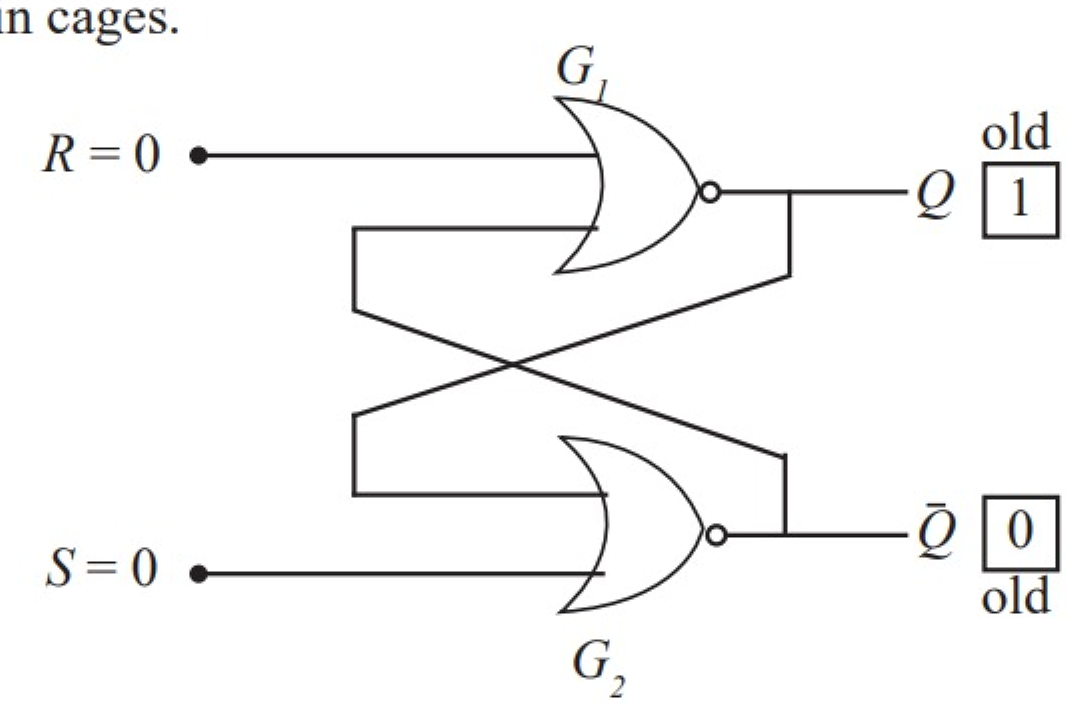


Figure 4.28

Considering G_1 - NOR gate,

One input is 0 (as $R = 0$) and the other input is also 0 (as $\bar{Q}_{old} = 0$). Therefore new output of G_1 is $Q_{new} = 1$.

Considering G_2 - NOR gate,

One input is 0 (as $S = 0$) and the other input is 1 (as $Q_{new} = 1$). Therefore new output of G_2 is $Q_{new} = 0$.

According to above, the new outputs of the flip-flop are steady at $Q_{new} = 1$ and $Q_{new} = 0$. Therefore the previous outputs remain unchanged.

In the case where the previous outputs were $Q_{old} = 0$ and $Q_{old} = 1$, it can be shown that the Q and $\bar{Q}$ outputs are unchanged.

This result can be shown in a simple table as follows.

S	R	Q_{n+1}	$\bar{Q}_{n+1}$
0	0	Q_n	$\bar{Q}_n$

Outputs are unchanged
This is the unchanged state of the outputs.

Where $Q_n, \bar{Q}_n$ are the previous outputs and $Q_{n+1}, \bar{Q}_{n+1}$ are the new outputs. Therefore the table says that when the inputs are as $S = 0$ and $R = 0$ the new outputs of the flip - flop are same as the previous outputs (Q_{n+1} is given as Q_n and Q_{n+1} is given as Q_n).

The input state of *S = 0, R = 0* is the "unchanged state" of the outputs.

2. Applying inputs as *S = 1* and *R = 0* (Set state)

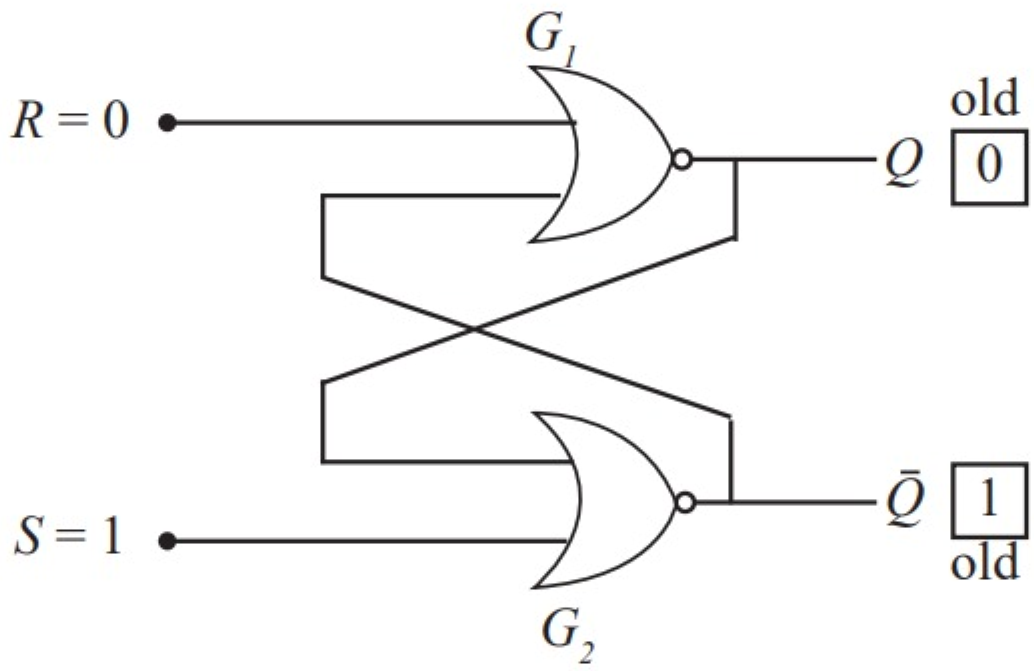


Figure 4.29

As shown in Figure 4.29, one input of G_2 - NOR gate is 1 (as $S = 1$). The other input is 0 (as $Q_{old} = 0$). Therefore, its new output $Q_{new} = 0$.

Now consider G_1-NOR gate. One input of it is 0 (as $R = 0$). The other input is also 0 (as $Q_{new} = 0$). Therefore its new output $Q_{new} = 1$. Accordingly the new outputs of the flip-flop are $Q_{new} = 1$ and $Q_{new} = 0$.

It can be shown that even in the case where the previous output were $Q_{old} = 1$ and $\bar{Q}_{old} = 0$, the new outputs will be $Q_{new} = 1$ and $Q_{new} = 0$ if the inputs are $S=1$ and $R = 0$.

This result can be represented by a simple table as shown below.

S	R	Q_{n+1}	$\bar{Q}_{n+1}$
1	0	1	0

Irrespective of the previous outputs Q_n and $\bar{Q}_n$, the new outputs Q_{n+1} and $\bar{Q}_{n+1}$ are 1 and 0 respectively. **This is the set state of the outputs.**

According to the above, when the inputs to the flip-flop are $S = 1$ and $R = 0$, irrespective of the previous outputs, the flip-flop will come to the set state. The set state of the flip-flop is the state where the new outputs are $Q_{n+1} = 1$ and $\bar{Q}_{n+1} = 0$.

3. Applying inputs as $S = 0$ and $R = 1$ (Reset state)

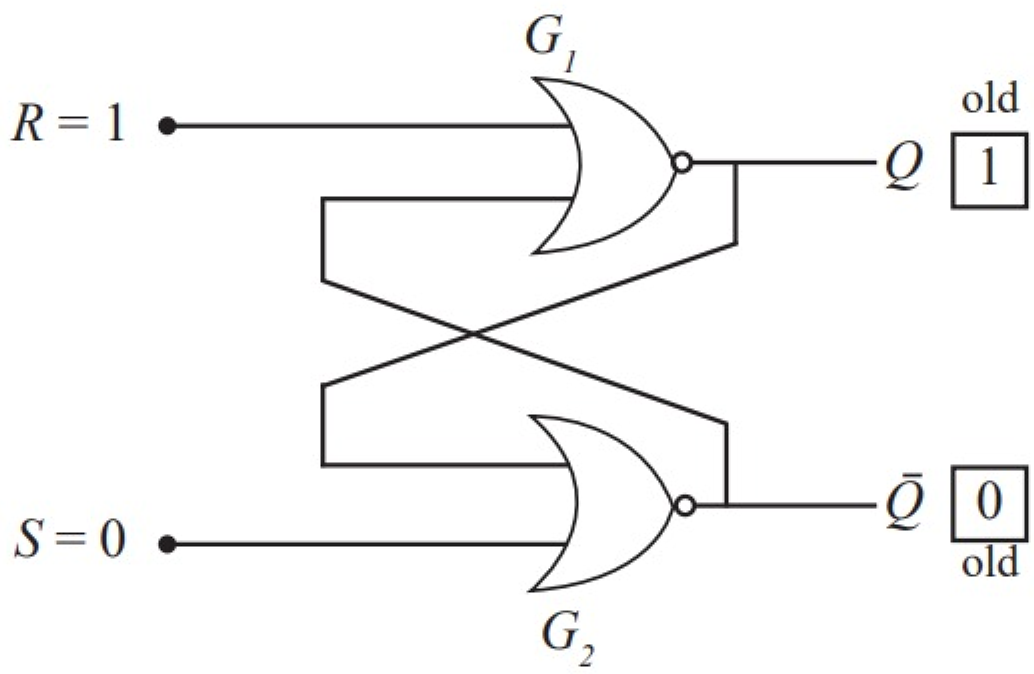


Figure 4.30

Consider the circuit shown in figure 4.30. One input of the G_1 - NOR gate is 1 (as $R = 1$). The other input is 0 (as $\bar{Q}_{old} = 0$). Therefore, the new output of G_1 is $Q_{new} = 0$.

Now one input of the G_2 - NOR gate is 0 (as S=0) and the other output is also 0 (as $Q_{new} = 0$). Therefore the new output of G_2 is $\bar{Q}_{new} = 1$. So, the new stable outputs of the flip-flop are $Q_{new} = 0$ and $\bar{Q}_{new} = 1$ (under the condition $\bar{Q}_{new} = 1$ also, $Q_{new} = 0$ condition holds).

Even in the case where the previous outputs of the flip-flop are $Q_{old} = 0$ and $\bar{Q}_{old} = 1$, it can be shown that the new outputs are steading at $Q_{new} = 0$ and $\bar{Q}_{new} = 1$.

This result can be given by a simple table as below.

S	R	Q_{n+1}	$\bar{Q}_{n+1}$
0	1	0	1

Irrespective of the previous outputs, the new outputs Q_{n+1} and $\bar{Q}_{n+1}$ are 0 and 1 respectively. **This is the reset state of the outputs.**

Accordingly, when the inputs to the flip-flop are applied as $S = 0$ and $R = 1$, irrespective of the previous outputs, the flip-flop comes to the reset state. The reset state of the flip-flop is the state where its new outputs are $Q_{new} = 0$ and $\bar{Q}_{new} = 1$.

4. Applying inputs as S = 1 and R = 1 (Invalid state)

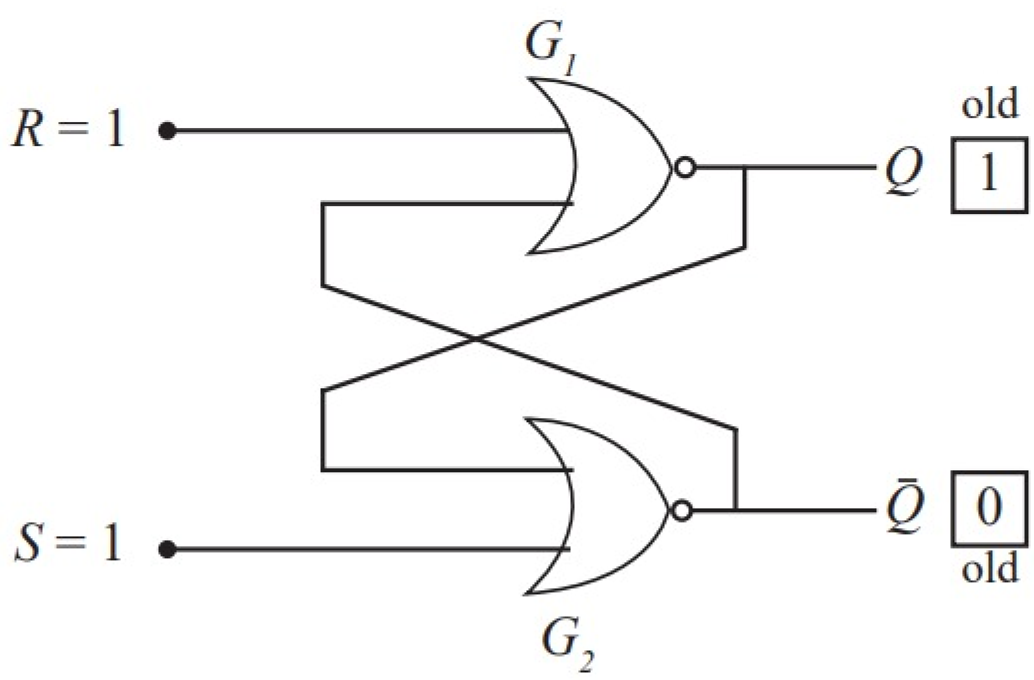


Figure 4.31

Consider the circuit shown in Figure 4.31. One input of the G_1 - NOR gate is 1 (as $R = 1$) and the other input is 0 (as $Q_{old} = 0$). Therefore the new output of G_1 is $Q_{new} = 0$. Now, one input of the G_2 - NOR gate is 1 (as $S = 1$) and the other input is 0 (as $Q_{new} = 0$). Therefore, the new output of G_2 is $\bar{Q}_{new} = 0$. It can be shown that, even in the case where previous outputs of the flip-flop are $Q_{old} = 0$ or $Q_{old} = 1$, the new outputs appear as $Q_{new} = 0$ and $\bar{Q}_{new} = 0$.

There is a major problem in this state. That is, it does not satisfy the essential condition of the flip-flop that the outputs should be with inversion to each other. The condition $Q_{new} = 0$ and $\bar{Q}_{new} = 0$ is contrary to this. Therefore, the input state $S = 1$ and $R = 1$ which is leading to this unwanted situation, is never applied to the flip-flop. Therefore, this input state is known as the **"invalid state"**.

Using the four simple tables presented on the occasions 1,2,3 and 4 above, the truth table for the *SR* flip-flop can be given as below.

Inputs		Outputs		Prevailing state
S	R	Q_{n+1}	$\bar{Q}_{n+1}$	
0	0	Q_n	$\bar{Q}_n$	Unchanged
0	1	1	0	Set
1	0	0	1	Reset
1	1	-	-	Invalid

Where Q_{n+1} and $\bar{Q}_{n+1}$ are the **new outputs** given by the flip-flop when the relevant inputs are applied. Q_n and $\bar{Q}_n$ are the **previous outputs** which were there in the previous state before appearing the new outputs.

4.16.2 Making a flip-flop using NAND gates.

In place of NOR gates, NAND gates can be used to make a flip-flop. However in that case the inputs applied to S and R should be inverted using two NOT gates, before they proceed. Such a flip-flop can be made as shown in the logic gate circuit given in figure 4.32.

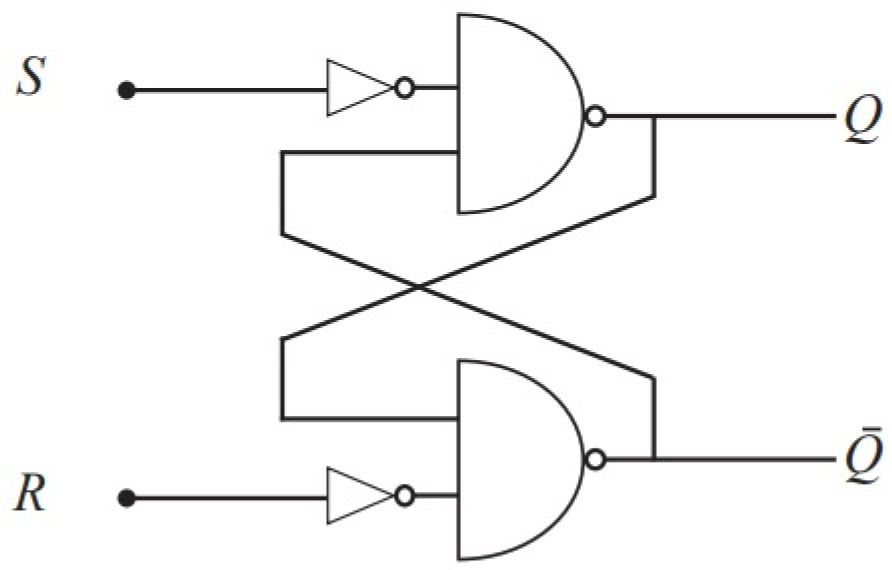


Figure 4.32

The operation of this is same as that of the flip-flop made using NOR gates. Try whether you can obtain the outputs of this flip-flop by applying appropriate inputs (except the invalid inputs $S = 1$ and $R = 1$).

To understand the operation of the flip-flop further more, we can use the circuit given in figure 4.33.

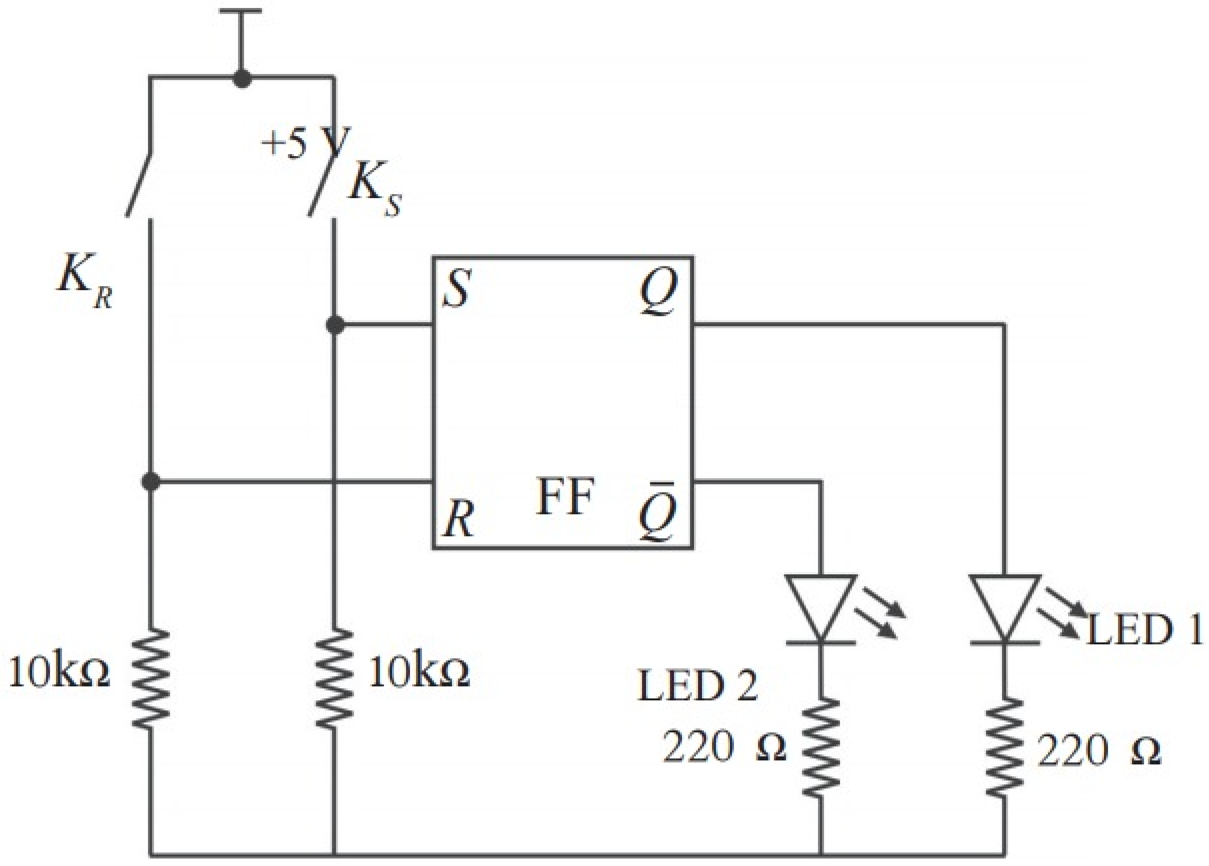


Figure 4.33

The signals supplied to the flip-flop should be definite. The inputs should not be kept open (floating). When the keys K_S and K_R are open (OFF) the S and R inputs get zero potential via 10 kΩ resistor. That is, they are at $S = 0$ and $R = 0$ definitely. When a key K_S or K_R is closed (ON) then the relevant S or R input gets voltage +5 V. That is, the input can be applied definitely as $S = 1$ or $R = 1$. Therefore the S and R inputs can be made definitely as follows.

When,

K_S is open (OFF), $S = 0$

K_S is closed (ON), $S = 1$

Similarly when,

K_R is open (OFF), $R = 0$

K_R is closed (ON), $R = 1$

It is important to take care not to close (ON) the two keys K_S and K_R at the same time. Otherwise the invalid state $S = 1$ and $R = 1$ arises.

When the flip-flop shown in this circuit gives output as $Q = 1$, the LED 1 lights and when $Q = 0$ it does not light. Also, when $\bar{Q} = 1$, the LED 2 lights and when $\bar{Q} = 0$ it does not light.

By applying appropriate inputs for S and R using keys K_S and K_R, observe the lighting condition of LED 1 and LED 2. Then the output levels of Q and $\bar{Q}$ can be decided.

Considering the above facts, the following table can be prepared for you to understand the operation of the flip-flop better.

© 2020 National Institute of Education. All Rights Reserved.

K_S	K_R	S	R	LED 1	LED 2	Q_n	$\bar{Q}_n$	
						1	0	Previous state
OFF	OFF	0	0	lights	doesn't light	1	0	Unchanged state
OFF	ON	0	1	doesn't light	lights	0	1	Reset state
ON	OFF	1	0	lights	doesn't light	1	0	Set state
OFF	OFF	0	0	lights	doesn't light	1	0	Unchanged state
ON	OFF	1	0	lights	doesn't light	1	0	Set state
OFF	ON	0	1	doesn't light	lights	0	1	Reset state
OFF	OFF	0	0	doesn't light	lights	0	1	Unchanged state
OFF	ON	0	1	doesn't light	lights	0	1	Reset state

Worked example

After a careful study of the above table, complete the following table. In that table the relevant Q and $\bar{Q}$ outputs can be given according to the change of inputs applied to the *S*-*R* flip-flop.

	S	R	Q	$\bar{Q}$	
1			1	0	← Previous outputs
2	0	0			
3	0	1			
4	0	0			
5	1	0			
6	0	0			
7	0	1			
8	0	0			
9	0	1			
10	1	0			

Answer

In the table,

The row (1) gives the previous output levels as $Q = 1$ and $\bar{Q} = 0$.

Consider the row (2)

There, $S = 0$ and $R = 0$. Therefore it is the **unchanged state**. So, the previous outputs remain unchanged. $\therefore \quad Q = 1, \ \bar{Q} = 0$

Consider the row (3)

There, $S = 0$ and $R = 1$. Therefore it is the **reset state.** $\therefore \quad Q = 0, \ \bar{Q} = 1$

Consider the row (4)

There, $S = 0$ and $R = 0$. Therefore it is the **unchanged state.** So, the previous outputs (in the row (3)) remain unchanged. $\therefore \quad Q = 0, \ \bar{Q} = 1$

Consider the row (5)

There, $S = 1$ and $R = 0$. This is the **set state**. $\therefore \quad Q = 1, \bar{Q} = 0$

Consider the row (6)

There, $S = 0$ and R=0. This is the **unchanged state**. Therefore , the previous outputs (in the row (5)) remain unchanged. $\therefore \quad Q = 1, \bar{Q} = 0$

Consider the row (7)

There, $S = 0$ and $R = 1$. This is the **reset state.** $\therefore \quad Q = 0, \bar{Q} = 1$

Consider the row (8)

There, $S = 0$ and $R = 0$. This is the **unchanged state**. Therefore, the previous outputs (in the row (7)) remain unchanged. $\therefore \quad Q = 0, \bar{Q} = 1$

Consider the row (9)

There, $S = 0$ and $R = 1$. This is the **reset state**. $\therefore \quad Q = 0, \bar{Q} = 1$

Consider the row (10)

There, $S = 1$ and $R = 0$. This is the **set state**. Therefore, the previous outputs (in the row (5)) remain unchanged. $\therefore \quad Q = 1, \bar{Q} = 0$

Now considering the relevant Q and $\bar{Q}$ values obtained on each occasion, the table can be completed as given below.

© 2020 National Institute of Education. All Rights Reserved.

	S	R	Q	$\bar{Q}$	
1)			1	0	← Previous outputs
2)	0	0	1	0	← Previous outputs unchanged
3)	0	1	0	1	← Reset
4)	0	0	0	1	← Previous outputs unchanged
5)	1	0	1	0	← Set
6)	0	0	1	0	← Previous outputs unchanged
7)	0	1	0	1	← Reset
8)	0	0	0	1	← Previous outputs unchanged
9)	0	1	0	1	← Reset
10)	1	0	1	0	← Set

4.16.3 Timing diagrams

Timing diagrams are used to represent the logic levels of the inputs and the relevant outputs when the valid signals are applied to the *S* - *R* flip-flop.

In the timing diagram the equal time intervals are concerned. In each time interval, the relevant *S* and *R* input levels and the *Q* output level which is obtained according to the present inputs and the previous output, are shown in the timing diagram.

To understand how a timing diagram can be drawn, let us represent the relevant inputs *S* and *R* and the output *Q* in the table we completed above, in a timing diagram as shown in Figure 4.34.

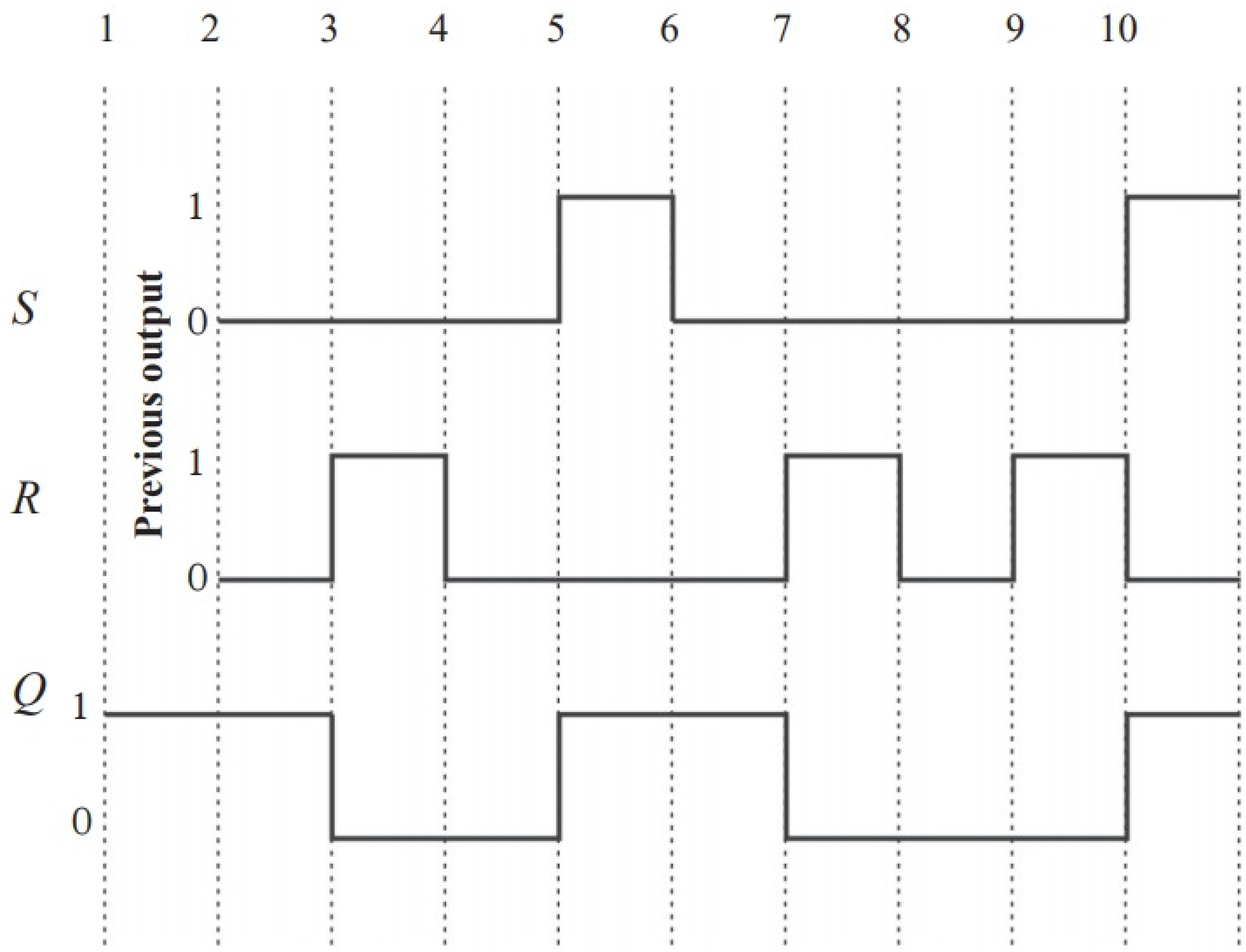


Figure 4.34

The way, how the output *Q* is obtained in each interval is as follows.

Interval	Output Q
1	Previous output of Q is given. $\therefore\ Q = 1$
2	$S = 0, R = 0$ This is the unchanged state. $\therefore\ Q = 1$
3	$S = 0, R = 1$ This is the reset state. $\therefore\ Q = 0$
4	$S = 0, R = 0$ This is the unchanged state. $\therefore\ Q = 0$
5	$S = 1, R = 0$ This is the set state. $\therefore\ Q = 1$
6	$S = 0, R = 0$ This is the unchanged state. $\therefore\ Q = 1$
7	$S = 0, R = 1$ This is the reset state. $\therefore\ Q = 0$
8	$S = 0, R = 0$ This is the unchanged state. $\therefore\ Q = 0$
9	$S = 0, R = 1$ This is the reset state. $\therefore\ Q = 0$
10	$S = 1, R = 0$ This is the set state. $\therefore\ Q = 1$

If you are given a timing diagram showing the operation of a flip-flop, you can easily read the input and output logic levels relevant to each internal. In a timing diagram only the Q output of the S - R flip-flop is shown usually. However if it is required, the logic levels of $\bar{Q}$ can be determined using Q, and then it can also be drawn in the timing diagram as shown in figure 4.35.

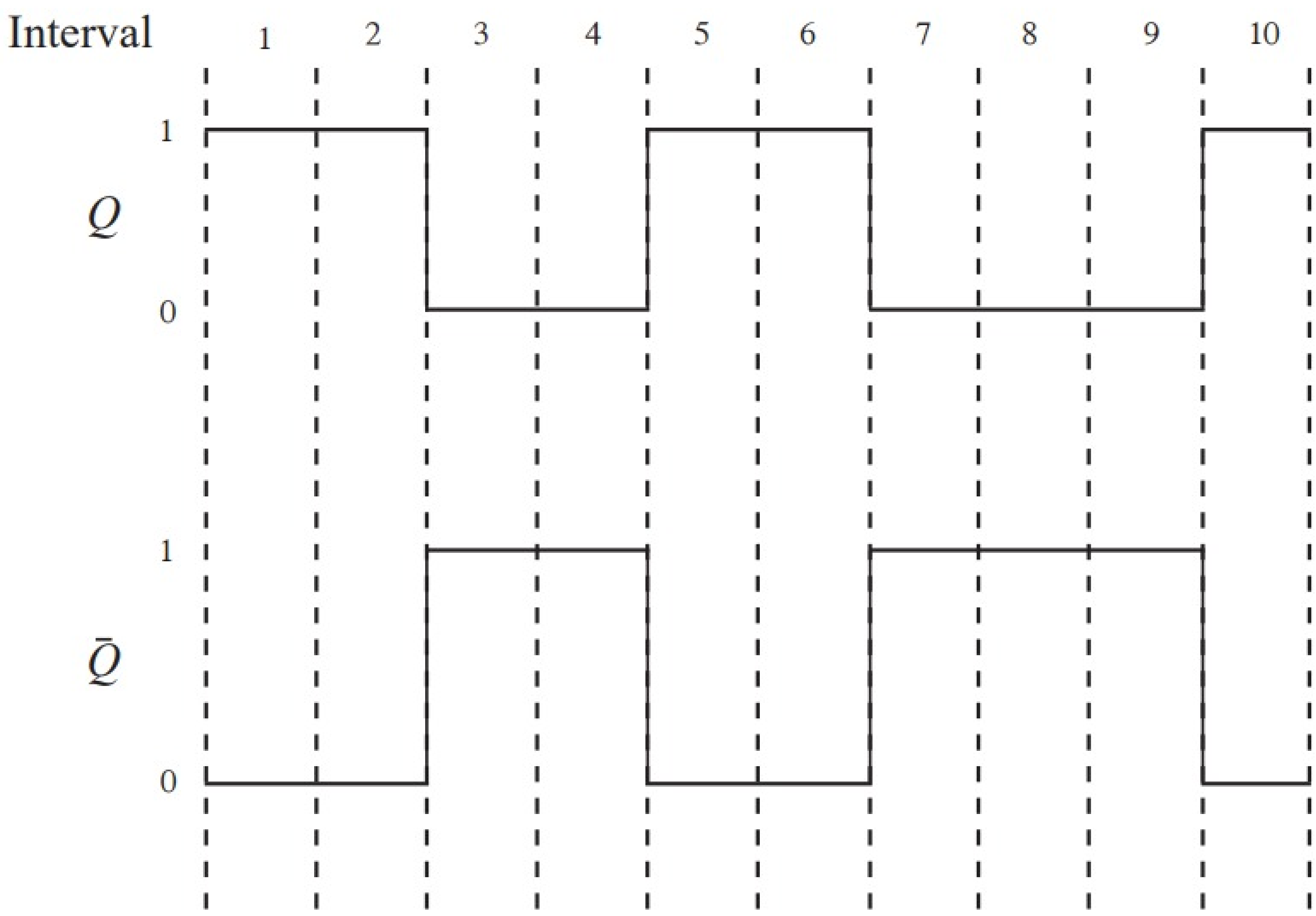


Figure 4.35

4.16.4 *S* - *R* flip-flop as a basic memory element

The electronic memory is used for the operation of many electronic devices such as computers and mobile phones. The electronic memory is an electronic circuit, made using a large number of flip-flops which are functioning as memory elements or memory cells. To know, how an *S-R* flip-flop is used in such a circuit, let us first discuss the memory unit.

4.17 Memory Units

In an electronic memory a place where one binary digit (0 or 1) can be stored, is called a memory cell or memory element. So, in a memory cell, a single binary digit (a bit) can be stored. The term "bits" is used as a shortened form for binary digits. That is, **bi***nary digi***ts** → bits.

The memory unit is a combination of eight memory cells. Such a combination can store eight bits and it is called a **"byte"**. In general a byte is represented as follows.

1	2	3	4	5	6	7	8

In each memory cell from 1 to 8, a binary digit 1 or 0 can be stored.

According to what has been described above, it can be understood how a memory unit of this kind can be made using eight *S - R* flip-flops. There, one flip-flop is used to store one bit (0 or 1). Using 8 such flip-flops a byte (8 bits) can be stored.

Let us now assign binary digits arbitrarily to each place in the byte (memory unit).

1	0	0	1	0	1	1	0

It can be considered that, the 8 bits represent the binary number 10010110. In that number, the bit having the highest positional value is at the left end, and here it is 1. The digit which has the highest positional value is known as the **Most Significant Digit.** It is written as **MSB**.

The bit having the lowest positional value in the byte is at the right end, and here it is 0. This bit is known as the **Least Significant Bit**. It is written as **LSB**.

Eg.

(i)

0	1	1	1	0	0	1	1

↑ MSB = 0
Digit at the left end = 0
∴ MSB = 0

↑ LSB = 1
Digit at the right end = 1
∴ LSB = 1

(ii)

0	0	1	0	1	1	1	0

MSB = 0
Digit at the left end = 0
∴ MSB = 0

LSB = 0
Digit at the right end = 0
∴ LSB = 0

(iii)

1	0	1	1	1	0	0	1

MSB = 0
Digit at the left end = 1
∴ MSB = 1

LSB = 0
Digit at the right end = 1
∴ LSB = 1

A memory circuit capable of storing two bits in its memory can be designed using two *S* - *R* flip-flops as shown in figure 4.36.

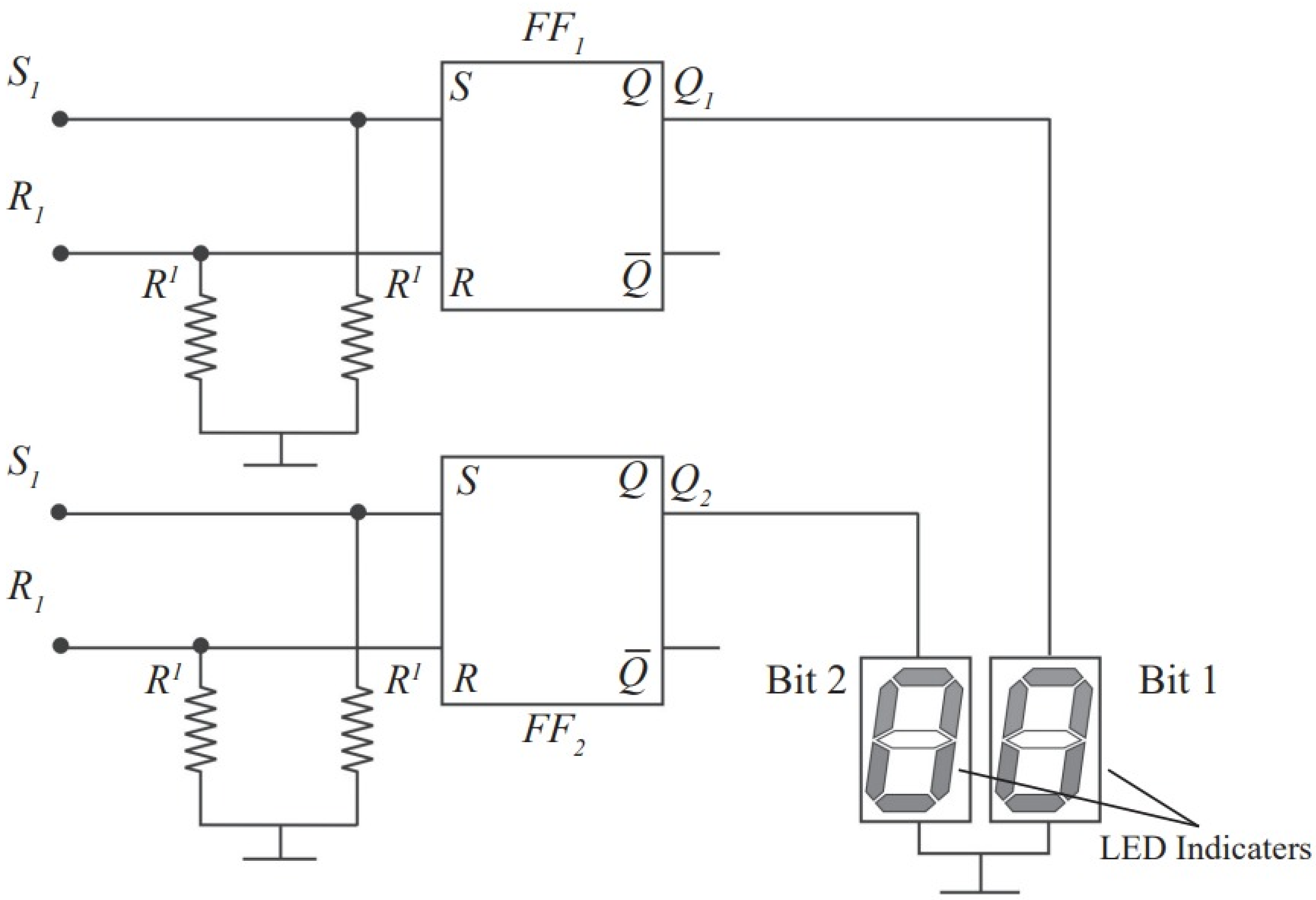


Figure 4.36

When no external input signal is applied to the *S* - *R* flip-flops FF_1 and FF_2 in this circuit, each flip-flop gets its inputs as $S = 0$ and $R = 0$ because the ground voltage (zero volts) is applied on both *S* and *R* inputs via the resistors *R*' (about 10 kΩ).

Now, let us apply $S_1=0$ and $R_1=1$ for the inputs of FF_1. Then the FF_1 is **reset** and output becomes $Q_1 = 0$. That displays 0 on the display **Bit 1**. Even though the input logic levels are removed *R*' resistors make the condition $S_1 = 0$ and $R_1 = 0$, which is the **unchanged** state. Therefore the output is unchanged and the output $Q = 0$ is still there. It can be said that the flip-flop has memorized its output $Q_1 = 0$. It is because of this reason that even when the input signals are removed the relevant output is there without change.

Therefore unless external signals are applied and S_1 and R_1 are changed, the "0" appeared on the display **Bit 1** will remain as it is.

Similarly by applying suitable input signals to S_2 and R_2 inputs of the FF_2 flip - flop its output can be made 0 or 1. The bit relevant to that output is displayed on the display **Bit 2**. Now even though the inputs are removed the bit on the display **Bit 2** is unchanged as FF_2 has memorized it. When it is required, the Q_2 output can be changed by applying appropriate logic levels to S_2 and R_2 inputs. Until then the bit that appeared on the display **Bit 2** remains unchanged. That is, it can be considered as the flip-flop has kept it in its memory.

So, in this manner it can be said that the above simple memory circuit consisting of two S - R flip-flops, is able to keep any two bit binary number (00,01,10 or 11) in its memory.

References

1. Duncan, T. (1997). *Success in Electronics - Second Edition*. Trans-Atlantic Publications, Hodder Education, UK.

2. Bandara, K. (2017). *Electronics - Second Edition*, Printer Related Express Service Supplies (Press), Kandy.

3. Eggleston, D. L. (2011). *Basic Electronics for Scientists and Engineers*. Occidental College, Los Angeles, USA.

4. Tooley, M. (2006). *Electronic Circuits Fundamentals and Applications*. Newnes, Kingston University, UK.

5. Bhargava, N.N., Gupta, S. C., & Kulshreshtha, D. C., (1983). *Basic Electronics and Linear Circuits*. Tata McGraw-Hill Education, New York, USA.