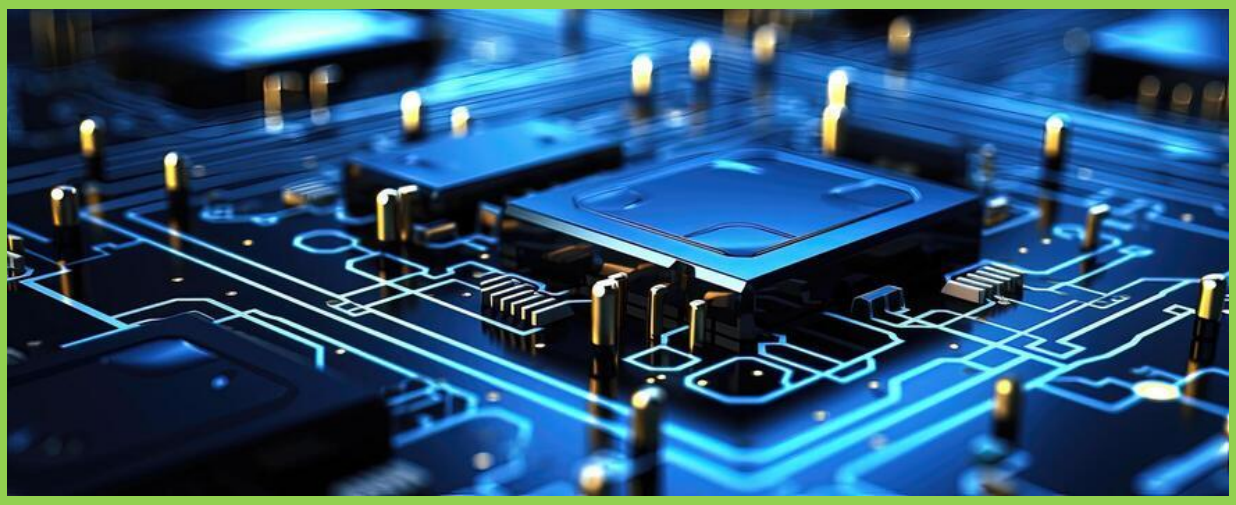
PHY (N)-350 & PHY (N)-350 L

B. Sc. VI Semester

BASIC ELECTRONICS
&
LABORATORY COURSE



DEPARTMENT OF PHYSICS
SCHOOL OF SCIENCES
UTTARAKHAND OPEN UNIVERSITY
HALDWANI

Board of Studies and Programme Coordinator

Board of Studies

Department of Physics (School of Sciences)

Unit Writing and Editing

Editing

Writing

Govt. MB PG College, Haldwani
Nainital, Uttarakhand
Dr. Meenakshi Rana
Department of Physics
School of Sciences, Uttarakhand Open University
Dr. Rajesh Mathpal
Department of Physics
School of Sciences, Uttarakhand Open University

Unit writing and Editing (LABORATORY COURSE)

Editing

Prof. Kamal Devlal
Professor, Department of Physics
School of Sciences,
Uttarakhand Open University

Dr. Meenakshi Rana
Department of Physics
Uttarakhand Open University,
Haldwani

Writing

Dr. Vijendra Lingwal
Department of Physics
Pt. L.M.S. Government PG College Rishikesh
Dehradun, Uttarakhand, India.

Dr. D K Upreti
Department of Physics,
Govt PG College, Ranikhet
Almora

Course Title and Code	**:**	**Basic Electronics (PHY (N)-350) & Laboratory Course (PHY (N)-350L)**
ISBN	**:**	
Copyright	**:**	**Uttarakhand Open University**
Edition	**:**	**2025**
Published By	**:**	**Uttarakhand Open University, Haldwani, Nainital- 263139**
Printed By	**:**	**htpp://uou.ac.in**

PHY (N)-350 & PHY (N)-350 L

BASIC ELECTRONICS
&
LABORATORY COURSE

DEPARTMENT OF PHYSICS
SCHOOL OF SCIENCES
UTTARAKHAND OPEN UNIVERSITY
Phone No. 05946-261122, 261123
Toll free No. 18001804025
Fax No. 05946-264232, E. mail info@uou.ac.in
htpp://uou.ac.in

Contents

Course: Basic Electronics **Course code: PHY (N)-350**

Credit: 3

Course 1: Laboratory Course

Course code: PHY (N)-350L
Credit: 1

UNIT 1 NETWORK THEOREMS

Structure

1.1 INTRODUCTION

A system of interconnected electronic components or circuits is called network. Sometimes solving for currents and voltages in multi-loop network can be quite complicated. We may apply the voltage law and current law, but using them it may lead to long systems of equations which will be difficult to solve. These difficulties can be overcome using certain network theorems which help us with network analysis. These theorems may be applied in particular circumstances. In this chapter, we will study the four network theorems i.e. Superposition theorem, Thevenin's theorem, Norton's theorem and Maximum Power Transfer theorem.

1.2 OBJECTIVES

- Learning the superposition theorem and applying it in analysis of network having two or more that two sources.
- Learning the Thevenin's theorem and applying it to simplify the resistive network into a simple circuit having a voltage source in series with network resistance and load.
- Learning the Norton's theorem and applying it to simplify the resistive network into a simple circuit having a current source in parallel with network resistance and load.
- Learning the maximum power transfer theorem and applying it to solve problems involving variations in load for a given source.

1.3 NETWORK THEOREMS

In electronics, many times we have combinations of various components like resistors, voltages sources, filters, inductors etc. Such combinations are called networks. Sometime, circuits are so large and difficult to solve easily. So there are some network theorems which can simplify the different complicated circuits. We will study now the four network theorems along with some practical examples.

1.4 THE SUPERPOSITION THEOREM

The superposition theorem can be stated as under:

In a linear network containing more than one source of voltage or current, the current through any branch is the algebraic sum of the currents produced by each source acting independently.

The statement simply means that the current flowing in a given branch of a circuit can be calculated for each voltage and current source acting alone, then add all the currents to get the actual current

in that branch. This theorem is applicable for linear networks only. In simple words, the linear circuit means the circuit for which the values of electronic components (like resistance, capacitance, inductance etc.) do not change with the change in voltage or current in the circuit.

Procedure for Applying the Superposition Theorem:

1. In a circuit containing many sources, select one source (voltage or current) and replace all other sources by their internal resistances.
2. If the internal resistance of sources is zero or not shown, after selecting one source, just replace the voltage source by short circuit and current source by open circuit.
3. Determine the current that flows through the desired branch due to the single voltage or current source. Find the current in that particular branch due to all sources.
4. Algebraically, sum the all obtained current to find the total current in the branch for original circuit.

Example 1.1: Use the superposition theorem to calculate the current I_3 in the circuit shown in Fig. 1.1(a). Given $E_1 = 6$ V & $E_2 = 12$ V

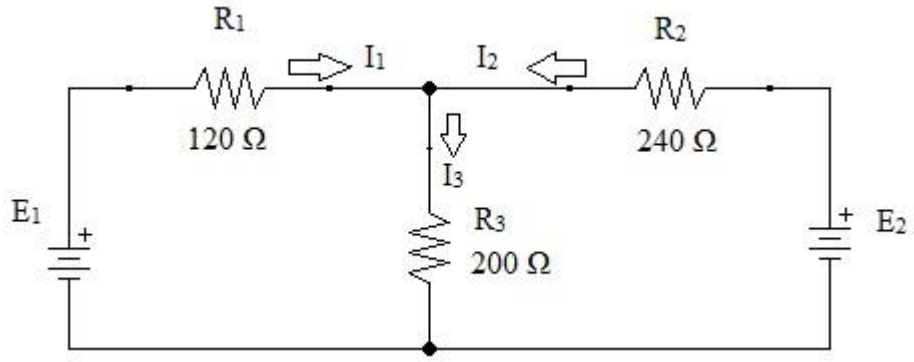


Figure 1.1(a)

Solution: In this circuit, first consider the E_1 source only and short circuit the other source E_2 [Fig. 1.1(b)]. Find the current in R_3 resistor due to E_1 only.

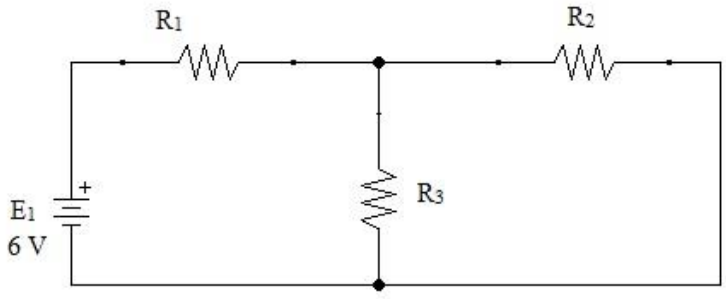


Figure 1.1(b)

In above circuit, the resistances R_2 and R_3 are in parallel and the combination is in series with R_1. So, the equivalent resistance is

$R = R_1 + R_2 \parallel R_3 = 120\ \Omega + 240\ \Omega \parallel 200\Omega = 229.09\ \Omega$

The total current due to E_1 source only = $E_1 / R = 26.19$ mA.

To find the current in R_3 (let say I_A) due to E_1, apply KVL in first loop,

So, $I_A = (6\ V - 26.19\ mA \times 120\ \Omega) / 200\Omega$

$I_A = 14.3$ mA (1)

Now, short circuit the first source E_1 [Fig. 1.1(c)] and find the current in R_3 due to E_2 source only.

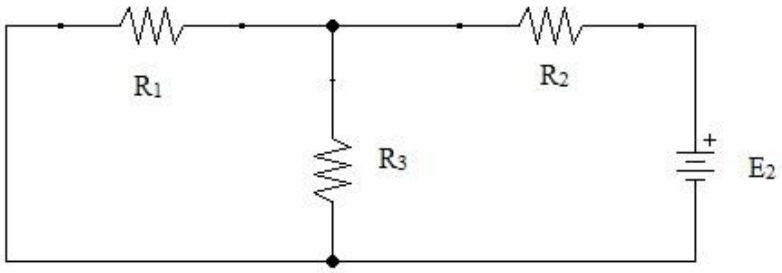


Figure 1.1(c)

Proceeding with the similar manner, the current in R_3 branch (let say I_B) due to source E_2,

$I_B = 14.3$ mA (2)

From equations (1) & (2), the total current I_3 for original circuit,

$I_3 = I_A + I_B = 14.3\ mA + 14.3\ mA = 28.6\ mA$

This is the required answer.

Now, for the second example, consider a circuit in which we have the voltage source and current source both.

Example 1.2: Determine the current in the 23 Ω resistor of Fig. 1.2(a) by applying the superposition theorem.

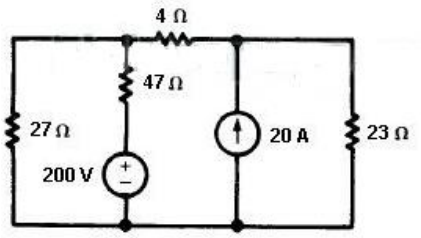


Figure 1.2(a)

Solution: First consider the voltage source only and the current source is replaced by open circuit [Fig. 1.2(b)],

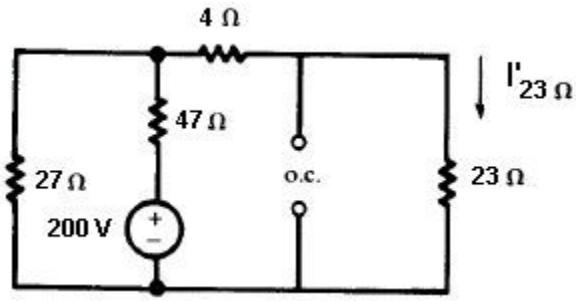


Figure 1.2(b)

You should have learnt by now that how to find the current in a branch due to voltage source only. For the circuit of Fig. 1.2(b), 4Ω and 23 Ω are in series. The 27 Ω resistor is in parallel with the combination of 4Ω & 23Ω resistances. In the last, 47Ω resistance is in series with the combination of 23Ω, 4Ω & 27Ω resistances.

So,

$$R_{eq} = 47 + \frac{(27)(4+23)}{54} = 60.5\ \Omega$$

Total current due to voltage source only = 200 / 60.5 = 3.31 A

The current I' flowing through 23Ω resistor (applying KVL to right loop),

$$I' = 1.65\ A \qquad (3)$$

To find the current in 23 Ω resistor due to current source, short circuit the voltage source [Fig. 1.2(c)].

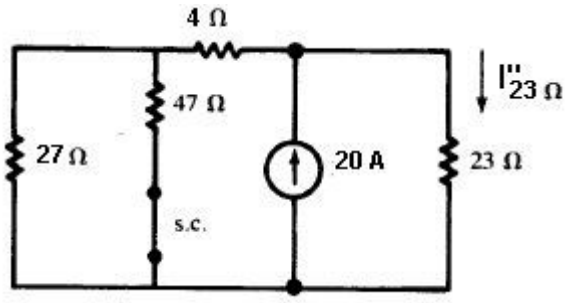


Figure 1.2 (c)

For the circuit left to current source, 27 Ω & 47Ω resistances are in parallel and their resultant is in series with 4Ω resistance.

So the equivalent resistance to the left of source= $4 + \frac{(27)(47)}{74} = 21.15\ \Omega$.

So the current through the 23Ω resistor due to current source

$$I'' = \left(\frac{21.15}{21.15+23}\right)\times 20 = 9.58\ A$$ [applying the current division, see appendix A]

Therefore, the total current through 23Ω resistor

$$I = I' + I'' = 11.23\ A$$

This is the required answer.

Now, you have understood that how to use superposition theorem in analyzing the circuit consisting of two or more voltage/current sources.

1.5 THEVENIN'S THEOREM

It stated as:

Any two-terminal network containing resistances and voltage sources and/or current sources may be replaced by a single voltage source in series with a single resistance.

Procedure for Applying Thevenin's Theorem:

1. Determine the Thevenin voltage V_{TH}. For this, open the load and find the terminal voltage from load end (output terminals).
2. Redraw the circuit with voltage source short and current source open.
3. Find the Thevenin resistance as seen from output terminals.

Example 1.3: Derive the Thevenin equivalent circuit for the network shown in Fig. 1.3(a). Also find the output voltage when R_L= 12 kΩ.

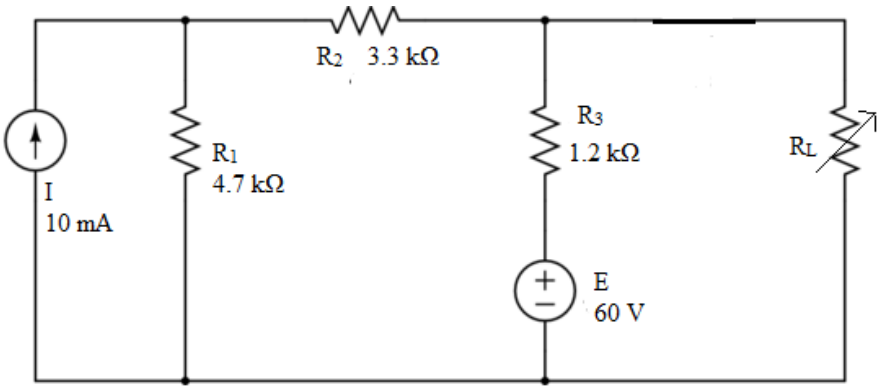


Figure 1.3(a)

Solution: To find the Thevenin voltage, disconnect the load R_L and find the terminal voltage V_{AB} from this end [Fig 1.3(b)].

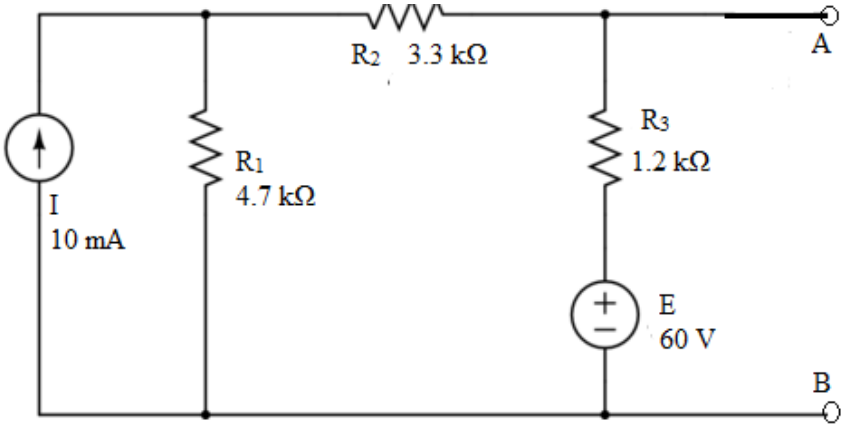


Figure 1.3(b)

The V_{AB} voltage will be the Thevenin voltage V_{TH}. To find this terminal voltage first finds the voltage across R_3 resistor. Now we will use the Superposition theorem to find the current in R_3. Short circuit the voltage source and find the current through R_3 resistor using current division rule.

So, current through R_3 due to current source = (4.7×10)/ (4.7+3.3+1.2) = 5.11 mA (in downward direction)

To find the current through R_3 due to voltage source, open circuit the current source. So the value of current due to voltage source = 60/ (4.7+3.3+1.2) = 6.52 mA (in upward direction).

Therefore, the total current through R_3 resistor = 6.52 mA – 5.11 mA = 1.41 mA.

The Thevenin voltage is now, $V_{AB} = V_{TH} = 60\ V - (1.41mA \times 1.2\ k\Omega) = 58.3\ V$

To find Thevenin resistance, disconnect the load, open the current source and short the voltage source [Fig. 1.3(c)].

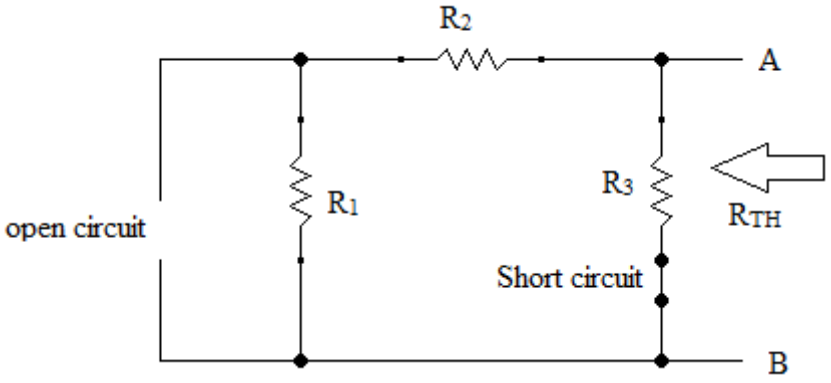


Figure 1.3(c)

Then find the equivalent resistance from load end.

Therefore, $R_{TH} = (R_1 + R_2) \| R_3 = (3.3\ k\Omega + 4.7\ k\Omega) \| 1.2\ k\Omega = 1.04\ k\Omega$

Thevenin equivalent circuit is now as shown in Fig. 1.3(d).

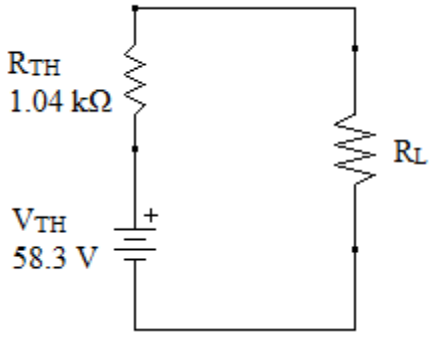


Figure 1.3(d)

The output voltage for $R_L = 12\ k\Omega$ can be calculated as [using Fig. 1.3(d)]

$$V_L = V_{TH}\, R_L / (R_L + R_{TH})$$

$$= (58.3 \times 12)/ (12+1.04) = 53.65\ V$$

So, this is the full procedure to simplify the circuits with Thevenin theorem.

1.6 NORTON'S THEOREM

The theorem states:

Any two terminal network containing resistances and voltage sources and/or current sources may be replaced by a single current source in parallel with a single resistance.

If you have learned that how to analyze the circuit using Thevenin theorem then applying this, Norton theorem is easy since the outcomes of Thevenin theorem can directly be used here.

Procedure for Applying Norton's Theorem (Indirect Method):

Now we will learn the indirect method to analyze the circuit with Norton's theorem.

1. Analyze the circuit as mentioned in Thevenin theorem to find Thevenin resistance R_{TH} and Thevenin voltage V_{TH}.
2. Norton's resistance is equal to Thevenin resistance i.e. $R_N = R_{TH}$.
3. Find the Norton current (I_N) using formula $I_N = V_{TH}/ R_N$
4. Draw the equivalent circuit with current source I_N connected in parallel with both the Norton's resistance R_N and load resistance R_L.

Example 1.4: Let us take the Example 1.3 again.

Now we know after analyzing in Thevenin way that R_{TH}= 1.04 kΩ and V_{TH} = 58.3 V,

Therefore, Norton resistance $R_N = R_{TH}$ = 1.04 kΩ

and Norton current $I_N = V_{TH}/ R_N$= 58.3/1.04 ~ 56 mA

Norton's equivalent circuit:

Using the values calculated above, the circuit may be redrawn as shown in Fig. 1.4.

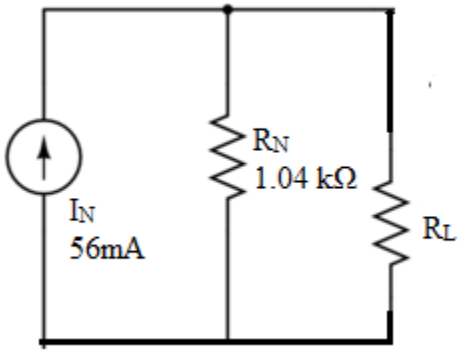


Figure 1.4

Procedure for Applying Norton's Theorem (direct Method):

1. Find the short-circuit current I_N at the network terminals. That means disconnect the load short the terminals.
2. Short circuit the voltage source and open the current source.
3. Find the Norton's resistance R_N in exactly the same way as used to find the Thevenin resistance R_N.

Example 1.5: Apply Norton's theorem to analyze the circuit of Fig. 1.5(a) to determine the load current for 5.6kΩ load resistance.

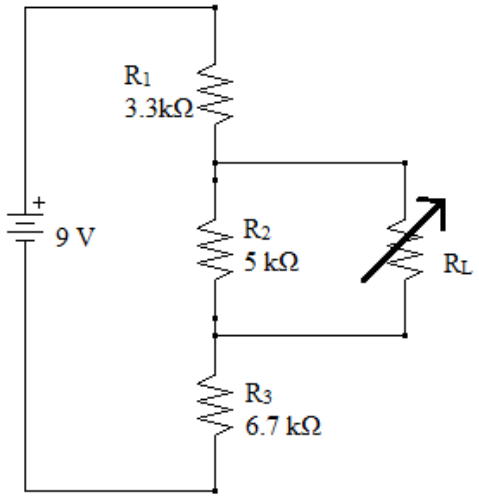


Figure 1.5(a)

Solution: To find the Norton current I_N, disconnect the load and short circuit the terminal as shown in Fig. 1.5(b).

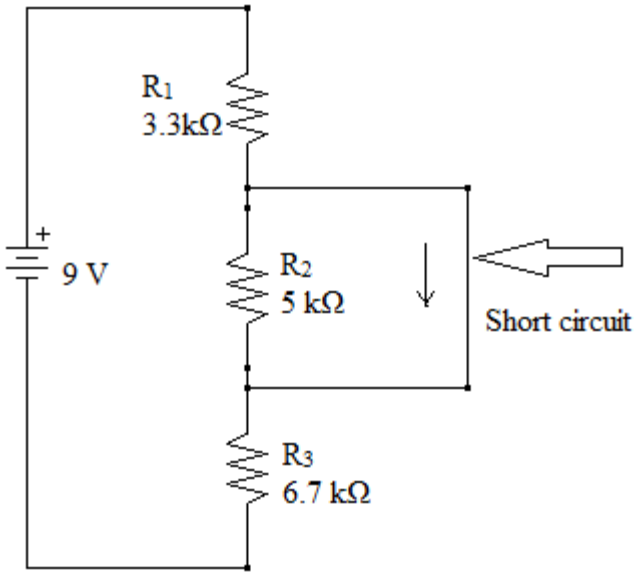


Figure 1.5(b)

The short circuit current I_N can be obtained by applying KVL to short circuited loop,

$$I_N = 9/ (3.3+6.7) = 900\ \mu A$$

The Norton resistance is $R_N = R_2 \| (R_1 + R_3)$ [in similar way as Thevenin resistance]

$$= 3.3\ k\Omega$$

The Norton equivalent circuit may be drawn as shown in Fig. 1.5(c).

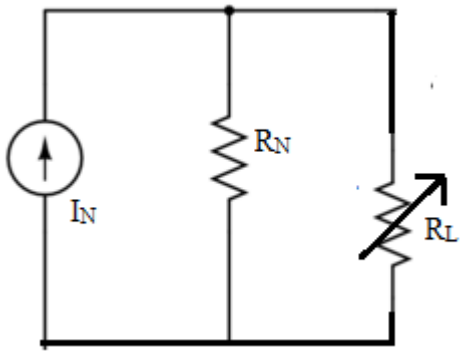


Figure 1.5(c)

Using this circuit, the current through load may be find as

$I_L = I_N R_N / (R_L + R_N) = 333\ \mu A$ [using current division]

1.7 MAXIMUM POWER TRANFER THEOREM

Maximum power is obtained from a network or source when the load resistance is equal to the resistance of the network or source as seen from load end.

We know that any network of sources and resistances may be reduced to a single voltage source in series with a resistance by applying Thevenin theorem. Consider the similar type of circuit as shown in Fig. 1.6.

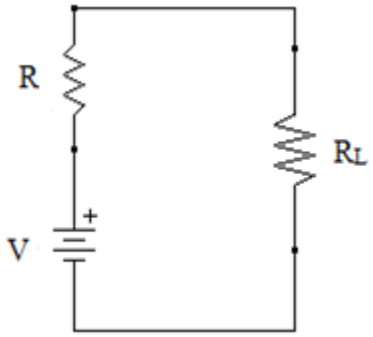


Figure 1.6

Here, R is the network resistance or internal resistance of the source

V is the voltage source or Thevenin equivalent of voltage source

R_L is the load resistance

The Maximum Power Transfer theorem says that maximum power will be transferred to load when load resistance R_L is equal to the network resistance or internal resistance of the source R.

We can prove this theorem in a simple manner.

From Fig. 1.6,

The current though the load $I = V/ (R+R_L)$, and so power absorbed by the load is

$$P_L = I^2R_L = \frac{V^2R_L}{(R+R_L)^2}$$

To determine the maximum power that is transferred to load, differentiate P_L with respect to R_L and put it equal to zero, i.e. $\frac{dP_L}{dR_L} = 0$

$$\frac{dP_L}{dR_L} = \frac{(R - R_L)V^2}{(R + R_L)^3} = 0$$

That gives $R = R_L$; and that is the exact statement of theorem.

Example 1.5: In the circuit of Fig. 1.7, determine the value of load resistance when the load draws maximum power. Also find the value of maximum power.

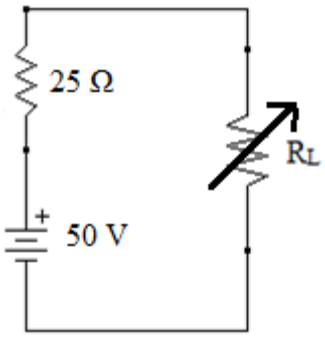


Figure 1.7

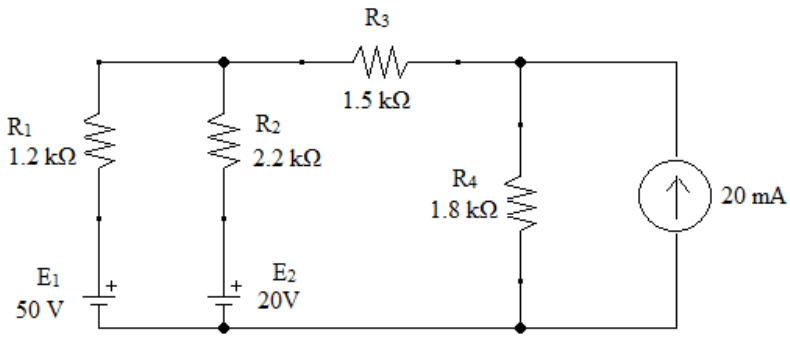
Solution: The source delivers the maximum power when load resistance is equal to source resistance.

So $R_L = 25\Omega$

The current flowing through load I = 50 / (25+25) = 1 A

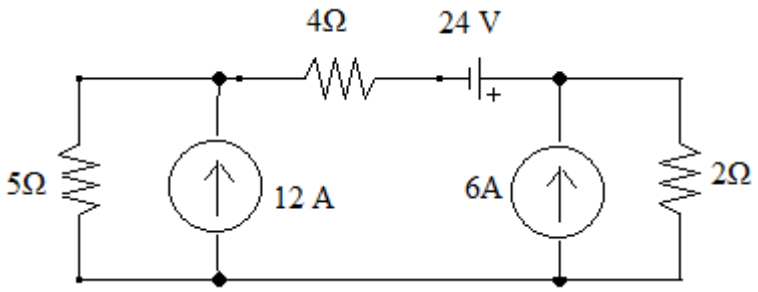
The maximum power delivered $P = I^2 R_L = 1 \times 25 = 25$ W

Self Assessment Question (SAQ) 1: For the circuit as shown below, determine the current through the resistor R_3.



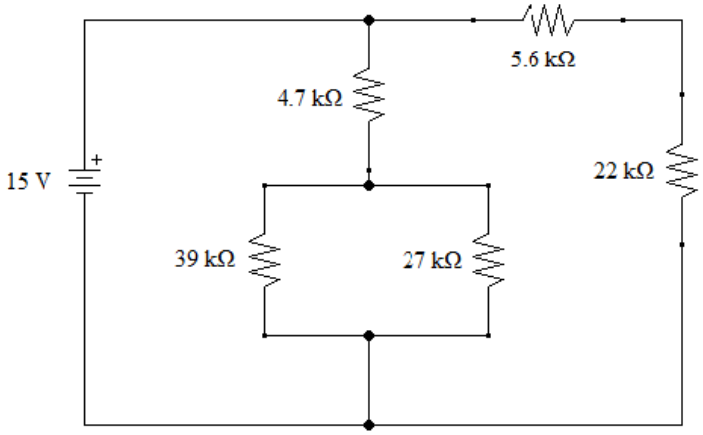
Ans: 0.84 mA

Self Assessment Question (SAQ) 2: Find the current through 5Ω resistor using superposition theorem.



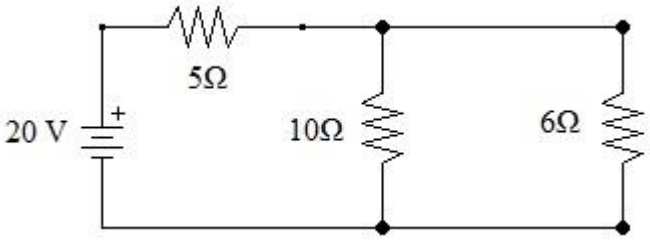
Ans:5.45 A

Self Assessment Question (SAQ) 3: For the circuit in Figure below, use Thevenin's theorem to determine the current through 22 kΩ resistor.



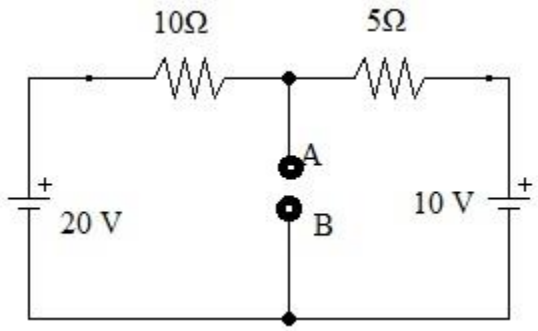
Ans: 543 μA

Self Assessment Question (SAQ) 4: Using Norton's theorem, determine the current through 6Ω resistance.



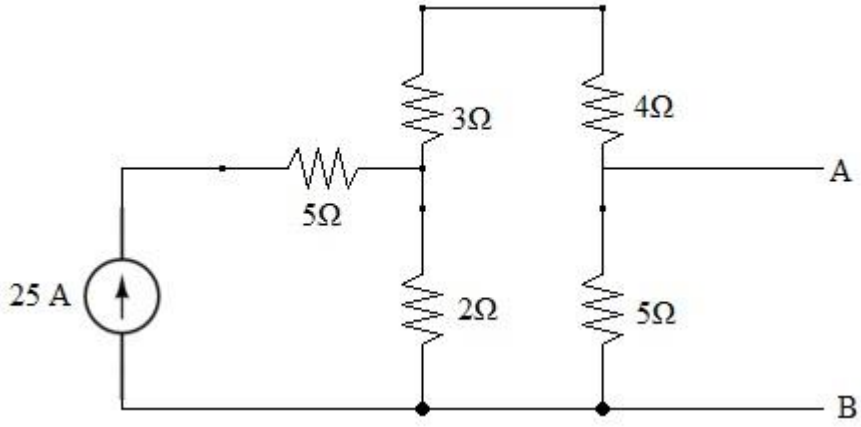
Ans: 1.43 A

Self Assessment Question (SAQ) 5: Determine Norton's equivalent circuit at terminals AB for the circuit shown below:



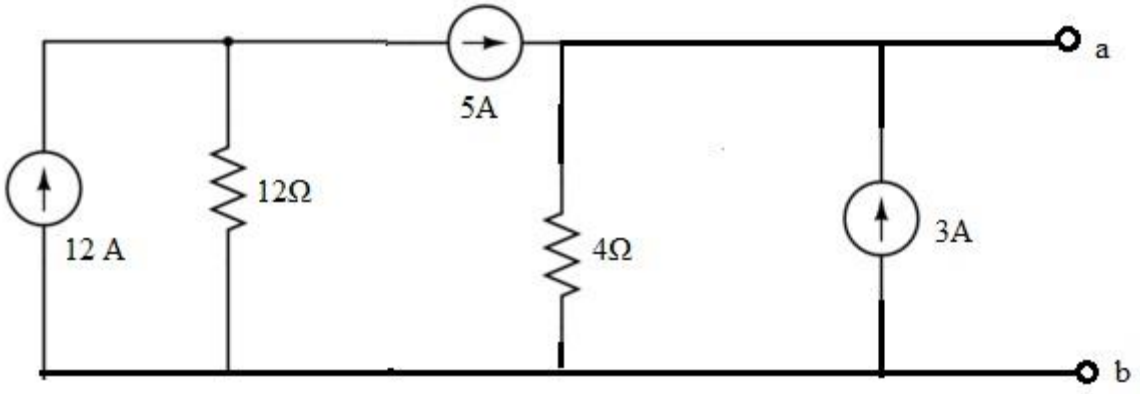
Ans: $I_N = 4$ A, $R_N = 3.33\ \Omega$

Self Assessment Question (SAQ) 6: Determine the Norton's equivalent circuit for the circuit given below:



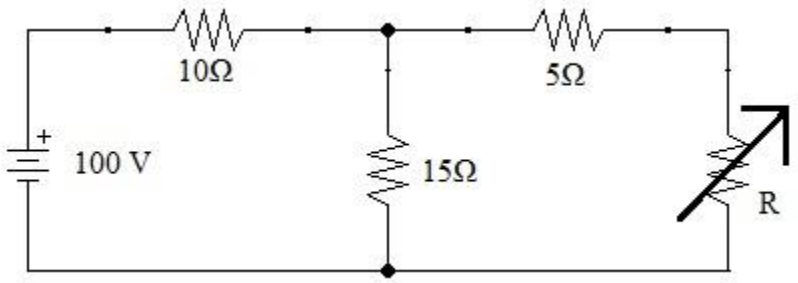
Ans: $R_{AB} = R_N = 3.21\Omega$; $I_N = 5.55$ A

Self Assessment Question (SAQ) 7: Determine the Thevenin and Norton equivalent circuits with respect to terminals ab for the circuit:



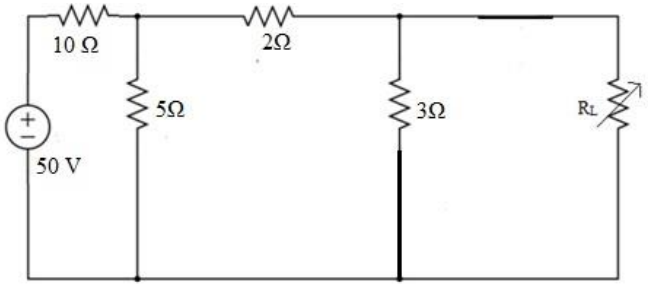
Ans: $V_{ab} = V_{TH} = 32$ V; $R_{ab} = R_{TH} = R_N = 4\Omega$; $I_{ab} = I_N = 8$A

Self Assessment Question (SAQ) 8: Find the value of the adjustable load resistance R which results in maximum power transfer to the load.



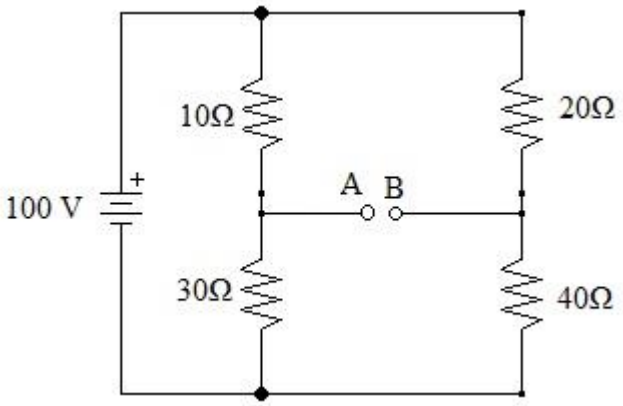
Ans: 11Ω

Self Assessment Question (SAQ) 9: Find the value of maximum power that is being delivered to the load in circuit below:



Ans: 4.67 W [Hint: Use the Thevenin theorem first to simplify the circuit and then use maximum power transfer theorem]

Self Assessment Question (SAQ) 10: Find the value of load resistance that receive maximum power from the source. Also find the maximum power delivered to the load.



Ans: R_L = 20.83Ω, Maximum Power = 0.833 W

[Hint: Use the Thevenin theorem first. The 10Ω resistance and 30Ω resistance are in parallel. Similarly, 20Ω & 40 Ω resistances are in parallel. Both combinations are in series then.]

1.8 SUMMARY

We have applied the four network theorems to analyze the circuits to find the desired current and/or voltage value. It is important to learn that how to apply these theorems. Complicated resistive networks are really hard to analyze without using theses theorems. All theorems have their unique applicability for different problems. You may have understood till now that in one theorem we may require the other theorem to apply. That is why, we have studied these theorems in a definite sequence. Like in applying Thevenin theorem, you may require the superposition theorem to apply first. Similarly, for Norton's theorem application, you may require superposition and Thevenin theorems both to apply first. To apply maximum power transfer theorem, you may need to apply all other three theorems first. Therefore, it is essential to learn that how to apply all the theorems, we discussed in text.

1.9 GLOSSARY

Resistance: Property of a substance due to which it opposes the flow of electricity through it.

Circuit: Combination of different electric components

Network: Combination of large number electric components or large circuit

Potentiometer: A 3-terminal variable resistor and is used to adjust the resistance in a circuit.

Rheostat: a variable resistor commonly used for handling higher currents and voltages

1.10 REFERENCES

1. Principles of electronics by V.K.Mehta

2. Basic Electronics by B.L.Thereja

3. Basic Electronics by D.C.Tayal

1.11 SUGGESTED READINGS

1. Electronic devices and circuits by Jacob Milliman and C.C. Halkais

2. The Feynman Lectures on Physics by Richard Feynman

3. Electronic devices and circuit theory by R.L. Boylestad and Louis N

1.12 TERMINAL QUESTIONS

1.12.1 Short answer type questions

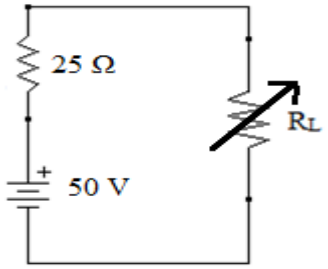
1. State superposition theorem.
2. What are linear and non-linear circuits?
3. State Thevenin's theorem drawing a Thevenin equivalent circuit.
4. State Norton's theorem drawing a Norton equivalent circuit.
5. State maximum power transfer theorem. Prove that the maximum power will be transferred to load when the load resistance is equal to source resistance.
6. Prove that the efficiency of power transfer when maximum transfer of power occurs is 50%.
7. State the Kirchhoff voltage law by taking suitable circuit diagram.
8. State the Kirchhoff current law with the help of suitable diagram.
9. What is voltage division? How is it achieved?
10. What is current division? How is it achieved?

1.12.2 Objective type questions

1. Superposition theorem is valid only for
   (a) Linear circuit (b) Non-linear circuits
   (c) Both linear and nonlinear circuits (d) Neither of the two
2. Thevenin equivalent circuit consist of
   (a) Voltage source in parallel with resistance
   (b) Voltage source in series with resistance
   (c) Current source in parallel with resistance
   (d) Current source in series with resistance
3. Thevenin resistance R_{TH} is found
   (a) by removing voltage sources along with their internal resistances
   (b) by short-circuiting the given two terminals
   (c) between any two 'open' terminals
   (d) between same open terminals as for V_{TH}
4. Norton equivalent circuit consist of
   (a) Voltage source in parallel with resistance
   (b) Voltage source in series with resistance
   (c) Current source in parallel with resistance
   (d) Current source in series with resistance
5. Maximum power is transferred when the load resistance is equal to
   (a) Source resistance (b) Half of the source resistance
   (c) Two-third of source resistance (d) zero
6. Superposition theorem is not applicable to circuits consisting of
   (a) Non-linear elements (b) Dependent current source

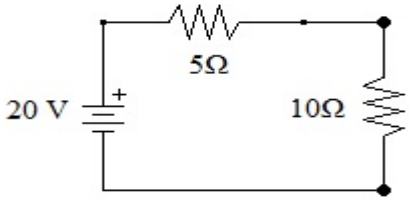
(c) Dependent voltage source (d) Linear elements

7. In the circuit shown in fig. below, what is the maximum power transferred to load?



(a) 5 W (b) 15 W
(c) 25 W (d) 50 W

8. For the circuit given below, 10Ω is the load resistance. The Norton's current I_{N9} through terminal of load is



(a) 1.33 A (b) 2 A
(c) 4 A (d) 5 A

9. Which one of these is a non-linear circuit component?
(a) Inductor (b) Condenser
(c) Resistor (d) Transistor

10. Efficiency of power transfer when maximum transfer of power occurs is
(a) 100 % (b) 80 %
(c) 75 % (d) 50 %

1.13 ANSWERS OF OBJECTIVE TYPE QUESTIONS

1.a, 2.b, 3.d ,4.c, 5.a, 6.a, 7.c, 8.c, 9.a, 10.d

Appendix I: Current Division Rule

Consider a parallel arrangement of resistors as shown in Fig. A1.

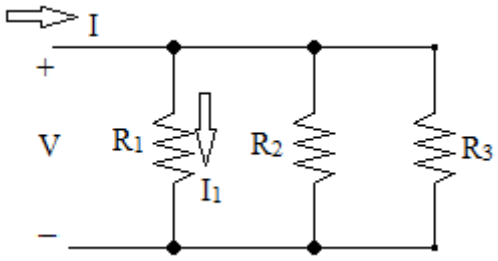


Fig. A1

In the circuit V is the voltage source, I is the total current and we are interested to find the current I_1 through the R_1 resistor.

The total current I will be equal to the sum of all branch currents (Kirchhoff current law),

So, $I = V/R_1 + V/R_2 + V/R_3$; here $I_1 = V/R_1$ and so on.

$$\frac{I_1}{I} = \frac{\frac{1}{R_1}}{\frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3}} = \frac{R_2R_3}{R_1R_2+R_1R_3+R_2R_3}$$

Therefore, remember this, when we want to find the current through any resistance in a parallel network, take the multiplication of rest of resistances in numerator and sum of all resistances multiplicative term in numerator. For each multiplicative term, take (n-1) resistances if there are total of n resistances. Take all combinations of resistances for multiplicative term and then finally sum all terms.

Now it is clear that if there are only two resistances R_1 and R_2, then the ratio of current I_1 to total current I may be given as

$$\frac{I_1}{I} = \frac{\frac{1}{R_1}}{\frac{1}{R_1}+\frac{1}{R_2}} = \frac{R_2}{R_1+R_2}$$

and

$$\boxed{I_1 = \frac{R_2}{R_1+R_2} \times I}$$

For two branch parallel circuit, we may say that the ratio of current in one branch to the total current is equal to the ratio of the resistance of other branch resistance to the sum of two resistances.

UNIT 2 NETWORK'S PASSIVE ELEMENTS AND FILTERS

Structure

2.1 INTRODUCTION

In this unit, all charges whether free or bound, were considered to be at rest. Charges in motion constitute an electric current. Such currents occur naturally in many situations. Lightning is one such phenomenon in which charges flow from the clouds to the earth through the atmosphere, sometimes with disastrous results. The flow of charges in lightning is not steady, but in our everyday life we see many devices where charges flow in a steady manner, like water flowing smoothly in a river. A torch and a cell-driven clock are examples of such devices. In the present unit, we shall study some of the basic laws concerning steady electric currents. An electric charge will experience a force if an electric field is applied. If it is free to move, it will thus move contributing to a current. In nature, free charged particles do exist like in upper strata of atmosphere called the ionosphere. However, in atoms and molecules, the negatively charged electrons and the positively charged nuclei are bound to each other and are thus not free to move. Bulk matter is made up of many molecules, a gram of water, for example, contains approximately 10^{22} molecules. These molecules are so closely packed that the electrons are no longer attached to individual nuclei. In some materials, the electrons will still be bound, i.e., they will not accelerate even if an electric field is applied. In other materials, notably metals, some of the electrons are practically free to move within the bulk material. These materials, generally called conductors, develop electric currents in them when an electric field is applied. If we consider solid conductors, then of course the atoms are tightly bound to each other so that the current is carried by the negatively charged electrons. There are, however, other types of conductors like electrolytic solutions where positive and negative charges both can move. In our discussions, we will focus only on solid conductors so that the current is carried by the negatively charged electrons in the background of fixed positive ions. Consider first the case when no electric field is present. The electrons will be moving due to thermal motion during which they collide with the fixed ions. An electron colliding with an ion emerges with the same speed as before the collision. However, the direction of its velocity after the collision is completely random. At a given time, there is no preferential direction for the velocities of the electrons. Thus on the average, the number of electrons travelling in any direction will be equal to the number of electrons travelling in the opposite direction. So, there will be no net electric current. A basic law regarding flow of currents was discovered by G.S. Ohm in 1828, long before the physical mechanism responsible for flow of currents was discovered. Imagine a conductor through which a current I is flowing and let V be the potential difference between the ends of the conductor. Then Ohm's law states that $V = RI$ Where the constant of proportionality R is called the resistance of the conductor. The resistance R not only depends on the material of the conductor but also on the dimensions of the conductor.

2.2 OBJECTIVES

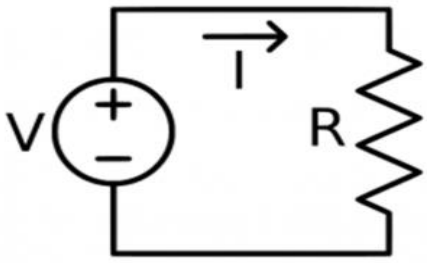
After studying this unit, you should be able to-

- Define resistance
- Understand different types of resistors
- Know various types of resistances
- Get how drift of charge carriers takes place in electric and magnetic fields
- Understand what variable resistors and rheostats
- Understand what is potentiometers
- Understand what is resistor color coding
- Understand what is power rating of resistors
- Understand how combination of resistors
- Understand what is Inductors and types of Inductance
- Define self and mutual inductance, variable inductance, stray inductance
- Understand how does inductances in combination
- Understand what is energy stored in an inductor
- Understand what is capacitor and its types
- Understand how does capacitors in combination
- Understand what is energy in electrostatic field of a capacitor,
- Understand Filter circuits,
- Low-pass and high-pass filters

2.3 RESISTANCE, OHM'S LAW AND RESISTIVITY

When there is no free flow of the electrons in a material due to obstruction created by the electrons during their movement, it is known as electrical resistance and is denoted by R. For electrical resistance to be high, the length of the conductor and the cross-sectional area should be more. The resistance offered by the copper is low whereas the resistance offered by the wood is high. According to Ohm's law, there is a relation between the current flowing through a conductor and the potential difference across it. It is given by, $V \propto I$ and $V = IR$. Where,

- **V** is the potential difference measured across the conductor (in volts)
- **I** is the current through the conductor (in amperes)
- **R** is the constant of proportionality called resistance (in ohms)



The electrical resistance of a circuit is defined as the ratio between the voltage applied to the current flowing through it. Rearranging the above relation, R = V/I. The unit of electrical resistance is ohms. 1ohm = 1 volt /1 ampere. Electric charge flows easily through some materials than others. The electrical resistance measures how much the flow of this electric charge is restricted within the circuit.

The Unit of Resistance

The SI units of resistance is ohm, and is denoted by the symbol Ω.A conductor is said to have a resistance of one ohm if it permits one ampere current to flow through it when one volt is impressed across its terminals. For insulators whose resistances are very high, a much bigger unit is used i.e., mega-ohm = 10^6 ohm (the prefix 'mega' meaning a million) or kilo-ohm = 10^3 ohm (kilo means thousand). In the case of very small resistances, smaller units like milli-ohm = 10^{-3} ohm or micro-ohm = 10^{-6} ohm are used.

Resistivity

Electric resistivity is defined as the electrical resistance offered per unit length and unit cross-sectional area at a specific temperature and is denoted by ρ. Electrical resistivity is also known as specific electrical resistance. The SI unit of electrical resistivity is ohms-metre (Ω.m). Following is the formula of electrical resistivity: J= ρ E Where,

- ρ is the resistivity of the material in Ω.m
- E is the electric field in $V.m^{-1}$
- J is the current density in $A.m^{-2}$

Consider now the situation when an electric field is present. Electrons will be accelerated due to this field by

$$a = \frac{F}{m} = \frac{-eE}{m} \tag{1}$$

where –e is the charge and m is the mass of an electron. Consider again the i^{th} electron at a given time t. This electron would have had its last collision some time before t, and let t be the time elapsed after its last collision. If $\mathbf{v}_i$ was its velocity immediately after the last collision, then its velocity V_t at time t is–

$$V_t = V_i + a\,t \tag{2}$$

$$V_t = V_i + \frac{-eE}{m}t \tag{3}$$

Since starting with its last collision it was accelerated with an acceleration given by Eq. (1) for a time interval t. The average velocity of the electrons at time t is the average of all the Vi. The average of V_i is zero. Since immediately after any collision, the direction of the velocity of an electron is completely random. The collisions of the electrons do not occur at regular intervals but at random times. Let us denote by τ, the average time between successive collisions. The average value of t then is τ (known as relaxation time). Thus, averaging Eq. (3) over the N-electrons at any given time t gives us for the average velocity $\mathbf{v}_d$. Substituting $V_i = 0$ and $t = \tau$ in Eq. (3), we get

$$v_d = \frac{-eE}{m}\tau \tag{4}$$

It tells us that the electrons move with an average velocity which is independent of time, although electrons are accelerated. This is the phenomenon of drift and the velocity $\mathbf{v}_d$ in Eq. (4) is called the drift velocity. Because of the drift, there will be net transport of charges across any cross section area in the presence of electric field **E**. Consider a planar area A, located inside the conductor such that the normal to the area is parallel to **E** (Fig. 2.1). Then because of the drift, in an infinitesimal amount of time Δt, all electrons to the left of the area at distances upto $|\mathbf{v}_d|.\Delta t$ would have crossed the area. If n is the number of free electrons per unit volume in the metal, then there are $n\,\Delta t|\mathbf{v}_d|A$ electrons.

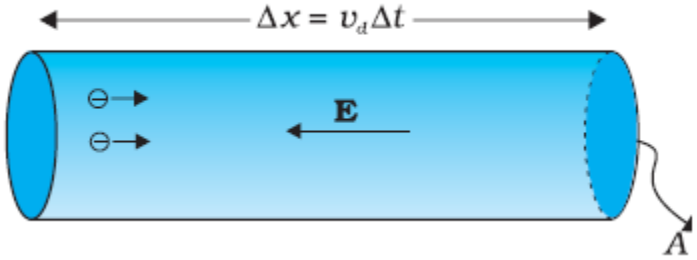


Figure 2.1: Current in a metallic conductor. The magnitude of current density in a metal is the magnitude of charge contained in a cylinder of unit area and length v_d.

Since each electron carries a charge –e, the total charge transported across this area A to the right in time Δt is $-ne\,A|\mathbf{v}_d|\Delta t$. is directed towards the left and hence the total charge transported along **E** across the area is negative of this. The amount of charge crossing the area A in time Δt is by definition is $q = I\,\Delta t$, where I is the magnitude of the current. Hence,

$$q = Ne = I\,\Delta t = n\,e\,A\,\mathbf{v}_d\Delta t \tag{5}$$

$$I = \frac{q}{\Delta t} = n\,e\,A\,\mathbf{v}_d \tag{6}$$

Substituting the value of $|\mathbf{v}_d|$ from Eq. (4)

$$I = \frac{q}{\Delta t} = n\, e\, A \frac{eE}{m}\tau \tag{7}$$

$$I = \frac{ne^2\tau\, A\, E}{m} \tag{8}$$

By definition I is related to the magnitude $|\mathbf{j}|$ of the current density by $I = |\mathbf{j}|\, A$. Hence

$$J = \frac{I}{A} = \frac{ne^2\tau\ E}{m} \tag{9}$$

The vector **j** is parallel to **E** and hence we can write Eq. (9) in the vector form

$$\mathbf{J} = \frac{ne^2\tau\ \mathbf{E}}{m} \tag{10 a}$$

However, **J** is also define by Ohm's law, $J = \sigma\, \mathbf{E}$. Comparison with Eq. (10 a) is exactly the Ohm's law, if we identify the conductivity σ as

$$\boldsymbol{\sigma} = \frac{ne^2\tau}{m} \tag{10 b}$$

We thus see that a very simple picture of electrical conduction reproduces Ohm's law. We have, of course, made assumptions that τ and n are constants, independent of E. Resistivity is reciprocal of conductivity as $\rho = \frac{1}{\sigma}$, hence

$$\boldsymbol{\rho} = \frac{\mathbf{1}}{\boldsymbol{\sigma}} = \frac{m}{ne^2\tau} \tag{10 c}$$

2.3.1 Resistance and Types of Resistors

The resistance R not only depends on the material of the conductor but also on the dimensions of the conductor. In general, the resistance R is inversely proportional to the cross-sectional area and proportional to the length of the conductor. Therefore,

$$R = \rho \frac{l}{A} \tag{11}$$

Substituting the value of ρ from equation (2k) we can write

$$R = \frac{m}{ne^2\tau}\frac{l}{A} \qquad (12)$$

Equation (4) shows the dependency of R in various factors like dimensions, material and other physical parameters (temperature etc.) In general resistance may be defined as the property of a substance due to which it opposes (or restricts) the flow of electricity (i.e., electrons) through it. Metals (as a class), acids and salts solutions are good conductors of electricity. Amongst pure metals, silver, copper and aluminum are very good conductors in the given order. This is due to the presence of a large number of free or loosely-attached electrons in their atoms. These vagrant electrons assume a directed motion on the application of an electric potential difference. These electrons while flowing pass through the molecules or the atoms of the conductor collide and other atoms and electrons, thereby producing heat. Those substances which offer relatively greater difficulty or hindrance to the passage of these electrons are said to be relatively poor conductors of electricity like bakelite, mica, glass, rubber, p.v.c. (polyvinyl chloride) and dry wood etc. Amongst good insulators can be included fibrous substances such as paper and cotton when dry, mineral oils free from acids and water, ceramics like hard porcelain and asbestos and many other plastics besides p.v.c. It is helpful to remember that electric friction is similar to friction in Mechanics.

Types of Resistors

(a) Carbon Composition

It is a combination of carbon particles and a binding resin with different proportions for providing desired resistance. Attached to the ends of the resistive element are metal caps which have axial leads of tinned copper wire for soldering the resistor into a circuit. The resistor is enclosed in a plastic case to prevent the entry of moisture and other harmful elements from outside. Billions of carbon composition resistors are used in the electronic industry every year. They are available in power ratings of 1/8, 1/4, 1/2, 1 and 2 W, in voltage ratings of 250, 350 and 500 V. They have low failure rates when properly used. Such resistors have a tendency to produce electric noise due to the current passing from one carbon particle to another. This noise appears in the form of a hiss in a loudspeaker connected to a hi-fi system and can overcome very weak signals. That is why carbon composition resistors are used where performance requirements are not demanding and where low cost in the main consideration. Hence, they are extensively used in entertainment electronics although better resistors are used in critical circuits.

(b) Deposited Carbon

Deposited carbon resistors consist of ceramic rods which have a carbon film deposited on them. They are made by placing a ceramic rod in a methane-filled flask and heating it until, by a gas cracking process, a carbon film is deposited on them. A helix-grinding process forms the resistive path. As compared to carbon composition resistors, these resistors offer a major improvement in lower current noise and in closer tolerance. These resistors are being replaced by metal film and metal glaze resistors.

(c) High-Voltage Ink Film

These resistors consist of a ceramic base on which a special resistive ink is laid down in a helical band. These resistors are capable of withstanding high voltages and find extensive use in cathode-ray circuits, in radar and in medical electronics. Their resistances range from 1 kΩ to 100,000 MΩ with voltage range upto 1000 kV.

(d) Metal Film

Metal film resistors are made by depositing vaporized metal in vacuum on a ceramic-core rod. The resistive path is helix-ground as in the case of deposited carbon resistors. Metal film resistors have excellent tolerance and temperature coefficient and are extremely reliable. Hence, they are very suitable for numerous high grade applications as in low-level stages of certain instruments although they are costlier.

(e) Metal Glaze

A metal glaze resistor consists of a metal glass mixture which is applied as a thick film to a ceramic substrate and then fired to form a film. The value of resistance depends on the amount of metal in the mixture. With helix-grinding, the resistance can be made to vary from 1 Ω to many mega ohms. Another category of metal glaze resistors consists of a tinned oxide film on a glass substrate.

(f) Wire-wound

Wire-wound resistors are different from all other types in the sense that no film or resistive coating is used in their construction. They consist of a ceramic-core wound with a drawn wire having accurately-controlled characteristics. Different wire alloys are used for providing different resistance ranges. These resistors have highest stability and highest power rating. Because of their bulk, high-power ratings and high cost, they are not suitable for low-cost or high-density, limited-space applications. The completed wire-wound resistor is coated with an insulating material such as baked enamel.

(g) Cermet (Ceramic Metal)

The cermet resistors are made by firing certain metals blended with ceramics on a ceramic substrate. The value of resistance depends on the type of mix and its thickness. These resistors have very accurate resistance values and show high stability even under extreme temperatures. Usually, they are produced as small rectangles having leads for being attached to printed circuit boards (PCB).

2.2 VARIAABLE RESISTORS AND RHEOSTATS

Resistors can be broadly of two types Fixed Resistors and Variable Resistors.

Fixed Resistors:

A fixed resistor is one for which the value of its resistance is specified and cannot be varied in general.

Variable Resistors:

As the name suggests, the resistance of a variable resistor can change. It can be used as a three terminal as well as a two terminal device. Mostly they are used as a three terminal device. Variable resistors are mostly used for device calibration. There are two general ways in which variable resistors are used. One is the variable resistor whose value is easily changed, like the volume adjustment of Radio. The other is semi-fixed resistor that is not meant to be adjusted by anyone but a technician. It is used to adjust the operating condition of the circuit by the technician.

Working of Variable Resistor

As shown in the figure, a variable resistor consists of a track which provides the resistance path. Two terminals of the device are connected to both the ends of the track. The third terminal is connected to a wiper that decides the movement of the track. The motion of the wiper through the track helps in increasing and decreasing the resistance.

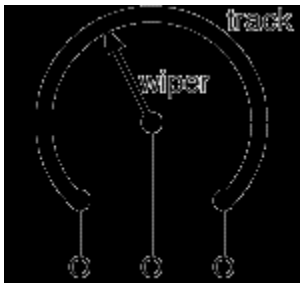


Figure 2.2: Variable resistor working.

The track is usually made of a mixture of ceramic and metal or can be made of carbon as well. As a resistive material is needed, carbon film type variable resistors are mostly used. They find applications in radio receiver circuits, audio amplifier circuits and TV receivers. For applications of small resistances, the resistance track may just be a coil of wire. The track can be in both the rotary as well as straight versions. In a rotary track, some of them may include a switch. The switch will have an operating shaft which can be easily moved in the axial direction with one of its ends moving from the body of variable resistor switch.

The rotary track resistor with has two applications. One is to change the resistance. The switch mechanism is used for the electric contact and non-contact by on/off operation of the switch. There are switch mechanism variable resistors with an annular cross-section which are used for the control of equipment. Even more, components are added to this type of a variable resistor so as to make them compatible with complicated electronic circuits. A high-voltage variable resistor such as a focus pack is an example. This device is capable of producing a variable focus voltage as well as a screen voltage. It is also connected to a variable resistance circuit and also a fixed resistance

circuit [bleeder resistor] to bring a change in the applied voltage. For this, both the fixed and variable resistor are connected in series.

A track made in a straight path is called a slider. As the position of a slider cannot be seen or confirmed according to the adjustment of resistance, a stopping mechanism is usually included to prevent the hazards caused due to over-rotation.

Types of variable Resistors

There are mainly three types of variable resistors. They are

1. Potentiometer
2. Rheostat
3. Presets

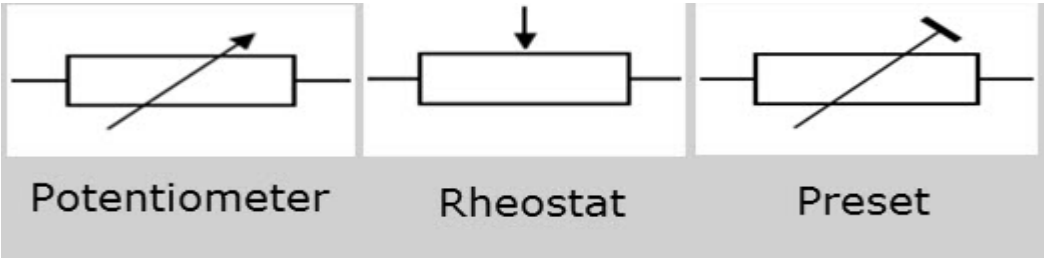


Figure 2.3: Circuit Symbols for various types of Variable resistors.

2.3 COLOR CODING OF RESISTORS

There are many types of resistors, both fixed and variable. The most common type for electronics use is the carbon resistor. They are made in different physical sizes with power dissipation limits commonly from 1 watt down to 1/8 watt. The resistance value and tolerance can be determined from the standard resistor color code.

In this color code system, the bands of different colors are used to identify the value of resistance and tolerance ratings. Also the power rating is determined by the physical size of the resistor. Generally, there are two common systems of the color code designation: a four band system and a five bands system. However, the four bands color code system is more popular and is widely used.

In four band system, reading the resistor from left to right, the first two color bands represent significant digits, the third band represents the decimal multiplier, and the fourth band represents the tolerance.

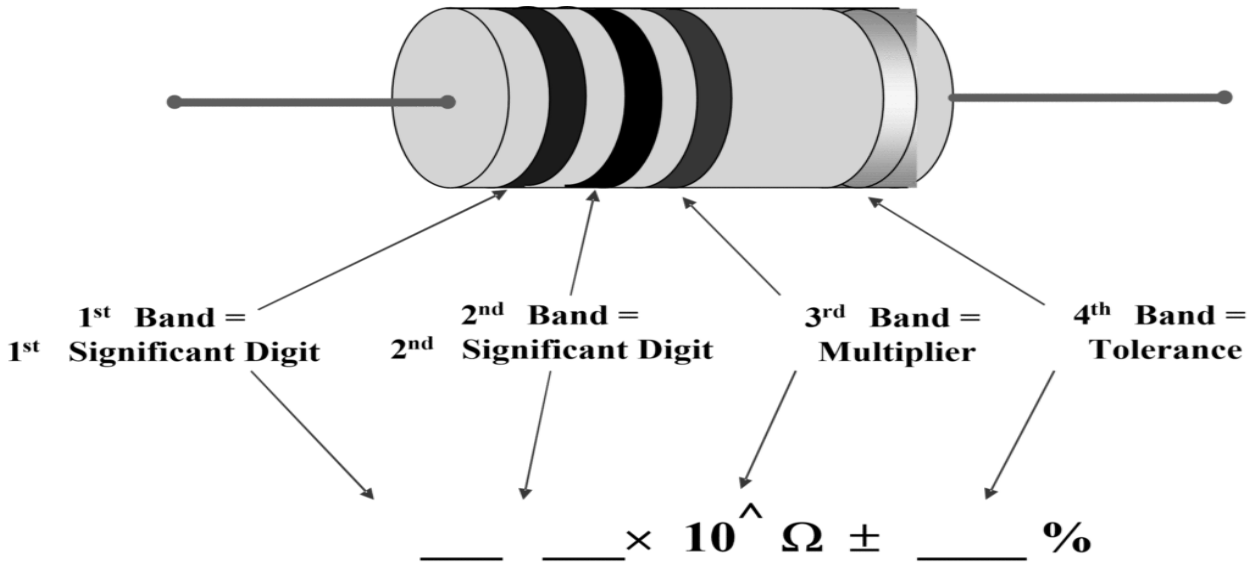


Figure 2.4: Interpretation of each band in four band color code.

In five bands system the first three color bands represent significant digits, the fourth band represents the decimal multiplier, and the fifth band represents the tolerance.

Codes for the different colors have been shown in table below

Color	Digit	Multiplier	Tolerance (%)
Black	0	10^0 (1)	
Brown	1	10^1	1
Red	2	10^2	2
Orange	3	10^3	
Yellow	4	10^4	
Green	5	10^5	0.5
Blue	6	10^6	0.25
Violet	7	10^7	0.1
Grey	8	10^8	
White	9	10^9	
Gold		10^{-1}	5
Silver		10^{-2}	10
(none)			20

Table: Color code for resistor designation.

For example in a given resistor let first strip be brown, second strip be red and third be orange and fourth be gold then resistance of the resistor would be $12 \times 10^3 \pm 5\%$.

2.4 POTENTIOMETER

The Potentiometer is an electric instrument that used to measure the EMF (electro motive force) of a given cell, compare EMFs of different cells, potential difference across two points of the circuit and internal resistance of a cell. It can also use as a variable resistor in most of the applications. These potentiometers are used in huge quantities in the manufacture of electronics equipment that provides a way of adjusting electronic circuits so that the correct outputs are obtained. Although their most obvious use must be for volume controls on radios and other electronic equipment used for audio.

Construction

The potentiometer consists of a long resistive wire L made up of magnum or with constantan and a battery of known EMF V volt. This voltage is called as **driver cell voltage**. Connect the two ends of the resistive wire L to the battery terminals as shown below; let us assume this is a primary circuit arrangement. One terminal of another cell (whose EMF E is to be measured) is at one end of the primary circuit and another end of the cell terminal is connected to any point on the resistive wire through a galvanometer G. Now let us assume this arrangement is a secondary circuit. The arrangement of the potentiometer circuit is as shown below.

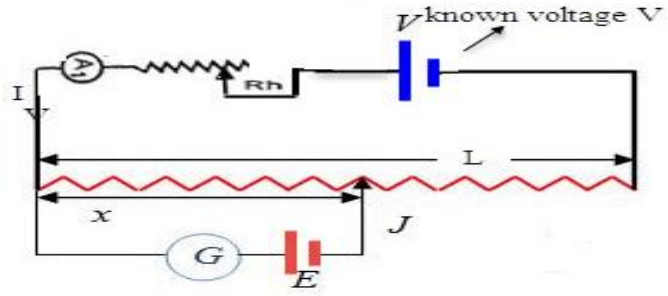


Figure 2.5: Construction of potentiometer.

The basic working principle of this is based on the fact that the fall of the potential across any portion of the wire is directly proportional to the length of the wire, provided wire has uniform cross-sectional area and the constant current flowing through it. The potentiometer wire is actually a wire with high resistivity (ṕ) with uniform cross-sectional area A. Thus, throughout the wire, it has uniform resistance. Now this potentiometer terminal connected to the cell of high EMF V (neglecting its internal resistance) called driver cell or the voltage source. Let the current through

the potentiometer is I and R is the total resistance of the potentiometer. **Following are the terms used to describe types of potentiometers:**

- It is defined as a three-terminal resistor having either sliding or rotating contact that forms an adjustable voltage divider.
- In order to use the potentiometer as a rheostat or variable resistor, it should have only two terminals with one end and the wiper.
- Slider pot or slide pot: This can be adjusted by sliding the wiper right or left with a finger or thumb.
- Thumbwheel pot or thumb pot: This can be adjusted infrequently with the help of small thumbwheel which is a small rotating potentiometer.
- Trimmer pot or trimpot: This can be adjusted once for fine tuning of an electric signal.

Working Principle

A potentiometer, also called as POT, is a 3-terminal variable resistor and is used to adjust the resistance in a circuit. This is a versatile instrument. It is basically a long piece of uniform wire, sometimes a few meters in length across which a standard cell is connected. In actual design, the wire is sometimes cut in several pieces placed side by side and connected at the ends by thick metal strip. (Fig. 2.6). In the figure, the wires run from A to C. The small vertical portions are the thick metal strips connecting the various sections of the wire. A current I flows through the wire which can be varied by a variable resistance (rheostat, R) in the circuit. Since the wire is uniform, the potential difference between A and any point at a distance l from A is $\varepsilon = K\,l$ where K is potential gradient or potential drop per unit length.

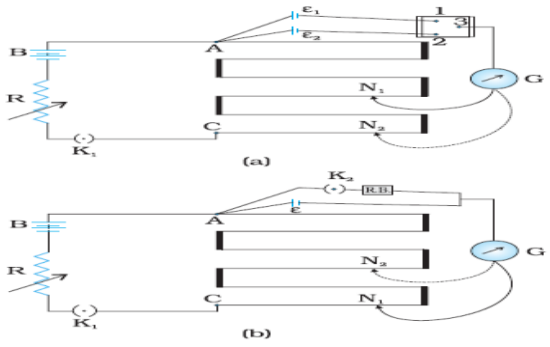


Figure 2.6: A potentiometer. G is a galvanometer and Ra variable resistance (rheostat). 1, 2, 3 are terminals of a two-way key (a) circuit for comparing emfs of two cells; (b) circuit for determining internal resistance of a cell.

Figure 2.6, shows an application of the potentiometer to compare the emf of two cells of emfε_1 and ε_2. The points marked 1, 2, 3 form a two-way key. Consider first a position of the key where 1 and 3 are connected so that the galvanometer is connected to ε_1. The jockeys moved along the wire till at a point N_1, at a distance $l_1$from A, there is no deflection in the galvanometer. We can apply Kirchhoff's loop rule to the closed loop$AN_1G_{31}A$ and get,

$$\mathcal{E}_1 = k\, l_1 \text{and} \mathcal{E}_2 = k\, l_2 \quad \text{Hence} \quad \frac{\varepsilon_1}{\varepsilon_2} = \frac{l_1}{l_2} \tag{13}$$

This simple mechanism thus allows one to compare the emf of any two sources. In practice one of the cells is chosen as a standard cell whose emf is known to a high degree of accuracy. The emf of the other cell is then easily calculated from Eq. (13). We can also use a potentiometer to measure internal resistance of a cell [Fig. 2.6]. For this the cell (emf ε) whose internal resistance (r) is to be determined is connected across a resistance box through a key K_2, as shown in the figure. With key K2 open, balance is obtained at length l1 (AN1). Then,

$$\mathcal{E} = k\, l_1 \tag{14}$$

When key K2 is closed, the cell sends a current (I) through the resistance box (R). If V is the terminal potential difference of the cell and balance is obtained at length l_2 (AN2), then

$$V = K\, l_2 \tag{15}$$

From Equation 14 and 15, we have

$$\frac{\varepsilon}{V} = \frac{l_1}{l_2} \tag{16}$$

But, $\varepsilon = I(r + R)$ and $V = IR$. This gives

$$\frac{\varepsilon}{V} = \frac{r+R}{R} = \frac{l_1}{l_2} \tag{17}$$

$$\frac{r+R}{R} = \frac{l_1}{l_2}$$

$$r = \left\{\frac{l_1}{l_2} - 1\right\} R \tag{18}$$

Using Eq. (18), we can find the internal resistance of a given cell. The potentiometer has the advantage that it draws no current from the voltage source being measured. As such it is unaffected by the internal resistance of the source.

In some other application the working of potentiometer is the same as that of a variable resistor. The construction is also the same. It has a resistive element as the track and a sliding contact called the wiper. The wiper is connected with the help of another sliding contact to another terminal. The position of the wiper depends on the type of POT used. For a panel POT, the wiper is kept in the middle. The resistive element has a terminal on both the ends and can be linear or logarithmic. It is usually made up of carbon or a mixture of ceramic and metal or even graphite. The most common form of the potentiometer is the single turn rotary potentiometer. This type of potentiometer is often used in audio volume control (logarithmic taper) as well as many other applications. Different materials are used to construct potentiometers, including carbon composition, cermet, conductive plastic, and the metal film. Rotary potentiometers are the most common type of potentiometers, where the wiper moves along a circular path.

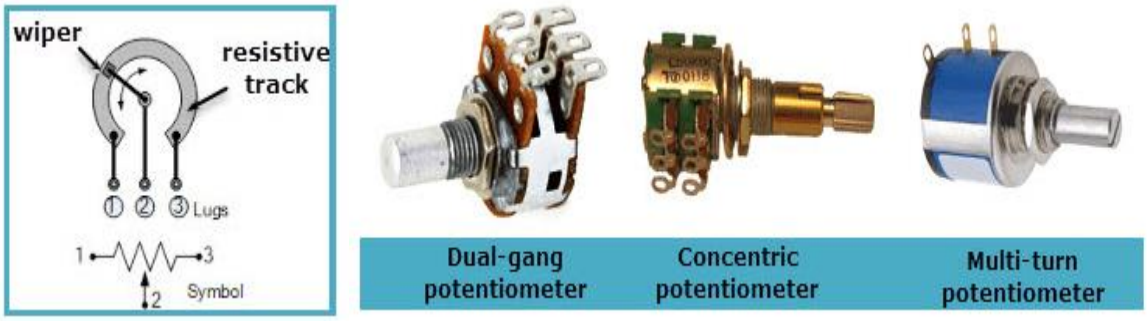


Figure 2.7: (a) Rotary potentiometer.

There is single turn POTs which changes its entire resistance in one rotation. More accurate POTs called multi-turn POTs are also present. They need about 20 to 30 rotations to change the entire resistance. They are much more accurate than the former. Pots are differentiated according to the track used. There are mainly two of them. They are

1. Logarithmic Potentiometer

2. Linear Potentiometer

1. Logarithmic Potentiometer: This type of a POT is shortly designated with the letter “A” in the component. For example, a POT with 5 kilo ohm resistance will be given as “5k A”. The track of this material will be a resistive material that may be tapered from one end to the other. It may also be a material whose resistivity varies from one end to the other. As like a logarithmic variable resistor this type of POT will have a logarithmic output. Due to this nature, they are always used in audio circuits. They are much costlier than other types of POTS.

2. **Linear Potentiometer:** This type of a POT is shortly designated with the letter "B" in the component. For example, a POT with 5 kilo ohm resistance will be given as "5k B". In this device, the track element has a constant cross-section causing a proportional resistance change between the wiper and one end of the terminal. The device depends on the electrical feature and not the resistive feature. This POT is used for a proportional change like adjusting the centering of a CRO.

Uses of a Potentiometer

- POTs are used for controlling the signal level of a circuit and not the power of the circuit
- Volume control on TV's and other audio equipments.
- Used in joysticks as a position transducer.
- TRIAC switching applications.
- Voltage divider circuits.

Rheostat

A rheostat is also a variable resistor and is a 2-terminal device. It is commonly used for handling higher currents and voltages. One terminal will be connected to the end of the track and the other to a moveable wiper. When the wiper moves from one end to the other then the resistance changes from zero to maximum.

Figure 2.7: (b) Rheostat

A rheostat can be made out of a potentiometer. The same mechanism is used except that the terminal that is not used will be connected to the wiper. This helps in reducing the variation in resistance. It also helps in gaining more mechanical strength when connected to a PCB. The track is usually made from a resistance wire which is wound on a heat resisting cylinder. They are wound together to form the shape of a toroid coil. The slider moves from one phase of the coil to another, thus varying the resistance. The sliders are made in the shape of metal fingers and they move across the tracks through tapping method.

Difference Between Potentiometer and Rheostat

- A potentiometer is a three terminal variable resistor, but a rheostat is a two terminal variable resistor.
- A potentiometer can be used as a rheostat but a rheostat cannot be used as a potentiometer.
- Potentiometers are often used to vary voltage and rheostats are used to vary current.

2.4 POWER RATING OF RESISTORS

When an electrical current passes through a resistor due to the presence of a voltage across it, electrical energy is lost by the resistor in the form of heat and the greater this current flow the hotter the resistor will get. This is known as the Resistor Power Rating. Resistors are rated by the value of their resistance and the electrical power given in watts, (W) that they can safely dissipate based mainly upon their size. Every resistor has a maximum power rating which is determined by its physical size as generally, the greater its surface area the more power it can dissipate safely into the ambient air or into a heatsink. A resistor can be used at any combination of voltage (within reason) and current so long as its "Dissipating Power Rating" is not exceeded with the resistor power rating indicating how much power the resistor can convert into heat or absorb without any damage to itself. The Resistor Power Rating is sometimes called the Resistors Wattage Rating and is defined as the amount of heat that a resistive element can dissipate for an indefinite period of time without degrading its performance. The power rating of resistors can vary a lot from less than one tenth of a watt to many hundreds of watts depending upon its size, construction and ambient operating temperature. Most resistors have their maximum resistive power rating given for an ambient temperature of +70^{o}C or below. Electrical power is the rate in time at which energy is used or consumed (converted into heat). The standard unit of electrical power is the Watt, symbol W and a resistors power rating is also given in Watts. As with other electrical quantities, prefixes are attached to the word "Watt" when expressing very large or very small amounts of resistor power. Some of the more common of these are provided in the table below:

The electrical power dissipation of any resistor in a DC circuit can be calculated using one of the following three standard formulas:

$$P = VI = I^2R = \frac{V^2}{R}$$

Where V is the voltage across the resistor in Volts. I is in current flowing through the resistor in Amperes R is the resistance of the resistor in Ohm's (Ω). As the dissipated resistor power rating is linked to their physical size, a 1/4 (0.250) W resistor is physically smaller than a 1W resistor, and resistors that are of the same ohmic value are also available in different power or wattage ratings.

Carbon resistors, for example, are commonly made in wattage ratings of 1/8 (0.125) W, 1/4 (0.250) W, 1/2 (0.5) W, 1W, and 2 Watts.

Type	Power Rating	Stability
Metal Film	Very low at less than 3 Watts	High 1%
Carbon	Low at less than 5 Watts	Low 20%
Wirewound	High up to 500 Watts	High 1%

Table: Power Rating of Different Types of Resistors

Generally speaking, the larger their physical size, the higher its wattage rating. However, it is always better to select a particular size resistor that is capable of dissipating two or more times the calculated power. When resistors with higher wattage ratings are required, wire wound resistors are generally used to dissipate the excessive heat.

2.5 COMBINATION OF RESISTORS

Resistances in Series

When some conductors having resistances R_1, R_2 and R_3 etc. are joined end-on-end as in Fig.2.8, they are said to be connected in series. It can be proved that the equivalent resistance or total resistance between points A and D is equal to the sum of the three individual resistances. Being a series circuit, it should be remembered that

- Current is the same through all the three conductors
- Voltage drop across each is different due to its different resistance and is given by Ohm's Law and
- Sum of the three voltage drops is equal to the voltage applied across the three conductors.

There is a progressive fall in potential as we go from point A to D as shown in Fig.

Therefore, $V = V_1 + V_2 + V_3 = IR_1 + IR_2 + IR_3$

But $V = IR$

Where, R is the equivalent resistance of series combination.

Therefore we have $IR = IR_1 + IR_2 + IR_3$

$$\text{Or } R = R_1 + R_2 + R_3 \qquad (19)$$

Resistances in Parallel

Three resistances, as joined in Fig. are said to be connected in parallel. In this case

- p.d. across all resistances is the same
- current in each resistor is different and is given by Ohm's Law and
- the total current from the source is the sum of the three separate currents in parallel paths.

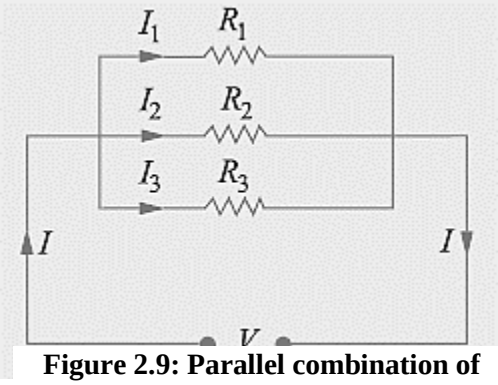


Figure 2.9: Parallel combination of resistances.

$$I = I_1 + I_2 + I_3 = \frac{V}{R_1} + \frac{V}{R_2} + \frac{V}{R_3}$$

Now $I = \frac{V}{R}$ where V is the applied voltage and R = equivalent resistance of parallel combination

Therefore we have, $\frac{V}{R} = \frac{V}{R_1} + \frac{V}{R_2} + \frac{V}{R_3}$

or $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$ ---------- (20)

2.6 INDUCTOR AND INDUCTANCE

The inductor is a passive component which stores the electrical energy in the magnetic field when the electric current passes through it. Or we can say that the inductor is an electrical device which possesses the inductance. The inductor is made of wire which has the property of inductance, i.e., opposes the flow of current. The inductance of wire increases by increasing the number of turns. The alphabet 'L' is used for representing the inductor, and it is measured in Henry. The inductance characterizes the inductor. The figure below shows the symbolic representation of inductor.

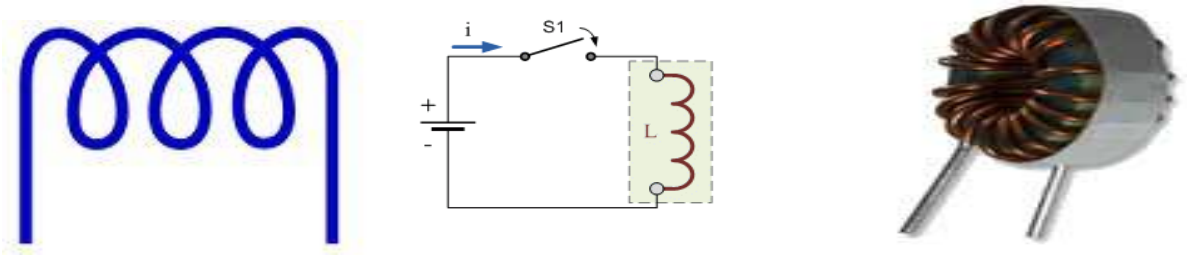


Figure 2.10: Inductor coils.

The electric current I flows through the coil generates the magnetic field around it. Consider the magnetic field generates the flux Φ when current flows through it. The ratio of the flux and the

current gives inductances L= $\frac{\Phi}{I}$. The inductance of the circuit depends on the current paths and the magnetic permeability of the nearer material. The magnetic permeability shows the ability of the material to forms the magnetic field.

The **Inductance** is the property of a material by virtue of which it opposes any change of magnitude and direction of electric current passing through it. In other words, the inductance is the property of the coil in which the emf is induced in the coil because of the variation of flux. The inductance is added in the circuit through the inductor. The inductor is basically the coil of wires which concentrates the magnetic field into the circuit. Josheph Henry [1797 – 1878] was American experimental physicist professor at Princeton University and first director of the Smithsonian Institution. He made important improvements in electromagnets by winding coils of insulated wire around iron pole pieces and invented an electromagnetic motor and a new, efficient telegraph. He discovered self-induction and investigated how currents in one circuit induce currents in another. Inductance is said to be one Henry, when current pass through a coil of conductor changes of one ampere per second and the voltage induces one volt across the coil.

From the experimental observations, Faraday arrived at a conclusion that an emf is induced in a coil when magnetic flux through the coil changes with time. Faraday stated experimental observations in the form of a law called Faraday's law of electromagnetic induction. The law is stated below.

The magnitude of the induced emf in a circuit is equal to the time rate of change of magnetic flux through the circuit. Mathematically, the induced emf is given by

$$\varepsilon = - \frac{d\Phi}{dt} \tag{21}$$

The negative sign indicates the direction of ε and hence the direction of current in a closed loop. In the case of a closely wound coil of N turns, change of flux associated with each turn, is the same. Therefore, the expression for the total induced emf is given by

$$\varepsilon = - N \frac{d\Phi}{dt} \tag{22}$$

The induced emf can be increased by increasing the number of turns N of a closed coil.

Explanation and Types of Inductors

An electric current can be induced in a coil by flux change produced by another coil in its vicinity or flux change produced by the same coil. These two situations are described separately in the next two sub-sections. However, in both the cases, the flux through a coil is proportional to the current. That is, $\Phi \ \alpha I$. There are two types of induction phenomena are commonly defined:

(i) Self-Induction phenomenon (ii) Mutual Induction phenomenon

2.7 SELF INDUCTANCE

The inductor is formed by a wire when a wire of finite length is twisted into a coil. When the current flows through the coil, an electromagnetic field is formed. The electromagnetic field changes if the direction of flow of current changes. This change in the electromagnetic field induces a voltage (ε) across the coil and is given by the equation shown below

$$\varepsilon = - \frac{d\Phi}{dt} \tag{23}$$

Flux associated with the coil is proportional to current, where constant of proportionality L is called self-inductance of the coil. It is also called the coefficient of self-induction of the coil. When the current is varied, the flux linked with the coil also changes and an emf is induced in the coil. Therefore, $\Phi \alpha I$ or $\Phi = L I$, substituting this, we get

$$\varepsilon = - L \frac{dI}{dt} \tag{24}$$

Where, I is the current flowing through the inductor in ampere. The voltage across the inductor will be zero if the current flowing through it remains constant. This means that when direct, steady current flows through the inductor, it behaves as a short-circuited coil in a steady state condition. If there is any small change in direction or strength of the current, the inductance will appear. Thus, the self-induced emf always opposes any change (increase or decrease) of current in the coil. If we put the value of dt as zero (dt=0) in equation (24) it is seen that for a minute change in current within zero time, gives an infinite voltage across the inductor which is not a feasible condition and thus in an inductor the current cannot be changed abruptly. Thus after switching the DC voltage, the inductors act as an open circuited coil.

It is possible to calculate the self-inductance for circuits with simple geometries. Let us calculate the self-inductance of a long solenoid of cross-sectional area A and length l, having n turns per unit length. The magnetic field due to a current I flowing in the solenoid having n number of turns per unit length is $B = \mu_0 n I$ (neglecting edge effects). The total flux linked with the solenoid is,

$$\Phi = NBA \tag{25}$$

Substituting the value $B = \mu_0 n I$, and $N = n l$, we get

$$\Phi = (nl)\ (\mu_0 n I)\ A \tag{26}$$

$$\Phi = \mu_0 n^2\ Al\ \ I \tag{27}$$

Comparing equation (27) with $\Phi = L I$, we can get coefficient of self-inductance as

$$L = \frac{\Phi}{I} = \mu_0 n^2\ Al \tag{28}$$

Or $$L = \mu_0 n^2\ Al \tag{29}$$

Where nl is the total number of turns equal to N. Thus, the self-inductance is

$$L = \mu_0 N n A \quad (30)$$

Or $$L = \mu_0 \frac{N^2}{L} A \quad (31)$$

The self-inductance of the coil depends on its geometry and on the relative permeability (μ_r) of the medium.

$$L = \mu_r \mu_0 \frac{N^2}{L} A \quad (32)$$

2.8 MUTUAL INDUCTANCE

Mutual Inductance between the two coils is defined as the property of the coil due to which it opposes the change of current in the other coil, or you can say in the neighboring coil. When the current in the neighboring coil is changing, the flux sets up in the coil and because of this changing flux emf is induced in the coil called Mutually Induced emf and the phenomenon is known as Mutual Inductance.

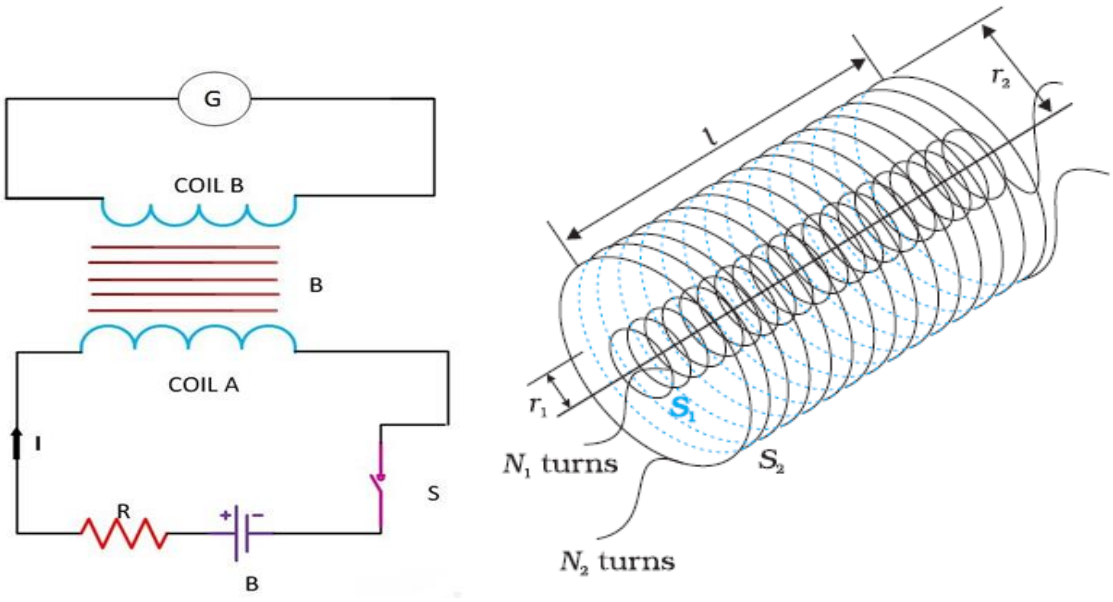


Figure 2.11: Two long co-axial solenoids of same length l.

Let us understand the phenomenon of Mutual Inductance by considering an example as shown in the above figure. Two coils namely coil A and coil B is placed nearer to each other. When the switch S is closed, and the current flows in the coil it sets up the flux Φ in the coil A and emf is induced in the coil and if the value of the current is changed by varying the value of the resistance (R), the flux linking with the coil B also changes because of this changing current. Thus this phenomenon of the linking flux of the coil A with the other coil, B is called Mutual Inductance. For determining the Mutual Inductance between the two coils, the following expression is used

$$\varepsilon = M\frac{dI_1}{dt} \tag{33}$$

or
$$M = \frac{\varepsilon}{\frac{dI_1}{dt}} \tag{34}$$

This expression is used when the magnitude of mutually induced emf in the coil and the rate of change of current in the neighboring coil is known. Where M is the mutual inductance coefficient which is defined as; If$\mathcal{E}$ = 1 volt and dI_1/dt = 1 ampere per second then putting this value in the equation (34) we get the value of mutual inductance as M=1 Henry.

Hence, from the above statement, you can define Mutual Inductance as "the two coils are said to have a mutual inductance of one Henry if an emf of 1 volt is induced in one coil or say primary coil when the current flowing through the other neighboring coil or secondary coil is changing at the rate of 1 ampere/second". The value of Mutual Inductance (M) depends upon the following factors

- Number of turns in the secondary or neighboring coil
- Cross-sectional area
- Closeness of the two coils

Consider Fig. 2.11 which shows two long co-axial solenoids each of length l. We denote the radius of the inner solenoid S_1 by r_1 and the number of turns per unit length by n_1. The corresponding quantities for the outer solenoid S_2 are r_2 and n_2, respectively. Let N_1 and N_2 be the total number of turns of coils S_1 and S_2, respectively. When a current I_2 is set up through S_2, it in turn sets up a magnetic flux through S_1. Let us denote it by Φ_1. The corresponding flux linkage with solenoid S_1 is

$$N_1\Phi_1 = M_{12}I_2$$

M_{12} is called the mutual inductance of solenoid S_1 with respect to solenoid S_2. It is also referred to as the coefficient of mutual induction. For these simple co-axial solenoids, it is possible to calculate M_{12}. The magnetic field due to the current I_2 in S_2 is $\mu_0 n_2 I_2$. The resulting flux linkage with coil S_1 is,

$$N_1\Phi_1 = NBA = (n_1 l)\,(\mu_0 n_2 I_2)\,(\pi r_1^2) = M_{12}I_2 \tag{35}$$

$$N_1\Phi_1 = NBA = (\boldsymbol{\mu_0 \pi\, n_1 n_2 l\, r_1^2})I_2 = M_{12}I_2 \tag{36}$$

where $n_1 l$ is the total number of turns in solenoid S_1. Therefore, mutual induction coefficient is

$$M_{12} = \boldsymbol{\mu_0 \pi\, n_1 n_2 l r_1^2} \quad (37)$$

Note that we neglected the edge effects and considered the magnetic field $\mu_0 n_2 I_2$ to be uniform throughout the length and width of the solenoid S_2. This is a good approximation keeping in mind that the solenoid is long, implying $l >> r_2$.We now consider the reverse case. A current I_1 is passed through the solenoid S1 and the flux linkage with coil S2 is,

$$N_2\Phi_2 = M_{21} I_1 \quad (38)$$

M_{21} is called the mutual inductance of solenoid S2 with respect to solenoid S_1.The flux due to the current I_1 in S_1 can be assumed to be confined solely inside S_1 since the solenoids are very long. Thus, flux linkage with solenoid S_2 is

$$N_2\Phi_2 = NBA = (n_2 l)\,(\mu_0 n_1 I_1)\,(\pi r_1^2) = M_{21} I_1 \quad (39)$$

$$\text{Or} \qquad N_2\Phi_2 = NBA = (\,\boldsymbol{\mu_0 \pi\, n_1 n_2 l\, r_1^2}\,)\, I_1 = M_{21} I_1 \quad (40)$$

where$n_2 l$ is the total number of turns in solenoid S_2. Therefore, mutual induction coefficient is

$$M_{21} = \boldsymbol{\mu_0 \pi\, n_1 n_2 l r_1^2} \quad (41)$$

It is clear from equation (37) and (41) that $M_{21} = M_{12}$. Above result have demonstrated this equality for long co-axial solenoids. However, the relation is far more general. In case if the inner solenoid was much shorter than (and placed well inside) the outer solenoid, then we could still have calculated the flux linkage $N_1\Phi_1$because the inner solenoid is effectively immersed in a uniform magnetic field due to the outer solenoid. In this case, the calculation of M_{12} would be easy. However, it would be extremely difficult to calculate the flux linkage with the outer solenoid as the magnetic field due to the inner solenoid would vary across the length as well as cross section of the outer solenoid. Therefore, the calculation of M_{21} would also be extremely difficult in this case. The equality $M_{12}=M_{21}$ is very useful in such situations. It is also important to know that the mutual inductance of a pair of coils, solenoids, etc., depends on their separation as well as their relative orientation.

2.9 VARIABLE INDUCTANCE

Variable inductance transducer uses the inductive effect in its operation. Various physical causes like- pressure, displacement, force, sound etc often transform and change the materials' self-inductance [L] or mutual inductance [M]. Inductance are handy with formula $L = \mu_0 n^2$ Al and M=

$\mu_0 \pi n_1 n_2 l r_1^2$. Here μ is permeability, l is length, A is cross section area where flux is established, l is length of concerned part and I is current in coil. They can be briefly categorized into four class:

Magnetic Circuit Transducer.

- Principle of operation: L and M of ac excited coil is varied by changes in magnetic circuits [or flux].
- Typical Applications: Measurement of pressure, displacement.

Reluctance pick up.

- Principle of operation: Reluctance, R is carried by changing the position of iron core of coil.
- Typical Applications: Measurement of pressure, displacement, vibrations, positions.

Differential Transformer [popularly known as LVDT].

- Principle of operation: Differential voltage of two secondary windings is varied by positioning the magnetic core through externally applied force.
- Typical Applications: Measurement of pressure, displacement, positions.

Magnetostriction gauge

- Principle of operation: Magnetic properties are varied by pressure and stress.
- Typical Applications: Measurement of pressure, sound.

STRAY INDUCTANCE

Stray inductance is unintended and unwanted inductance in a circuit. Inductance does not exist only within inductors. In fact, any wires or component leads that have current flowing through them create magnetic fields. When these magnetic fields are created, they can produce an inductive effect. Thus, wires or components leads can act as inductors if they are long enough. Such effects are often present within circuits (for example, between conductive runs of wire traces or components with long leads such as capacitors), even though they are not intended. This unintended inductance is referred to as stray inductance, and it can result in a disruption of normal current flow within a circuit.

2.13 INDUCTANCE IN COMBINATION

Like resistance and capacitor, inductors are also connected in series and parallel combination in varies circuit as desire in the circuit.

2.13.1 Series Inductor Circuit

In the series inductor circuit, the number of inductors is connected in series in the circuit, and the same amount of current will flow in each of the inductors connected. For example, if L_1, L_2, L_3…… inductors are connected in series and current I flow through the circuit as shown in the figure below

Figure 2.12: Inductor in series.

The current across inductor L_1, L_2, L_3 will be same I. The value of voltage across L_1, L_2, L_3 will be $\varepsilon_1, \varepsilon_2, \varepsilon_3$ respectively. However total voltage across MN is the sum of voltage across each inductor, thus

$$\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$$

Substituting the values, this becomes

$$L_{eq}\frac{dI}{dt} = L_1\frac{dI}{dt} + L_2\frac{dI}{dt} + L_3\frac{dI}{dt}$$

The total or equivalent inductance will be given by the equation

$$L_{eq} = L_1 + L_2 + L_3$$

2.13.2 Parallel Inductor Circuit

If the number of inductors is connected parallel with each other than the circuit is known as a parallel inductor circuit. In this type of circuit, the circuit is divided into each branch of the circuit as shown in the figure below

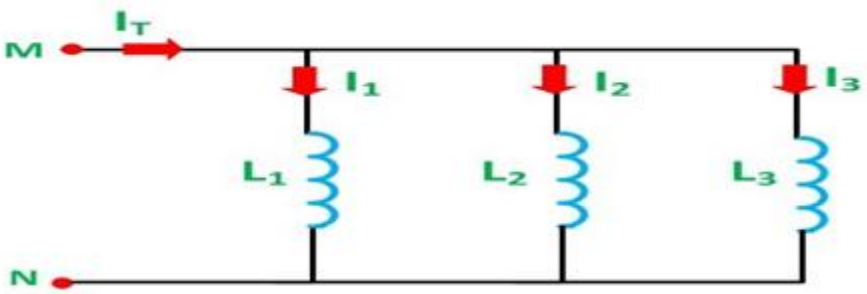


Figure 2.13: Inductor in parallel.

Since voltage across each coil in a parallel circuit remains same. Thus

$$\varepsilon = \varepsilon_1 = \varepsilon_2 = \varepsilon_3$$

Using equation (24) we can write :$L_{eq}\frac{dI}{dt} = \varepsilon$ or $\frac{dI}{dt} = \frac{\varepsilon}{L_{eq}}$, similarly$\frac{dI_1}{dt} = \frac{\varepsilon_1}{L_1} = \frac{\varepsilon}{L_1}$, $\frac{dI_2}{dt} = \frac{\varepsilon_2}{L_2} = \frac{\varepsilon}{L_2}$ and $\frac{dI_3}{dt} = \frac{\varepsilon_3}{L_3} = \frac{\varepsilon}{L_3}$.

Current I_1 is flowing in the inductor L_1, and similarly, current I_2 in L_2 and I_3 in L_3inductor and I is the total amount of current flowing in the circuit and it is sum of I_1, I_2 and I_3. Hence,

$$I = I_1 + I_2 + I_3$$

Differentiating this equation with time, we get

$$\frac{dI}{dt} = \frac{dI_1}{dt} + \frac{dI_2}{dt} + \frac{dI_3}{dt}$$

Or$\frac{\varepsilon}{L_{eq}} = \frac{\varepsilon}{L_1} + \frac{\varepsilon}{L_2} + \frac{\varepsilon}{L_3}$

The equivalent inductance is given by the equation shown below

$$\frac{1}{L_{eq}} = \frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3}$$

2.14 ENERGY STORED IN AN INDUCTOR

The self-induced emf is also called the back emf as it opposes any change in the current in a circuit. Physically, the self-inductance plays the role of inertia. It is the electromagnetic analogue of mass in mechanics. So, work needs to be done against the back emf (ε) in establishing the current. This work done is stored as magnetic potential energy. For the current I at an instant in a circuit, the rate of work or The power absorbed by the inductor is given by the equation shown below

$$P = \frac{dW}{dt} = \varepsilon . I = (L\frac{dI}{dt}) . I = LI . \frac{dI}{dt}$$

The energy absorbed by the inductor is given as

$$W = \int_0^t P\, dt = \int_0^t L\, I\, . \frac{dI}{dt}\, dt = \int_0^t L\, I\, .\, dI = \frac{1}{2} L\, I^2$$

This expression reminds us of mv2/2 for the (mechanical) kinetic energy of a particle of mass m, and shows that L is analogous to m (i.e., L is electrical inertia and opposes growth and decay of current in the circuit). The inductor stores a finite amount of energy even though the voltage across it may be negligible. The inductors are classified depending upon various factors like the size, core material used, type of windings, etc. The core plays an important role in the selection of the inductor. Based on the core material the various types of inductor are as follows

- Ferromagnetic or iron core inductor
- Air core inductor
- Toroidal core inductor
- Laminated core inductor
- Powered core inductor

This energy is actually stored in the magnetic field generated by the current flowing through the inductor. In a pure inductor, the energy is stored without loss, and is returned to the rest of the circuit when the current through the inductor is ramped down, and its associated magnetic field collapses. Consider the case of simple solenoid whose $L = \mu_0 \frac{N^2}{L}$ Aand magnetic field inside is $B = \mu_0 n I$.Using this in equation (41) and can be combined to give

$$W = \frac{1}{2} L\, I^2 = \frac{1}{2}\mu_0 \frac{N^2}{l} A\left(\frac{Bl}{\mu 0 N}\right)^2$$

This reduces to

$$W = \frac{B^2}{2\mu 0} Al$$

This represents the energy stored in the magnetic field of the solenoid. However, the volume of the field-filled core of the solenoid is Al, so the magnetic energy density (*i.e.*, the energy per unit volume) inside the solenoid is,

$$U = \frac{W}{volume} = \frac{B^2}{2\mu 0}$$

It turns out that this result is quite general. Thus, we can calculate the energy content of any magnetic field by dividing space into little cubes (in each of which the magnetic field is approximately uniform), applying the above formula to find the energy content of each cube, and summing the energies thus obtained to find the total energy. When electric and magnetic fields exist together in space, then both can be combined to give an expression for the total energy stored in the combined fields per unit volume

$$U = \frac{electrical\ field\ energy + magnetic\ field\ energy}{volume}$$

$$U = \frac{1}{2}\varepsilon_0 E^2 + \frac{B^2}{2\mu 0}$$

2.15 CAPACITOR AND ITS TYPES

An arrangement of two conducting surfaces separated by an insulating material is called a **capacitor** or **condenser.** Its purpose is to store charge in a small space. The conducting surfaces are called the plates of the capacitor and the insulating material is called the dielectric. The most commonly used dielectrics are air, mica, waxed paper, ceramics etc. The following points may be noted carefully

- The ability of a capacitor to store charge (i.e. its capacitance) depends upon the area of plates, distance between plates and the nature of insulating material (or dielectric).
- A capacitor is generally named after the dielectric used e.g. air capacitor, paper capacitor, mica capacitor etc.
- The capacitor may be in the form of parallel plates, concentric cylinder or other arrangement.

How does a capacitor store charge?

Fig. 2.14 shows how a capacitor stores charge when connected to a d.c. supply. The parallel plate capacitor having plates A and B is connected across a battery of V volts as shown in Fig. 2.14(i). When the switch S is open as shown in Fig. 2.14(i), the capacitor plates are neutral i.e. there is no charge on the plates. When the switch is closed as shown in Fig. 2.14 (ii), the electrons from plate A will be attracted by the +ve terminal of the battery and these electrons start accumulating on plate B. The result is that plate A attains more and more positive charge and plate B gets more and more negative charge. This action is referred to as charging a capacitor because the capacitor plates are becoming charged. This process of electron flow or charging (i.e. detaching electrons from

plate A and accumulating on B) continues till p.d. across capacitor plates becomes equal to battery voltage V. When the capacitor is charged to battery voltage V, the current flow ceases as shown in Fig. 2.14(iii). If now the switch is opened as shown in Fig. 2.14 (iv), the capacitor plates will retain the charges. Thus the capacitor plates which were neutral to start with now have charges on them. This shows that a capacitor stores charge. The following points may be noted about the action of a capacitor:

- When a d.c. potential difference is applied across a capacitor, a charging current will flow until the capacitor is fully charged when the current will cease. This whole charging process takes place in a very short time, a fraction of a second. Thus a capacitor once charged, prevents the flow of direct current.

- The current does not flow through the capacitor i.e. between the plates. There is only transference of electrons from one plate to the other.

- When a capacitor is charged, the two plates carry equal and opposite charges (say + Q and –Q). This is expected because one plate loses as many electrons as the other plate gains. Thus charge on a capacitor means charge on either plate.

- The energy required to charge the capacitor (i.e. transfer of electrons from one plate to the other) is supplied by the battery.

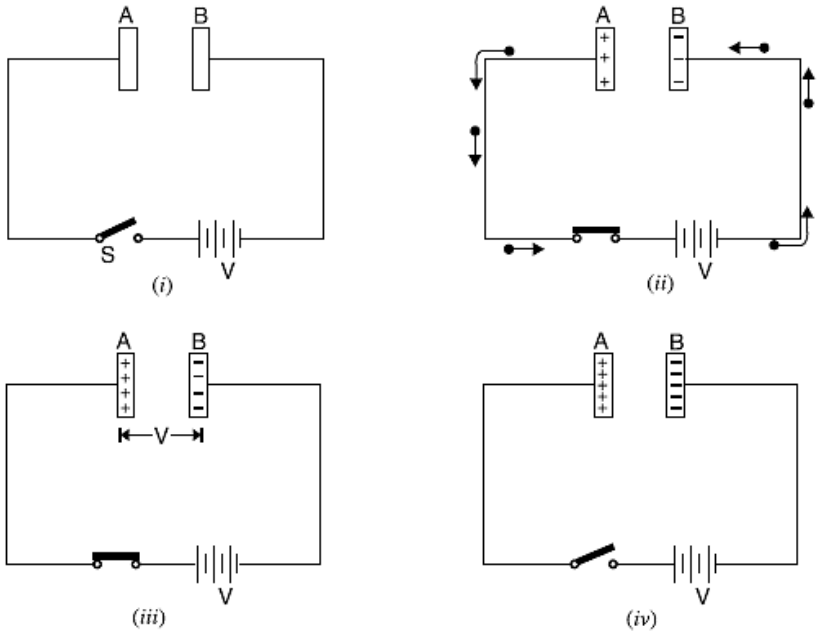


Figure 2.14: Mechanism of capacitor holding charge.

Capacitance

The property of a capacitor to 'store electricity' may be called its capacitance. As we may measure the capacity of a tank, not by the total mass or volume of water it can hold, but by the mass in kg of water required to raise its level by one metre, similarly, the capacitance of a capacitor is defined as "the amount of charge required to create a unit p.d. between its plates."

Suppose we give Q coulomb of charge to one of the two plate of capacitor and if a p.d. of V volts is established between the two, then its capacitance is

$$C = \frac{Q}{V} = \frac{charge}{potential\ difference}$$

Hence, capacitance is the charge required per unit potential difference. By definition, the unit of capacitance is coulomb/volt which is also called farad (in honor of Michael Faraday)

1 farad = 1 coulomb/volt

One farad is defined as the capacitance of a capacitor which requires a charge of one coulomb to establish a p.d. of one volt between its plates. One farad is actually too large for practical purposes. Hence, much smaller units like microfarad(μF), Nano farad (nF) and micro-microfarad (μμF) or picofarad (pF) are generally employed.

$$1\mu F = 10^{-6}F\ ;\ 1nF = 10^{-9}F\ ;\ 1pF = 10^{-12}F$$

Factors Affecting Capacitance

The ability of a capacitor to store charge (i.e. its capacitance) depends upon the following factors:

(i) Area of plate. The greater the area of capacitor plates, the larger is the capacitance of the capacitor and vice-versa. It is because larger the plates, the greater the charge they can hold for a given p.d. and hence greater will be the capacitance.

(ii) Thickness of dielectric. The capacitance of a capacitor is inversely proportional to the thickness (i.e. distance between plates) of the dielectric. The smaller the thickness of dielectric, the greater the capacitance and vice-versa. When the plates are brought closer, the electrostatic field is intensified and hence capacitance increases.

(iii) Relative permittivity of dielectric. The greater the relative permittivity of the insulating material (i.e., dielectric), the greater will be the capacitance of the capacitor and vice-versa. It is

because the nature of dielectric affects the electrostatic field between the plates and hence the charge that accumulates on the plates.

Capacitance of Parallel-Plate Capacitor with Uniform Medium

We have seen that the capacitance of a capacitor can be determined from its electrical properties using the relation C = Q/V. However, it is often desirable to determine the capacitance of a capacitor in terms of its dimensions and relative permittivity of the dielectric. Although there are many forms of capacitors, the most important arrangement is the parallel-plate capacitor. Consider a parallel plate capacitor consisting of two plates, each of area A square meters and separated by a uniform dielectric of thickness d meters and relative permittivity $\varepsilon_r$as shown in Fig. Let a p.d. of V volts applied between the plates place a charge of +Q and − Q on the plates. With reasonable accuracy, it can be assumed that electric field between the plates is uniform.

Electric displacement between plates is D = Q/A coulomb/m^2

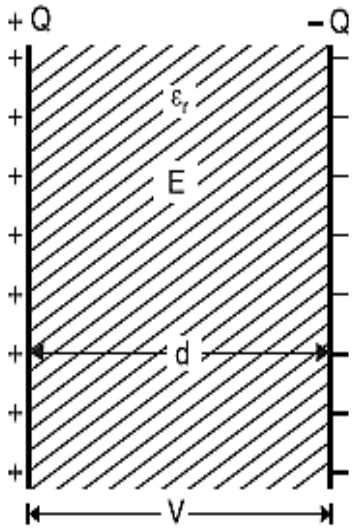


Figure 2.15 (a) Capacitor with uniform medium.

Electric intensity between plates is E = V/d

But$D = \varepsilon_o \varepsilon_r E$

or$\frac{Q}{A} = \varepsilon_o \varepsilon_r \frac{V}{d}$

or$\frac{Q}{V} = \varepsilon_o \varepsilon_r \frac{A}{d}$

The ratio Q/V is the capacitance C of the capacitor

$$C = \varepsilon_o \varepsilon_r \frac{A}{d} \quad in\ medium$$

$$C = \varepsilon_o \frac{A}{d} \qquad in\ air$$

Parallel-Plate Capacitor with Composite Medium

Suppose the space between the plates is occupied by three dielectrics of thicknessesd_1, d_2 and d_3 metres and relative permittivities ε_{r1}, $\varepsilon_{r2}$and ε_{r3} respectively as shown in Fig. Theelectric flux density or electric displacement D in the dielectrics remains the same and is equal to Q/A. However, the electric intensities in the three dielectrics will be different and are given by;

$$E_1 = \frac{D}{\varepsilon_o \varepsilon_{r1}};\ E_2 = \frac{D}{\varepsilon_o \varepsilon_{r2}};\ E_3 = \frac{D}{\varepsilon_o \varepsilon_{r3}}$$

If V is the total p.d. across the capacitor and V_1, V_2 and V_3 the p.d.s. across the three dielectricsrespectively, then

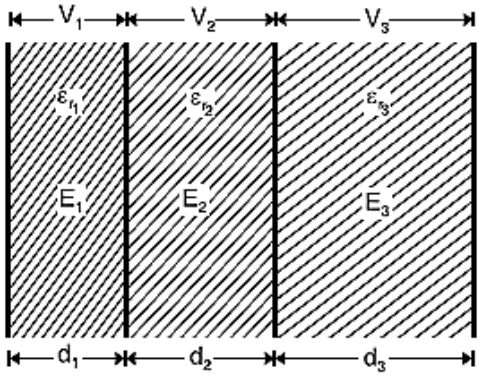


Figure 2.15: (b) Capacitor with composite medium.

$$V = V_1 + V_2 + V_3$$

$$= E_1 d_1 + E_2 d_2 + E_3 d_3$$

$$= \frac{D}{\varepsilon_o \varepsilon_{r1}} d_2 + \frac{D}{\varepsilon_o \varepsilon_{r2}} d_2 + \frac{D}{\varepsilon_o \varepsilon_{r3}} d_3$$

$$= \frac{D}{\varepsilon_o}\left[\frac{d_1}{\varepsilon_{r1}} + \frac{d_2}{\varepsilon_{r2}} + \frac{d_3}{\varepsilon_{r3}}\right]$$

$$= \frac{Q}{\varepsilon_o A}\left[\frac{d_1}{\varepsilon_{r1}} + \frac{d_2}{\varepsilon_{r2}} + \frac{d_3}{\varepsilon_{r3}}\right] \quad \left(since\ D = \frac{Q}{A}\right)$$

Or $$C = \frac{Q}{V} = \frac{\varepsilon_o A}{\left[\frac{d_1}{\varepsilon_{r1}} + \frac{d_2}{\varepsilon_{r2}} + \frac{d_3}{\varepsilon_{r3}}\right]}$$

But Q/V is the capacitance C of the capacitor. Therefore,

$$C = \frac{\varepsilon_o A}{\left[\frac{d_1}{\varepsilon_{r1}} + \frac{d_2}{\varepsilon_{r2}} + \frac{d_3}{\varepsilon_{r3}}\right]}\ farad$$

Spherical Capacitor

(a) When outer sphere is earthed

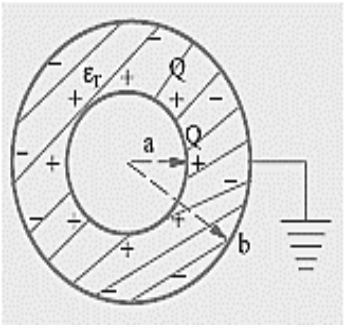


Figure 2.15: (C) Spherical Capacitor with outer sphere earthed.

Consider a spherical capacitor consisting of two concentric spheres of radii 'a' and 'b' metres as shown in Fig. Suppose, the inner sphere is given a charge of + Q coulombs. It will induce a charge of − Q coulombs on the inner surfaces which will go to earth. If the dielectric medium between the two spheres has a relative permittivity of ε_r, then the free surface potential of the innersphere due to its own charge $Q/4\pi\varepsilon_o\varepsilon_r$a volts. The potential of theinner sphere due to − Q charge on the outer sphere is − $Q/4\pi\varepsilon_o\varepsilon_r$b (remembering that potential anywhere inside a sphere is the same as thatits surface).
Total potential difference between two surfaces is

$$V = \frac{Q}{4\pi\varepsilon_o\varepsilon_r a} - \frac{Q}{4\pi\varepsilon_o\varepsilon_r b}$$

$$= \frac{Q}{4\pi\varepsilon_o\varepsilon_r}\left[\frac{1}{a} - \frac{1}{b}\right] = \frac{Q}{4\pi\varepsilon_o\varepsilon_r}\left[\frac{b-a}{ab}\right]$$

$$\frac{Q}{V} = \frac{4\pi\varepsilon_o\varepsilon_r ab}{b-a}$$

Therefore, we have,

$$C = \frac{4\pi\varepsilon_o\varepsilon_r ab}{b-a}\ farad$$

(b) When inner sphere is earthed

Such a capacitor is shown in Fig. If a charge of + Q coulombs is given to the outer sphere A, it will distribute itself over both its inner and outer surfaces. Some charge Q_2coulomb will remain on the outer surface of A because it is surrounded by earth all around. Also, some charge + Q_1 coulombs will shift to its inner side because there is an earthed sphere B inside A.Obviously, $Q = Q_1 + Q_2$. The inner charge +Q_1 coulomb on A induces$-Q_1$coulomb on B but the other induced charge of + Q_1coulomb goes to earth. Now, there are two capacitors connected in parallel :

(i) One capacitor consists of the inner surface of A and the outer surface of B. Its capacitance, as found earlier, is

$$C_1 = \frac{4\pi\varepsilon_o\varepsilon_r ab}{b-a}$$

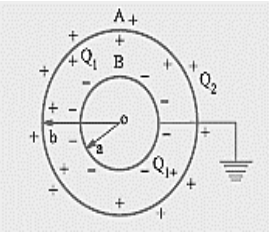


Figure 2.15: (d) Spherical capacitor with inner sphere grounded

(ii) The second capacitor consists of outer surfaces of B and earth. Its capacitance is C_2= 4 $\pi\varepsilon_o b-$ if surrounding medium is air. Total capacitance C = $C_1 + C_2$.

Cylindrical Capacitor

A cylindrical capacitor consists of two co-axial cylinders separated by an insulating medium. This is an important practical case since a single core cable is in effect a capacitor of this kind. The conductor (or core) of the cable is the inner cylinder while the outer cylinder is represented by lead sheath which is at earth potential. The two co-axial cylinders have insulation between them.

Consider a single core cable with conductor diameter d metres and inner sheath diameter D metres (Fig.).Let the charge per metre axial length of the cable be Q coulombs and $\varepsilon_r$be the relative permittivity of the insulating material. Consider a cylinder of radius x metres. According to Gauss's theorem, electric flux passing through this cylinder is Q coulombs. The surface area of this cylinder is

$$= 2\pi x \times 1 = 2\pi x\ m^2$$

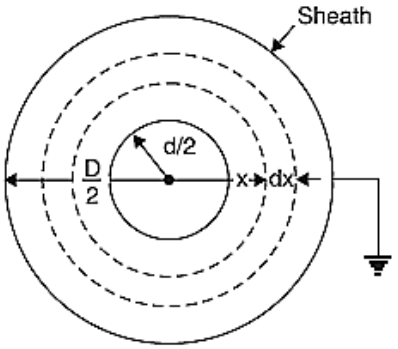


Figure 2.15: (e) Cylindrical capacitor.

Electric flux density at any point P on the considered cylinder is given by;

$$D_x = \frac{Q}{2\pi x}\ C/m^2$$

Electric intensity at point P is given by;

$$E_x = \frac{D_x}{\varepsilon_0 \varepsilon_r} = \frac{Q}{2\pi x \varepsilon_0 \varepsilon_r}\ V/m$$

The work done in moving a unit positive charge from point P through a distance dx in the direction of electric field is $E_x$dx. Hence the work done in moving a unit positive charge from conductor to sheath, which is the p.d. V between the conductor and sheath, is given by;

$$V = \int_{d/2}^{D/2} E_x dx = \int_{d/2}^{D/2} \frac{Q}{2\pi x \varepsilon_0 \varepsilon_r} dx = \frac{Q}{2\pi x \varepsilon_0 \varepsilon_r} log_e \frac{D}{d}$$

Capacitance is given by

$$C = \frac{Q}{V} = \frac{Q}{\frac{Q}{2\pi x \varepsilon_0 \varepsilon_r} log_e \frac{D}{d}} = \frac{2\pi x \varepsilon_0 \varepsilon_r}{log_e \frac{D}{d}}\ farad\ per\ m$$

If the cable has a length of l metres, then capacitance of the cable is

$$C = \frac{2\pi x \varepsilon_0 \varepsilon_r}{log_e \frac{D}{d}}\ l \quad farad$$

Multiple and Variable Capacitors

Multiple capacitors are shown in Fig.{2.15(g)} (a) and Fig. {2.15(g)} (b). The arrangement of Fig. {2.15(g)} (a)is equivalent to two capacitors joined in parallel. Hence, its capacitance is double that of a single capacitor. Similarly, the arrangement of Fig. {2.15(g)} (b) has four times the

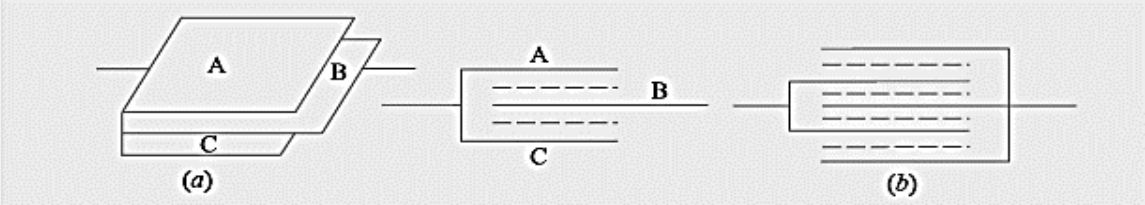


Figure 2.15: (g) Multiple capacitors.

capacitance of single capacitor.

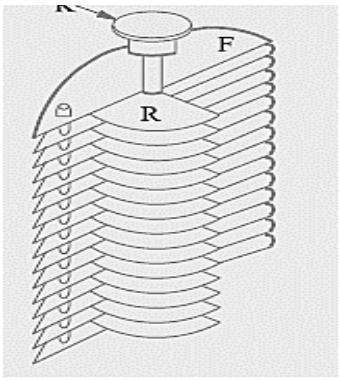


Figure 2.15: (f)
Variable capacitor.

If one set of plates is fixed and the other is capable of rotation, then capacitance of such a multiplate capacitor can be varied. Such variable capacitance air capacitors are widely used in radio work (Fig). The set of fixed plates F is insulated from the other set R which can be rotated by turning the knob K. The common area between the two sets is varied by rotating K, hence the capacitance between the two is altered. Minimum capacitance is obtained when R is completely rotated out of F and maximum when R is completely rotated in i.e. when the two sets of plates completely overlap each other. The capacitance of such a capacitor is

$$C = \frac{(n-1)\varepsilon_o \varepsilon_r A}{d}$$

Where n is the number of plates which means that (n − 1) is the number of capacitors.

2.16 COMBINATION OF CAPACITORS

2.16.1 Series Combination

With reference of Fig. 2.16, let$C_1,\ C_2, C_3$arecapacitances of three capacitors and $V_1,\ V_2,\ V_3$ are p.d. across these three capacitors, where

V = applied voltage across combination

C = combined or equivalent capacitance

In series combination, charge on all capacitors is the same but p.d. across each is different.

$$V = V_1 + V_2 + V_3$$

or

$$\frac{Q}{C} = \frac{Q}{C_1} + \frac{Q}{C_2} + \frac{Q}{C_3}$$

or

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_1} + \frac{1}{C_1}$$

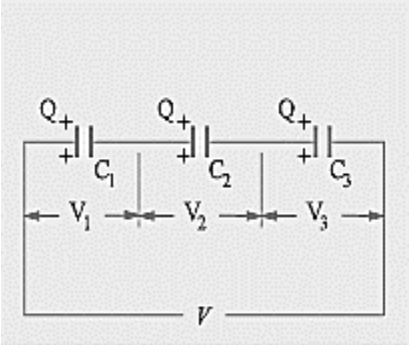


Figure 2.16: (a) Series combination.

2.16.2 Parallel Combination

Consider three capacitors, having capacitancesC_1, C_2 and C_3 farad respectively, connected inparallel across a p.d. of V volts [See Fig. 2.16 (b)]. In parallel connection, p.d. across each capacitor is the same but charge on each is different. Therefore, total charge

$$Q = Q_1 + Q_2 + Q_3 = C_1V + C_2V + C_3V$$

$$= V(C_1 + C_2 + C_3)$$

or

$$\frac{Q}{V} = (C_1 + C_2 + C_3)$$

But Q/V is the total capacitance of the parallel combination

$$C = C_1 + C_2 + C_3$$

Thus capacitors in parallel are treated in the same manner as are resistors in series.

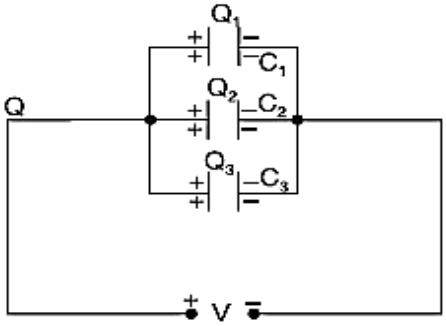


Figure 2.16: (b) Parallel combination of capacitors.

2.17 ENERGY IN ELECTROSTATIC FIELD OF CAPACITOR

Charging a capacitor means transferring electrons from one plate of the capacitor to the other. This involves expenditure of energy because electrons have to be moved against the opposing forces.

This energy is stored in the electrostatic field set up in the dielectric medium. On discharging the capacitor, the field collapses and the stored energy is released.

Consider a capacitor of C farad being charged from a d.c.source of V volts as shown in Fig. Suppose at any stage of charging, the charge on the capacitor is q coulomb and p.d. across the plates is v volts

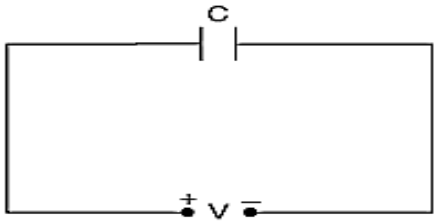


Figure 2.17: Energy stored capacitor.

Then

$$C = \frac{q}{V}$$

At this instant, v joules (by definition of v) of work will be done in transferring 1 C of charge from one plate to the other. If further small charge dq is transferred, then work done is

$$dW = V\ dq$$
$$= CV\ dV$$

Total work done in raising the potential of uncharged capacitor to V volts is

$$W = \int_0^V C\ V\ dV = C\ \frac{V^2}{2}$$

This work done is stored in the electrostatic field set up in the dielectric.

Energy stored in the capacitor in terms of Q and C are also express as,

$$E = \frac{1}{2}\ CV^2 = \frac{1}{2}QV = \frac{Q^2}{2C}\ joules$$

2.18 FILTER CIRCUIT

The reactance of inductors and capacitors depend on the frequency of the a.c. signal applied to them. That is why these devices are known as frequency-selective. By using various combinations of resistors, inductors and capacitors, we can make circuits that have the property of passing or rejecting either low or high frequencies or bands of frequencies. These frequency selective networks, which alter the amplitude and phase characteristics of the input a.c. signal, are called filters. Their performance is usually expressed in terms of how much attenuation a band of frequencies experiences by passing through them. Attenuation is commonly expressed in terms of decibels (dB).

Types of filters

A.C. filter networks are divided into two major categories: (i) active networks and (ii) passive networks. Active filter networks usually contain transistors and/or operational amplifiers in combination with R, L and C elements to obtain the desired filtering effect. On the other side passive filter networks which usually consist of series-parallel combinations of R, L and C elements. There are four types of such passive networks, as described below:

1. Low-Pass Filter. As the name shows, it allows only low frequencies to pass through, but attenuates (to a lesser or greater extent) all higher frequencies. The maximum frequency which it allows to pass through, is called cutoff frequency f_c(also called break frequency). There are RL and RClow-pass filters.

2. High-Pass Filter. It allows signals with higher frequencies to pass from input to output while rejecting lower frequencies. The minimum frequency it allows to pass is called cutoff frequency f_c. There are $RL$and RChigh pass filters.

3. Bandpass Filter. It is a resonant circuit which is tuned to pass a certain band or range of frequencies while rejecting all frequencies below and above this range (called passband).

4. Bandstop Filter. It is a resonant circuit that rejects a certain band or range of frequencies while passing all frequencies below and above the rejected band. Such filters are also called wave traps, notch filters or band-elimination, band-separation or band rejection filters.

Applications. A.C. filters find application in audio systems and television etc. Bandpass filters are used to select frequency ranges corresponding to desired radio or television station channels. Similarly, bandstop filters are used to reject undesirable signals that may contaminate the desirable signal. For example, low-pass filters are used to eliminate undesirable hum in d.c. power supplies.

The Decibel System

These systems of logarithmic measurement are widely used in audio, radio, TV and instrument industry for comparing two voltages, currents or power levels. These levels are measured in a unit called bel (B) or decibel (dB) which is 1/10th of a bel.

Suppose we want to compare the output power P_o of a filter with its input power P_i. The power level change is

$$= 10\, log_{10} \frac{P_o}{P_i}\ dB$$

It should be noted that dB is the unit of power change (i.e. increase or decrease) and not of power itself. Moreover, 20 dB is not twice as much power as 10 dB.

However, when voltage and current levels are required, then the expressions are:

$$current\ level = 20\,log_{10}\frac{I_o}{I_i}\ dB$$

$$voltage\ level = 20\,log_{10}\frac{V_o}{V_i}\ dB$$

2.19 LOW PASS AND HIGH PASS FILTERS

Low Pass RC Filter

A simple low-pass RC filter is shown in Fig.2.19 (a). As stated earlier, it permits signals of low frequencies upto f_c to pass through while attenuating frequencies above f_c. The range offrequencies upto$f_c$is called the passband of the filter. Fig. 2.19 (b) shows the frequency response curve of such a filter. It shows how the signal output voltage V_o varies with the signal frequency. As seen at f_c, output signal voltage is reduced to 70.7% of the input voltage. The output is said to be – 3 dB at f_c. Signal outputs beyond $f_c$roll-off or attenuate at a fixed rate of – 6 dB/octave or – 20 dB/decade. As seen from the frequency-phase response curve of Fig. 2.19 (c), the phase angle between $V_o$and $V_i$is 45° at cutoff frequency f_c.

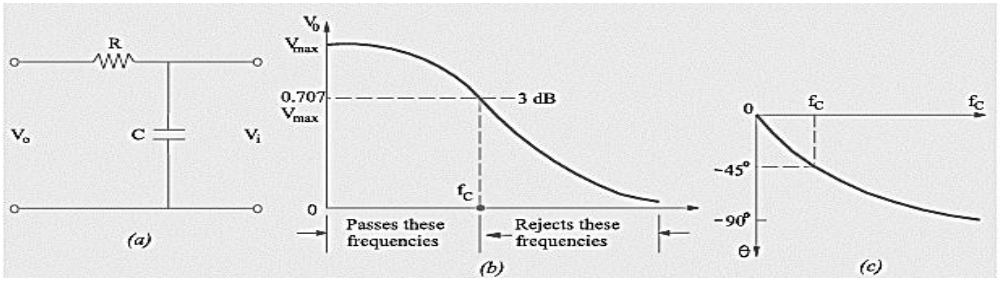


Figure 2.19: Low pass RC filter.

By definition cutoff frequency f_c occurs where **(a)** V_o = 70.7% V_i i.e. V_o is – 3 dB down from V_i **(b)** R = X_c and V_R = V_C in magnitude. **(c)** The impedance phase angle θ = – 45°. The same is the angle between V_o and V_i.

As seen, the output voltage is taken across the capacitor. Resistance R offers fixed opposition to frequencies but the reactance offered by capacitor C decreases with increase in frequency. Hence, low-frequency signal develops over C whereas high-frequency signals are grounded. Signal frequencies above f_c develop negligible voltage across C. Since R and C are in series, we can findthe low-frequency output voltage V_o developed across C by using the voltage-divider rule.

$$V_0 = V_i \frac{jX_c}{R + jX_c} \; and \; f_c = \frac{1}{2CR}$$

Low Pass RL Filter

It is shown in Fig. 2.20 (a). Here, coil offers high reactance to high frequencies and low reactance to low frequencies. Hence, low frequencies upto fc can pass through the coil without much opposition. The output voltage is developed across R. Fig. 2.20 (b) shows the frequency output response curve of the filter. As seen at fc, Vo = 0.707 Vi and its attenuation level is –3 dB with respect to Vo i.e. the voltage at f = 0. However, it may be noted that being an RL circuit, the

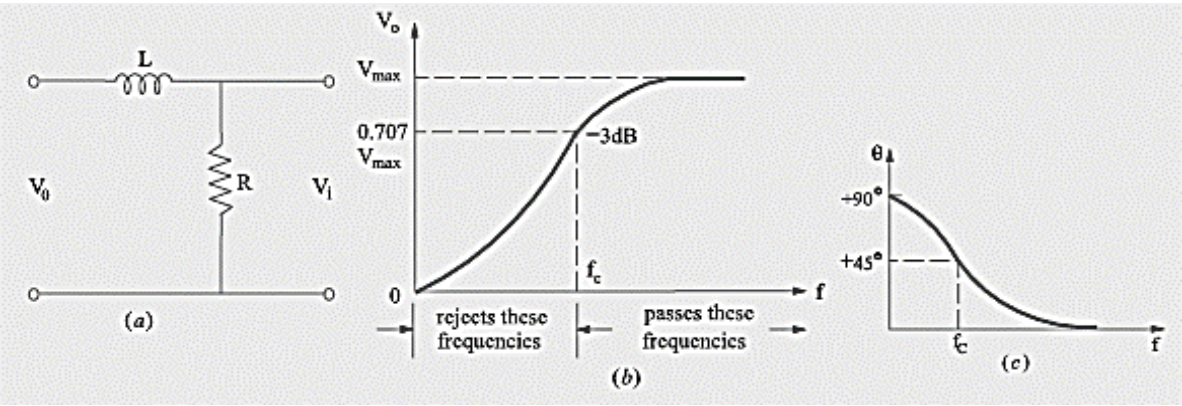


Figure 2.20: Low pass RL filter.

impedance phase angle is +45° (and not –45° as in low-pass RC filter). Again at f_c, R = X_L.Using the voltage-divider rule, the output voltage developed across R is given by

$$V_0 = V_i \frac{jX_L}{R + jX_L} \; and \; f_c = \frac{R}{2\pi L}$$

Other Types of Low-Pass Filters

There are many other types of low-pass filters in which instead of pure resistance, series chokes are commonly used along with capacitors.

(i) Inverted–L Type. It is shown in Fig. 2.21 (a). Here, inductive reactance of the choke blocks higher frequencies and C shorts them to ground. Hence, only low frequencies below fc (for which X is very low) are passed without significant attenuation.

(ii) T-Type. It is shown in Fig. 2.21 (b). In this case, a second choke is connected on the output side which improves the filtering action.

(iii) π-Type. It is shown in Fig. 2.21 (c). The additional capacitor further improves the filtering action by grounding higher frequencies.

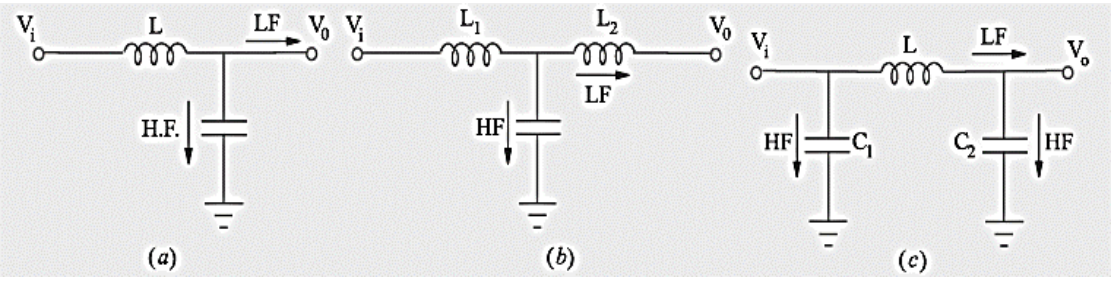


Figure 2.21: Different types of low pass filter.

It would be seen from the above figures that choke is always connected in series between the input and the output and capacitors are grounded in parallel. The output voltage is taken across the capacitor.

High-Pass RC Filter

It is shown in Fig. 2.22 (a). Lower frequencies experience considerable reactance by the capacitor and are not easily passed. Higher frequencies encounter little reactance and are easily passed. The high frequencies passing through the filter develop output voltage V_o across R. As seen from the frequency response of Fig. 2.22 (b), all frequencies above f_c are passed whereas those below it are attenuated. As before, f_c corresponds to –3 dB output voltage or half-power point. At f_c, R = X_C and the phase angle between V_o and V_i is +45° as shown in Fig. 2.22 (c). It may be noted that high-pass RC filter can be obtained merely by interchanging the positions of R and C in the low-pass RC filter of Fig. 2.22 (a).

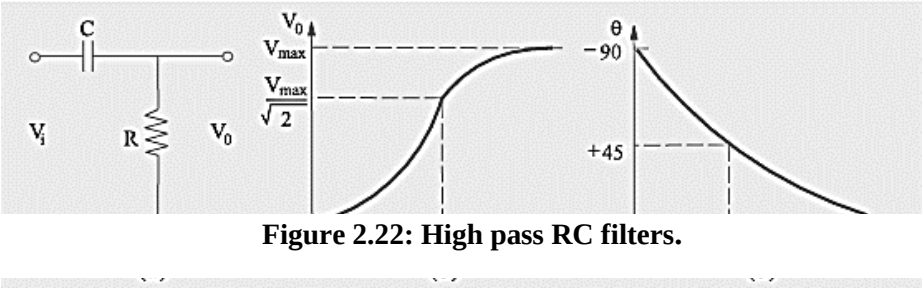


Figure 2.22: High pass RC filters.

Since R and C are in series across the input voltage, the voltage drop across R, as found by the voltage-divider rule, is

$$V_0 = V_i \frac{jX_c}{R + jX_c} \; and \; f_c = \frac{1}{2CR}$$

A very common application of the series capacitor high-pass filter is a coupling capacitor between two audio amplifier stages. It is used for passing the amplified audio signal from one stage to the next and simultaneously block the constant d.c. voltage.

High-Pass RL Filter

It is shown in Fig. 2.23 and can be obtained by 'swapping' position of R and L in the low-pass RL circuit of Fig. 2.20 (a). Its response curves are the same as for high-pass RC circuit and are as shown above. As usual, its output voltage equals the voltage which drops across XL. It is given by

$$V_0 = V_i \frac{jX_L}{R + jX_L} \; and \; f_c = \frac{R}{2\pi L}$$

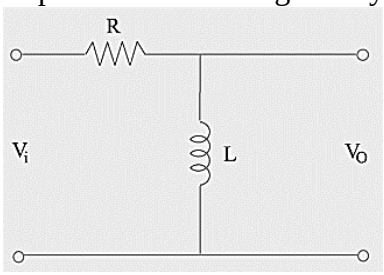


Figure 2.23: High pass RL filters.

2.20 SUMMARY

- Current through a given area of a conductor is the net charge passing per unit time through the area.
- Ohm's law: The electric current I flowing through a substance is proportional to the voltage V across its ends, i.e., V α I or V = RI, where R is called the resistance of the substance. The unit of resistance is ohm: 1Ω = 1 V A^{-1}.
- The resistance R of a conductor depends on its length l and constant cross-sectional area A through the relation, $R = \rho\frac{l}{A}$, where ρ called resistivity is a property of the material and depends on temperature and pressure.

- Electrical resistivity of substances varies over a very wide range. Metals have low resistivity, in the range of 10^{-8} Ω m to 10^{-6} Ω m.
- Insulators like glass and rubber have 10^{22} to 10^{24}times greater resistivity.
- Semiconductors like Si and Ge lie roughly in the middle range of resistivity on a logarithmic scale.
- In most substances, the carriers of current are electrons; in some cases, for example, ionic crystals and electrolytic liquids, positive and negative ions carry the electric current.
- Current density **j** gives the amount of charge flowing per second per unit area normal to the flow, $\mathbf{j} = nq\mathbf{v}_d$ where n is the number density (number per unit volume) of charge carriers each of charge q, and $\mathbf{v}_d$ is the drift velocity of the charge carriers. For electrons q = – e. If **j** is normal to a cross-sectional area **A** and is constant over the area, the magnitude of the current I through the area is nev_dA.
- Using E = V/l, I = nev_d A, and Ohm's law, one obtains $\frac{eE}{m} = \rho \frac{ne^2}{m} v_d$.
- The proportionality between the foreseen the electrons in a metal due to the external field E and the drift velocity v_d (not acceleration) can be understood, if we assume that the electrons suffer collisions with ions in the metal, which deflect them randomly. If such collisions occur on an average at a time interval τ,v_d= a τ□= eE τ/m where a is the acceleration of the electron. This gives $\boldsymbol{\rho} = \frac{\mathbf{1}}{\boldsymbol{\sigma}} = \frac{m}{ne^2\tau}$.
- The Potentiometer is an electric instrument that used to measure the EMF (electro motive force) of a given cell, compare EMFs of different cells, potential difference across two points of the circuit and internal resistance of a cell.
- There are many types of resistors, both fixed and variable. The most common type for electronics use is the carbon resistor. They are made in different physical sizes with power dissipation limits commonly from 1 watt down to 1/8 watt. The resistance value and tolerance can be determined from the standard resistor color code.
- A resistor can be used at any combination of voltage (within reason) and current so long as its "Dissipating Power Rating" is not exceeded with the resistor power rating indicating how much power the resistor can convert into heat or absorb without any damage to itself.
- Equivalent resistance R of series combination is $R = R_1 + R_2 + R_3$
- Equivalent resistance R of parallel combination is $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$
- Faraday's laws of induction imply that the emf induced in a coil of N turns is directly related to the rate of change of flux through it, $\varepsilon = -\frac{d\Phi}{dt}$
- Here Φ is the flux linked with one turn of the coil. If the circuit is closed, a current I = Ɛ/R is set up in it, where R is the resistance of the circuit.
- The polarity of the induced emf is such that it tends to produce a current which opposes the change in magnetic flux that produces it. The negative sign in the expression for Faraday's law indicates this fact.
- Inductance is the ratio of the flux-linkage to current. It is equal to NΦ/I.

- A changing current in a coil (coil 2) can induce an emf in a nearby coil (coil 1). This relation is given by, $\varepsilon = M \frac{dI_1}{dt}$ or $M = \frac{\varepsilon}{\frac{dI_1}{dt}}$
- The quantity M_{12} is called mutual inductance of coil 1 with respect to coil 2. One can similarly define M_{21}. There exists a general equality, $M_{12} = M_{21}$.
- When a current in a coil changes, it induces a back emf in the same coil. The self-induced emf is given by, $\varepsilon = L\left(\frac{dI}{dt}\right)$, where L is the self-inductance of the coil. It is a measure of the inertia of the coil against the change of current through it.
- The self-inductance of a long solenoid, the core of which consists of a magnetic material of permeability μ_r, is given by $L = \mu_r \mu_0 n_2 Al$, where A is the area of cross-section of the solenoid, l its length and n the number of turns per unit length.
- Equivalent Inductance L of series combination is $L_{eq} = L_1 + L_2 + L_3$
- Equivalent Inductance L of parallel combination is $\frac{1}{L_{eq}} = \frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3}$
- The ability of a capacitor to store charge (i.e. its capacitance) depends upon the area of plates, distance between plates and the nature of insulating material (or dielectric).
- Capacitance C of a parallel plate capacitor is expressed as $C = \frac{\varepsilon_o A}{\left[\frac{d_1}{\varepsilon_{r1}} + \frac{d_2}{\varepsilon_{r2}} + \frac{d_3}{\varepsilon_{r3}}\right]}\ farad$ and for spherical capacitor is $C = \frac{4\pi\varepsilon_o\varepsilon_r ab}{b-a}\ farad$.
- Equivalent capacitance C of series combination is $\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_1} + \frac{1}{C_1}$
- Equivalent capacitance C of parallel combination is $C = C_1 + C_2 + C_3$
- A.C. filter networks are divided into two major categories: (i) active networks and (ii) passive networks. Active filter networks usually contain transistors and/or operational amplifiers in combination with R, L and C elements to obtain the desired filtering effect.
- These system of logarithmic measurement is widely used in audio, radio, TV and instrument industry for comparing two voltages, currents or power levels. These levels are measured in a unit called bel (B) or decibel (dB) which is $1/10^{th}$ of a bel.
- Suppose we want to compare the output power P_o of a filter with its input power P_i. The power level change is $= 10\ log_{10}\frac{P_o}{P_i}\ dB$ which should be noted that dB is the unit of power change (i.e. increase or decrease) and not of power itself. Moreover, 20 dB is not twice as much power as 10 dB.
- The cutoff frequency f_c for a low pass RC filter occurs where **(a)** V_o = 70.7% V_i i.e. V_o is – 3 dB down from V_i **(b)** R = X_c and $V_R = V_C$ in magnitude. **(c)** The impedance phase angle $\theta = -45°$. The same is the angle between V_o and V_i. However, it may be noted that being a low pass RL filter circuit, the impedance phase angle is +45° (and not –45° as in low-pass RC filter).
- It may be noted that high-pass RC filter can be obtained merely by interchanging the positions of R and C in the low-pass RC filter.
- It can be obtained high-pass RL filter by 'swapping' position of R and L in the low-pass RL circuit.

2.21 GLOSSARY

1**. Resistance**: It may be defined as the property of a substance due to which it opposes (or restricts) the flow of electricity (i.e., electrons) through it.

2.**Unit of Resistance (Ω)**: The practical unit of resistance is ohm. A conductor is said to have a resistance of one ohm if it permits one ampere current to flow through it when one volt is impressed across its terminals.

3.**Types of Resistors**: Various types of resistors are available for practical use as follows:
(a) Carbon Composition(e) Deposited Carbon
(b) High-Voltage Ink Film(f) Metal Film
(c) Metal Glaze(g) Wire-wound
(d) Cermet (Ceramic Metal)

4**. Fixed Resistors**: A fixed resistor is one for which the value of its resistance is specified and cannot be varied in general.

5. **Variable Resistors**: The resistance of a variable resistor can change. It can be used as a three terminal as well as a two terminal device. Mostly they are used as a three terminal device. Variable resistors are mostly used for device calibration.

6.**Types of variable resistors**: There are mainly three types of variable resistors. They are Potentiometer Rheostat and Presets.

7**.Potentiometer:** A potentiometer, also called as POT, is a 3-terminal variable resistor and is used to adjust the resistance in a circuit. There are mainly two of them logarithmic potentiometer and linear potentiometer.

8. **Rheostat**: A rheostat is also a variable resistor and is a 2-terminal device. It is commonly used for handling higher currents and voltages. One terminal will be connected to the end of the track and the other to a moveable wiper. When the wiper moves from one end to the other, the resistance changes from zero to maximum.

9. **Resistor color coding**: In this color code system, the bands of different colors are used to identify the value of resistance and tolerance ratings. Also the power rating is determined by the physical size of the resistor. There are two common systems of the color code designation: a four band system and a five bands system. However, the four bands color code system is more popular and is widely used.

10**. Four band system**: In four band system, reading the resistor from left to right, the first two color bands represent significant digits, the third band represents the decimal multiplier, and the fourth band represents the tolerance.

11.**Five band system**: In five bands system the first three color bands represent significant digits, the fourth band represents the decimal multiplier, and the fifth band represents the tolerance.

12. **Power rating of resistors**: When an electrical current passes through a resistor due to the presence of a voltage across it, electrical energy is lost by the resistor in the form of heat and the greater this current flow the hotter the resistor will get. This is known as the Resistor Power Rating.

13. **Power dissipated in a resistor**: The electrical power dissipation of any resistor in a DC circuit can be calculated using one of the following three standard formulas:

$$P = VI = I^2R = \frac{V^2}{R}$$

14. **Series combination of resistors**: When some conductors having resistances R1, R2 and R3 etc. are joined end-on-end, they are said to be connected in series. The equivalent resistance of series combination is given by

$$R = R_1 + R_2 + R_3$$

15**. Parallel combination of resistors**: When some conductors having resistances R1, R2 and R3 etc. are connected such that one of their end is connected to a common point and the remaining ends are connected to other common point. The equivalent resistance of parallel combination is given by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

16. **Capacitor**: Any two conducting surfaces separated by an insulating material is called a capacitor or condenser**.** Its purpose is to store charge in a small space. The conducting surfaces are called the plates of the capacitor and the insulating material is called the dielectric.

17. **Capacitance**: The property of a capacitor to 'store electricity' may be called its capacitance. The capacitance of a capacitor is defined as "the amount of charge required to create a unit p.d. between its plates. "Suppose we give Q coulomb of charge to one of the two plate of capacitor and if a p.d. of V volts is established between the two, then its capacitance is

$$C = \frac{Q}{V} = \frac{charge}{potential\ difference}$$

18. **Unit of Capacitance**: The unit of capacitance is coulomb/volt which is also called farad (in honor of Michael Faraday)1 farad = 1 coulomb/volt. One farad is defined as the capacitance of a capacitor which requires a charge of one coulomb to establish a p.d. of one volt between its plates.

19. **Factors affecting capacitance**: The ability of a capacitor to store charge (i.e. its capacitance) depends upon the following factors:
(i) Area of plate(ii) Thickness of dielectric(iii) Relative permittivity of dielectric

20. **Series combination of capacitors**: For the capacitancesC_1, C_2, C_3 connected in series, the equivalent capacitance C is given by

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_1} + \frac{1}{C_1}$$

21.**Parallel combination of capacitors**: For the capacitancesC_1, C_2, C_3 connected in parallel, the equivalent capacitance C is given by

$$C = C_1 + C_2 + C_3$$

22. **Energy stored in the capacitor**: Charging a capacitor means transferring electrons from one plate of the capacitor to the other. This involves expenditure of energy because electrons have to be moved against the opposing forces. This energy is stored in the electrostatic field set up in the dielectric medium. On discharging the capacitor, the field collapses and the stored energy is released. Energy stored in the capacitor is

$$\mathrm{E} = \frac{1}{2}\,\mathrm{CV}^2 = \frac{1}{2}\mathrm{QV} = \frac{\mathrm{Q}^2}{2\mathrm{C}} \text{ joules}$$

23. **Capacitance of Parallel-Plate Capacitor**: The capacitance of a parallel plate capacitor with plate area A, plate separation d and medium od dielectric constant $\varepsilon_r$in between the plates is given by

$$\mathrm{C} = \varepsilon_o \varepsilon_r \frac{\mathrm{A}}{\mathrm{d}}$$

24. **Capacitance of a cylindrical capacitor**: A cylindrical capacitor consists of two co-axial cylinders separated by an insulating medium. This is an important practical case since a single core cable is in effect a capacitor of this kind. The conductor (or core) of the cable is the inner cylinder

while the outer cylinder is represented by lead sheath which is at earth potential. The capacitance per unit length is given by

$$C = \frac{2\pi x \varepsilon_0 \varepsilon_r}{\log_e \frac{D}{d}} \text{ farad per m}$$

25. **Capacitance of a spherical capacitor**: Consider a spherical capacitor consisting of two concentric spheres of radii 'a' and 'b' metres. When the outer sphere is earthed the capacitance is given by

$$C = \frac{4\pi\varepsilon_o\varepsilon_r ab}{b-a} \text{ farad}$$

When the inner sphere is earthed the capacitance: In this situation we have two capacitances in parallel given by

i) One capacitor consists of the inner surface of A and the outer surface of B. Its capacitance, as found earlier, is

$$C_1 = \frac{4\pi\varepsilon_o\varepsilon_r ab}{b-a}$$

(ii) The second capacitor consists of outer surfaces of B and earth. Its capacitance is $C_2 = 4\pi\varepsilon_o b-$ if surrounding medium is air. Total capacitance $C = C_1 + C_2$.

26. **Filter circuits:** The reactance of inductors and capacitors depend on the frequency of the a.c. signal applied to them. That is why these devices are known as frequency-selective. By using various combinations of resistors, inductors and capacitors, we can make circuits that have the property of passing or rejecting either low or high frequencies or bands of frequencies. These frequency selective networks, which alter the amplitude and phase characteristics of the input a.c. signal, are called filters.

27. **Types of filter circuits**: A.C. filter networks are divided into two major categories: (i) active networks and (ii) passive networks. Active filter networks usually contain transistors and/or operational amplifiers in combination with R, L and C elements to obtain the desired filtering effect. On the other side passive filter networks which usually consist of series-parallel combinations of R, L and C elements. There are four types of such passive networks as follow

1. Low-Pass Filter2. High-Pass Filter.3. Bandpass Filter4. Bandstop Filter

28. **The Decibel (dB) system**: This system of logarithmic measurement is widely used in audio, radio, TV and instrument industry for comparing two voltages, currents or power levels. These levels are measured in a unit called bel (B) or decibel (dB) which is $1/10^{th}$ of a bel. Suppose we want to compare the output power P_o of a filter with its input power P_i. The power level change is

$$= 10 \log_{10} \frac{P_o}{P_i} \text{ dB}$$

2.22 REFERENCES

1. Boylestad, R.L., Nashelsky ,L., "Electronic Devices and Circuit Theory", 11th edition, Pearson Publication.
2. Theraja, B.L., Theraja, A.K. 2005, "A Textbook of Electrical Technology Vol. I", S. Chand Publication.
3. Mehta, V.K., Mehta, R, "Principles of Electronics", S. Chand Publication.
4. Eggleston, D.L 2011, "Basic Electronics for Scientists and Engineers", Cambridge University Press.
5. "Physics Part –II A textbook for class XII", 1st edition, NCERT.
6. P. Kumar, N. K. Lohani and L. P. Verma, Kumaun Physics, S. J. Publications, India, 2018.
7. Ghatak, Optics, McGraw-Hill Education India Pvt. Ltd., 2017.
8. S. L Gupta and V. Kumar, Practical Physics, PragatiPrakashan, 2014.

2.23 TERMINAL QUESTIONS

1. An air-cored solenoid with length 30 cm, area of cross-section 25 cm^2 and number of turns 500, carries current of 2.5 A. The current is suddenly switched off in a brief time of 10^{-3} s. How much is the average back emf induced across the ends of the open switch in the circuit? Ignore the variation in magnetic field near the ends of the solenoid.
2. Answer the following questions:

(a) A steady current flows in a metallic conductor of non-uniform cross-section. Which of these quantities is constant along the conductor: current, current density, electric field, drift speed?
(b) Is Ohm's law universally applicable for all conducting elements? If not, give examples of elements which do not obey Ohm's law.
(c) A low voltage supply from which one needs high currents must have very low internal resistance. Why?
(d) A high tension (HT) supply of, say, 6 kV must have a very large internal resistance. Why?
(e) Given n resistors each of resistance R, how will you combine them to get the (i) maximum (ii) minimum effective resistance? What is the ratio of the maximum to minimum resistance?

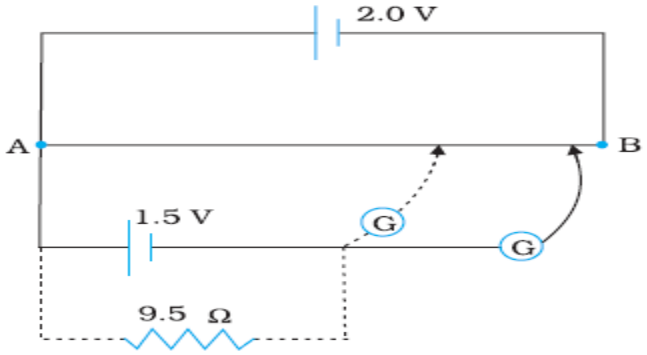
3.Answer the following questions:

(a) Given the resistances of 1 Ω, 2 Ω, 3 Ω, how will be combine them to get an equivalent resistance of (i) (11/3) Ω (ii) (11/5) Ω, (iii) 6Ω, (iv) (6/11) Ω?

(b) Obtain an expression for the mutual inductance between two concentric circular rigs of different radius.

(c) Current in a circuit falls from 5.0 A to 0.0 A in 0.1 s. If an average emf of 200 V induced, give an estimate of the self-inductance of the circuit. A pair of adjacent coils has a mutual inductance of 1.5 H. If the current in one coil changes from 0 to 20 A in 0.5 s, what is the change of flux linkage with the other coil?

(d) Figure shows a 2.0 V potentiometer used for the determination of internal resistance of a 1.5 V cell. The balance point of the cell in open circuit is 76.3 cm. When a resistor of 9.5 Ω is used in the external circuit of the cell, the balance point shifts to 64.8 cm length of the potentiometer wire. Determine the internal resistance of the cell.



UNIT 3 SEMICONDUCTOR AND SEMICONDUCTOR DEVICES

Structure

3.1 INTRODUCTION

Devices in which a controlled flow of electrons can be obtained are the basic building blocks of all the electronic circuits. Before the discovery of transistor in 1948, such devices were mostly vacuum tubes (also called vacuum valves) like the vacuum diode which has two electrodes, viz., anode (often called plate) and cathode; triode which has three electrodes – cathode, plate and grid; tetrode and pentode (respectively with 4 and 5 electrodes). In a vacuum tube, the electrons are supplied by a heated cathode and the controlled flow of these electrons in vacuum is obtained by varying the voltage between its different electrodes. Vacuum is required in the inter-electrode space; otherwise the moving electrons may lose their energy on collision with the air molecules in their path. In these devices the electrons can flow only from the cathode to the anode (i.e., only in one direction). Therefore, such devices are generally referred to as valves. These vacuum tube devices are bulky, consume high power, operate generally at high voltages (~100 V) and have limited life and low reliability. The seed of the development of modern solid-state semiconductor electronics goes back to 1930's when it was realized that some solid-state semiconductors and their junctions offer the possibility of controlling the number and the direction of flow of charge carriers through them. Simple excitations like light, heat or small applied voltage can change the number of mobile charges in a semiconductor. Note that the supply and flow of charge carriers in the semiconductor devices are within the solid itself, while in the earlier vacuum tubes/valves, the mobile electrons were obtained from a heated cathode and they were made to flow in an evacuated space or vacuum. No external heating or large evacuated space is required by the semiconductor devices. They are small in size, consume low power, operate at low voltages and have long life and high reliability. Even the Cathode Ray Tubes (CRT) used in television and computer monitors which work on the principle of vacuum tubes are being replaced by Liquid Crystal Display (LCD) monitors with supporting solid state electronics. Much before the full implications of the semiconductor devices was formally understood, a naturally occurring crystal of galena (Lead sulphide, PbS) with a metal point contact attached to it was used as detector of radio waves. In the following sections, we will introduce the basic concepts of semiconductor physics and discuss some semiconductor devices like junction diodes (a 2-electrode device).

3.2 OBJECTIVES

After studying this unit, you should be able to-

- Get definition of semiconductor
- Understand different types of semiconductor
- Know various energy bands and charge carriers in semiconductor
- Get how drift of carriers takes place in electric and magnetic fields
- Understand what Hall Effect is
- Understand working, characteristic and function of p-n junction diode
- P-n junction fabrication (basic idea) and its characteristics

- Understand junction breakdown- zener and avalanche breakdowns
- Understand working, characteristic and function of zener diode
- Get knowledge about varies diodes; tunnel diode, varactor diode, photo diode and LED.

3.3 CLASSIFICATION OF METALS, CONDUCTORS AND SEMICONDUCTORS

On the basis of the relative values of electrical conductivity (σ) or resistivity (ρ = 1/σ), the solids are broadly classified as: (i) Metals: They possess very low resistivity (or high conductivity). $\rho \sim 10^{-2} - 10^{-8}$ Ω m and $\sigma \sim 10^{2} - 10^{8}$ Ω^{-1} m^{-1} (ii) Semiconductors: They have resistivity or conductivity intermediate to metals and insulators. $\rho \sim 10^{-5} - 10^{6}$ Ω m and $\sigma \sim 10^{5} - 10^{-6}$ Ω^{-1} m^{-1} (iii) Insulators: They have high resistivity (or low conductivity). $\rho \sim 10^{11} - 10^{19}$ Ω m and $\sigma \sim 10^{-11} - 10^{-19}$ Ω^{-1} m^{-1}. The values of ρ and σ given above are indicative of magnitude and could well go outside the ranges as well. Relative values of the resistivity are not the only criteria for distinguishing metals, insulators and semiconductors from each other. There are some other differences, which will become clear as we go along in this chapter. Our interest in this chapter is in the study of semiconductors, which could be: (i) Elemental semiconductors: Si and Ge (ii) Compound semiconductors: Examples are: (a) Inorganic: CdS, GaAs, CdSe, InP, etc. (b) Organic: anthracene, doped pthalocyanines, etc. In this chapter, we will restrict ourselves to the study of inorganic semiconductors, particularly elemental semiconductors Si and Ge. The general concepts introduced here for discussing the elemental semiconductors, by-and-large, apply to most of the compound semiconductors as well.

On the basis of energy bands According to the Bohr atomic model, in an isolated atom the energy of any of its electrons is decided by the orbit in which it revolves. But when the atoms come together to form a solid they are close to each other. So the outer orbits of electrons from neighboring atoms would come very close or could even overlap. This would make the nature of electron motion in a solid very different from that in an isolated atom. Inside the crystal each electron has a unique position and no two electrons see exactly the same pattern of surrounding charges. Because of this, each electron will have a different energy level. These different energy levels with continuous energy variation form what are called energy bands. The energy band which includes the energy levels of the valence electrons is called the valence band. The energy band above the valence band is called the conduction band. With no external energy, all the valence electrons will reside in the valence band. If the lowest level in the conduction band happens to be lower than the highest level of the valence band, the electrons from the valence band can easily move into the conduction band. Normally the conduction band is empty. But when it overlaps on the valence band electrons can move freely into it. This is the case with metallic conductors. If there is some gap between the conduction band and the valence band, electrons in the valence band

all remain bound and no free electrons are available in the conduction band. This makes the material an insulator. But some of the electrons from the valence band may gain external energy to cross the gap between the conduction band and the valence band. Then these electrons will move into the conduction band. At the same time, they will create vacant energy levels in the valence band where other valence electrons can move. Thus the process creates the possibility of conduction due to electrons in conduction band as well as due to vacancies in the valence band. Let us consider what happens in the case of Si or Ge crystal containing N atoms. For Si, the outermost orbit is the third orbit (n = 3), while for Ge it is the fourth orbit (n = 4). The number of electrons in the outermost orbit is 4 (2s and 2p electrons). Hence, the total number of outer electrons in the crystal is 4N. The maximum possible number of electrons in the outer orbit is 8 (2s + 6p electrons). So, for the 4N valence electrons there are 8N available energy states. These 8N discrete energy levels can either form a continuous band or they may be grouped in different bands depending upon the distance between the atoms in the crystal (see box on Band Theory of Solids). At the distance between the atoms in the crystal lattices of Si and Ge, the energy band of these 8N states is split apart into two which are separated by an energy gap E_g.

3.4 WHAT ARE SEMICONDUCTORS?

Semiconductors are a group of materials having conductivities between those of metals and insulators. Two general classifications of semiconductors are the elemental semiconductor materials, found in group IV of the periodic table, and the compound semiconductor materials, most of which are formed from special combinations of group III and group V elements. Table 1 shows the portion of the periodic table which the more common semiconductors are found and Table 2 lists a few of the semiconductor materials. (Semiconductors can also be formed from combinations of group II and group VI elements, but in general these will not be considered in this text.)

The elemental materials, those that are composed of single species of atoms, are silicon and germanium. Silicon is by far the most common semiconductor used in integrated circuits and will be emphasized to a great extent. The two-element, or *binary,* compounds such as gallium arsenide or gallium phosphide are formed by combining one group III and one group V element. Gallium arsenide is one of the more common of the compound semiconductors. Its good optical properties make it useful in optical devices. GaAs is also used in specialized applications in which, for example, high speed is required.

III	IV	V
5 **B** Boron	6 **C** Carbon	
13 **Al** Aluminum	14 **Si** Silicon	15 **P** Phosphorus
31 **Ga** Gallium	32 **Ge** Germanium	33 **As** Arsenic
49 **In** Indium		51 **Sb** Antimony

Table 3.1: A Portion of the periodic table.

Elemental semiconductors	
Si	Silicon
Ge	Germanium
Compound semiconductors	
AlP	Aluminum phosphide
AlAs	Aluminum arsenide
GaP	Gallium phosphide
GaAs	Gallium arsenide
InP	Indium phosphide

Table 3.2: A list of some semiconducting materials.

The three semiconductors used most frequently in the construction of electronic devices are Ge, Si, and GaAs.

3.5 TYPES OF SEMICONDUCTORS

On the basis of the absence or presence of impurity atoms the semiconducting materials may be classified as intrinsic or extrinsic semiconductors respectively.

3.5.1 Intrinsic Semiconductors

To understand why Si, Ge, and GaAs are the semiconductors of choice for the electronics industry requires some understanding of the atomic structure of each and how the atoms are bound together to form a crystalline structure. The fundamental components of an atom are the electron, proton, and neutron. In the lattice structure, neutrons and protons form the nucleus and electrons appear in fixed orbits around the nucleus. The Bohr model for the three materials is provided in Figure 1.

As indicated in Figure1, silicon has 14 orbiting electrons, germanium has 32 electrons, gallium has 31 electrons, and arsenic has 33 orbiting electrons. For germanium and silicon there are four electrons in the outermost shell, which are referred to as *valence electrons*. Gallium has three valence electrons and arsenic has five valence electrons. Atoms that have four valence electrons are called *tetravalent*, those with three are called *trivalent*, and those with five are called *pentavalent.*

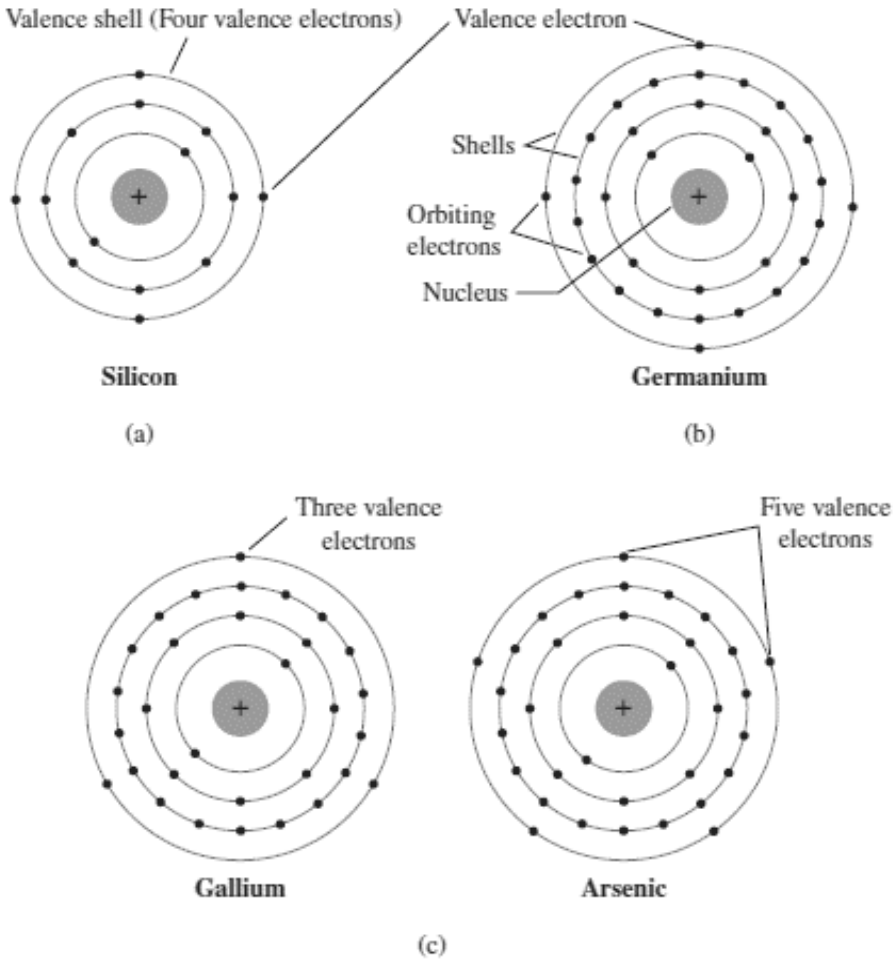


Figure 3.1: Atomic structure of (a) Silicon; (b) Germanium and (c) Gallium and arsenic.

In a pure silicon or germanium crystal the four valence electrons of one atom form a bonding arrangement with four adjoining atoms, as shown in Figure 2. This bonding of atoms, strengthened by the sharing of electrons, is called covalent bonding.

Because GaAs is a compound semiconductor, there is sharing between the two different atoms, as shown in Figure 3. Each atom, gallium or arsenic, is surrounded by atoms of the complementary

type. There is still a sharing of electrons similar in structure to that of Ge and Si, but now five electrons are provided by the As atom and three by the Ga atom. This picture of crystal structure exists only at low temperatures. As the temperature increases, more thermal energy becomes available to these electrons and some of these electrons may break–away (becoming *free* electrons contributing to conduction).

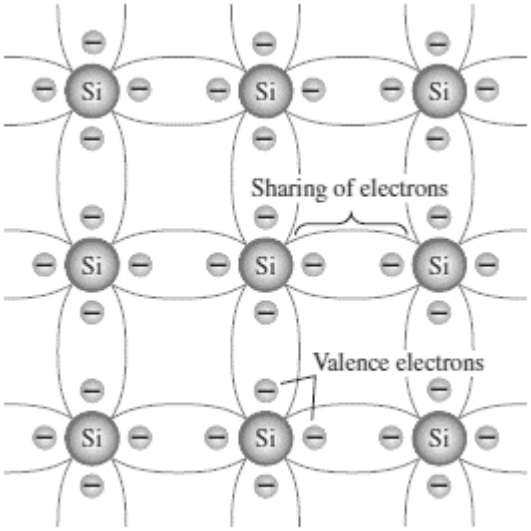


Figure 3.2: Covalent bonding in silicon.

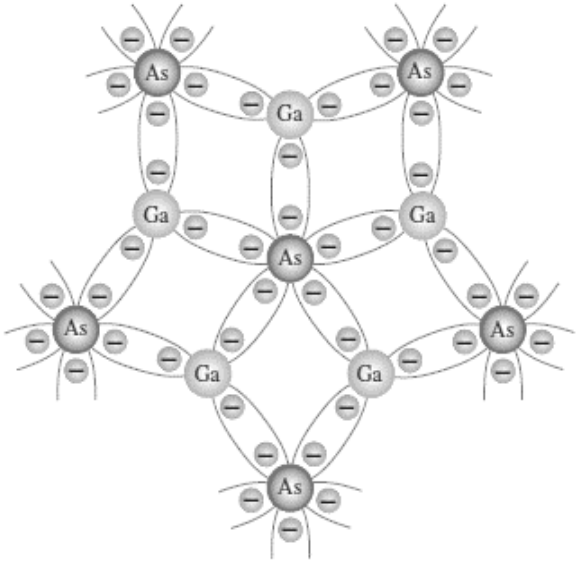


Figure 3.3: Covalent bonding in GaAs crystal.

At room temperature there are approximately 1.5×10^{10} free carriers in 1 cm^3 of *intrinsic* silicon material, that is, 15,000,000,000 (15 billion) electrons in a space smaller than a small sugar cube—an enormous number. The free electrons in a material due only to external causes are referred to as *intrinsic carriers.* Table 3 compares the number of intrinsic carriers per cubic centimeter (abbreviated n_i) for Ge, Si, and GaAs.

Semiconductor	Intrinsic Carriers (per cubic centimeter)
GaAs	1.7×10^6
Si	1.5×10^{10}
Ge	2.5×10^{13}

Table 3.3: Intrinsic Carriers n_i.

The term intrinsic is applied to any semiconductor material that has been carefully refined to reduce the number of impurities to a very low level—essentially as pure as can be made available through modern technology.

The thermal energy effectively ionizes only a few atoms in the crystalline lattice and creates a *vacancy* in the bond. The neighborhoods, from which the free electron (with charge –*q*) has come out leaves a vacancy with an effective charge (+*e*). This *vacancy* with the effective positive electronic charge is called a *hole*. The hole behaves as an *apparent free particle* with effective positive charge.

In intrinsic semiconductors, the number of free electrons, n_e is equal to the number of holes, n_h. That is $n_i = n_h = n_e$ where n_i is called intrinsic carrier concentration.

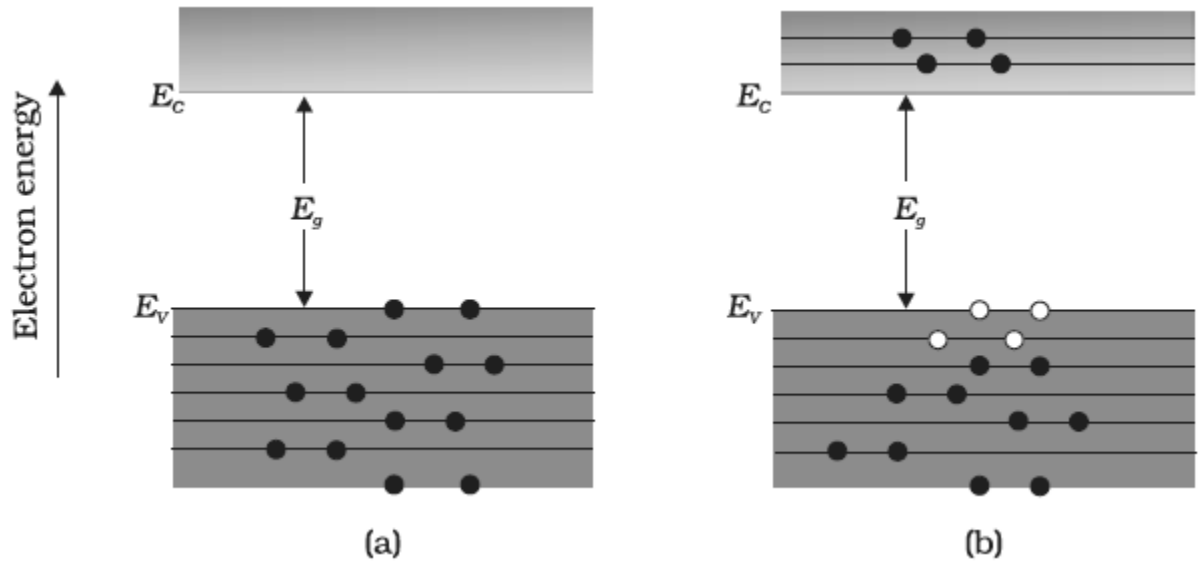


Figure 3.4: a) An intrinsic semiconductor at T = 0 K behave like insulator; b) At T > 0 K four thermally generated electron hole pairs. The filled circles represent (•) represents electrons and empty circles (○) represents holes.

An intrinsic semiconductor will behave like an insulator at $T = 0$ K as shown in Figure 4(a). It is the thermal energy at higher temperatures ($T > 0$K), which excites some electrons from the valence band to the conduction band. These thermally excited electrons at $T > 0$ K, partially occupy the conduction band. Therefore, the energy-band diagram of an intrinsic semiconductor will be as shown in Figure 4(b). Here, some electrons are shown in the conduction band. These have come from the valence band leaving equal number of holes there.

3.5.2 Extrinsic Semiconductors

The conductivity of an intrinsic semiconductor depends on its temperature, but at room temperature its conductivity is very low. As such, no important electronic devices can be developed using these semiconductors. Hence there is a necessity of improving their conductivity. This can be done by making use of impurities.

When a small amount, say, a few parts per million (ppm), of a suitable impurity is added to the pure semiconductor, the conductivity of the semiconductor is increased manifold. Such materials are known as e*xtrinsic semiconductors* or *impurity semiconductors*. The deliberate addition of a desirable impurity is called *doping* and the impurity atoms are called *dopants*. Such a material is also called a *doped semiconductor*. There are two extrinsic materials of immeasurable importance to semiconductor device fabrication: *n* -type and *p* -type materials.

3.5.3 Types of Extrinsic Semiconductors

(a) n - type semiconductor

An *n* -type material is created by introducing impurity elements that have *five* valence electrons (*pentavalent*), such as *antimony*, *arsenic*, and *phosphorus.* Each is a member of a subset group of elements in the Periodic Table of Elements referred to as Group V because each has five valence electrons. The effect of such impurity elements is indicated in Figure 5 (using antimony as the impurity in a silicon base). Note that the four covalent bonds are still present. There is, however, an additional fifth electron due to the impurity atom, which is *unassociated* with any particular covalent bond. This remaining electron, loosely bound to its parent (antimony) atom, is relatively free to move within the newly formed *n* -type material. As a result, the ionization energy required to set this electron free is very small and even at room temperature it will be free to move in the lattice of the semiconductor. For example, the energy required is ~ 0.01 eV for germanium, and 0.05 eV for silicon, to separate this electron from its atom.

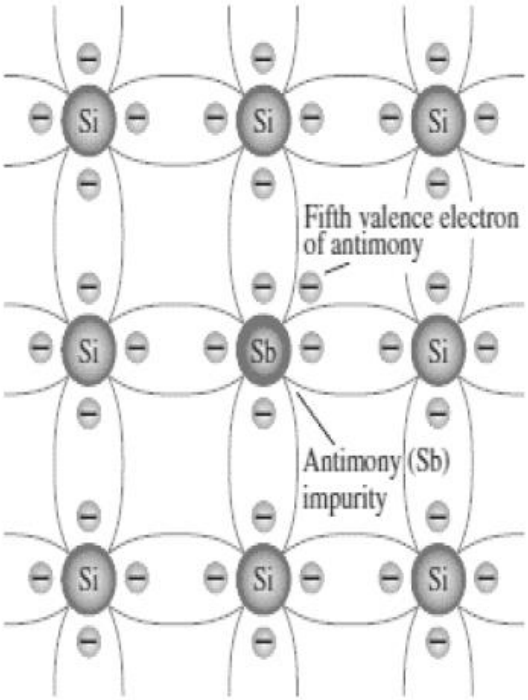


Figure 3.5: Antimony impurity in *n* type material.

This is in contrast to the energy required to jump the forbidden band (about 0.72 eV for germanium and about 1.1 eV for silicon) at room temperature in the intrinsic semiconductor.

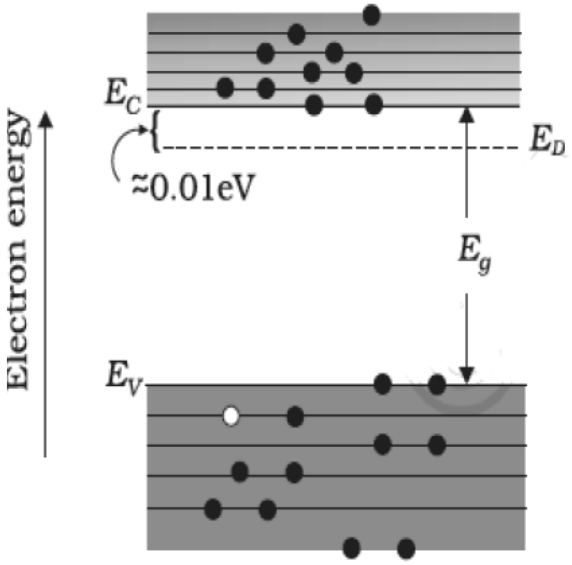


Figure 3.6: Energy Bands of n type semiconductor at T > 0K.

Thus, the pentavalent dopant is donating one extra electron for conduction and hence is known as *donor* impurity. Since the inserted impurity atom has donated a relatively "free" electron to the structure: Diffused impurities with five valence electrons are called donor atoms. It is important to realize that even though a large number of free carriers have been established in the *n* -type material, it is still electrically *neutral* since ideally the number of positively charged protons in the nuclei is still equal to the number of free and orbiting negatively charged electrons in the structure.

In a doped semiconductor the total number of conduction electrons n_e is due to the electrons contributed by donors and those generated intrinsically, while the total number of holes n_h is only due to the holes from the intrinsic source. But the rate of recombination of holes would increase due to the increase in the number of electrons. As a result, the number of holes would get reduced further. Thus, with proper level of doping the number of conduction electrons can be made much larger than the number of holes. These semiconductors are, therefore, known as n-*type semiconductors*. For n-type semiconductors, we have, $n_e >> n_h$.

The effect of this doping process on the relative conductivity can best be described through the use of the energy-band diagram of Figure 6. Note that a discrete energy level (called the *donor level*) appears in the forbidden band with an E_g significantly less than that of the intrinsic material. Those free electrons due to the added impurity sit at this energy level and have less difficulty absorbing a sufficient measure of thermal energy to move into the conduction band at room temperature. The result is that at room temperature, there are a large number of carriers (electrons) in the conduction level, and the conductivity of the material increases significantly.

(b) p - type Semiconductor

The *p* -type material is formed by doping a pure germanium or silicon crystal with impurity atoms having *three* valence electrons. The elements most frequently used for this purpose are *boron*, *gallium*, and *indium*. Each is a member of a subset group of elements in the Periodic Table of Elements referred to as Group III because each has three valence electrons. The effect of one of these elements, boron, on a base of silicon is indicated in Figure 7.

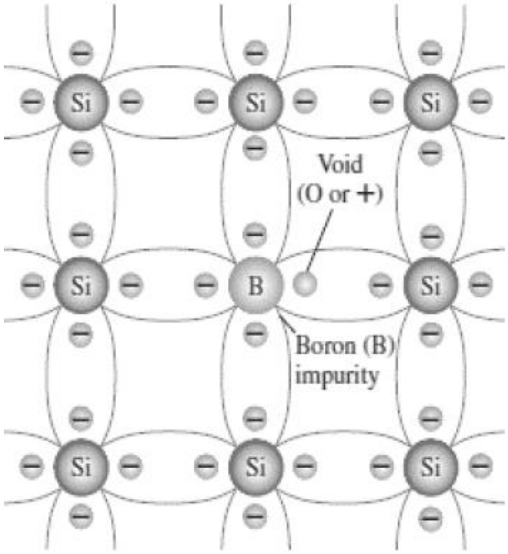


Figure 3.7: Boron impurity in p type material.

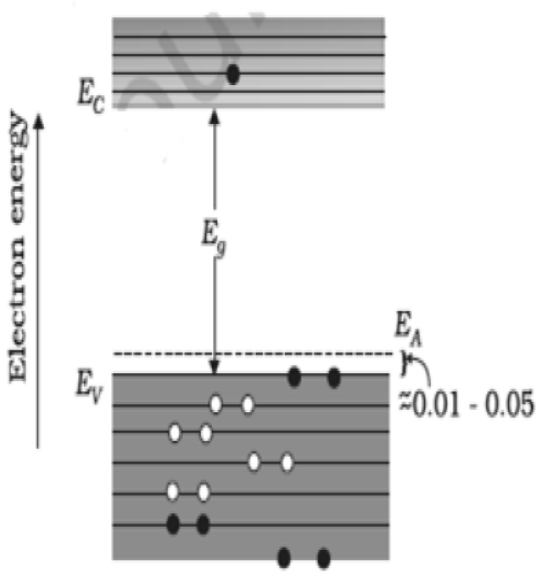


Figure 3.8: Energy bands of p type semiconductor at T > 0 K.

Note that there is now an insufficient number of electrons to complete the covalent bonds of the newly formed lattice. The resulting vacancy is called a *hole* and is represented by a small circle or a plus sign, indicating the absence of a negative charge. Since the resulting vacancy will readily *accept* a free electron: The diffused impurities with three valence electrons are called acceptor atoms. The resulting *p* -type material is electrically neutral, for the same reasons described for the *n* -type material.

It is obvious that one *acceptor* atom gives one *hole*. These holes are in addition to the intrinsically generated holes while the source of conduction electrons is only intrinsic generation. For p-type semiconductors, the recombination process will reduce the number (n_i)of intrinsically generated electrons to n_e. We have, for p-type semiconductors $n_h >> n_e$.

For p-type semiconductor, the acceptor energy level E_A is slightly above the top E_V of the valence band as shown in Figure 8. With very small supply of energy an electron from the valence band can jump to the level E_A and ionize the acceptor negatively. (Alternately, we can also say that with very small supply of energy the hole from level E_A sinks down into the valence band. Electrons rise up and holes fall down when they gain external energy.) At room temperature, most of the acceptor atoms get ionized leaving holes in the valence band. Thus at room temperature the density of holes in the valence band is predominantly due to impurity in the extrinsic semiconductor.

3.5.4 Conduction through p-type and n-type semiconductors

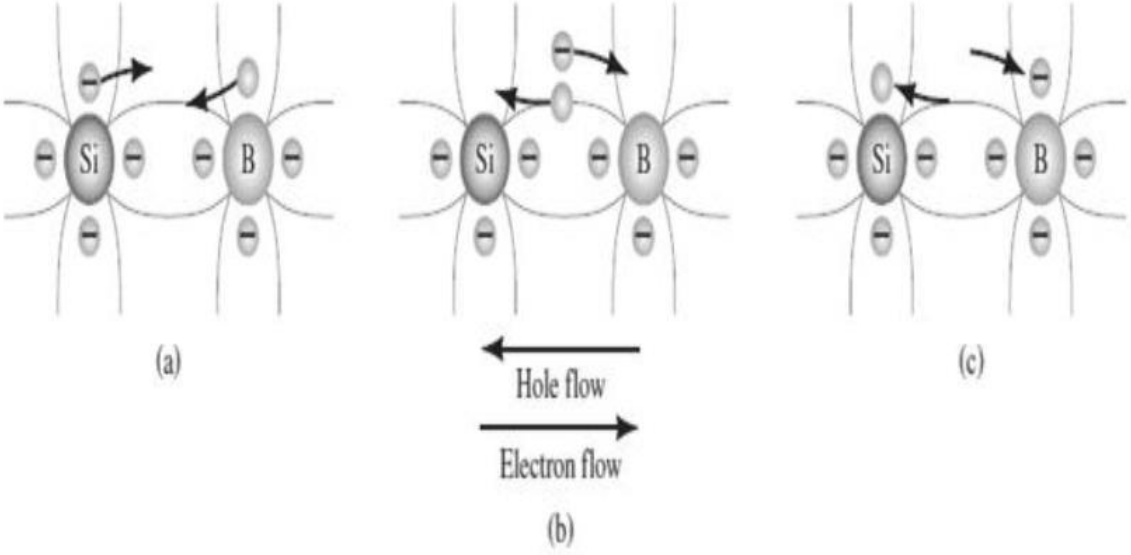


Figure 3.9: Electron versus hole flow.

The effect of the hole on conduction is shown in Figure 9. If a valence electron acquires sufficient kinetic energy to break its covalent bond and fills the void created by a hole, then a vacancy, or hole, will be created in the covalent bond that released the electron. There is, therefore, a transfer of holes to the left and electrons to the right, as shown in Figure 9. The direction to be used in this text is that of *conventional flow*, which is indicated by the direction of hole flow.

3.5.5 Majority and Minority Carriers

In the intrinsic state, the number of free electrons in Ge or Si is due only to those few electrons in the valence band that have acquired sufficient energy from thermal or light sources to break the covalent bond or to the few impurities that could not be removed. The vacancies left behind in the covalent bonding structure represent our very limited supply of holes. In an *n* -type material, the number of holes has not changed significantly from this intrinsic level.

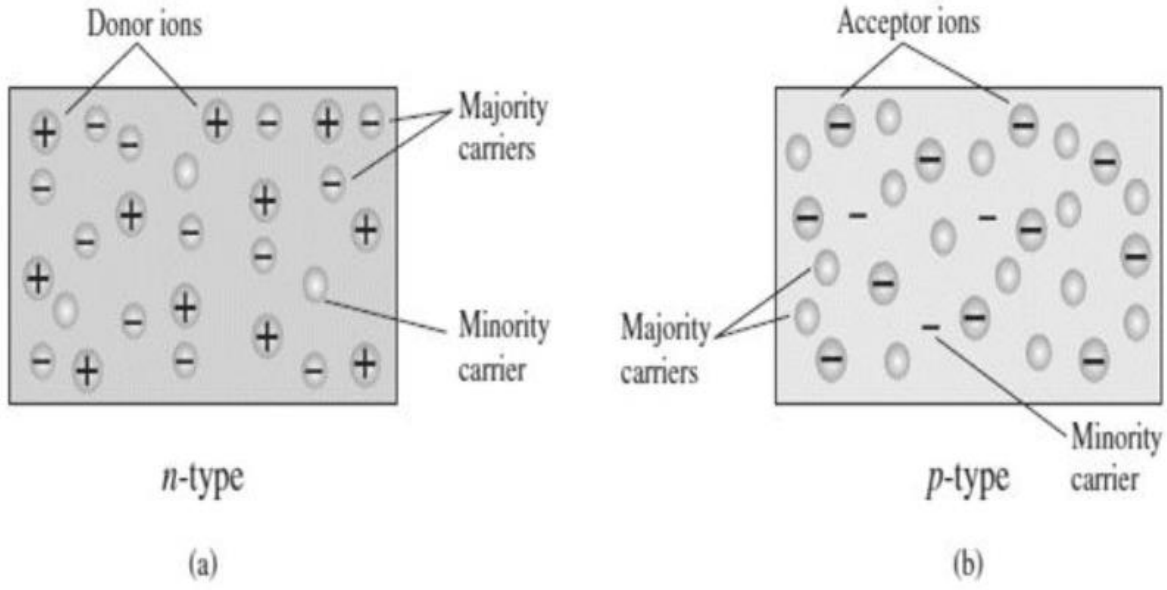


Figure 3.10: (a) n-type material; (b) p-type material.

The net result, therefore, is that the number of electrons far outweighs the number of holes. For this reason: *In an n-type material (Figure 10 a) the electron is called the majority carrier and the hole the minority carrier.* For the *p* -type material the number of holes far outweighs the number of electrons, as shown in Figure 10b. Therefore: *In a p-type material the hole is the majority carrier and the electron is the minority carrier.*

When the fifth electron of a donor atom leaves the parent atom, the atom remaining acquires a net positive charge: hence the plus sign in the donor-ion representation. For similar reasons, the minus sign appears in the acceptor ion. The *n* - and *p* -type materials represent the basic building blocks of semiconductor devices.

3.6 ENERGY BANDS AND CHARGE CARRIERS

On the basis of energy bands According to the Bohr atomic model, in an isolated atom the energy of any of its electrons is decided by the orbit in which it revolves. But when the atoms come together to form a solid they are close to each other. So the outer orbits of electrons from

neighboring atoms would come very close or could even overlap. This would make the nature of electron motion in a solid very different from that in an isolated atom. In semiconductors we are primarily interested in the valence band and conduction band. Moreover, for most applications we are interested in what happens near the top of the valence band and the bottom of the conduction band. These states originate from the atomic levels of the valence shell in the elements making up the semiconductor.

IV Semiconductors

C $1s^2 2s^2 2p^2$

Si $1s^2 2s^2 2p^6 3s^2 3p^2$

Ge $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^2$

III-V Semiconductors

Ga $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^1$

As $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^3$

Inside the crystal each electron has a unique position and no two electrons see exactly the same pattern of surrounding charges. Because of this, each electron will have a different energy level. These different energy levels with continuous energy variation form what are called energy bands. The energy band which includes the energy levels of the valence electrons is called the valence band. The energy band above the valence band is called the conduction band. With no external energy, all the valence electrons will reside in the valence band. If the lowest level in the conduction band happens to be lower than the highest level of the valence band, the electrons from the valence band can easily move into the conduction band. Normally the conduction band is empty. But when it overlaps on the valence band electrons can move freely into it. This is the case with metallic conductors. If there is some gap between the conduction band and the valence band, electrons in the valence band all remain bound and no free electrons are available in the conduction band. This makes the material an insulator. But some of the electrons from the valence band may gain external energy to cross the gap between the conduction band and the valence band. Then these electrons will move into the conduction band. At the same time, they will create vacant energy levels in the valence band where other valence electrons can move. Thus the process creates the possibility of conduction due to electrons in conduction band as well as due to vacancies in the valence band. Let us consider what happens in the case of Si or Ge crystal containing N atoms. For Si, the outermost orbit is the third orbit (n = 3), while for Ge it is the fourth orbit (n = 4). The number of electrons in the outermost orbit is 4 (2s and 2p electrons). Hence, the total number of outer electrons in the crystal is 4N. The maximum possible number of electrons in the outer orbit is 8 (2s + 6p electrons). So, for the 4N valence electrons there are 8N available energy states. These 8N discrete energy levels can either form a continuous band or they may be grouped in different bands depending upon the distance between the atoms in the crystal. At the distance between the atoms in the crystal lattices of Si and Ge, the energy band of these 8N states is split apart into two which are separated by an energy gap E_g. The lower band which is completely occupied by the $4N$ valence electrons at temperature of absolute zero is the *valence band.* The other band consisting of $4N$ energy states, called the *conduction band*, is completely empty at absolute zero.

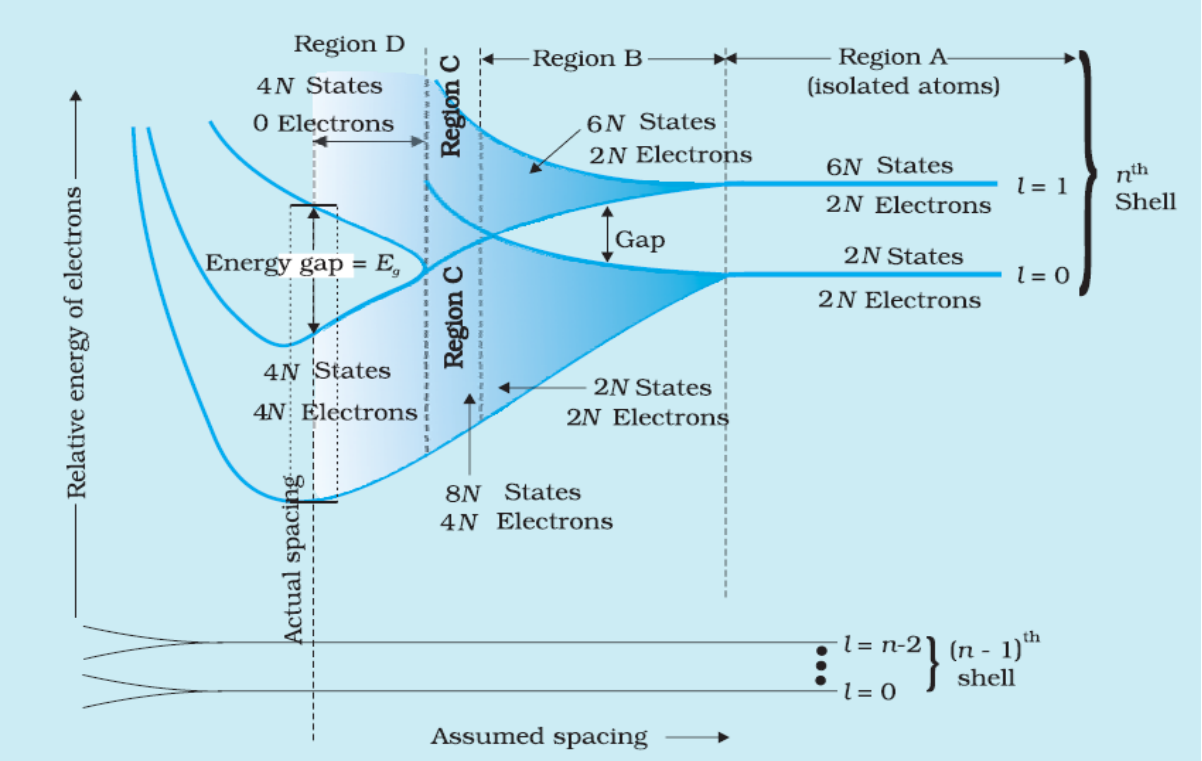


Figure 3.11: At the distance between the atoms in the crystal lattices of Si and Ge, the energy band of these 8*N* states is split apart into two which are separated by an *energy gap Eg*.

Consider that the Si or Ge crystal contains *N* atoms. Electrons of each atom will have discrete energies in different orbits. The electron energy will be same if all the atoms are *isolated*, i.e., separated from each other by a large distance. However, in a crystal, the atoms are close to each other (2 to 3 Å) and therefore the electrons interact with each other and also with the neighbouring atomic cores. The overlap (or interaction) will be more felt by the electrons in the outermost orbit while the inner orbit or core electron energies may remain unaffected. Therefore, for understanding electron energies in Si or Ge crystal, we need to consider the changes in the energies of the electrons in the outermost orbit only. For Si, the outermost orbit is the third orbit ($n = 3$), while for Ge it is the fourth orbit ($n = 4$). The number of electrons in the outermost orbit is 4 ($2s$ and $2p$ electrons). Hence, the total number of outer electrons in the crystal is $4N$. The maximum possible number of outer electrons in the orbit is 8 ($2s + 6p$ electrons). So, out of the $4N$ electrons, $2N$ electrons are in the $2N$ *s-states* (orbital quantum number $l = 0$) and 2N electrons are in the available $6N$ *p-states*. Obviously, some *p*-electron states are empty as shown in the extreme right of Figure 11. This is the case of well separated or isolated atoms [region A of Figure]. Suppose these atoms start coming nearer to each other to form a solid. The energies of these electrons in the outermost orbit may change (both increase and decrease) due to the interaction between the electrons of different atoms. The $6N$ states for $l = 1$, which originally had identical energies in the isolated atoms, spread out and form an *energy band* [region B in Figure]. Similarly, the $2N$ states for $l = 0$,

having identical energies in the isolated atoms, split into a second band (carefully see the region B of Figure) separated from the first one by an *energy gap*. At still smaller spacing, however, there comes a region in which the bands merge with each other. The lowest energy state that is a split from the upper atomic level appears to drop below the upper state that has come from the lower atomic level. In this region (region C in Figure), *no energy gap exists where the upper and lower energy states get mixed.* Finally, if the distance between the atoms further decreases, the energy bands again split apart and are separated by an *energy gap Eg* (region D in Figure 11).

The total number of available energy states 8*N* has been *re-apportioned* between the two bands (4*N* states each in the lower and upper energy bands). Here the significant point is that there are exactly as many states in the lower band (4*N*) as there are available valence electrons from the atoms (4*N*). Therefore, this band (called the *valence band*) is completely filled while the upper band is completely empty. The upper band is called the *conduction band*. The lowest energy level in the conduction band is shown as E_C and highest energy level in the valence band is shown as E_V. Above E_C and below E_V there are a large number of closely spaced energy levels, as shown in Fig (12). The gap between the top of the valence band and bottom of the conduction band is called the *energy band gap* (Energy gap *Eg*). It may be large, small, or zero, depending upon the material. These different situations are depicted in Fig. 12 and discussed below:

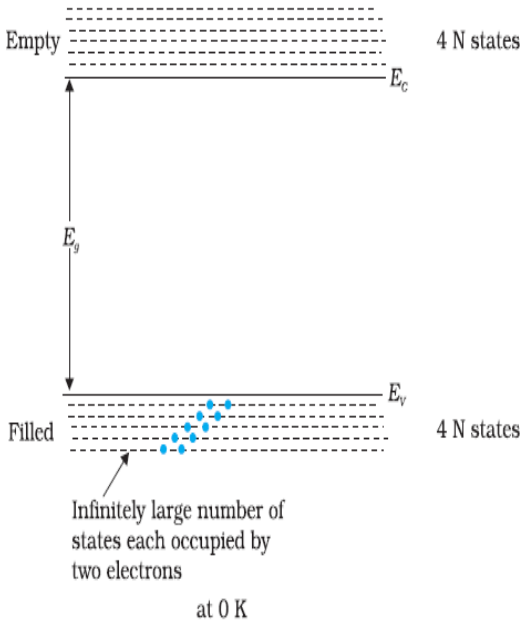


Figure 3.12: The energy band positions in a semiconductor at 0 K.

The upper band, called the conduction band, consists of infinitely large number of closely spaced energy states. The lower band, called the valence band, consists of closely spaced completely filled energy states.

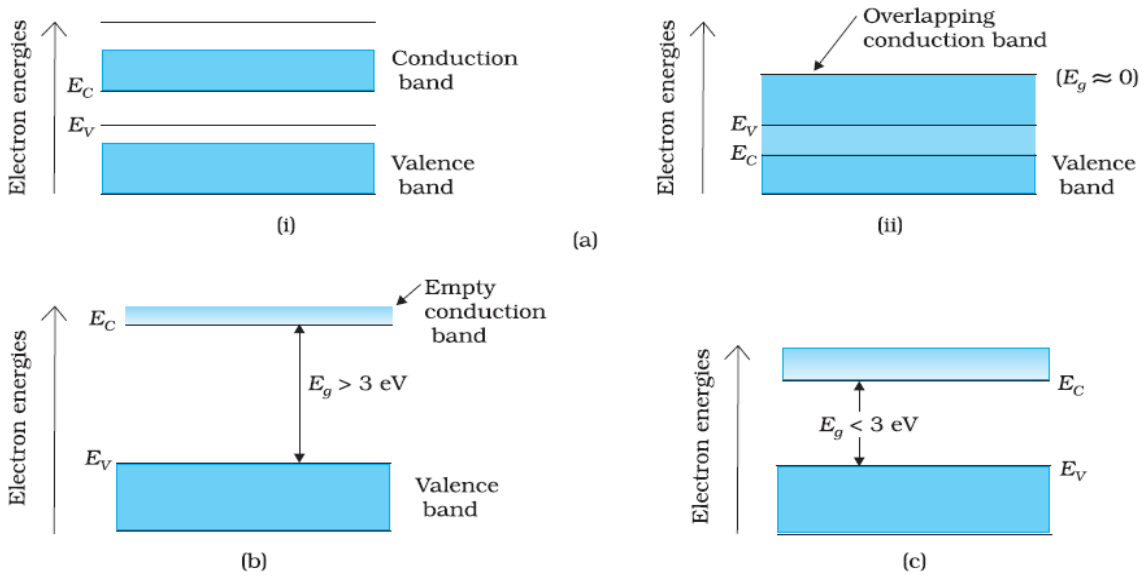


Figure 3.13: Difference between energy bands of (a) metals, (b) insulators and (c) semiconductors.

Case I: This refers to a situation, as shown in Fig. 13(a). One can have a metal either when the conduction band is partially filled and the balanced band is partially empty or when the conduction and valance bands overlap. When there is overlap electrons from valence band can easily move into the conduction band. This situation makes a large number of electrons available for electrical conduction. When the valence band is partially empty, electrons from its lower level can move to higher level making conduction possible. Therefore, the resistance of such materials is low or the conductivity is high.

Case II: In this case, as shown in Fig. 13 (b), a large band gap *Eg* exists (*Eg* > 3 eV). There are no electrons in the conduction band, and therefore no electrical conduction is possible. Note that the energy gap is so large that electrons cannot be excited from the valence band to the conduction band by thermal excitation. This is the case of *insulators*.

Case III: This situation is shown in Fig. 13 (c). Here a finite but small band gap (*Eg* < 3 eV) exists. Because of the small band gap, at room temperature some electrons from valence band can acquire enough energy to cross the energy gap and enter the *conduction band*. These electrons (though small in numbers) can move in the conduction band. Hence, the resistance of *semiconductors* is not as high as that of the Insulators. In this section we have made a broad classification of metals, conductors and semiconductors. In the section which follows you will learn the conduction process in semiconductors.

3.7 DRIFTS OF CARRIERS IN ELECTRIC AND MAGNETIC FIELDS

The movement of an empty electron state, i,e,. a hole under an electric field. The electrons move in the direction opposite to the electric field so that the hole moves in the direction of the electric field thus behaving as if it were positively charged. The velocities and currents due to electrons and holes. The current flow is in the same direction, even though the electron and holes have opposite velocities. The electron effective mass in the valence band is negative, but the hole behaves as if it has a positive mass.

3.8 HALL EFFECT

In 1879 E. H. Hall observed that when an electrical current passes through a sample placed in a magnetic field, a potential proportional to the current and to the magnetic field is developed across the material in a direction perpendicular to both the current and to the magnetic field. This effect is known as the Hall effect, and is the basis of many practical applications and devices such as magnetic field measurements, and position and motion detectors.

With the measurements he made, Hall was able to determine for the first time the sign of charge carriers in a conductor. Even today, Hall effect measurements continue to be a useful technique for characterizing the electrical transport properties of metals and semiconductors.

Theory of Hall Effect

Consider a conducting slab as shown in Figure_ with length L in the x direction, width w in the y direction and thickness t in the z direction. Assume the conductor to have charge carrier of charge q (can be either positive or negative or both, but we take it to be of just one sign here), charge carrier number density n (i.e., number of carriers per unit volume), and charge carrier drift velocity v_x when a current I_x flows in the positive x direction. The drift velocity is an average velocity of the charge carriers over the volume of the conductor; each charge carrier may move in a seemingly random way within the conductor, but under the influence of applied fields there will be a net transport of carriers along the length of the conductor. The current I_x is the current density J_x times the cross-sectional area of the conductor *wt.*

The current density J_x is the charge density *nq* times the drift velocity v_x. In other words

$$I_x = J_x wt = nqv_x wt$$

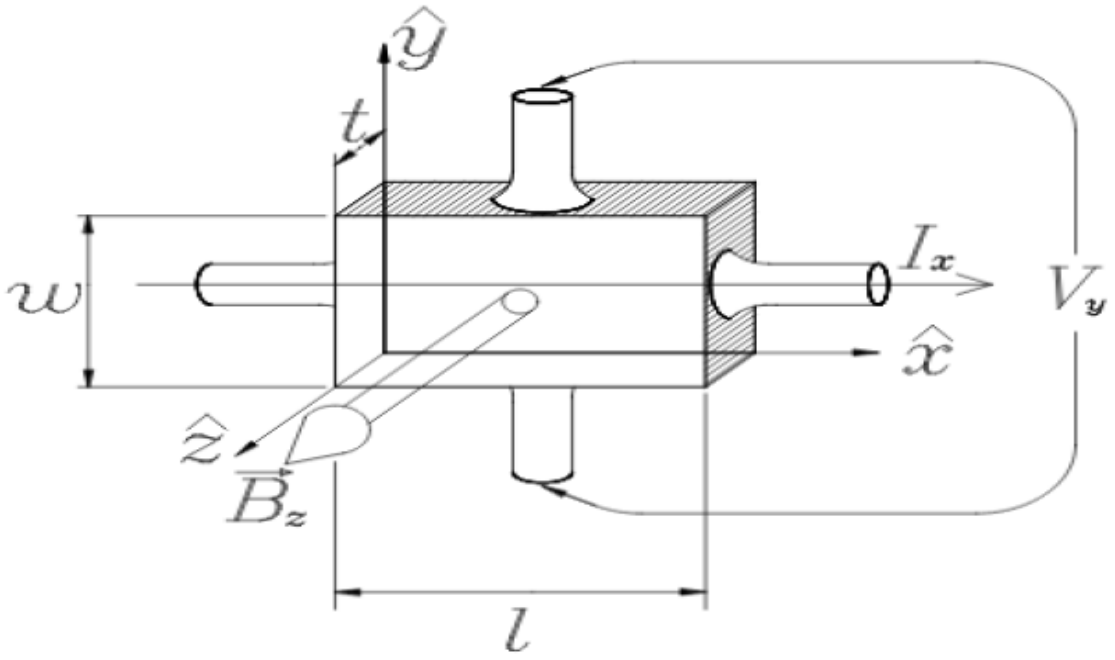


Figure 3.114: Geometry of fields and sample in Hall effect experiment.

The current I_x is caused by the application of an electric field along the length of the conductor E_x. In the case where the current is directly proportional to the field, we say that the material obeys Ohm's law, which may be written

$$J_x = \sigma E_x$$

where σ is the conductivity of the material in the conductor. Now assume that the conductor is placed in a magnetic field perpendicular to the plane of the slab. The charge carriers will experience a Lorentz force $q\vec{v} \times \vec{B}$ that will deflect them toward one side of the slab. The result of this deflection is to cause an accumulation of charges along one side of the slab which creates a transverse electric field E_y that counteracts the force of the magnetic field. When steady state is reached, there will be no net flow of charge in the y direction, since the electrical and magnetic forces on the charge carriers in that direction must be balanced. Assuming these conditions, it is easy to show that

$$E_y = v_x B_z$$

where E_y is the electric field, called the Hall field, in the y direction and B_z the magnetic field in the z direction.

In an experiment, we measure the potential difference across the sample, the Hall voltage V_H which is related to the Hall field by

$$V_H = -\int_0^w E_y \, dy = -E_y w$$

Thus from the above equation we have

$$V_H = -\left(\frac{1}{nq}\right)\frac{I_x B_z}{t}$$

The term in parenthesis is known as the Hall coefficient:

$$R_H = \frac{1}{nq}$$

It is positive if the charge carriers are positive, and negative if the charge carriers are negative. In practice, the polarity of V_H determines the sign of the charge carriers. The SI units of the Hall coefficient are m^3/C.

Applications The Hall effect measurements provide the following information about the solid

1. The sign (electrons or holes) of charge carrier is determined.

2. The carrier concentration (number of charge carriers per unit volume) is determined.

3. The mobility of charge carriers is measured directly.

4. We can decide whether a material is a metal, semiconductor or insulator.

5. From the knowledge of measured Hall voltage, the unknown magnetic field can be measured provided the value of Hall constant for the slab is known.

3.9 p-n JUNCTIONS

When a p-type semiconductor is suitably joined to n-type semiconductor, the contact surface is called **p-n junction.** Most semiconductor devices contain one or more *p-n* junctions. The *p-n* junction is of great importance because it is in effect, the *control element* for semiconductor devices. A thorough knowledge of the formation and properties of *p-n* junction can enable the reader to understand the semiconductor devices. A clear understanding of the junction behavior is important to analyse the working of other semiconductor devices. We will now try to understand how a junction is formed and how the junction behaves under the influence of external applied voltage.

3.10 p-n JUNCTION FABRICATION AND PROPERTIES

In actual practice, the characteristic properties of *p*-*n* junction will not be apparent if a *p*-type block is just brought in contact with *n*-type block. In fact, *p*-*n* junction is fabricated by special techniques. One common method of making *p*-*n* junction is called *alloying*. In this method, a small block of indium (trivalent impurity) is placed on an *n*-type germanium slab as shown in Figure (15 *i*). The system is then heated to a temperature of about 500ºC. The indium and some of the germanium melt to form a small puddle of molten germanium-indium mixture as shown in Figure (15 *ii*). The temperature is then lowered and puddle begins to solidify. Under proper conditions, the atoms of indium impurity will be suitably adjusted in the germanium slab to form a single crystal. The addition of indium overcomes the excess of electrons in the *n*-type germanium to such an extent that it creates a *p*-type region. As the process goes on, the remaining molten mixture becomes increasingly rich in indium. When all germanium has been redeposited, the remaining material appears as indium button which is frozen on to the outer surface of the crystallized portion as shown in Figure (15 *iii*). This button serves as a suitable base for soldering on leads.

Properties of *p*-*n* Junction

At the instant of *p*-*n*-junction formation, the free electrons near the junction in the *n* region begin to diffuse across the junction into the *p* region where they combine with holes near the junction.

The result is that *n* region loses free electrons as they diffuse into the junction. This creates a layer of positive charges (pentavalent ions) near the junction. As the electrons move across the junction, the *p* region loses holes as the electrons and holes combine. The result is that there is a layer of negative charges (trivalent ions) near the junction.

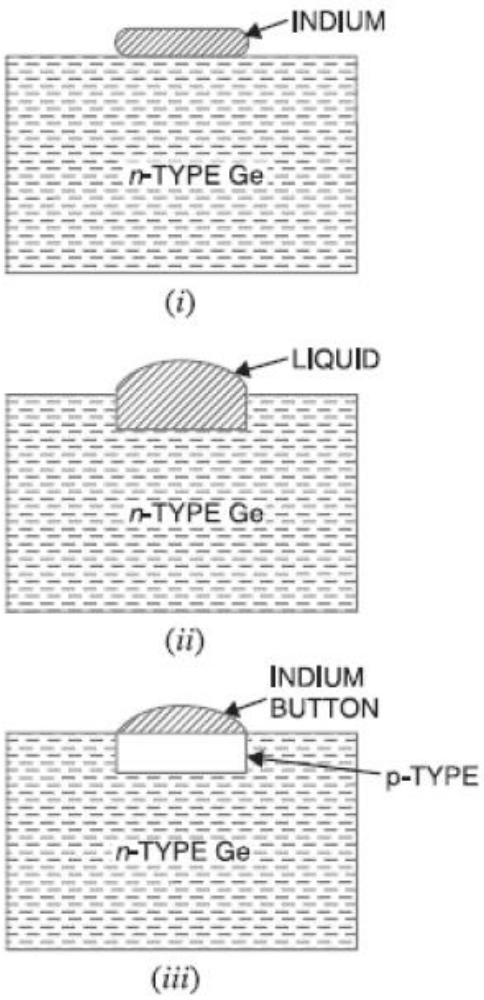


Figure 3.15: Fabrication of p-n junction.

These two layers of positive and negative charges form the *depletion region* (or *depletion layer*). The term depletion is due to the fact that near the junction, the region is depleted (*i.e.* emptied) of *charge carries* (free electrons and holes) due to diffusion across the junction.

It may be noted that depletion layer is formed very quickly and is very thin compared to the *n* region and the *p* region. For clarity, the width of the depletion layer is shown exaggerated.

Once *p*-*n* junction is formed and depletion layer created, the diffusion of free electrons stops. In other words, the depletion region acts as a barrier to the further movement of free electrons across the junction. The positive and negative charges set up an electric field. This is shown by a black arrow in Figure (16 *i*).

The electric field is a barrier to the free electrons in the *n*-region. There exists a potential difference across the depletion layer and is called **barrier potential** (V_0). The barrier potential of a *p*-*n*

junction depends upon several factors including the type of semiconductor material, the amount of doping and temperature. The typical barrier potential is approximately: For silicon, V_0 = 0.7 V ; For germanium, V_0 = 0.3 V.

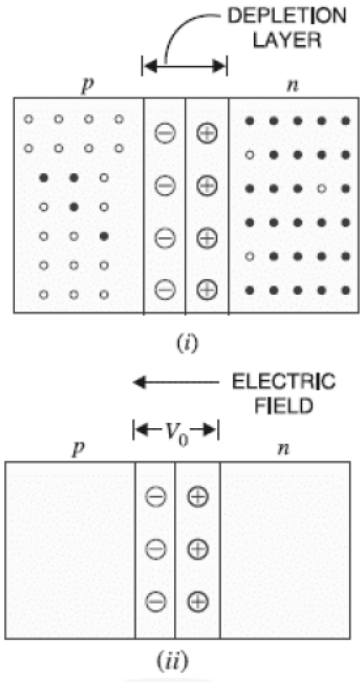


Figure 3.126: p-n junction.

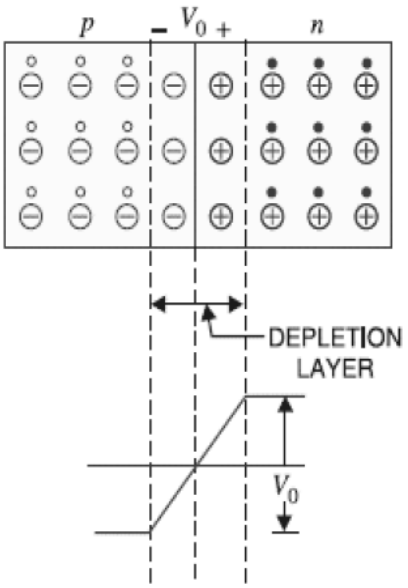


Figure 3.137: Potential distribution curve.

3.11 p-n JUNCTION DIODE

Biasing of p-n junction diode: In electronics, the term bias refers to the use of dc voltage to establish certain operating conditions for an electronic device. In relation to a *p-n* junction, there are following two bias conditions

1. Forward Biasing
2. Reverse Biasing

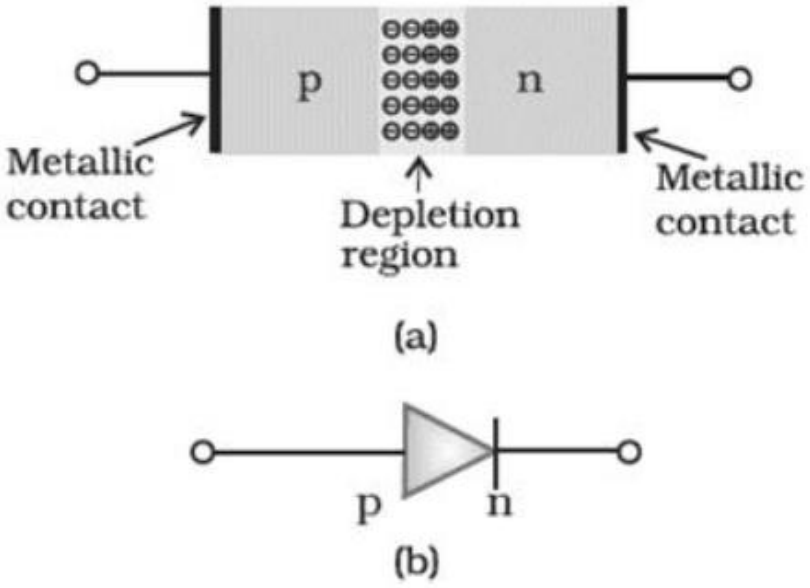


Figure 3.18: (a) Semiconductor diode (b) Symbol for p-n junction diode.

1. Forward Biasing

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

To apply forward bias, connect positive terminal of the battery to *p*-type and negative terminal to *n*-type as shown in Figure. The applied forward potential establishes an electric field which acts against the field due to potential barrier. Therefore, the resultant field is weakened and the barrier height is reduced at the junction as shown in Figure. As potential barrier voltage is very small (0.1 to 0.3 V), therefore, a small forward voltage is sufficient to completely eliminate the barrier.

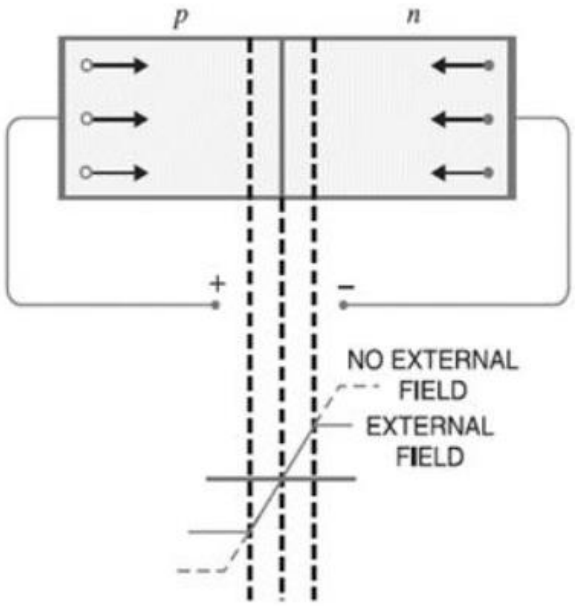


Figure 3.149: Forward Biasing of p-n junction diode.

Once the potential barrier is eliminated by the forward voltage, junction resistance becomes almost zero and a low resistance path is established for the entire circuit. Therefore, current flows in the circuit. This is called *forward current*. With forward bias to *p-n* junction, the following points are worth noting:

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

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

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

2. Reverse Biasing

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

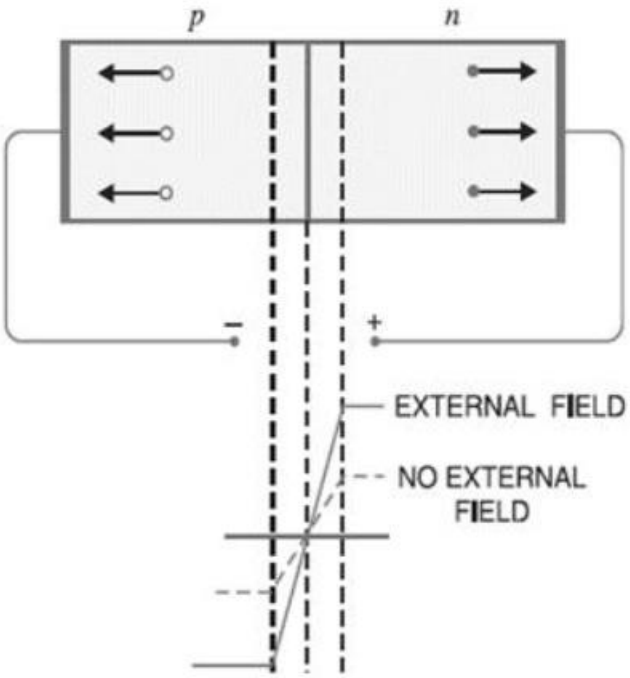


Figure 3.20: Reverse Biasing of p-n junction diode.

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

1. The potential barrier is increased.

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

3. No current flows in the circuit due to the establishment of high resistance path.

Electronic Circuit for Voltage Current (V-I) characteristics of a p-n junction diode

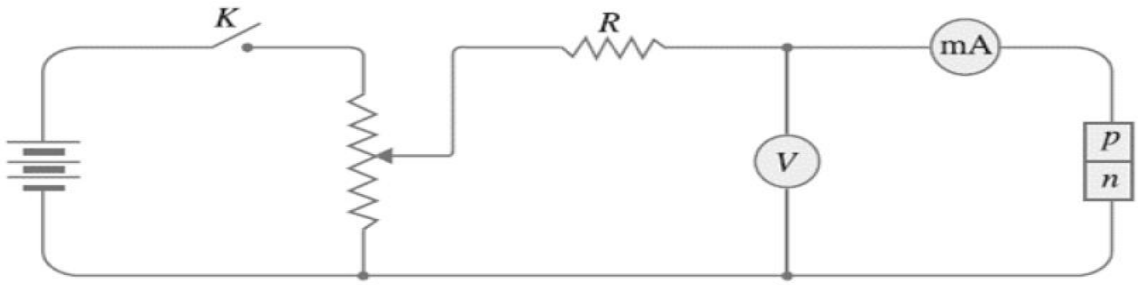


Figure 3.215: Circuit for the V-I characteristics of p-n junction diode.

Features of Voltage Current (V-I) characteristics of a p-n junction diode

Volt-ampere or *V-I* characteristic of a *p-n* junction (also called a *crystal or semiconductor diode*) is the curve between voltage across the junction and the circuit current. Usually, voltage is taken along *x* axis and current along *y*-axis. Figure 21 shows the circuit arrangement for determining the *V-I* characteristics of a *p-n* junction. The characteristics can be studied under three heads, namely; *zero external voltage*, *forward bias* and *reverse bias*.

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

2**. Forward bias.** With forward bias to the *p-n* junction *i.e. p*-type connected to positive terminal and *n*-type connected to negative terminal, the potential barrier is reduced. At some forward voltage (0.7 V for Si and 0.3 V for Ge), the potential barrier is altogether eliminated and current starts flowing in the circuit. From now onwards, the current increases with the increase in forward voltage.

Thus, a rising curve *OB* is obtained with forward bias as shown in Figure. From the forward characteristic, it is seen that at first (*region OA*), the current increases very slowly and the curve is non-linear. It is because the external applied voltage is used up in overcoming the potential barrier. However, once the external voltage exceeds the potential barrier voltage, the *p-n* junction behaves like an ordinary conductor. Therefore, the current rises very sharply with increase in external voltage (*region AB on the curve*). The curve is almost linear.

3. **Reverse bias**. With reverse bias to the *p-n* junction *i.e. p*-type connected to negative terminal and *n*-type connected to positive terminal, potential barrier at the junction is increased. Therefore, the junction resistance becomes very high and practically no current flows through the circuit. However, in practice, a very small current (of the order of μ*A*) flows in the circuit with reverse bias as shown in the reverse characteristic. This is called *reverse saturation current* (I_s) and is due to the minority carriers. It may be recalled that there are a few free electrons in *p*-type material and a few holes in *n*-type material. These undesirable free electrons in *p*-type and holes in *n*-type are called *minority carriers*. These minority carriers, the applied reverse bias appears as forward bias. Therefore, a small current flows in the reverse direction. If reverse voltage is increased continuously, the kinetic energy of electrons (minority carriers) may become high enough to knock out electrons from the semiconductor atoms. At this stage *breakdown* of the junction occurs, characterized by a sudden rise of reverse current and a sudden fall of the resistance of barrier region. This may destroy the junction permanently.

3.12 p-n JUNCTION BREAKDOWN

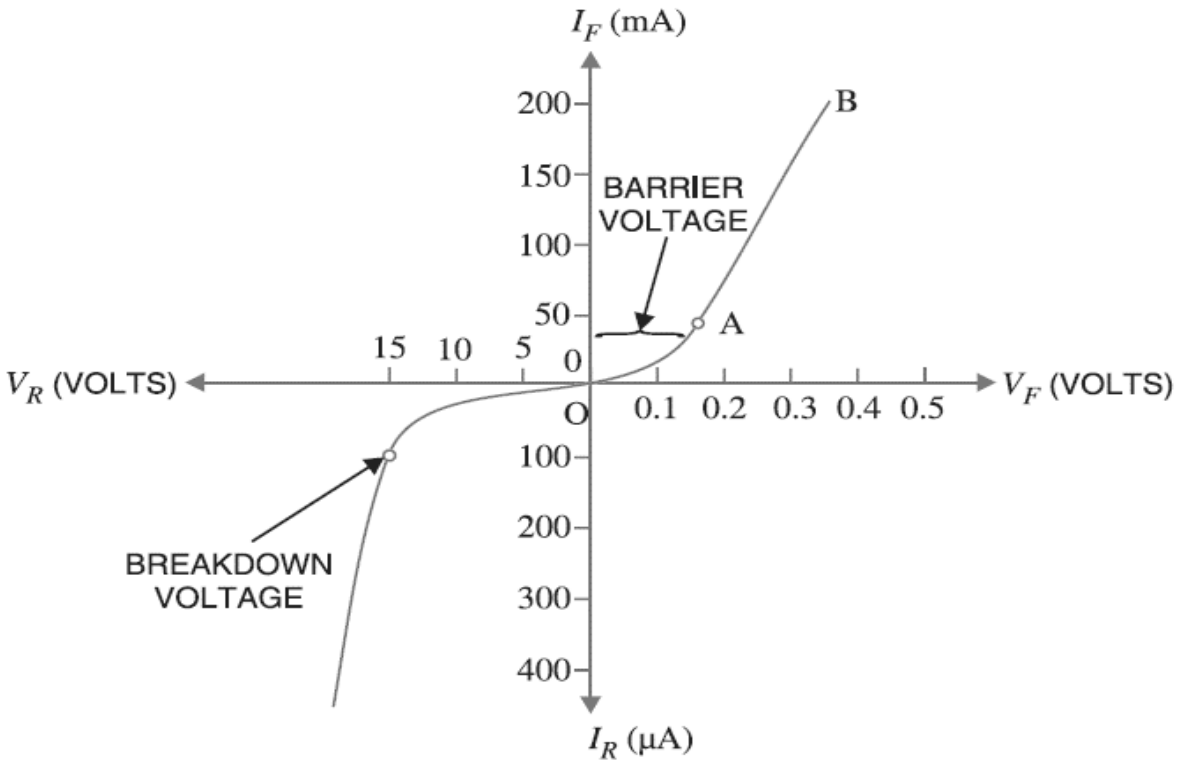


Figure 3.22: V-I characteristics of a p-n junction diode.

The process by which a depletion region at the p-n junction is destroyed and allows a large reverse current is called depletion region breakdown.

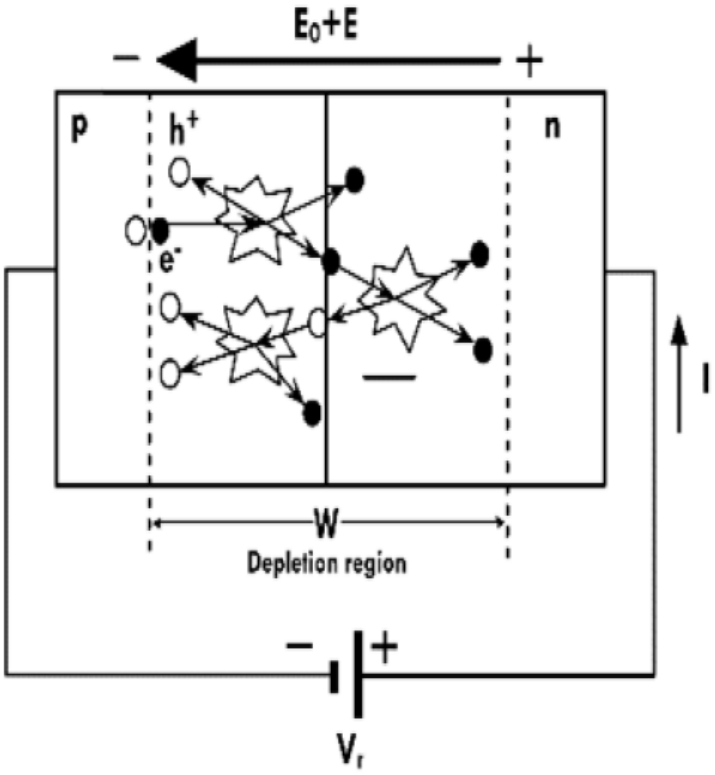


Figure 3.23: Avalanche breakdown.

Consider the V-I characteristics of a p-n junction in reverse bias, as shown in Figure. There is initially a small reverse saturation current due to thermally generated electron and holes in the depletion region. This current is called drift current, since this is due to movement of the thermally generated carriers under the applied electric field. With increase in voltage there is a particular value, called the breakdown voltage, beyond which the current increases rapidly. This is called junction breakdown. There are two main mechanisms of junction breakdown, depending on the dopant concentration levels.

1. Avalanche breakdown

2. Zener breakdown

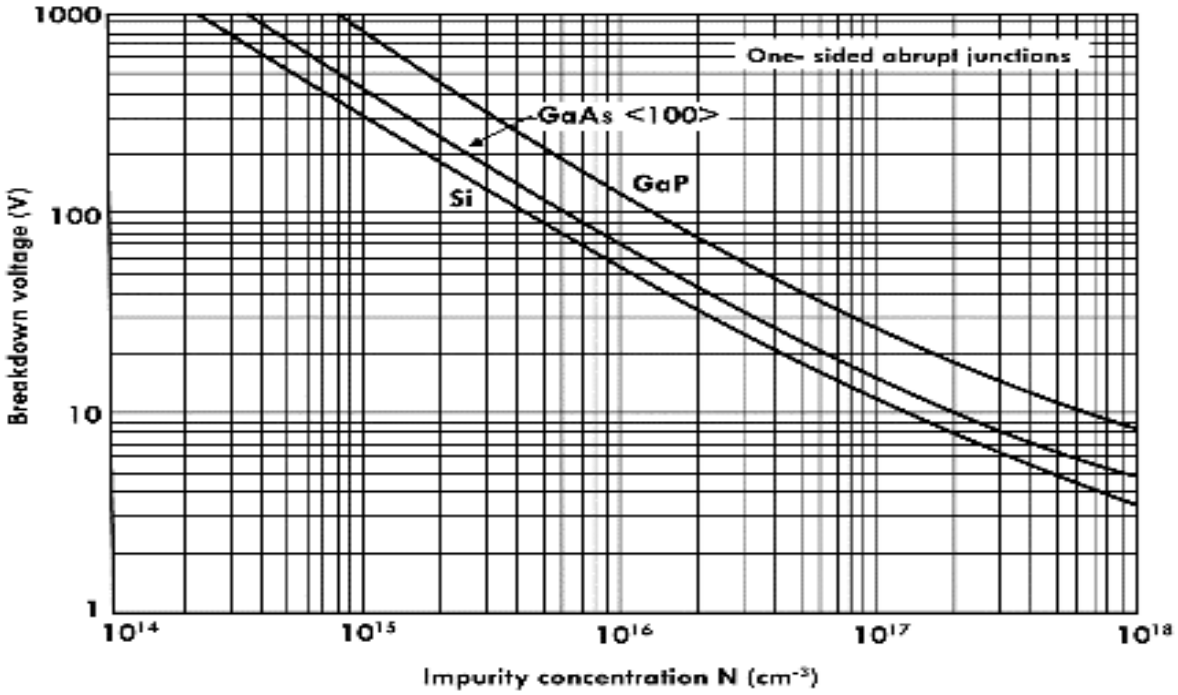


Figure 24: Avalanche breakdown voltage as a function of dopant concentration.

Avalanche Breakdown: Avalanche breakdown occurs in moderately and lightly doped p-n junctions with a wide depletion region. The schematic of the process is shown in Figure 23. Electron hole pairs thermally generated in the depletion region are accelerated by the external reverse bias. Electrons are accelerated towards then side and holes towards the p side. These electrons can interact with other Si atoms and if they have sufficient energy can knock off electrons from these Si atoms. This process is called impact ionization and leads to production of a large number of electrons. This causes the rapid rise in current. The breakdown voltage decreases with increase in dopant concentration, as shown in Figure 24.

Zener breakdown: With increase in doping concentration the breakdown mechanism, changes from Avalanche to a tunneling mechanism. This is called a Zener breakdown. This is because the depletion width decreases with dopant concentration. Also, the reverse bias causes an offset in the bands such that it is possible for carriers to tunnel across the narrow depletion region. This tunneling process is shown schematically in Figure, where electrons tunnel from the valence band on the p side to the conduction band on the n side, driven by the externally applied reverse bias. Tunneling also leads to a large increase in current. The transition from avalanche to Zener as the primary breakdown mechanism with dopant concentration is shown in Figure. Zener diodes are primarily used as surge protectors in circuits, since there is a rapid increase in current with a small change in voltage.

Prior to breakdown there is a high resistance (small reverse saturation current) but after breakdown the resistance is very small. This can be used as voltage regulators in circuits.

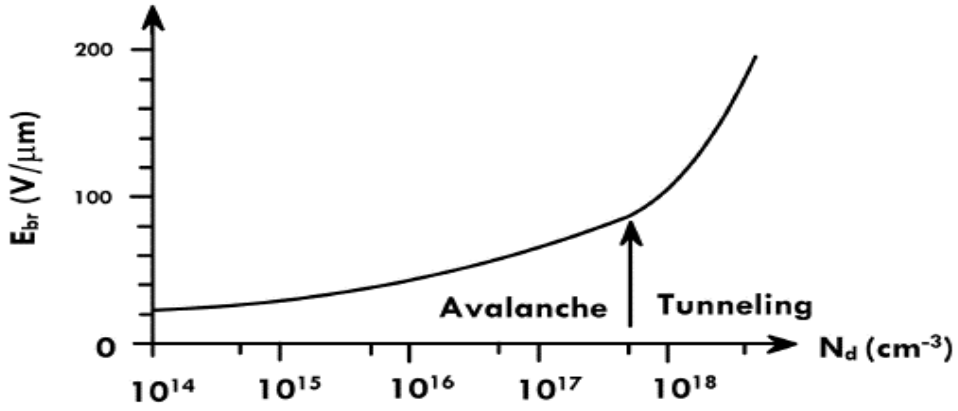


Figure 3.25: Tunneling of electrons from p to n side in a heavily doped p-n junction under reverse bias.

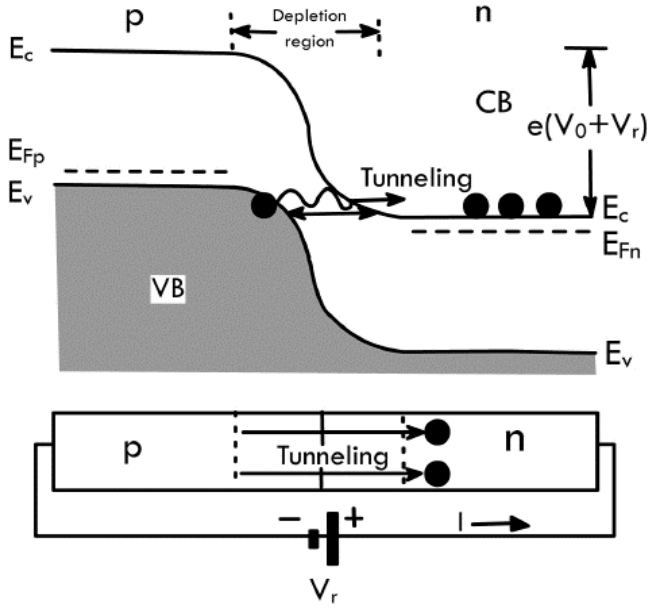


Figure 3.26: Breakdown field vs dopant concentration.

Key differences between avalanche & Zener breakdown

1. The breakdown which occurs because of the collision of the electrons inside the p-n junction is called avalanche breakdown, whereas the Zener breakdown occurs when the heavy electric field is applied across the p-n junction.
2. The avalanche breakdown occurs in the thick region, whereas the Zener breakdown occurs in the thin region.
3. After the avalanche breakdown, the junction of the diode will not regain its original position, whereas after the Zener breakdown the junction regains its original position.
4. The existence of the electric field is more on the Zener breakdown as compared to the avalanche breakdown. Because the mechanism of Zener breakdown occurs in the heavily doped region.
5. The avalanche breakdown produces the pairs of electrons and holes because of the thermal effects, whereas the Zener diode produces the electrons.
6. The avalanche breakdown occurs in low doping material, whereas the Zener breakdown occurs in high doping material.
7. The avalanche breakdown voltage causes because of high reverse potential because it is lightly doped whereas the Zener breakdown is because of low reverse potential.
8. The temperature coefficient of the avalanche breakdown is positive, whereas the temperature coefficient of Zener breakdown is negative. The positive temperature coefficient means the temperature of the material increases with the reverse voltage and negative temperature coefficient means temperature decreases with the potential differences.
9. In avalanche breakdown, the mechanism of ionization occurs because of collision of electrons, whereas in the Zener breakdown ionization occurs because of the electric field.
10. The avalanche breakdown voltage is directly proportional to the temperature, whereas the Zener breakdown voltage is inversely proportional to the temperature. The voltage of Zener breakdown is less than the Avalanche breakdown.

3.13 ZENER DIODE

It is a special purpose semiconductor diode, named after its inventor C. Zener. It is designed to operate under reverse bias in the breakdown region and used as a voltage regulator. The symbol for Zener diode is shown in Figure 27. Zener diode is fabricated by heavily doping both p-, and n-sides of the junction. Due to this, depletion region formed is very thin ($<10^{-6}$m) and the electric field of the junction is extremely high ($\sim 5 \times 10^{6}$ V/m) even for a small reverse bias voltage of about 5V.

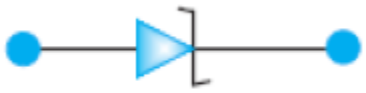


Figure 3.167: Circuit symbol of Zener diode.

The breakdown or zener voltage depends upon the amount of doping. If the diode is heavily doped, depletion layer will be thin and consequently the breakdown of the junction will occur at a lower reverse voltage. On the other hand, a lightly doped diode has a higher breakdown voltage. When an ordinary crystal diode is properly doped so that it has a sharp breakdown voltage, it is called a zener diode. A typical Zener diode characteristic is shown in **Figure 28**. The maximum reverse current, $I_{Z(\max)}$ which the Zener diode can withstand is dependent on the design and construction of the diode. A design guideline that the minimum Zener current, where the characteristic curve remains at V_z (near the knee of the curve), is $0.1/I_{Z(\max)}$.

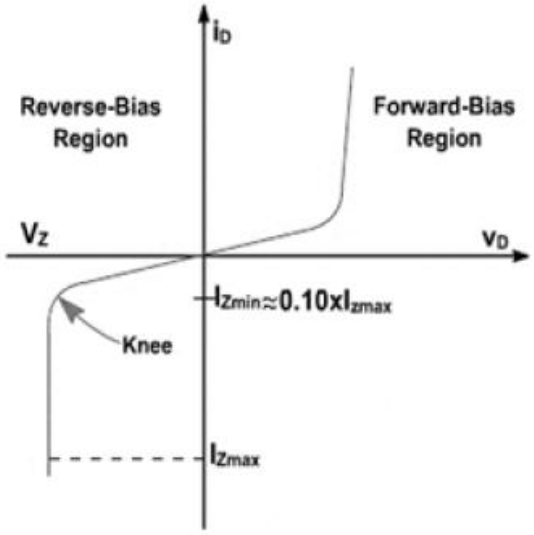


Figure 3.178: Zener diode characteristics.

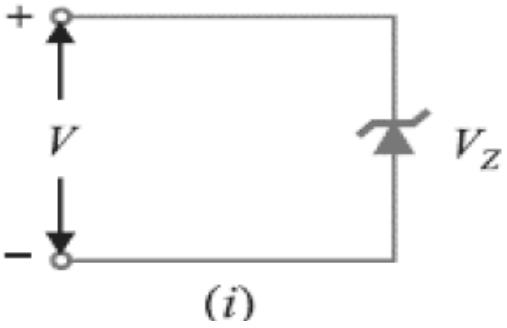


Figure 3.29: V ≥ V_Z

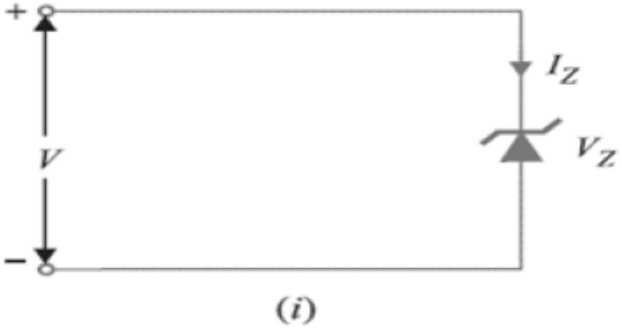


Figure 3.30: Equivalent circuit of Zener for "ON" state.

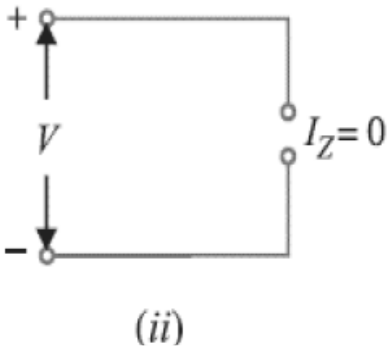


Figure 3.31:18 V_Z > V > 0.

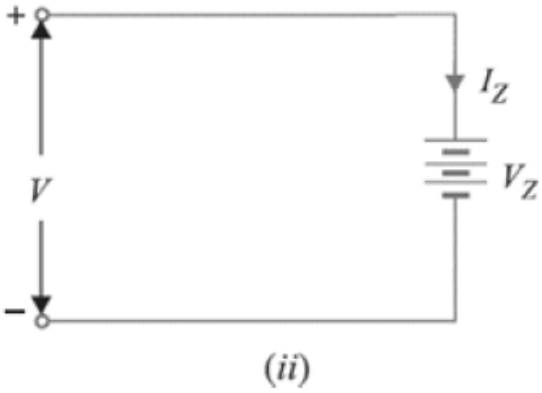


Figure 3.32: Equivalent circuit of Zener for "OFF" state.

The power handling capacity of these diodes is better. The power dissipation of a zener diode equals the product of its voltage and current.

$$P_Z = V_Z I_Z$$

The amount of power which the zener diode can withstand ($V_Z I_{Z(\max)}$) is a limiting factor in power supply design.

Equivalent Circuit of Zener Diode

The analysis of circuits using zener diodes can be made quite easily by replacing the zener diode by its equivalent circuit.

(i) **"On" state**. When reverse voltage across a zener diode is equal to or more than break down voltage V_Z, the current increases very sharply. In this region, the curve is almost vertical. It means that voltage across zener diode is constant at V_Z even though the current through it changes. Therefore, in the breakdown region, an ideal zener diode can be represented by a battery of voltage V_Z as shown in Figure (*i*). Under such conditions, the zener diode is said to be in the "ON" state.

(ii) **"OFF" state.** When the reverse voltage across the zener diode is less than V_Z but greater than 0 V, the zener diode is in the "OFF" state. Under such conditions, the zener diode can be represented by an open-circuit as shown in Fig.32.

Zener Diode as Voltage Stabilizer

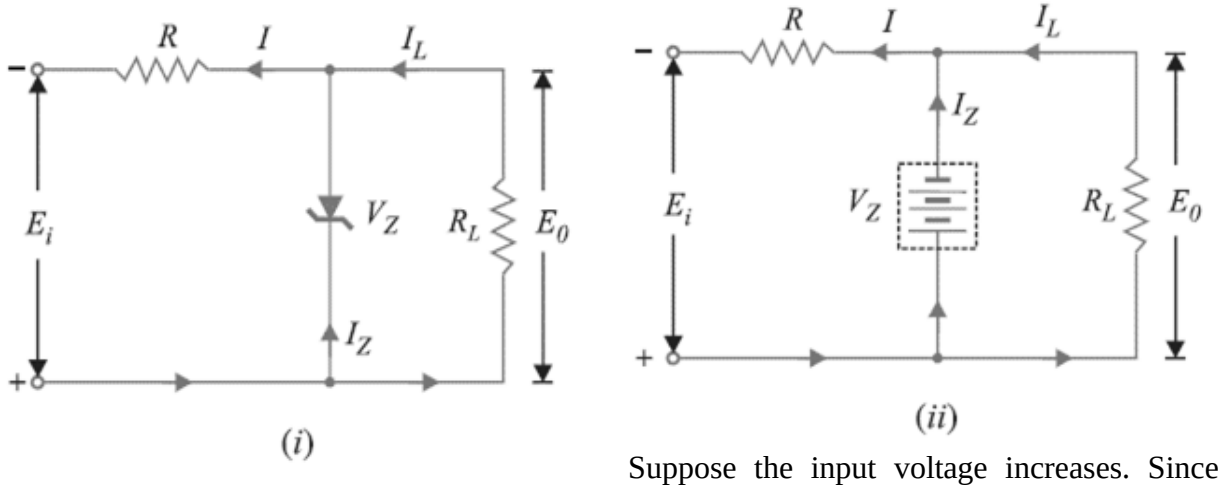
A zener diode can be used as a voltage regulator to provide a constant voltage from a source whose voltage may vary over sufficient range. The circuit arrangement is shown in Figure (*i*). The zener diode of zener voltage V_Z is reverse connected across the load R_L across which constant output is

desired. The series resistance *R* absorbs the output voltage fluctuations so as to maintain constant voltage across the load. It may be noted that the zener will maintain a constant voltage V_Z (=E_0) across the

Load so long as the input voltage does not fall below V_Z. When the circuit is properly designed, the load voltage E_0 remains essentially constant (equal to V_Z) even though the input voltage E_i and load resistance R_L may vary over a wide range.

Figure 3.33: Zener as a Voltage Regulator.

Figure 3.194: Equivalent circuit of zener as a voltage regulator.



Suppose the input voltage increases. Since the zener is in the breakdown region, the zener diode is equivalent to a battery V_Z as shown in Fig.34. It is clear that output voltage remains constant at V_Z(= E_0). The excess voltage is dropped across the series resistance *R*.

This will cause an increase in the value of total current *I*. The zener will conduct the increase of current in *I* while the load current remains constant. Hence, output voltage E_0 remains constant irrespective of the changes in the input voltage E_i. Now suppose that input voltage is constant but the load resistance R_L decreases. This will cause an increase in load current. The extra current cannot come from the source because drop in *R* (and hence source current *I*) will not change as the zener is within its regulating range. The additional load current will come from a decrease in zener current I_Z. Consequently, the output voltage stays at constant value.

Voltage drop across $R = E_i - E_0$

Current through *R*, $I = I_Z + I_L$

Applying Ohm's law, we have,

$$R = \frac{E_i - E_0}{I_Z + I_L}$$

3.14 TUNNEL DIODE

The tunnel diode was first introduced by Leo Esaki in 1958. Its characteristics, shown in Figure, are different from any diode discussed thus far in that it has a negative-resistance region. In this region, an increase in terminal voltage results in a reduction in diode current.

The tunnel diode is fabricated by doping the semiconductor materials that will form the *p-n* junction at a level 100 to several thousand times that of a typical semiconductor diode. This results in a greatly reduced depletion region, of the order of magnitude of 10^{-6}cm, or typically about 1/100 the width of this region for a typical semiconductor diode. It is this thin depletion region, through which many carriers can "tunnel" rather than attempt to surmount, at low forward-bias potentials that accounts for the peak in the curve of Figure. For comparison purposes, a typical semiconductor diode characteristic is superimposed on the tunnel-diode characteristic of Figure 35.

This reduced depletion region results in carriers "punching through" at velocities that far exceed those available with conventional diodes. The tunnel diode can therefore be used in high-speed applications such as in computers, where switching times in the order of nanoseconds or picoseconds are desirable. We know that an increase in the doping level reduces the Zener potential. Note the effect of a very high doping level on this region in Figure. The semiconductor materials most frequently used in the manufacture of tunnel diodes are germanium and gallium arsenide. The ratio $I_P > I_V$is very important for computer applications. For germanium, it is typically 10:1, and for gallium arsenide, it is closer to 20:1. The peak current I_P of a tunnel diode can vary from a few microamperes to several hundred amperes. The peak voltage, however, is limited to about 600 mV. For this reason, a simple VOM with an internal dc battery potential of 1.5 V can severely damage a tunnel diode if applied improperly.

Figure 3.205: Circuit symbol of Tunnel diode.

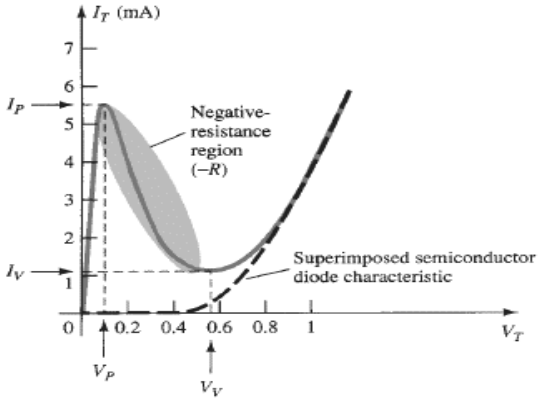


Figure 3.36: Tunnel diode characteristics.

V-I Characteristic. Above Figure shows the V-I characteristic of a typical tunnel diode.

(*i*) As the forward voltage across the tunnel diode is increased from zero, electrons from the *n* region "tunnel" through the potential barrier to the *p*-region. As the forward voltage increases, the diode current also increases until the *peak-point P* is reached. The diode current has now reached peak current I_P at about peak-point voltage V_P. Until now the diode has exhibited positive resistance.

(*ii*) As the voltage is increased beyond V_P, the tunneling action starts decreasing and the diode current decreases as the forward voltage is increased until *valley-point* is reached at valley-point voltage V_V. In the region between peak-point and valley-point the diode exhibits negative resistance *i.e*., as the forward bias is increased, the current decreases. This suggests that tunnel diode, when operated in the negative resistance region, can be used as an oscillator or a switch.

(*iii*) When forward bias is increased beyond valley-point voltage V_V, the tunnel diode behaves as a normal diode. In other words, from point V onwards, the diode current increases with the increase in forward voltage *i.e.*, the diode exhibits positive resistance once again. Figure shows the symbol of tunnel diode. It may be noted that a tunnel diode has a high reverse current but operation under this condition is not generally used.

3.15 VARACTOR (VARICAP) DIODES

A junction diode which acts as a variable capacitor under changing reverse bias is known as a **varactor diode.**

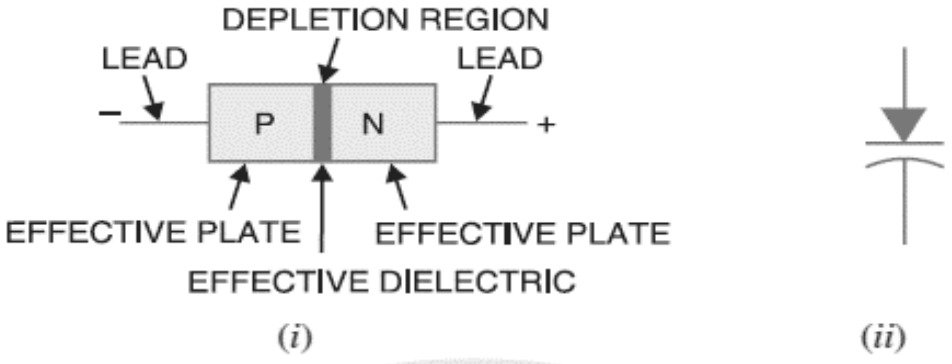


Figure 3.37: Varactor diode.

When a *p*-*n* junction is formed, depletion layer is created in the junction area. Since there are no charge carriers within the depletion zone, the zone acts as an insulator. The *p*-type material with holes (considered positive) as majority carriers and *n*-type material with electrons (negative

charge) as majority carriers act as charged plates. Thus the diode may be considered as a capacitor with *n*-region and *p* region forming oppositely charged plates and with depletion

zone between them acting as a dielectric. This is illustrated in Figure (*i*). A varactor diode is specially constructed to have high capacitance under reverse bias. Figure
(*ii*) shows the symbol of varactor diode. The values of capacitance of varactor diodes are in the picofarad (10^{-12} F) range.

Theory: For normal operation, a *varactor diode is always reverse biased.* The capacitance of varactor diode is found as

$$C_T = \varepsilon \frac{A}{W_d}$$

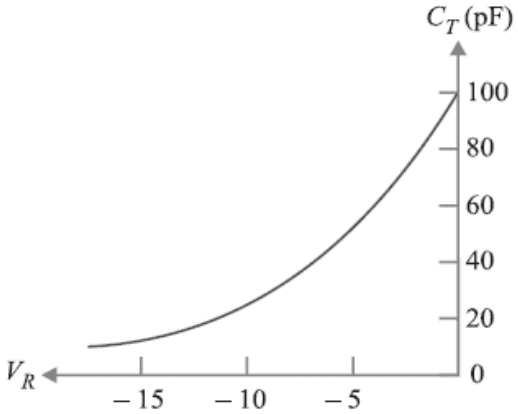


Figure 3.38. Varactor diode characteristics.

where

C_T = Total capacitance of the junction

ε = permittivity of the semiconducting material

A = Area of cross section of the junction

W_d = Width of the depletion layer

When reverse voltage across a varactor diode is increased, the width W_d of the depletion layer increases. Therefore, the total junction capacitance C_T of the junction decreases. On the other hand, if the reverse voltage across the diode is lowered, the width W_d of the depletion layer decreases. Consequently, the total junction capacitance C_T increases.

Figure shows the curve between reverse bias voltageV_R across varactor diode and total junction capacitance C_T. Note that C_T can be changed simply by changing the voltage V_R. For this reason, a varactor diode is sometimes called **voltage-controlled capacitor.**

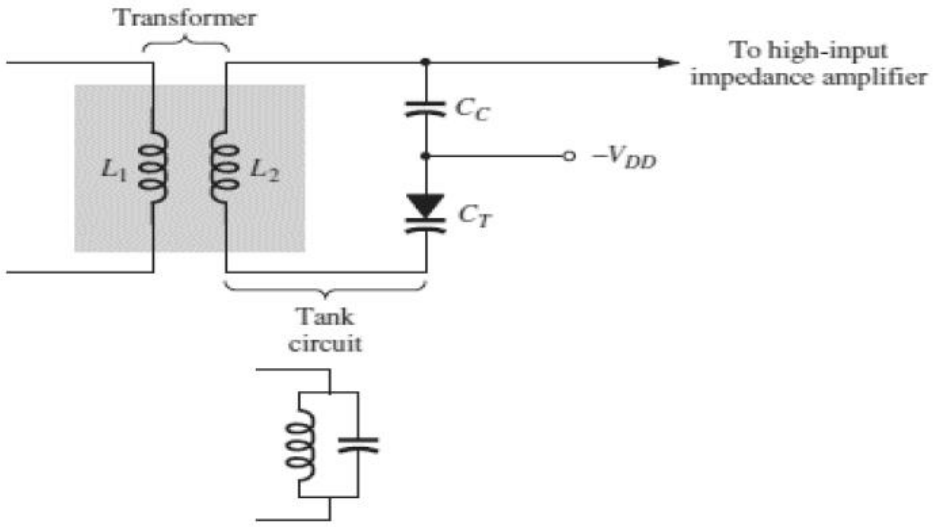


Figure 3.39: Tuning network using varactor diode.

Application: In Figure, the varactor diode is employed in a tuning network. That is, the resonant frequency of the parallel *LC* combination is determined by $f_P = 1/2\pi\sqrt{L_2 C_T'}$ (high- *Q* system) with the level of $C_T' = C_T + C_C$ determined by the applied reverse-bias potential V_{DD} The coupling capacitor C_C is present to provide isolation between the shorting effect of L_2 and the applied bias. The selected frequencies of the tuned network are then passed on to the high-input amplifier for further amplification.

3.16 PHOTO DIODE AND LED

We have seen so far, how a semiconductor diode behaves under applied electrical inputs. Now we learn about semiconductor diodes in which carriers are generated by photons (photo-excitation). All these devices are called *optoelectronic devices*. We shall study the functioning of the following optoelectronic devices:
(i) *Photodiodes* used for detecting optical signal (photo detectors).
(ii) *Light emitting diodes* (LED) which convert electrical energy into light.
(iii) *Photovoltaic devices* which convert optical radiation into electricity (*solar cells*).

(i) *Photodiode*

A Photodiode is again a special purpose p-n junction diode fabricated with a transparent window to allow light to fall on the diode. It is operated under reverse bias. When the photodiode is illuminated with light (photons) with energy (*h*v) greater than the energy gap (*Eg*) of the semiconductor, then electron-hole pairs are generated due to the absorption of photons. The diode

is fabricated such that the generation of *e-h* pairs takes place in or near the depletion region of the diode. Due to electric field of the junction, electrons and holes are separated before they recombine. The direction of the electric field is such that electrons reach n-side and holes reach p-side. Electrons are collected on n-side and holes are collected on p-side giving rise to an emf.
When an external load is connected, current flows. The magnitude of the photocurrent depends on the intensity of incident light (photocurrent is proportional to incident light intensity). It is easier to observe the change in the current with change in the light intensity, if a reverse bias is applied. Thus photodiode can be used as a photodetector to detect optical signals.

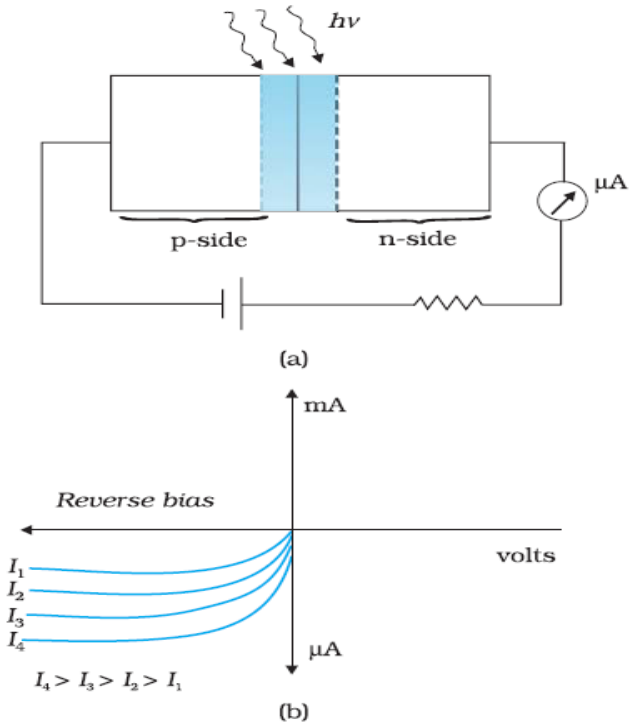


Figure 3.40: (a) An illuminated photodiode under reverse bias, (b) *I-V* characteristics of a Photodiode for different illumination intensity $I_4 > I_3 > I_2 > I_1$.

The circuit diagram used for the measurement of *I-V* characteristics of a photodiode is shown in Fig. 40(a) and a typical I-*V* characteristics in Fig. 40 (b).

A **photo-diode** is a reverse-biased silicon or germanium p-n junction in which reverse current increases when the junction is exposed to light. The reverse current in a photo-diode is directly proportional to the intensity of light falling on its p-n junction. This means that
greater the intensity of light falling on the p-n junction of photo-diode, the greater will be the reverse current.

Principle: When a rectifier diode is reverse biased, it has a very small reverse leakage current. The same is true for a photo-diode. The reverse current is produced by thermally generated electron hole pairs which are swept across the junction by the electric field created by the reverse voltage. In a rectifier diode, the reverse current increases with temperature due to an increase in the number of electron-hole pairs. A photo-diode differs from a rectifier diode in that when its p-n junction is exposed to light, the reverse current increases with the increase in light intensity and vice-versa. This is explained as follows. When light (photons) falls on the p-n junction, the energy is imparted by the photons to the atoms in the junction. This will create more free electrons (and more holes). These additional free electrons will increase the reverse current. As the intensity of light incident on the p-n junction increases, the reverse current also increases. In other words, as the incident light intensity increases, the resistance of the device (photo-diode) decreases.

Photo-diode package. Figure 41 (*i*) shows a typical photo-diode package. It consists of a *p-n* junction mounted on an insulated substrate and sealed inside a metal case. A glass window is mounted on top of the case to allow light to enter and strike the *p-n* junction. The two leads extending from the case are labelled anode and cathode. The cathode is typically identified by a tab extending from the side of the case. Figure 41 (*ii*) shows the schematic symbol of a photo-diode. The inward arrows represent the incoming light.

Figure 3.41: Photo diode.

Characteristics of Photo-diode There are two important characteristics of photodiode.

(*i*) Reverse current -Illumination curve. Figure 42 shows the graph between reverse current (I_R) and illumination (E) of a photo-diode. The reverse current is shown on the vertical axis and is measured in μA. The illumination is indicated on the horizontal axis and is measured in mW/cm^2. Note that graph is a straight line passing through the origin. $I_R = m\,E$ where m = slope of the straight line. The quantity m is called the *sensitivity* of the photo-diode.

(*ii*) Reverse voltage -Reverse current curve. Figure 43 shows the graph between reverse current (I_R) and reverse voltage (V_R) for various illumination levels. It is clear that for a given reverse-

biased voltage, the reverse current I_R increases as the illumination (E) on the *p-n* junction of photo-diode are increased.

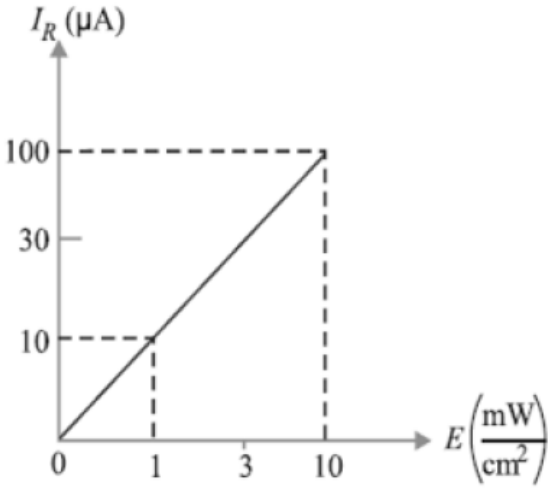


Figure 3.42: Reverse current I_R versus illumination E of photo diode.

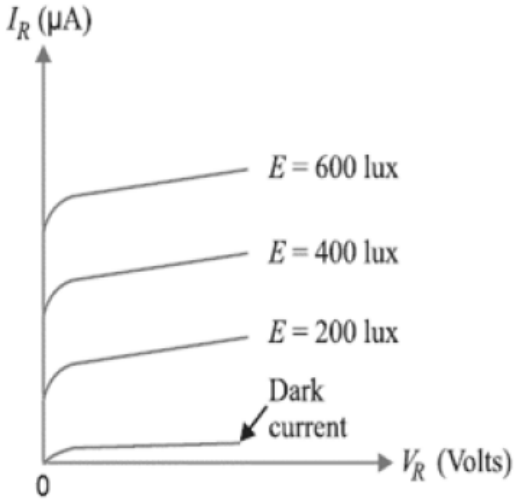


Figure 3.43: Reverse current I_R versus reverse voltage V_R.

Applications of Photo diode: Some common applications include alarm system and counter circuits.

(ii) *Light emitting diode*

It is a heavily doped p-n junction which under forward bias emits spontaneous radiation. The diode is encapsulated with a transparent cover so that emitted light can come out. When the diode is forward biased, electrons are sent from n → p (where they are minority carriers) and holes are sent from p → n (where they are minority carriers). At the junction boundary the concentration of minority carriers increases compared to the equilibrium concentration (i.e., when there is no bias). Thus at the junction boundary on either side of the junction, excess minority carriers are there which recombine with majority carriers near the junction. On recombination, the energy is released in the form of photons. Photons with energy equal to or slightly less than the band gap are emitted. When the forward current of the diode is small, the intensity of light emitted is small. As the forward current increases, intensity of light increases and reaches a maximum. Further increase in the forward current results in decrease of light intensity. LEDs are biased such that the light emitting efficiency is maximum. The *V-I* characteristics of a LED is similar to that of a Si junction diode. But the threshold voltages are much higher and slightly different for each colour. The reverse breakdown voltages of LEDs are very low, typically around 5V. So care should be taken that high reverse voltages do not appear across them. LEDs that can emit red, yellow, orange,

green and blue light are commercially available. The semiconductor used for fabrication of visible LEDs must at least have a band gap of 1.8 eV (spectral range of visible light is from about 0.4 μm to 0.7 μm, i.e., from about 3 eV to 1.8 eV). The compound semiconductor Gallium Arsenide – Phosphide ($GaAs_{1-x}P_x$) is used for making LEDs of different colours. $GaAs_{0.6}P_{0.4}$ (Eg ~ 1.9 eV) is used for red LED. GaAs (Eg ~ 1.4 eV) is used for making infrared LED. These LEDs find extensive use in remote controls, burglar alarm systems, optical communication, etc. Extensive research is being done for developing white LEDs which can replace incandescent lamps. LEDs have the following advantages over conventional incandescent low power lamps:

(i) Low operational voltage and less power.
(ii) Fast action and no warm-up time required.
(iii) The bandwidth of emitted light is 100 Å to 500 Å or in other words it is nearly (but not exactly) monochromatic.
(iv) Long life and ruggedness.
(v) Fast on-off switching capability.

The increasing use of digital displays in calculators, watches, and all forms of instrumentation has contributed to an extensive interest in structures that emit light when properly biased. The two types in common use to perform this function are the light-emitting diode (LED) and the liquid-crystal display (LCD). The light-emitting diode is a diode that gives off visible or invisible (infrared) light when energized. In any forward-biased $p - n$ junction there is, within the structure and primarily close to the junction, a recombination of holes and electrons. This recombination requires that the energy possessed by the unbound free electrons be transferred to another state. In all semiconductor $p - n$ junctions some of this energy is given off in the form of heat and some in the form of photons.

In Si and Ge diodes the greater percentage of the energy converted during recombination at the junction is dissipated in the form of heat within the structure, and the emitted light is insignificant. For this reason, silicon and germanium are not used in the construction of LED devices. On the other hand: Diodes constructed of GaAs emit light in the infrared (invisible) zone during the recombination process at the p–n junction.

Even though the light is not visible, infrared LEDs have numerous applications where visible light is not a desirable effect. These include security systems, industrial processing, optical coupling, safety controls such as on garage door openers, and in home entertainment centers, where the infrared light of the remote control is the controlling element.
The basic construction of an LED appears in Figure with the standard symbol used for the device. The external metallic conducting surface connected to the p -type material is smaller to permit the emergence of the maximum number of photons of light energy when the device is forward-biased. Note in the figure that the recombination of the injected carriers due to the forward-biased junction results in emitted light at the site of the recombination. There will, of course, be some absorption of the packages of photon energy in the structure itself, but a very large percentage can leave, as shown in the figure.

Advantages of LED

The light-emitting diode (LED) is a solid-state light source. LEDs have replaced incandescent lamps in many applications because they have the following advantages:

(*i*) Low voltage

(*ii*) Longer life (more than 20 years)

(*iii*) Fast on-off switching

Applications of LEDs

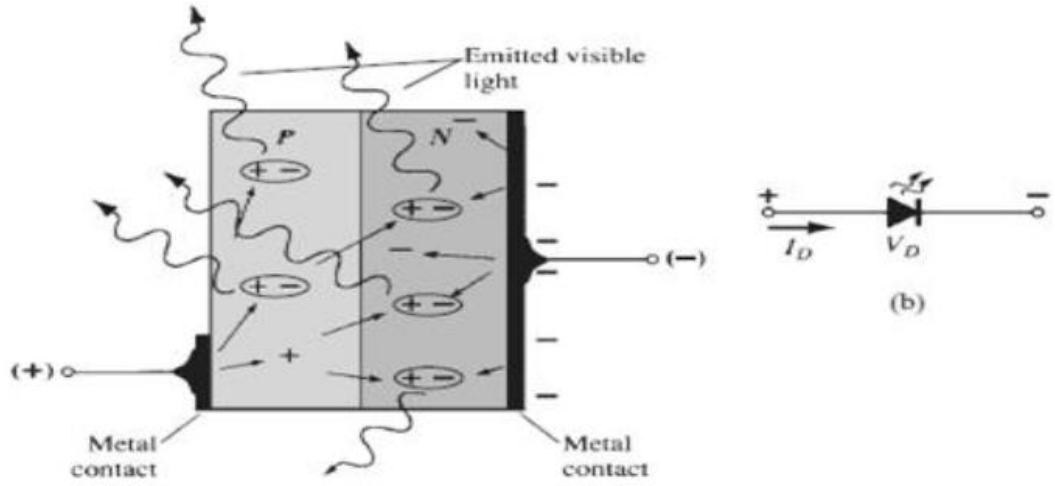


Figure 3.44: a) Process of electroluminescence in LED; b) Circuit symbol.

The LED is a low-power device. The power rating of a LED is of the order of milliwatts. This means that it is useful as an indicator but not good for illumination. Probably the two most common applications for visible LEDs are (*i*) as a power indicator (*ii*) seven-segment display.

(*i*) As a power indicator. A LED can be used to indicate whether the power is on or not. Figure 45 shows the simple use of the LED as a power indicator. When the switch *S* is closed, power is applied to the load. At the same time current also flows through the LED which lights, indicating power is on. The resistor R_S in series with the LED ensures that current rating of the LED is not exceeded.

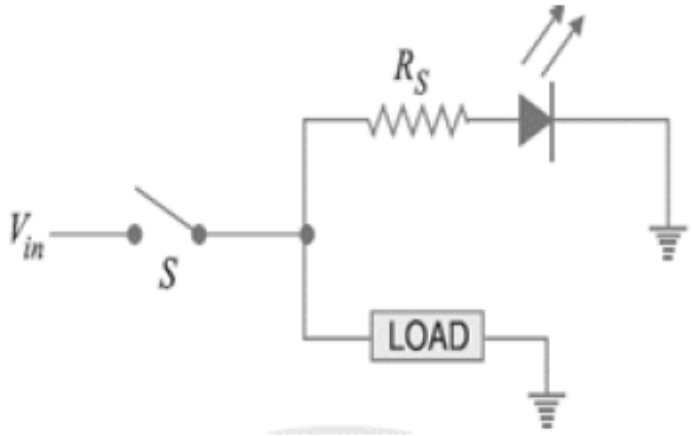


Figure 3.45: LED as power indicator.

***(ii)* Seven-segment display.** LEDs are often grouped to form seven-segment display. Figure (46 i) shows the front of a seven segment display. It contains seven LEDs (*A, B, C, D, E, F* and *G*) shaped in a figure of 8. Each LED is called a segment. If a particular LED is forward biased, that LED or segment will light and produces a bar of light. By forward biasing various combinations of seven LEDs, it is possible to display any number from 0 to 9. For example, if LEDs *A, B, C, D* and *G* are lit (by forward biasing them), the display will show
the number 3. Similarly, if LEDs *C, D, E, F, A* and *G* are lit, the display will show the number *6*. To get the number 0, all segments except *G* are lit. Figure (46 *ii*) shows the schematic diagram of seven-segment display. External series resistors are included to limit currents to

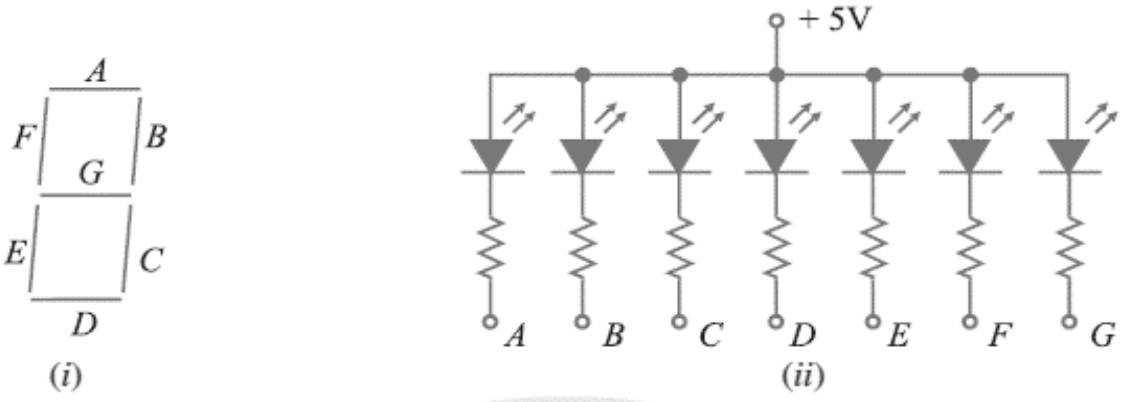


Figure 3.46: shows the front of a seven segment display.

safe levels. Note that the anodes of all seven LEDs are connected to a common positive voltage source of +5 V. This arrangement is known as *common-anode type*. In order to light a particular LED, say *A*, we ground the point *A* in Figure (46 *ii*). It forward biases the LED *A* which will be lit.

(iii) *Solar cell*

A solar cell is basically a p-n junction which generates emf when solar radiation falls on the p-n junction. It works on the same principle (photovoltaic effect) as the photodiode, except that no external bias is applied and the junction area is kept much larger for solar radiation to be incident because we are interested in more power. A simple p-n junction solar cell is shown in Fig. 38. A p-Si wafer of about 300 μm is taken over which a thin layer (~0.3 μm) of n-Si is grown on one-side by diffusion process. The other side of p-Si is coated with a metal (back contact). On the top of n-Si layer, metal finger electrode (or metallic grid) is deposited. This acts as a front contact. The metallic grid occupies only a very small fraction of the cell area (<15%) so that light can be incident on the cell from the top. The generation of emf by a solar cell, when light falls on, it is due to the following three basic processes: generation, separation and collection—

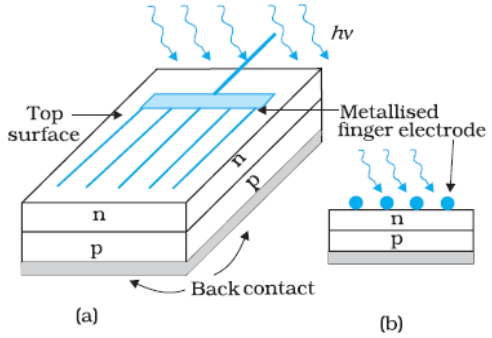


Figure 3.47 (a): Typical p-n junction solar cell; (b) Cross-sectional view.

(i) generation of e-h pairs due to light (with $h\nu > Eg$) close to the junction; **(ii) generation** of electrons and holes due to electric field of the depletion region. Electrons are swept to n-side and holes to p-side; (iii) the electrons reaching the n-side are collected by the front contact and holes reaching p-side are collected by the back contact. Thus p-side becomes positive and n-side becomes negative giving rise to *photo voltage*. When an external load is connected as shown in the Fig. 47 (a) a photocurrent *IL* flows through the load. A typical *I-V* characteristics of a solar cell is shown in the Fig. 48 (b). Note that the $I - V$ characteristics of solar cell is drawn in the fourth quadrant of the coordinate axes.

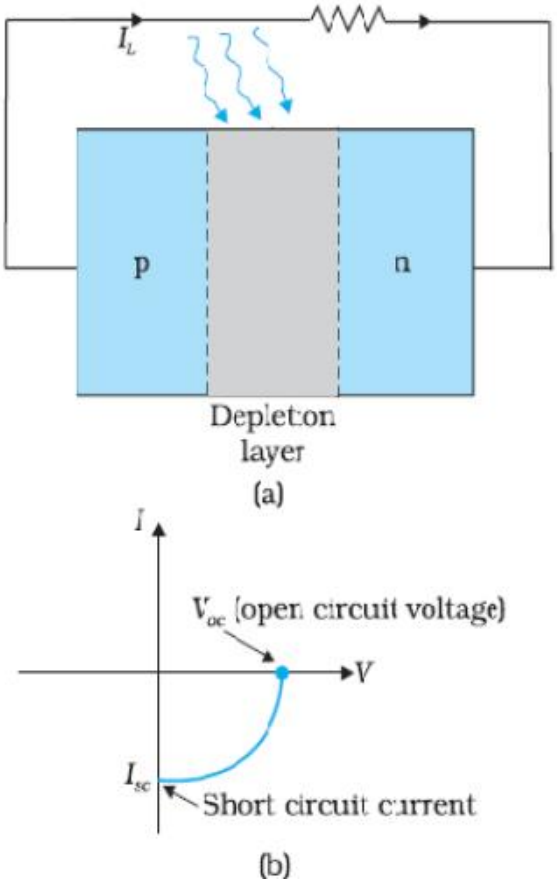


Figure 3.48 (a): A typical illuminated p-n junction solar cell; (b) *I-V* characteristics of a solar cell.

This is because a solar cell does not draw current but supplies the same to the load. Semiconductors with band gap close to 1.5 eV are ideal materials for solar cell fabrication. Solar cells are made with semiconductors like Si (*Eg* = 1.1 eV), GaAs (*Eg* = 1.43 eV), CdTe (*Eg* = 1.45 eV), CuInSe2 (*Eg* = 1.04 eV), etc. The important criteria for the selection of a material for solar cell fabrication are (i) band gap (~1.0 to 1.8 eV), (ii) high optical absorption (~104 cm^{-1}), (iii) electrical conductivity, (iv) availability of the raw material, and (v) cost. Note that sunlight is not always required for a solar cell. *Any* light with photon energies greater than the band gap will do. Solar cells are used to power electronic devices in satellites and space vehicles and also as power supply to some calculators. Production of low-cost photovoltaic cells for large-scale solar energy is a topic for research.

3.17 SUMMARY

1. Semiconductors are the basic materials used in the present solid state electronic devices like diode, transistor, ICs, etc.
2. Metals have low resistivity (10–2 to 10–8 Ωm), insulators have very high resistivity (>108 Ωm–1), while semiconductors have intermediate values of resistivity.
3. Semiconductors are elemental (Si, Ge) as well as compound (GaAs,CdS, etc.).
4. Pure semiconductors are called 'intrinsic semiconductors'. The presence of charge carriers (electrons and holes) is an 'intrinsic' property of the material and these are obtained as a result of thermal excitation. The number of electrons (n_e) is equal to the number of holes (n_h) in intrinsic conductors. Holes are essentially electron vacancies with an effective positive charge.
5. The number of charge carriers can be changed by 'doping' of a suitable impurity in pure semiconductors. Such semiconductors are known as extrinsic semiconductors. These are of two types (n-type and p-type).
6. In n-type semiconductors, $n_e >> n_h$ while in p-type semiconductors $n_h >> n_e$.
7. n-type semiconducting Si or Ge is obtained by doping with pentavalent atoms (donors) like As, Sb, P, etc., while p-type Si or Ge can be obtained by doping with trivalent atom (acceptors) like B, Al, In etc.
8. There are two distinct band of energies (called valence band and conduction band) in which the electrons in a material lie. Valence band energies are low as compared to conduction band energies. All energy levels in the valence band are filled while energy levels in the conduction band may be fully empty or partially filled. The electrons in the conduction band are free to move in a solid and are responsible for the conductivity. The extent of conductivity depends upon the energy gap (E_g) between the top of valence band (E_V) and the bottom of the conduction band E_C. The electrons from valence band can be excited by heat, light or electrical energy to the conduction band and thus, produce a change in the current flowing in a semiconductor.
9. For insulators $Eg > 3$ eV, for semiconductors Eg is 0.2 eV to 3 eV, while for metals $Eg \approx 0$.
10. p-n junction is the 'key' to all semiconductor devices. When such a junction is made, a 'depletion layer' is formed consisting of immobile ion-cores devoid of their electrons or holes. This is responsible for a junction potential barrier.
11. By changing the external applied voltage, junction barriers can be changed. In forward bias (n-side is connected to negative terminal of the battery and p-side is connected to the positive), the barrier is decreased while the barrier increases in reverse bias. Hence, forward bias current is more (mA) while it is very small (μA) in a p-n junction diode.
12. There are some special purpose diodes.
13. Zener diode is one such special purpose diode. In reverse bias, after a certain voltage, the current suddenly increases (breakdown voltage) in a Zener diode. This property has been used to obtain *voltage regulation*.
14. p-n junctions have also been used to obtain many photonic or optoelectronic devices where one of the participating entity is 'photon':(a) Photodiodes in which photon excitation results in a change of reverse saturation current which helps us to measure light intensity; (b) Solar cells which convert photon energy into electricity; (c) Light Emitting Diode and Diode Laser in which electron excitation by a bias voltage results in the generation of light.

3.18 GLOSSARY

1. **Solid** are **classified** in three categories metals, semiconductors and non-metals depending upon their conductivity.
2. **Semiconductors** have their conductivity in between those of metals and non-metals. Unlike metals they have a negative temperature coefficient of resistivity. This mean with the rise of temperature there is a decrease in their resistance and their conductivity increases.
3. **Intrinsic Semiconductors**: Semiconductors in their purest form possible are known as intrinsic semiconductors. Even intrinsic semiconductors have free electrons and vacancy (i.e. hole) because some of the electrons break their covalent bonds. This is because of the thermal energy acquired by these electrons. For intrinsic semiconductor density of holes is equal to the density of the electrons ($\boldsymbol{n_h = n_e}$).
4. **Holes**: When an electron gets free from a covalent bond it creates a vacancy in the bond. This vacancy (center of electron deficiency) acts as a positive center know as hole.
5. Intrinsic semiconductors have very **low conductivity** as the number density of electrons and holes in pure semiconductors is very low.
6. **Extrinsic semiconductors**: Semiconductors in their pure form are not very useful because of their low conductivity. In order to increase their conductivity, impurities are added to them.
7. **Doping**: The deliberate addition of a desirable impurity is called *doping* and the impurity atoms are called *dopants*. Such a material is also called a *doped semiconductor/extrinsic semiconductor.*
8. **n – type semiconductor**: An *n* -type material is created by introducing impurity elements that have *five* valence electrons (*pentavalent*), such as *antimony*, *arsenic*, and *phosphorus.* The pentavalent dopant is donating one extra electron for conduction and hence is known as *donor* impurity. For n-type semiconductors, density of electrons is much higher than the density of holes i.e. $\boldsymbol{n_e >> n_h}$**.**
9. **p – type semiconductor**: The *p* -type material is formed by doping a pure germanium or silicon crystal with impurity atoms having *three* valence electrons. The elements most frequently used for this purpose are *boron*, *gallium*, and *indium.* The trivalent dopant is deficient of one electron and thus creates a vacancy. The resulting vacancy will readily *accept* a free electron: thus diffused trivalent impurities are called acceptor atoms. For p-type semiconductors, density of holes is much higher than the density of electrons i.e. $\boldsymbol{n_h >> n_e}$**.**
10. Extrinsic semiconductor like intrinsic semiconductors as a whole are neutral in nature. This is because the total number of protons and the electrons are equal.
11. In an n-type material the electron is called the majority carrier and the hole the minority carrier. On the other-hand in a p-type material the hole is the majority carrier and the electron is the minority carrier.
12. **Hall Effect:** When an electrical current passes through a sample placed in a magnetic field, a potential proportional to the current and to the magnetic field is developed across the material in a direction perpendicular to both the current and to the magnetic field. This effect is known as the Hall effect.

13. **Hall Coefficient**: It is given by the relation $R_H = \frac{1}{nq}$ where n is the number density of charge carriers and q is the charge on each carrier.

14. **p-n junction**: When a p-type semiconductor is suitably joined to n-type semiconductor, the contact surface is called p-n junction. Most semiconductor devices contain one or more *p-n* junctions. The *p-n* junction is of great importance because it is in effect, the *control element* for semiconductor devices.

15. **Depletion Layer**: When a p-n junction is formed this cause electron hole recombination near the junction. As a result of this around the junction both p and n side are left with only immobile ions and all the charge carriers are gone. These two layers of positive and negative charges form the *depletion region* (or *depletion layer*). The term depletion is due to the fact that near the junction, the region is depleted (*i.e.* emptied) of *charge carries* (free electrons and holes) due to diffusion across the junction.

16. **Barrier Potential**: There exists a potential difference across the depletion layer and is called barrier potential (V_0). The barrier potential of a *p-n* junction depends upon several factors including the type of semiconductor material, the amount of doping and temperature. The typical barrier potential is approximately: For silicon, V_0 = 0.7 V ; For germanium, V_0 = 0.3 V.

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

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

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

20. **Junction Breakdown**: The process by which a depletion region at the p-n junction is destroyed and allows a large reverse current is called depletion region breakdown. There are two main mechanisms of junction breakdown, depending on the dopant concentration levels. 1. Avalanche breakdown 2. Zener breakdown

21. **Avalanche breakdown**: Avalanche breakdown occurs in moderately and lightly doped p-n junctions with a wide depletion region. Electron hole pairs thermally generated in the depletion region are accelerated by the external reverse bias. Electrons are accelerated towards then side and holes towards the p side. These electrons can interact with other Si atoms and if they have sufficient energy can knock off electrons from these Si atoms. This process is called impact ionization and leads to production of a large number of electrons. This causes the rapid rise in current.

22. **Zener Breakdown**: With increase in doping concentration the breakdown mechanism, changes from Avalanche to a tunneling mechanism. This is called a Zener breakdown. This is because the depletion width decreases with dopant concentration. Also, the reverse bias causes an offset in the bands such that it is possible for carriers to tunnel across the narrow depletion region.

23. **Zener Diode**: It is a special purpose semiconductor diode, named after its inventor C. Zener. It is designed to operate under reverse bias in the breakdown region and used as a voltage regulator.

24. **Tunnel Diode**: The tunnel diode was first introduced by Leo Esaki in 1958. Its characteristics, are different from any other diode as it has a negative-resistance region. In this region, an increase in terminal voltage results in a reduction in diode current. It has a greatly reduced depletion region, of the order of magnitude of 10^{-6}cm, or typically about 1/100 the width of this region for a typical semiconductor diode. It is this thin depletion region, through which many carriers can "tunnel" rather than attempt to surmount, at low forward-bias potentials.

25. **Varactor Diode**: A junction diode which acts as a variable capacitor under changing reverse bias is known as a varactor diode**.** The capacitance of varactor diode is found as

$C_T = \varepsilon \frac{A}{W_d}$,

where C_T = Total capacitance of the junction, ε = permittivity of the semiconducting material A = Area of cross section of the junction, W_d = Width of the depletion layer

26. **Photo diode**: It is a reverse-biased silicon or germanium p-n junction in which reverse current increases when the junction is exposed to light. The reverse current in a photo-diode is directly proportional to the intensity of light falling on its p-n junction. This means that greater the intensity of light falling on the p-n junction of photo-diode, the greater will be the reverse current.

27. **LED**: The light-emitting diode is a diode that gives off visible or invisible (infrared) light when energized. In any forward-biased $p - n$ junction there is, within the structure and primarily close to the junction, a recombination of holes and electrons. This recombination requires that the energy possessed by the unbound free electrons be transferred to another state. In all semiconductor $p - n$ junctions some of this energy is given off in the form of heat and some in the form of photons.

3.19 REFERENCES

1. Boylestad, R.L., Nashelsky ,L., "Electronic Devices and Circuit Theory", 11th edition, *Pearson Publication.*
2. Theraja, B.L., Theraja, A.K. 2005, "A Textbook of Electrical Technology Vol. I", *S. Chand Publication*.
3. Mehta, V.K., Mehta, R, "Principles of Electronics", *S. Chand Publication.*
4. Eggleston, D.L 2011, "Basic Electronics for Scientists and Engineers", *Cambridge University Press.*
5. "Physics Part –II A textbook for class XII", 1st edition, NCERT.
6. P. Kumar, N. K. Lohani and L. P. Verma, Kumaun Physics, S. J. Publications, India, 2018.
7. A. Ghatak, Optics, McGraw-Hill Education India Pvt. Ltd., 2017.
8. S. L Gupta and V. Kumar, Practical Physics, Pragati Prakashan, 2014.

3.19 TERMINAL QUESTIONS

1. The number of silicon atoms per m^3 is 5×10^{28}. This is doped simultaneously with 5×10^{22} atoms per m3 of Arsenic and 5×10^{20} per m^3 atoms of Indium. Calculate the number of electrons and holes. Given that $n_i = 1.5 \times 1016\ m^{-3}$. Is the material n-type or p-type?
2. In an intrinsic semiconductor the energy gap E_g is 1.2eV. Its hole mobility is much smaller than electron mobility and independent of temperature. What is the ratio between conductivity at 600K and that at 300K? Assume that the temperature dependence of intrinsic carrier concentration n_i is given by

$$n_i = n_0 \exp\left(-\frac{E_g}{2k_B T}\right)$$

where n_0 is a constant.

3. A p-n photodiode is fabricated from a semiconductor with band gap of 2.8 eV. Can it detect a wavelength of 6000 nm?
4. In a p-n junction diode, the current I can be expressed as $I = I_0 \exp\{(eV/2k_BT) - 1\}$ where I_0 is called the reverse saturation current, V is the voltage across the diode and is positive for forward bias and negative for reverse bias, and I is the current through the diode, k_B is the Boltzmann constant (8.6×10–5 eV/K) and T is the absolute temperature. If for a given diode $I0 = 5 \times 10–12$ A and T = 300 K, then

(a) What will be the forward current at a forward voltage of 0.6 V?
(b) What will be the increase in the current if the voltage across the diode is increased to 0.7 V?
(c) What is the dynamic resistance?
(d) What will be the current if reverse bias voltage changes from 1 V to 2 V?

UNIT 4 RECTIFIERS AND POWER SUPPLIES

Structure

4.1 INTRODUCTION

In this chapter, we will discuss about rectifier, filters and zenar diode as voltage regulator. Rectification is the process to convert alternating current (a.c.) into direct current (d.c.). A filter is a reactive network that blocks or attenuates unwanted signals and passes the desired signal. Zener diodes are widely used as voltage references and as shunt regulators to regulate the voltage across small circuits.

4.2 OBJECTIVES

- To learn about Rectifier
- To learn about Half-wave rectification and Full-wave rectification
- To discuss p-n junction diode with forward and reverse bias of operation
- To discuss Low Pass and High Pass Filters
- To discuss Zener diode, its operating conditions and application as voltage regulator

4.3 RECTIFIER

A rectifier is an electrical device which converts alternating current/signals (AC) to direct current/signals (DC). This process is known as rectification. Rectifiers have many uses including as components of power supplies and as amplitude modulation detectors of radio signals. Rectifiers are most commonly made using solid state diodes but other type of components can be used when very high voltages or currents are involved. When only a single diode is used to rectify AC (by blocking the negative or positive portion of the waveform), the difference between the term diode and the term rectifier is simply one of usage. The term rectifier describes a diode that is being used to convert AC to DC. Most rectifier circuits contain a number of diodes in a specific arrangement to more efficiently convert AC power to DC power.

4.3.1 Half-Wave Rectifier

In half wave rectification, either the positive or negative half of the AC wave is passed, while the other half is blocked. Because only one half of the input waveform reaches the output, it is only 50% efficient if used for power transfer. Half-wave rectification can be achieved with a single diode in a single phase supply as shown in figure 4.1, or with three diodes in a three-phase supply.

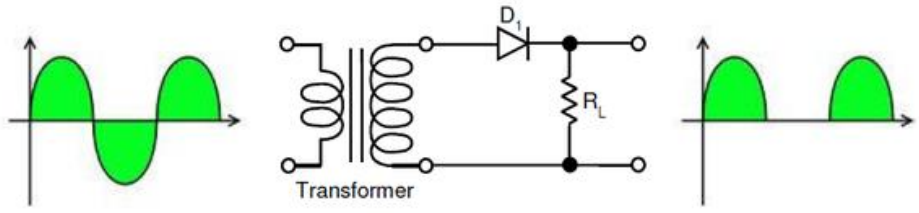


Figure 4.1: Half wave rectifier using one diode.

The output DC voltage of a half wave rectifier, given a sinusoidal input, can be calculated as:

$$V_{DC} = \frac{1}{\pi}\int_{0}^{\pi} V_{peak} \sin\theta d\theta$$

Or, $V_{DC} = \frac{V_{peak}}{\pi}$

And $V_{rms} = \frac{V_{peak}}{2}$, where V_{rms} is known as the root mean square value.

4.3.2 Full-Wave Rectifier

A full-wave rectifier converts both the positive and negative halves of the input waveform to a single polarity (positive or negative) at its output. By using both halves of the AC waveform, full-wave rectification is more efficient than half wave.

When a simple transformer without a center tapped secondary is used, four diodes are required instead of the one needed for half-wave rectification. Four diodes arranged this way are called a diode bridge or bridge rectifier as shown in figure 4.2. The bridge rectifier can also be used for translating a DC input of unknown or arbitrary polarity into an output of known polarity. This is generally required in electronic telephones or other telephony devices where the DC polarity on the two phone wires is unknown. There are also applications for protecting against accidental battery reversal in battery-powered circuits.

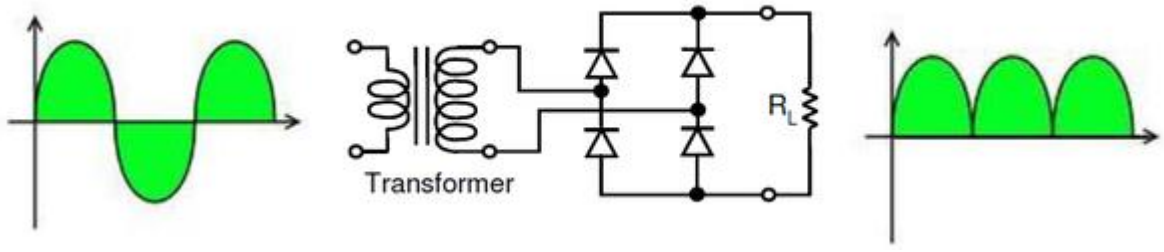


Figure 4.2: Bridge rectifier: a full-wave rectifier using four diodes.

For single-phase AC, if the transformer is center-tapped, then two diodes back-to-back (i.e. anode-to-anode or cathode-to-cathode) can form a full-wave rectifier. Twice as many windings are required on the transformer secondary to obtain the same output voltage compared to the bridge

rectifier above. This is not as efficient from the transformer perspective because current flows in only one half of the secondary during each positive and negative half cycle of the AC input.

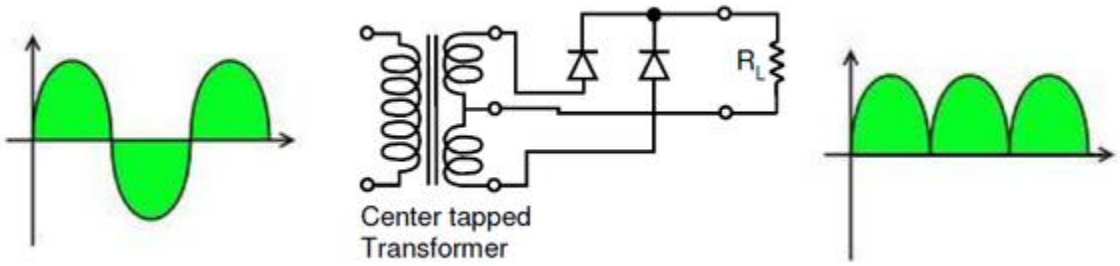


Figure 4.3: Full-wave rectifier using a center tapped transformer and two diodes.

If a second pair of diodes is included as in figure 4.3 then both positive and negative polarity voltages with respect to the transformer center tap can be generated. One can also view this arrangement to be the same as adding a center tap to the secondary winding in the full-wave bridge rectifier from figure 4.2.

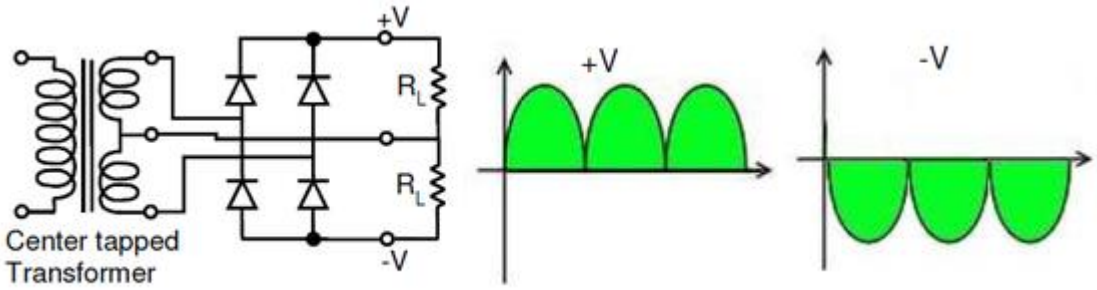


Figure 4.4: Dual polarity full-wave rectifier using a center tapped transformer and four diodes.

4.3.3 Rectifier Output Smoothing

Half-wave or full-wave rectification does not produce a constant-voltage DC as we have seen in the previous figures. In order to produce a steady DC voltage from a rectified AC source, a filter or smoothing circuit is needed. It can be done using a capacitor placed across the DC output of the rectifier. There will still remain an amount of AC ripple voltage where the voltage is not completely smoothed. The amplitude of the remaining ripple depends on how much the load discharges the capacitor between the peaks of the waveform.

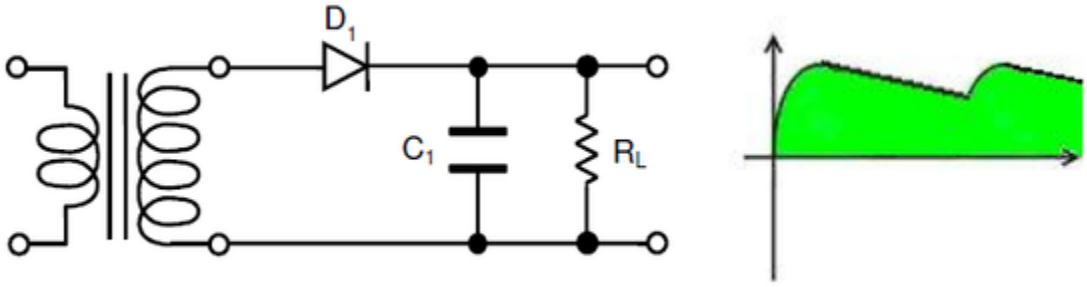


Figure 4.5(a): Half-wave Rectifier RC-Filter.

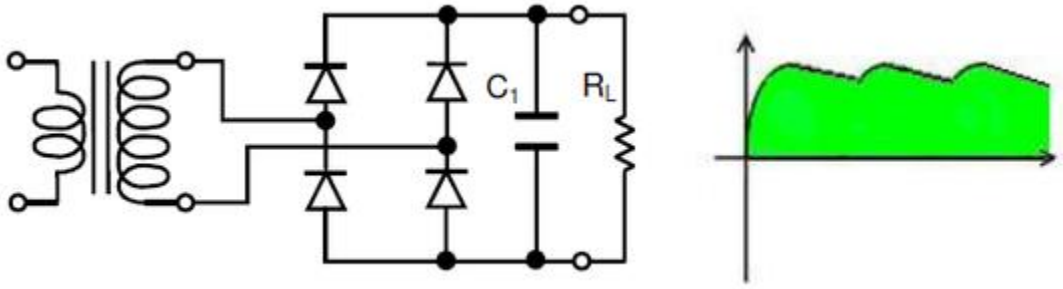


Figure 4.5(b): Full-wave rectifier RC-Filter.

Sizing of the filter capacitor, C1, represents a tradeoff. For a given load, RL, a larger capacitor will reduce ripple but will cost more and will create higher peak currents in the transformer secondary and in the supply feeding it. In extreme cases where many rectifiers are loaded onto a power distribution circuit, it may prove difficult for the power distribution grid to maintain a correctly shaped sinusoidal voltage waveform.

For a given tolerable ripple, the required capacitor size is proportional to the load current and inversely proportional to the supply frequency and the number of output peaks of the rectifier per input cycle. The load current and the supply frequency are generally outside the control of the designer of the rectifier system but the number of peaks per input cycle can be affected by the choice of rectifier design. The maximum ripple voltage present for a Full Wave Rectifier circuit is not only determined by the value of the smoothing capacitor but by the frequency and load current, and is calculated as:

$$V_{ripple} = \frac{I_{load}}{FC}$$

Where Vripple is the maximum ripple voltage on the DC output I_{load} is the DC load current

F is the frequency of the ripple (generally 2X the AC frequency)
C is the smoothing capacitor

A half-wave rectifier, figure 4.5(a) will only give one peak per cycle and for this and other reasons is only used in very small power supplies and where cost and complexity are of concern. A full wave rectifier, figure 4.5(b) achieves two peaks per cycle and this is the best that can be done with single-phase input. For three-phase inputs a three-phase bridge will give six peaks per cycle and even higher numbers of peaks can be achieved by using transformer networks placed before the rectifier to convert to a higher phase order.

To further reduce this ripple, a LC π-filter (pi-filter) such as shown in figure 4.6 can be used. This complements the reservoir capacitor, C_1 with a series inductor, L_1 and a second filter capacitor, C_2 so that a steadier DC output can be obtained across the terminals of the final filter capacitor. The series inductor presents a high impedance at the ripple current frequency.

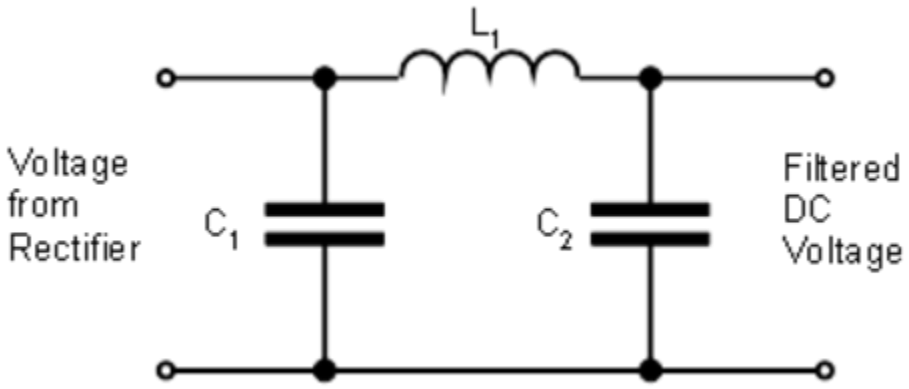


Figure 4.6: LC π-filter (pi-filter).

A more usual alternative to a filter, and essential if the DC load requires a very smooth supply voltage, is to follow the filter capacitor with a voltage regulator which we will discuss in section 4.6. The filter capacitor needs to be large enough to prevent the troughs of the ripple getting below the drop-out voltage of the regulator being used. The regulator serves both to remove the last of the ripple and to deal with variations in supply and load characteristics. It would be possible to use a smaller filter capacitor (which can be large for high-current power supplies) and then apply some filtering as well as the regulator, but this is not a common design strategy. The extreme of this approach is to dispense with the filter capacitor altogether and put the rectified waveform straight into an inductor input filter. The advantage of this circuit is that the current waveform is smoother and consequently the rectifier no longer has to deal with the current as a large current pulse just at the peaks of the input sine wave, but instead the current delivery is spread over more of the cycle.

The downside is that the voltage output is much lower - approximately the average of an AC half-cycle rather than the peak.

A half-wave rectifier is a circuit that allows only one half-cycle of the AC voltage waveform to be applied to the load, resulting in one non-alternating polarity across it. The resulting DC delivered to the load "pulsates" significantly.

A full-wave rectifier is a circuit that converts both half-cycles of the AC voltage waveform to an unbroken series of voltage pulses of the same polarity. The resulting DC delivered to the load doesn't "pulsate" as much.

Capacitors are used to smooth out or filter the ripple present in the rectified DC, and sometimes more complex filters using inductors as well as capacitors are used.

4.4 LOW PASS AND HIGH PASS FILTERS

A filter is a reactive network that blocks or attenuates unwanted signals and passes the desired signal. Filters are classified according to their frequency response. A low pass filter allows to pass the signals of low frequency and block or attenuates the signals of all frequencies higher that a selected cut-off frequency f_c. Similarly, a high pass filter permits the signals of frequencies higher than selected cut-off frequency blocking the signals of low frequencies. Other types of filters are band pass and band reject filters. In this chapter, we will study the high pass and low pass filters only.

Before we move to study the different type of low pass and high pass filters, let us understand 'decibel (dB)'. The change in output, by applying filter to input signal, is measured in decibel.

Decibel: A decibel is defined as ten times the common logarithms of the ratio of the input power to the output power,

$$\text{Decibel D} = 10 \log_{10} \frac{P_1}{P_2},$$

As power is proportional to square of voltage and square of current, decibel in terms of voltage or current ratio can be given as

$$D = 20 \log_{10} \left|\frac{V_1}{V_2}\right| = 20 \log_{10} \left|\frac{I_1}{I_2}\right|$$

Insertion Loss: When filter is connected to the circuit, there is a loss of gain which is called Insertion loss. It happens because the filter resistance effectively reduce the load resistance presented to signal source. (You will better understand it when you solve the exercise given in this chapter)

4.4.1 Low Pass Filter

The ideal response of a low-pass filter is given in Fig.4.7.

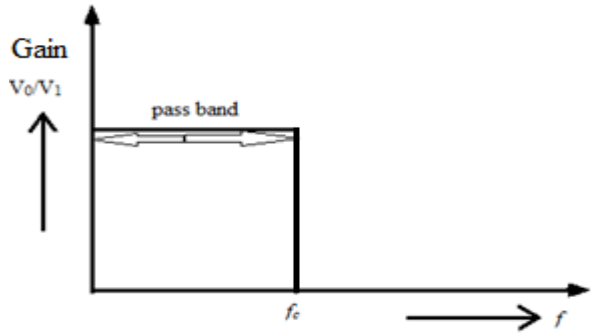


Figure 4.7: low-pass filter.

For ideal low pass filter, the gain V_0/V_1 is equal to 1 and the signals of frequency, higher than cut-off frequency f_c, are blocked completely. However, the response of practical filter circuit is slightly different from ideal one. The gain is slightly less than ideal value near the cut-off frequency f_c and falls rapidly for the frequency higher than cut-off. In low-pass filter section, we will study the RC low pass filters and RL low pass filters.

4.4.1.1 RC Low Pass Filter

The circuit of a basic low pass filter is given in Fig.4.8. Resistor 'R' and capacitor 'C' makes a voltage divider circuit.

The capacitive reactance of capacitor is $X_C = 1/ 2\pi f C$ (4.0)

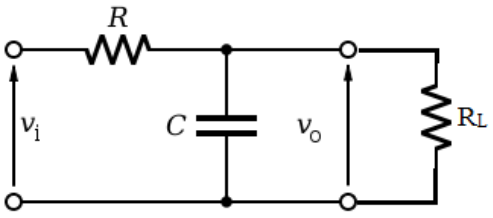


Figure 4.8: RC low-pass filter.

So the voltage across the load, $v_0 = \frac{X_c}{\sqrt{R^2+X_c^2}} \times v_i$ (4.1)

From equation (4.0), we can see that when frequency is low the value of X_C is much greater than R. In this case, the output voltage across load is nearly same as input voltage i.e. $v_0/v_i \approx 1$ [equation (4.1)]. In decibel form, v_0/v_i = 0 dB as $\log_{10} 1 = 0$. When frequency is high, resistance R becomes greater than X_C and output signals falls rapidly to zero value. It is obvious that cut-off frequency is the frequency when $R = X_C$. At $R = X_C$, from equation (4.1)

$$v_0 = \frac{R}{\sqrt{R^2 + R^2}} \times v_i; \frac{v_0}{v_i} = \frac{1}{\sqrt{2}} = 0.707$$

In decibel form,

$$\frac{v_0}{v_i} = 20log_{10}(0.707) \sim -3\ dB$$

Therefore, the cut-off frequency may be defined as the frequency at which the output voltage or gain falls by 3dB from its normal level.

When R = X_C; implies that R = 1/ 2π f_c C

Giving f_c = 1/ 2πRC (4.2)

Equation (4.2) is giving the formula of cut-off frequency for this low pass filter.

The gain/ frequency response graph for the RC low pass filter is given in Fig. 4.9. We can see that at cut-off frequency the gain falls to 3dB mark. For the frequencies higher than the cut-off frequency, the gain falls rapidly.

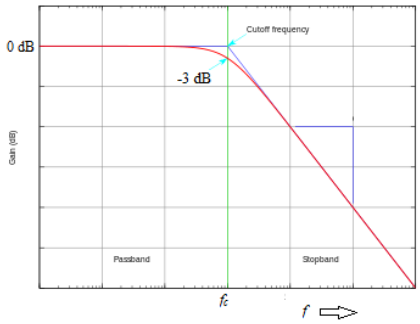


Figure 4.9

As it is clear from filter circuit [Fig. 4.9] that the output voltage is the voltage across capacitor, that is why there is the phase difference between the input and output voltages. The phase/frequency response is shown in Fig. 4.10.

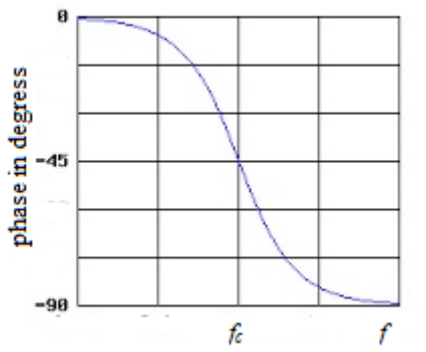


Figure 4.10

The phase angle (θ) between the input and filter output can be given by the formula

$$\theta = tan^{-1}\left[\frac{-R}{X_C}\right] \quad (4.3)$$

Initially as $X_C >> R$, the phase shift is almost zero. As the value of X_C decreases compare to R due to increase in frequency of input signal, the phase shift (magnitude) increases and reaches the maximum value of -90°. At cut-off frequency, when $X_C = R$, it is clear from equation (4.3) that the phase angle is θ = - 45°.

Combining the voltage gain value and input-output phase angle relationship [equations (2.1) & (4.3)], we have the low-pass filter transfer function

$$\frac{v_0}{v_i} = \frac{X_c}{\sqrt{R^2+X_c^2}} \angle tan^{-1}\left[\frac{-R}{X_C}\right] \quad (4.4)$$

First term on right hand side specify the voltage gain magnitude and second term specify the phase difference between input and output. Substituting $R = 1/2\pi f_c C$ and $X_C = 1/2\pi f C$ in equation (4.4),

$$\frac{v_0}{v_i} = \frac{1}{\sqrt{(f/f_c)^2+1}} \angle tan^{-1}\left[\frac{-f}{f_C}\right] \quad (4.5)$$

In decibel form

$$\frac{v_0}{v_i} = 20\ log\left(\frac{1}{\sqrt{(f/f_c)^2+1}}\right) \angle tan^{-1}\left[\frac{-f}{f_C}\right] \quad (4.6)$$

Example 1: Find the cut-off frequency for the filter circuit shown in Fig.4.11. Also find the phase shift at f_c.

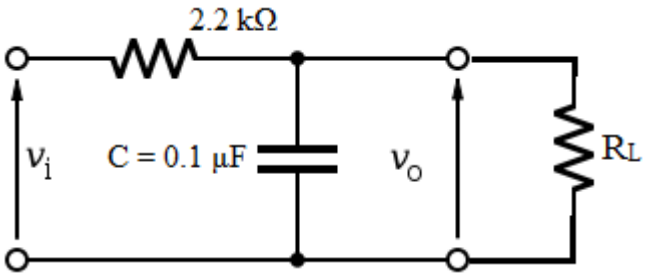


Figure 4.11

Solution:

From equation (2.12), the cut-off frequency

$f_c = 1/2\pi RC = 1/2\times3.14\times2.2\ k\Omega\times0.1\ \mu F = 720\ Hz$

At f_c, X_C = R, so phase shift at f_c [equation (2.13)]

$$\theta = tan^{-1}[-1] = -45^o$$

Example 2: Find the voltage gain (in dB) and phase angle for the low-pass filter circuit of Fig. 4.12 at 500 Hz frequency.

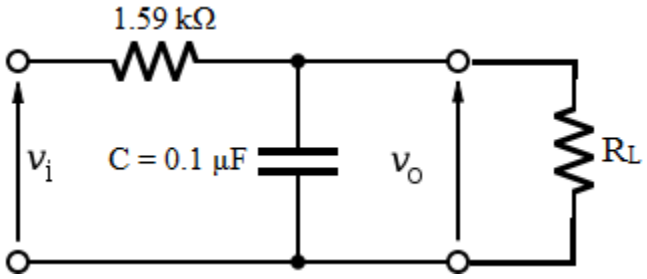


Figure 4.12

Solution:

f_c = 1/ 2πRC = 1/ 2×3.14×1.59 kΩ×0.1 μF = 1kHz; so f/f_c =0.5

using equation (2.16)

$$\frac{v_0}{v_i} = 20\ log\left(\frac{1}{\sqrt{(0.5)^2+1}}\right)\angle tan^{-1}[-0.5] = -1dB\angle -26.6^o$$

4.4.1.2 RL Low Pass Filter

The working principle and analysis of RL low-pass filter circuit is similar to RC low-pass filter. The resistor and inductor make a voltage divider circuit and output voltage is obtained across the resistor (Fig. 2.18).

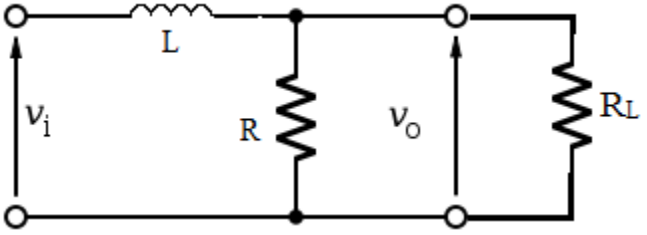


Figure 4.13

The inductive reactance of inductor is $X_L = 2\pi f L$ (4.7)

It is obvious from equation (4.7) that inductor offers a low resistance to low frequency signals and high resistance to high frequency signals. The result is that low frequency signals pass and high frequency signals attenuate.

the voltage across the load, $v_0 = \frac{R}{\sqrt{R^2+X_L^2}} \times v_i$ (4.8)

From equation (4.7), we can see that when frequency is low the value of $X_L << R$. In this case, we get the maximum voltage gain, nearly equal to 1. When frequency is high, X_L becomes greater than R and output signals falls rapidly to zero value. It is obvious that cut-off frequency is the frequency when $R = X_L$.

When $R = X_L$; implies that $R = 2\pi f_c L$

Giving $f_c = \frac{1}{2\pi\left(\frac{L}{R}\right)}$ (4.9)

This is the formula of cut-off frequency for RL low pass filter.

The phase angle (θ) between the input and filter output can be given by the formula

$\theta = tan^{-1}\left[\frac{-X_L}{R}\right]$ (4.10)

Combining the voltage gain value and input-output phase angle relationship [equations (4.8) & (4.10)], we have the low-pass filter transfer function for RL filter

$\frac{v_0}{v_i} = \frac{R}{\sqrt{R^2+X_L^2}} \angle tan^{-1}\left[\frac{-X_L}{R}\right]$ (4.11)

Substituting $R = 2\pi f_c L$ and $X_L = 2\pi f L$ in equation (4.11),

$\frac{v_0}{v_i} = \frac{1}{\sqrt{(f/f_c)^2+1}} \angle tan^{-1}\left[\frac{-f}{f_C}\right]$ (4.12)

In decibel form

$\frac{v_0}{v_i} = 20\ log\left(\frac{1}{\sqrt{(f/f_c)^2+1}}\right) \angle tan^{-1}\left[\frac{-f}{f_C}\right]$ (4.13)

This equation is same as for RC low pass filter.

Example 3: A RL low pass filter is constructed with R = 8.2 kΩ and L = 50 mH. Find the cut-off frequency and determine signal attenuation and phase shift at 50 kHz.

Solution:

Using equation (4.9),

$$f_c = \frac{1}{2\pi\left(\frac{50\text{ mH}}{8.2\text{ k}\Omega}\right)} = 26\text{ kHz}$$

using equation (4.13),

$$\frac{v_0}{v_i} = 20\,log\left(\frac{1}{\sqrt{(50\,kHz/26\,kHz)^2+1}}\right)\angle tan^{-1}\left[\frac{-50\,kHz}{26\,kHz}\right]$$

$$= -6.7\,dB\angle -62.5^o$$

4.4.2 High Pass Filter

This particular class of filter allow the signals above cut-off frequency to pass by attenuating the signals of low frequencies. The ideal response of a high-pass filter is given in Fig. 4.14.

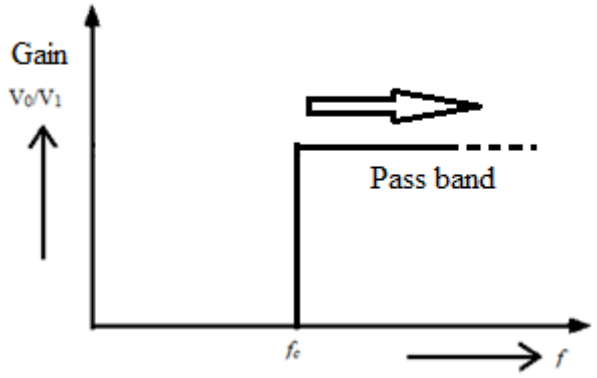


Figure 4.14

For ideal high pass filter, the gain V_0/V_1 is equal to 1 for signals of frequencies higher than cut-off frequency and the signals of frequency lower than cut-off frequency f_c, are blocked completely. Of course the response of practical filter circuit is slightly different from ideal one as we have learnt in previous section.

4.4.2.1 RC High-Pass Filter

You may easily understand that if we switch the position of R and C in low pass filter, we will obtain the circuit of high pass RC filter. The circuit of a basic RC high pass filter is given in Fig. 4.15

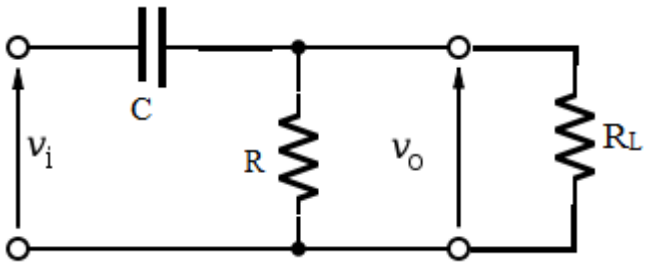


Figure 4.15

The voltage across the load, $v_0 = \frac{R}{\sqrt{R^2+X_c^2}} \times v_i$ (4.14)

From equation (4.14), we can see that when frequency is low the value of X_C is much greater than R. In this case, the output voltage across load is very less or zero. When frequency is high, resistance R becomes greater than X_C and output signals approaches toward the value of input signal. At $f = f_c$, when $R = X_C$

$$v_0 = \frac{R}{\sqrt{R^2+R^2}} \times v_i;\ \frac{v_0}{v_i} = \frac{1}{\sqrt{2}} = 0.707$$

In decibel form,

$$\frac{v_0}{v_i} = 20log_{10}(0.707) \sim -3\ dB$$

Therefore, the cut-off frequency, for high-pass filter, may be defined as the frequency at which the magnitude of output voltage or gain approaches to 3dB value. The cut-off frequency is same as for low-pass filter, and given by

$f_c = 1/ 2\pi RC$ (4.15)

The gain/ frequency and phase/frequency response is given a single Fig.4.16. We can see that at cut-off frequency the gain approaches to 3dB mark. For the frequencies higher than the cut-off frequency, the gain attains the maximum value.

The phase angle (θ) between the input and filter output can be given by the formula

$\theta = tan^{-1}\left[\frac{X_C}{R}\right]$ (4.16)

Initially as $X_C >> R$, the phase shift is almost 90^0. As the value of X_C decreases compare to R due to increase in frequency of input signal, the phase shift (magnitude) decreases and approaches to zero. At cut-off frequency, when $X_C = R$, it is clear from equation (4.16) that the phase angle is $\theta = -45^o$.

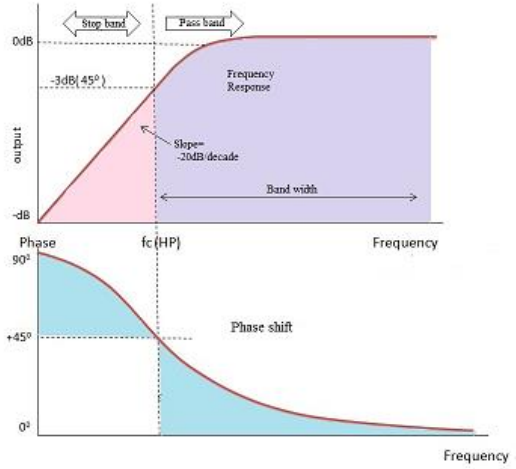


Figure 4.16

Combining the voltage gain value and input-output phase angle relationship [equations (4.14) & (4.16)], we have the transfer function

$$\frac{v_0}{v_i} = \frac{R}{\sqrt{R^2+X_c^2}} \angle tan^{-1}\left[\frac{X_C}{R}\right] \qquad (4.17)$$

Substituting R =1/ $2\pi f_c$C and X_C = 1/ 2π f C in equation (4.17),

$$\frac{v_0}{v_i} = \frac{1}{\sqrt{(f_c/f)^2+1}} \angle tan^{-1}\left[\frac{f_c}{f}\right] \qquad (4.18)$$

In decibel form

$$\frac{v_0}{v_i} = 20\ log\left(\frac{1}{\sqrt{(f_c/f)^2+1}}\right) \angle tan^{-1}\left[\frac{f_c}{f}\right] \qquad (4.19)$$

4.4.2.2 RL High-Pass Filter

Switching the position of R and L in RL low-pass filter provide you the circuit of RL high pass filter (Fig. 4.17).

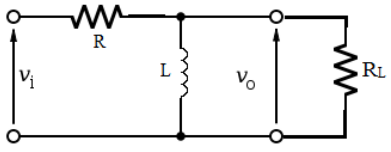


Figure 4.17

The inductor and resistor make the voltage divider circuit. The result is that high frequency signals pass and low frequency signals attenuate.

the voltage across the load, $v_0 = \frac{X_L}{\sqrt{R^2+X_L^2}} \times v_i$ (4.20)

From equation (4.20), we can see that when frequency is low the value of $X_L << R$. In this case, attenuation of signals occurs as frequency of signals decreases. When frequency is high, X_L becomes greater than R and output attain maximum value. The cut-off frequency is

$$f_c = \frac{1}{2\pi\left(\frac{L}{R}\right)} \quad (4.21)$$

The high-pass filter transfer function for RL filter

$$\frac{v_0}{v_i} = \frac{X_L}{\sqrt{R^2+X_L^2}} \angle tan^{-1}\left[\frac{R}{X_L}\right] \quad (4.22)$$

Substituting $R = 2\pi f_c L$ and $X_L = 2\pi f L$ in equation (4.22),

$$\frac{v_0}{v_i} = \frac{1}{\sqrt{(f_c/f)^2+1}} \angle tan^{-1}\left[\frac{f_c}{f}\right] \quad (4.23)$$

In decibel form

$$\frac{v_0}{v_i} = 20\ log\left(\frac{1}{\sqrt{(f_c/f)^2+1}}\right) \angle tan^{-1}\left[\frac{f_c}{f}\right] \quad (4.24)$$

This equation is same as for RC high pass filter.

4.5 ZENER DIODE AS VOLTAGE REGULATOR

Zener diodes are widely used as voltage references and as shunt regulators to regulate the voltage across small circuits. When connected in parallel with a varying voltage source, such as the diode rectifier we just discussed, so that it is reverse biased, the zener diode conducts when the voltage reaches the diode's reverse breakdown voltage. From that point on, the relatively low impedance of the diode keeps the voltage across the diode at that value.

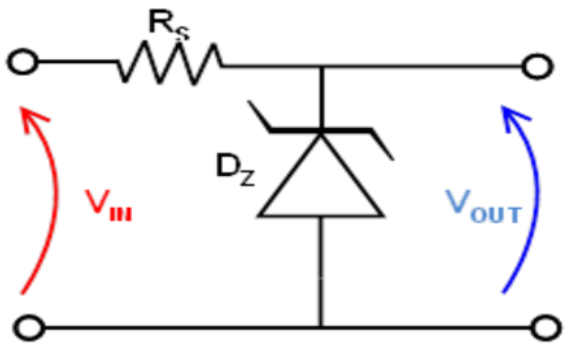


Figure 4.18: Zener diode voltage reference.

In the circuit shown in figure 4.18, a typical shunt regulator, an input voltage, V_{in} is regulated down to a stable output voltage V_{out}. The reverse bias breakdown voltage of diode D_z is stable over a wide current range and holds V_{out} relatively constant even though the input voltage may fluctuate over a fairly wide range. Because of the low impedance of the diode when operated like this, series resistor R_S is used to limit current through the circuit.

In the case of this simple reference, the current flowing in the diode is determined using Ohm's law and the known voltage drop across the resistor R_S.

$$I_{Diode} = \frac{\left(V_{IN} - V_{OUT}\right)}{R_S}$$

The value of R_S must satisfy two conditions:

R_S must be small enough that the current through D_Z keeps in reverse breakdown. If insufficient current exists through D_Z, then V_{out} will be unregulated, and less than the nominal breakdown voltage. When calculating R_S, allowance must be made for any current through any external load that might be connected to V_{out}, not shown in this diagram.

R_S must be large enough that the current through D_Z does not exceed the rated maximum and destroy the device. If the current through D_Z is I_D, its breakdown voltage V_B and its maximum power dissipation P_{MAX}, then:

$$I_D \times V_B < P_{MAX}$$

A load may be placed across the diode in this reference circuit, and as long as the zener stays in reverse breakdown, the diode will provide a stable voltage source to the load. Zener diodes in this configuration are often used as stable references for more complicated voltage regulator circuits involving buffer amplifier stages to supply large currents to the load.

Shunt regulators are simple, but the requirements that the ballast resistor, RS, be small enough to avoid excessive voltage drop during worst-case operation (low input voltage concurrent with high load current) tends to leave a lot of current flowing in the diode much of the time, making for a fairly inefficient regulator with high quiescent power dissipation, only suitable for smaller loads.

These devices are also encountered, typically in series with a base-emitter junction, in transistor stages where selective choice of a device centered around the avalanche or zener point can be used to introduce compensating temperature co-efficient balancing of the transistor PN junction. An example of this kind of use would be a DC error amplifier used in a regulated power supply circuit feedback loop system.

As a side note: zener diodes are also used in surge protectors to limit transient voltage spikes. Another notable application of the zener diode is the use of noise caused by its avalanche breakdown in a random number generator that never repeats.

Regulator Design Example:

An output voltage of 5V is required and the output current required is 60mA.

We first must choose a zener diode, $V_Z = 4.7V$ which is the nearest value available.

We need to determine the nominal input voltage and it must be a few volts greater than V_Z. For this example, we will use $V_{in} = 8V$.

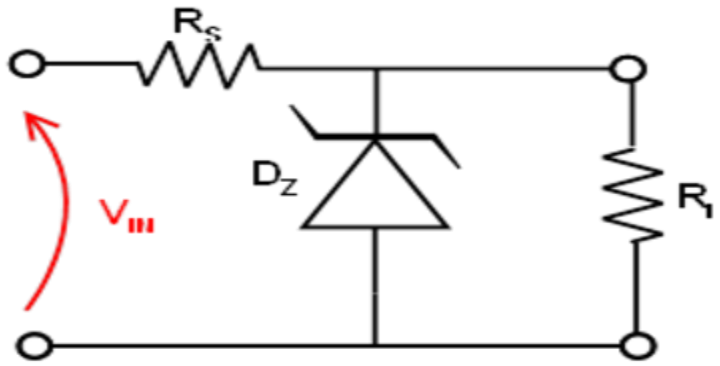
As a rule of thumb we choose the nominal current through the zener to be 10% of the required output load current or 6mA. This then determines the current Imax = 66mA which will flow through R_S (output current plus 10%).

The series resistor $R_S = (8V - 4.7V) / 66mA = 50\Omega$, we would choose $R_S = 47\Omega$ which is the nearest standard value.

The resistor power rating PRS > (8V - 4.7V) × 66mA = 218mW, so we choose PRS = 0.5W

The maximum power that could be dissipated in the zener when there is zero current in the output load can be calculated as $P_Z > 4.7V \times 66mA = 310mW$, so we would choose $P_Z = 400mW$.

Exercise 4: For the circuit shown, if the power supply voltage V_{in} increases, what will be the effect on voltage across the load resistor R_L ?



.

4.6 SUMMARY

In this chapter, we have learnt about half and full wave rectifier specially bridge and centre tap rectifier in details. It is explained that the ripples in full wave rectifier is quite small as compared to a half wave rectifier. Full wave rectifier circuits have better voltage regulation than half wave rectifier. We also studied Zener diode as voltage regulator. Low pass and high pass filters are enlightened in details.

4.7 TERMINAL QUESTIONS

1. What is the role of filter circuits?
2. Which one would you choose, out of L section filter and pi section filter?
3. The frequency of input ac signal in half wave rectifier is 50 Hz. What is the frequency of the chief ripple in output signal?
4. The frequency of ac signal in full wave rectifier is 50 hz. What is the frequency of the chief ripple output signal?

True and False Statements:

1. In a half wave rectifier, the rms ripple voltage exceeds the dc output voltage.
2. The efficiency of a half wave rectifier is about 40.5%.
3. The ripple factor of a half wave rectifier is much less than that of the full wave rectifier.
4. A full wave rectifier is more efficient than a half wave rectifier.
5. Ripple factor of full wave rectifier is more than 1.
6. Transformer is essential in full wave rectifier.
7. In a bridge rectifier, current flows through one diode only at any instant.
8. In a pi section filter, one inductor and two capacitors are used.

9. Capacitor filter is more efficient than choke input filter.
10. The filter circuits are used to minimise ac components from the rectified output.

Fill in the blanks:

1. ……. is a device that converts ac into dc.
2. a single pn junction diode is used as……rectifier.
3. The maximum efficiency of a full wave rectifier is…….
4. the bridge rectifier is not used for…. voltage applications.
5. in the full wave rectifier with shunt capacitance filter, the ripple factor……. With increasing load resistor.

4.8 ANSWERS

True and false Statements

1.T, 2.T, 3. F, 4. T, 5. F, 6. T, 7. F, 8. T, 9. F, 10. T

Fill in the blanks

1-Rectifier, 2-half wave, 3-81, 4-low, 5-decreases

UNIT 5 TRANSISTORS

Structure

5.1 INTRODUCTION

In previous chapter, we have discussed the p-n diode. We know that a p-n junction diode is a combination of an n-type and a p-type semiconductor, joined together with specific techniques. The bipolar junction transistor (BJT) is three-layer semiconductor device consisting of two n-type and one p-type or two p-type and one n-type material. When there exists a thin layer of p-type material between two n-type semiconductors, the transistor is called npn transistor. In the same way, when there exists a thin layer of n-type material between two p-type semiconductors, the transistor is called pnp transistor. A npn or pnp transistor can be operated in three configurations, i.e. common-base, common-emitter and common-collector. The word 'common' refers to the terminal which common to both input voltage and output voltage. We will study all three configurations of transistor in later sections. We will also discuss the current gain parameters α & β and the relation between them.

5.2 OBJECTIVES

- To learn about transistor, its construction and operation
- Discussing npn and pnp transistors
- To learn about various transistor configurations
- Discussing current gains of transistor and relations between them
- Briefly discussing the various transistor regions of operation

5.3 TRANSISTOR: CONSTRUCTION AND OPERATION

Transistor is a three-terminal device. The three terminals are called as emitter (E), base (B) and collector (C). E & C both may be n-type semiconductors and B p-type semiconductor making it npn transistor or E & C will be p-type semiconductors and B n-type semiconductor making it pnp transistor. The block diagrams of npn and pnp transistors are given in Fig. 5.1.

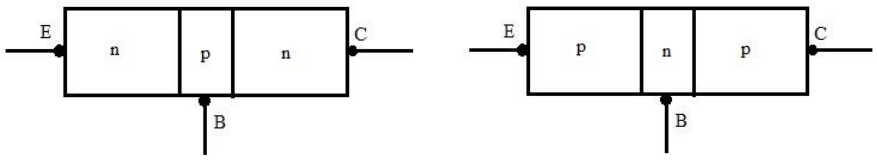


Figure 5.1: Block diagrams of npn and pnp transistors.

From Fig. 5.1, it is evident that transistor is a three terminal device. Three terminals are called emitter (indicated by E), base (indicated by B) and collector (indicated by C). The emitter is heavily doped, the base is lightly doped and the collector doping is slightly less than the emitter. The symbol of npn and pnp transistors is shown in Fig. 5.2.

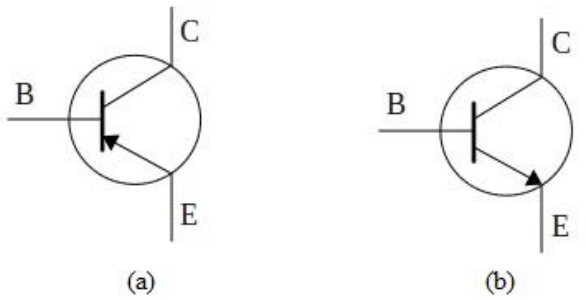


Figure 5.2: (a) Symbol of pnp transistor (b) Symbol of npn transistor.

In Fig. 5.2, the arrow sign shows the direction of flow of current. Transistors are mainly used for amplification of signals and for switching circuits. Based on power category, transistors are classified as low power and high power transistors. Low power transistors typically pass currents of 1 mA to 20 mA. However, the current levels for high power transistors range from 100 mA to several amps.

Now we will discuss the basic transistor operation of npn transistor. Consider the arrangement of Fig. 5.3, which shows a npn transistor connected with external voltage source for the normal operation. As transistors are made of using three extrinsic semiconductors (two n-type and one p-type for npn transistor), there exist two p-n junctions. One junction is called the emitter-base junction and another one is collector-base junction. For normal operation, the emitter-base junction is forward biased and collector-base junction is reverse biased. To make the emitter-base junction forward bias, a voltage source V_B is connected in such a way that positive and negative terminal of battery are connected to emitter and base, respectively.

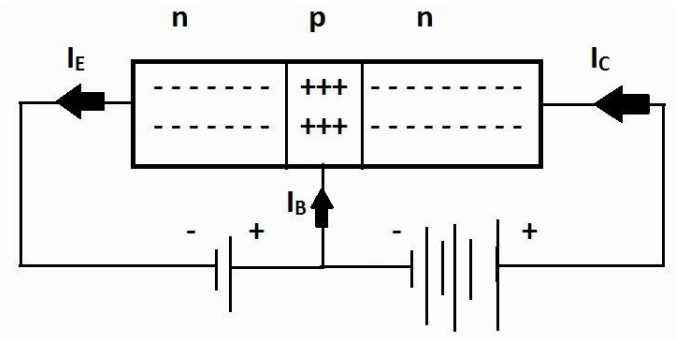


Figure 5.3: Normal active mode of operation for npn transistor.

Because of this forward biasing of emitter-base junction, the barrier potential is reduced and the electron from n-side flow towards the p-type base. The larger number of these electrons diffuse across the reverse-biased junction into the n-type material where the electrons are being collected

by applying the high positive voltage and we get the collector current I_C. Some of the electrons entering into base region from emitter region do not reach to collector. These electrons flow out via the base connection. However, a very small number of electrons take this path as the resistance of base terminal is high. We get a small base current I_B because of these electrons flowing out through the base. The value of base current is of the order of microamperes, although for the collector current this value is of order of milliamperes. The total current is the emitter current I_E, which is the sum of collector and base current, i.e.

$$I_E = I_C + I_B \qquad (5.1)$$

The electrons are the majority carriers and holes are minority carriers for npn transistor. For the pnp transistor, holes are the majority carriers. Generally, the collector current is the sum of current due to majority and minority carriers both. However, the current due to minority carriers is much smaller than the minority carriers and may be neglected under specific conditions.

5.4 TRANSISTOR: VOLTAGES AND CURRENTS

5.4.1 Terminal Voltages

Consider a npn transistor circuit of Fig. 5.4. The base and emitter junction is connected by a

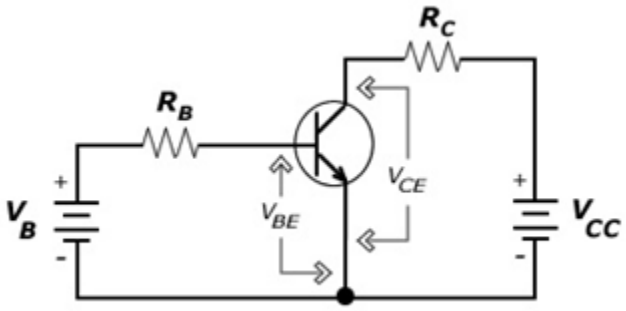


Figure 5.4: An npn transistor circuit.

battery V_B in forward biasing. The collector supply voltage V_{CC} is connected via collector resistance R_C. The value of V_{CC} is much larger than V_B to ensure that the collector-base junction is reverse biased. Base-emitter voltage is similar to diode forward bias voltage, i.e. V_{BE} is 0.7 V for silicon transistor and 0.3 V for germanium transistor. The transistor circuit in Fig. 5.4 is a common emitter circuit considering V_{CE} as output voltage. This type of circuit is used for signal amplification purpose. Normally, when base-emitter junction is forward biased and collector-base junction is reverse biased then the circuit is used for amplification purpose. We also can apply different biasing (forward or reverse) to both the junction but then the role of transistor circuit will be different.

5.4.2 Transistor currents and current gains (α and β)

As shown in Fig. 5.3, the emitter current is the sum of collector current and base current, i.e.

$$I_E = I_C + I_B$$

Practically, the base current is very small compared to the collector current and so the emitter current is nearly equal to the collector current. The relation between emitter and collector current is given by:

$$I_C = \alpha_{dc}\, I_E \tag{5.2}$$

Here, α_{dc} is the emitter-to-collector current gain. In equation (5.2), we have neglected the reverse current I_{CBO} due to reverse biasing of collector base junction. This current, I_{CBO}, is very small compared to collector and emitter current and so may be neglected. From equations (5.1) and (5.2),

$$I_C = \alpha_{dc}\,(I_B + I_C), \text{ which gives} \qquad I_C = (\alpha_{dc}\, I_B)/\,(1 - \alpha_{dc}) \tag{5.3}$$

There exists a relation between I_C and I_B, which is

$$I_C = \beta_{dc}\, I_B \tag{5.4}$$

β_{dc} is the base-to-collector current gain. It is the ratio of collector current to base current. As base current is much smaller than the collector current, β_{dc} typically ranges from 25 to 300.

Using equations (5.3) and (5.4), we can find the relation between α_{dc} and β_{dc},

$$\beta_{dc} = \alpha_{dc}\,/(1- \alpha_{dc}) \tag{5.5}$$

Another symbol of base-to-collector current gain is h_{fe}. The formulas of equations (5.1) to (5.5) are equally valid for pnp transistors also.

Example 5.1:

A transistor has a α_{dc} =0.98 and base current I_B = 100 μA. Find I_C, I_E and β_{dc}.

Solution:

We know by equation (5.4) and (5.5), $I_C = [\alpha_{dc}/(1- \alpha_{dc})]\, I_B$, so $I_c = (0.98\times 100\ \mu A)/\ 1\text{-}0.98$

I_C = 4.9 mA

$I_E = I_C + I_B$ = 4.9 mA + 0.1 mA = 5 mA

$\beta_{dc} = I_C/\ I_B$ = 4.9 mA/ 0.1 mA = 49

Another way to find β_{dc} is, $\beta_{dc} = \alpha_{dc}/(1- \alpha_{dc})$ = 49

Example 5.2:

What are the values of α and β of a transistor when $I_B = 40\ \mu A$ and $I_C = 3$ mA?

Solution:

$\beta = I_C/I_B = 3\ mA/ 40\ \mu A = 75$

From equation (4.5), $\alpha = \beta / (1+\beta) = 75/ 76 = 0.986$

Example 5.3:

Calculate α and I_C for the transistor if $\beta = 40$ and $I_B = 25\ \mu A$. Determine the new base current to give $I_C = 5$ mA.

Solution:

$I_C = \beta I_B = 40\times 25\mu A = 1$ mA

$\alpha = \beta /(1+\beta) = 40/41 = 0.976$

When $I_C = 5$ mA, $I_B = I_C/\beta = 5mA/ 40 = 125\ \mu A$

5.5 TRANSISTOR: CONFIGURATION AND CHARACTERISTICS

Transistor may be connected in three configurations, i.e. common base, common emitter and common collector configurations. The term ‘common’ refers to the terminal which is common to both input and output. For example, in common base configuration, the base is common to input and output. Now we will study all three configurations with their input and output characteristics.

5.5.1 Common-base Characteristics

In common base configuration, the base terminal is common to input and output. The characteristics are the graphs of currents and variable voltages. Input characteristic refers to the graph of input current and input variable voltage. In the same way, output characteristic refers to the graph of output current and output variable voltage. A simplified common base npn transistor configuration is shown in Fig. 5.5.

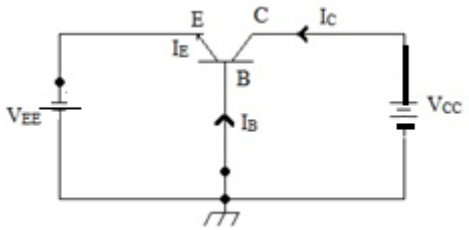


Figure. 5.5: npn common base configuration.

Fig. 5.5 is a simplified picture in which we have not shown the externally connected resistors. You may see from this figure that the base is common. In order to operate the transistor in active region, the emitter-base junction is forward biased and the collector-base junction is reverse biased.

Input characteristics:

The input terminal voltage is V_{EB} and input current is I_E. Input characteristics are drawn as a variation of I_E with V_{EB} keeping the output voltage (V_{CB}) constant. The graph of V_{EB} vs I_E for different constant values of V_{CB} is the input characteristics. Because the E-B junction is forward biased, so the input characteristics are same as the characteristics of a forward biased p-n junction. Input characteristics are shown in Fig. 5.6.

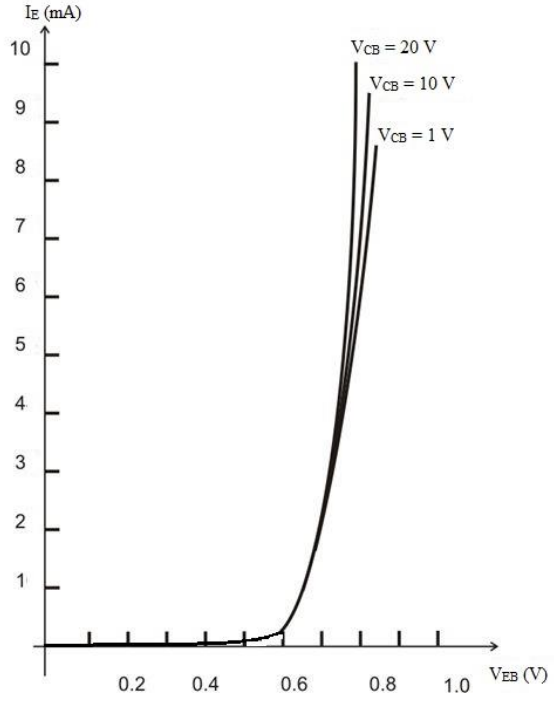


Figure 5.6: Common-base input characteristics.

The higher the values of V_{CB} result in the higher levels of I_E. This happens because of the reduction of depletion region width with increasing values of V_{CB}.

Output Characteristics:

Output characteristics are defined as the variation of V_{CB} with I_C for constant I_E. For the different fixed levels of I_E, the values of V_{CB} and current I_C are recorded and drawn (Fig. 5.7).

Three regions are mentioned in Fig. 5.7. As mentioned earlier also, the active region is normally employed for the amplification purpose. It is evident that the value of collector current I_C is almost same as emitter current I_E. We know already by current relations, as the base current is very small, the collector current and emitter current are nearly equal. Increasing V_{CB} will have a least effect on collector current.

A V_{CB} with opposite polarity is required to reduce the collector current to zero level. When V_{CB} is reduced to zero, a collector current I_C still flows. This is due to the barrier voltage existing at the CB junction even though the applied is zero. So, we have to apply forward biasing at CB junction to reduce the I_C to zero.

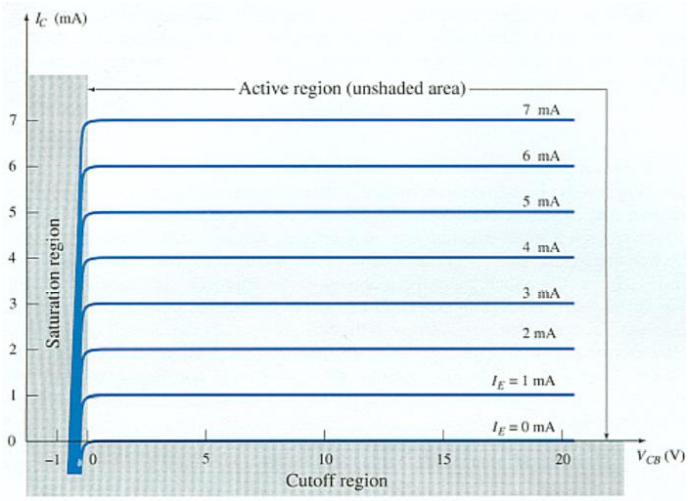


Figure 5.7: Output characteristics of common-base configuration.

When the reverse bias voltage at V_{CB} junction is exceeded to a high value, a breakdown may occur. During breakdown, the collector current rises sharply, and the transistor may destroy. It is advised to keep the voltage V_{CB} under the safe limit. Typical maximum V_{CB} levels range from 25 V to 80 V.

Current Gain Characteristics:

The current gain characteristics are the graph of output current I_C versus input current I_E. Graph of both currents are plotted keeping V_{CB} constant. By looking equation (5.2), we can expect the graph to be a straight line. Also, as we know, both currents are nearly equal, the graph will be a straight line showing similar values on both axes. Current gain characteristics graph is shown in Fig. 5.8.

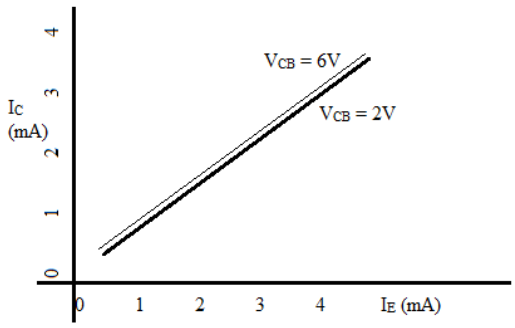


Figure 5.8: Common-base current gain characteristics.

From the figure above, we may observe that the V_{CB} has only a small effect on the current gain characteristics.

5.5.2 Common-Emitter Characteristics

The common-emitter transistor configuration, commonly used in amplifier applications, is shown in Fig. 5.9. The input is applied between base and emitter terminals, and the output is taken between collector and emitter terminal. Clearly, the emitter terminal is common to both input and output. The terminal input voltage is V_{BE} and the input current is base current I_B. Similarly, the output terminal voltage is V_{CE} and the output current is I_C. Based on these parameters, we will draw the input and output characteristics.

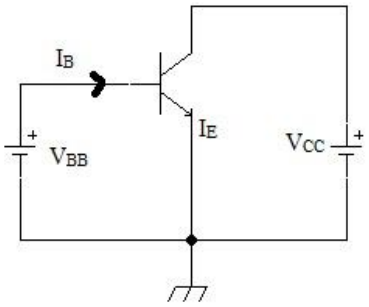


Figure 5.9: Common Emitter configuration.

Input Characteristics:

It is a graph of V_{BE} versus I_B, keeping V_{CE} constant. The graph is plotted for different constant V_{CE} levels. Input characteristics are shown in Fig. 5.10.

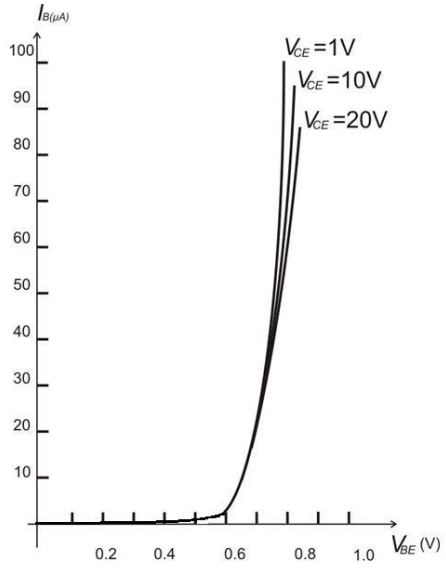


Figure 5.10: CE input characteristics.

Like we observed in common-base characteristics, the input characteristics of CE are the same as of forward bias p-n junction. We get an increase in current I_B when the V_{BE} value approach to knee voltage value.

For a given value of V_{BE}, I_B is decreased when the V_{CE} value is increased. When the V_{CE} value is increased then more charge carriers are collected at C terminal and the fewer charge carriers flow out via the base junction and so the base current is reduced.

Output Characteristics:

To draw the output characteristics, I_B is maintained constant at several levels and the values of collector current I_C are recorded for the increasing values of V_{CE}. The output characteristics curves are shown below (Fig. 5.11).

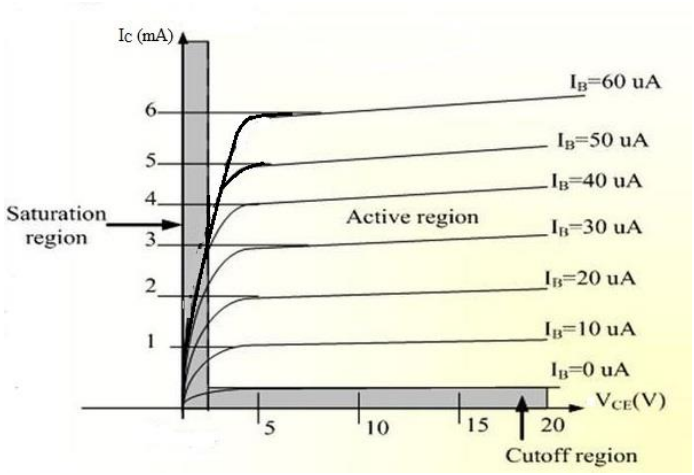


Figure 5.11: CE output characteristics.

For a constant I_B level, the collector current I_C slightly increases with the increase of collector-emitter voltage (V_{CE}). When V_{CE} is increased, more of the electrons from emitter flow across C-B junction and so the I_C level increases in small amount.

With the increased level of I_B, the I_C values increase many fold depending on the current gain β of transistor. For the active region, the graph of I_C is not a horizontal line as it was in case of C-B output characteristics curve. The reason is the effect of V_{CE} on the I_C as explained earlier.

In the output characteristics graph, for $I_B = 0$ level, a significant value of I_C is obtained. The reverse saturation current still exists at cut-off condition and this current increase many fold in case of C-E configuration. We may also prove that using current equations, in the following way:

If the reverse saturation current is I_{CBO}, then the relation between collector current and emitter current may be given as

$I_C = \alpha I_E + I_{CBO}$, replacing I_E with $I_C + I_B$ gives

$I_C = \alpha (I_C + I_B) + I_{CBO}$

Finally, $I_C = \frac{\alpha I_B}{1-\alpha} + \frac{I_{CBO}}{1-\alpha}$

If $I_B = 0$, then $I_C = \frac{I_{CBO}}{1-\alpha}$

For example, if I_{CBO} = 1 μA and α =0.996, then I_C = 0.25 mA, which is a significant value of collector current.

Current Gain Characteristics:

CE current gain characteristics are plotted between current I_C and current I_B for various fixed levels of V_{CE} (Fig. 5.12).

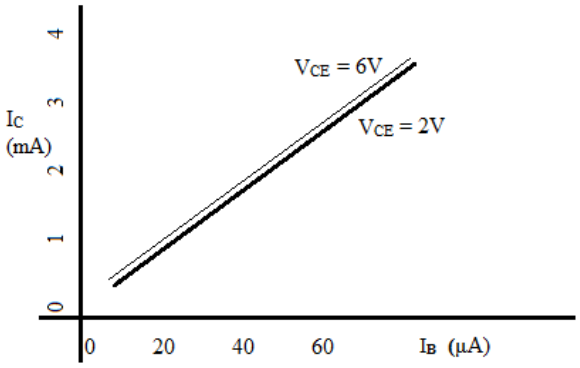


Figure 5.12: CE current gain characteristics.

5.5.3 Common-Collector Characteristics

For common collector configuration, the collector is common to both input voltage and output voltage, as shown in Fig. 5.13.

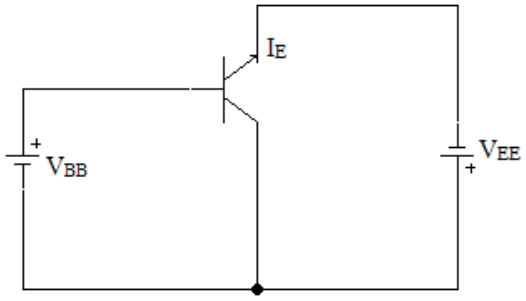


Figure 5.13: Common collector configuration.

The common collector configuration is mainly used for impedance matching. The circuit has a high input resistance and low output impedance. CC has the voltage gain of nearly equal to one so clearly this circuit cannot be used for signal amplification purpose. So, this circuit may be worked as coupler between the two stages of circuits or networks.

Input Characteristics:

The input characteristics for CC circuit are very different from CB or CE configuration. The input characteristics are plotted between the base-collector voltage V_{BC} and base current I_B.

The CC input characteristics are given in Fig. 5.14.

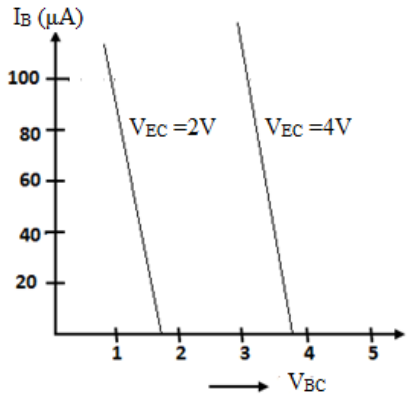


Figure 5.14: CC input characteristics.

It is evident from the Fig. 5.14 that V_{BC} is largely depending on the V_{EC} level. This can be understood as the V_{BC} may be written as

$V_{BC} = V_{EC} - V_{EB}$

Increasing the V_{CB} value with constant V_{EC} reduces the level of V_{EB} and so the value of base current I_B.

Output Characteristics:

The common collector output characteristics are I_E plotted versus V_{CE} for several fixed levels of I_B. The plot is very similar to the output characteristics of common emitter configuration (Fig. 5.15).

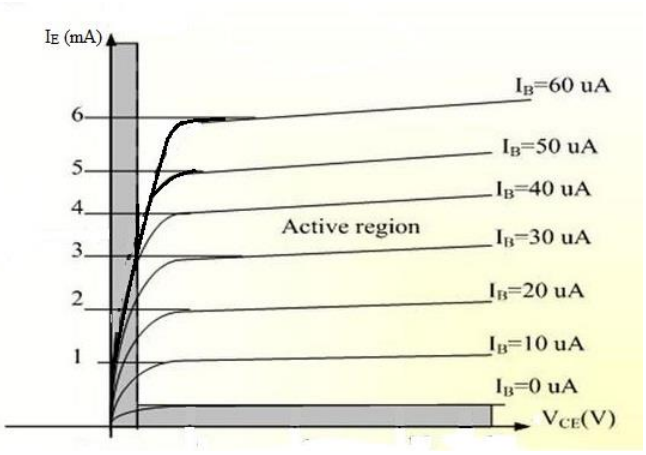


Figure 5.15: CC output characteristics.

Current Gain Characteristics:

The current gain characteristics are the graphs of I_E versus I_B for the several fixed value of V_{CE}. Characteristics are very similar to that of common-emitter characteristics. The current gain characteristics are shown in Fig. 5.16.

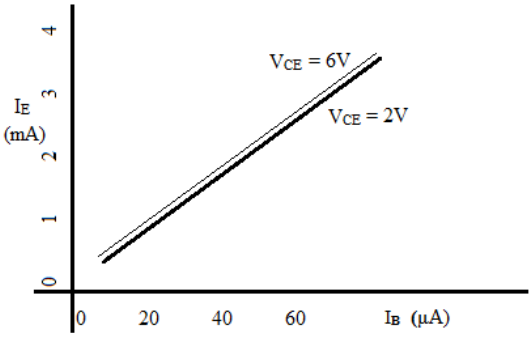


Figure 5.16: CC current gain characteristics.

5.5.4 Transistor Operating Regions

So far, we have discussed three transistor configurations. In the output characteristics, some operating regions are mentioned. Now we will discuss about theses operating regions in brief. Let us redraw the typical output characteristics, again in Fig. 5.17. The figure below belongs to the output characteristics of CE configuration.

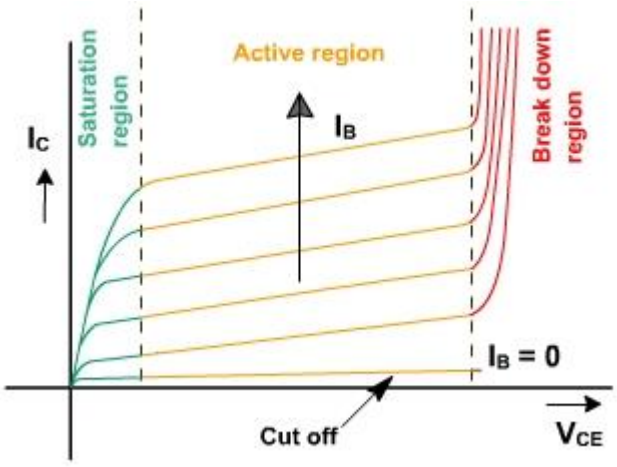


Figure 5.17: Output characteristics showing various operating regions.

The four regions, mentioned in the figure above, are Saturation region, active region, cut-off region and breakdown region.

Active Region:

When the transistor is used as an amplifier, it is operated in active region. In the active region, the base-emitter junction is forward biased while the collector-base junction is reversed biased. Working as amplifier in the active region, the transistor is analyzed using its equivalent model. These models are h-model, T-model, r_e model etc.

Cut-off Region:

A transistor is in the cut-off region when both the base-emitter and collector-base junctions are reverse-biased. Ideally, the transistor acts as open in cut-off region. However, practically a small reverse saturation current exists. At higher temperature, this reverse current becomes important as it increases rapidly with temperature.

Saturation Region:

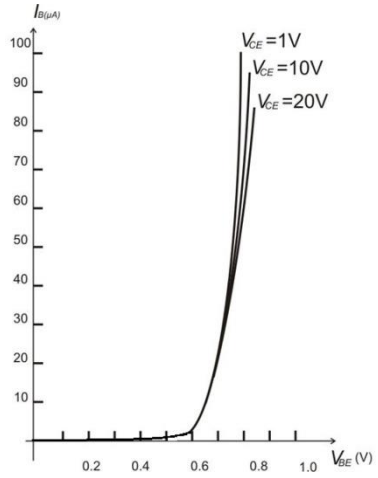
A transistor is working in the saturation region when both the base-emitter and collector-base junctions are forward-biased. Ideally, the transistor acts as short in the saturation region. It is important to note that the relation $I_C = \beta I_B$ does not hold for saturation region, where β is current gain of active region of operation. In fact, the β in the saturation region is much smaller than the β of active region of operation. The transistor is operated in saturation and cut-off region when it is to be work as a switch, mostly in digital electronics.

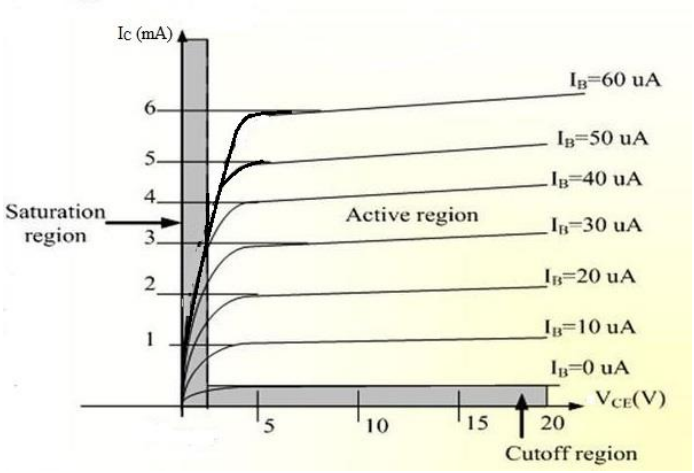
Breakdown Region:

When the voltage is high, the charge carriers are sufficiently energetic to do the ionization by collision when striking the other atoms. The electrons released in this way collide with other atoms to produce more free electrons. This process is called Avalanche breakdown. It may damage the device and this region of operation should be avoided.

Example 5.4:

For common emitter circuit, the input and output characteristics are given below. Determine I_B, I_C and β_{dc} when $V_{BE} = 0.7$ V and $V_{CE} = 6$V.





Solution:

From input characteristics, when V_{BE} = 0.7 V, the value of base current is $I_B \approx 20$ μA

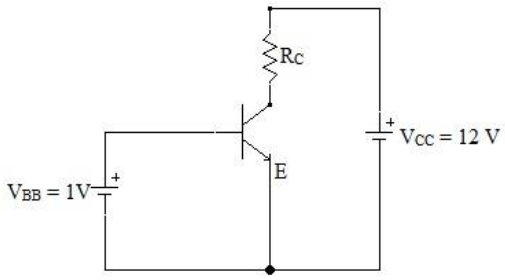
From output characteristics, when I_B = 20 μA & V_{CE} = 6 V,

Then $I_C \approx 2$ mA

$\beta_{dc} = I_C / I_B$ = 2 mA/ 20 μA = 100

Example 5.5:

For the CE circuit shown below, I_C = 12 mA, V_{CE} =4 V, β = 80 and V_{BE} = 0.7 V. Find I_B and collector resistance R_C.



Solution:

$I_B = I_C/\beta$ = 12 mA/80 = 0.15 mA

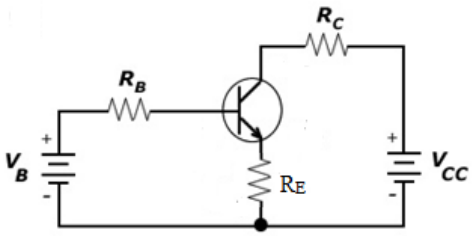
Applying the KVL to the RHS loop,

$V_{CC} - I_C R_C - V_{CE} = 0$,

$R_C = (V_{CC} – V_{CE})/ I_C = (12 -4)$ V/ 12 mA = 666 Ω

Example 5.6:

For the CE circuit given below, V_B =3V, V_{CC} =12V, I_C =12 mA, V_{CE} = 5.5 V, β =100, V_{BE} = 0.7 V and R_E = 50 Ω. Find I_B, R_B and R_C.



Solution:

$I_B = I_C/ \beta = 12/100 = 0.12$ mA

Applying KVL to base-emitter loop,

$R_B = V_B – V_{BE} / I_B = (3 – 0.7)/ 0.12 = 19.16$ kΩ

Applying KVL to collector-emitter circuit,

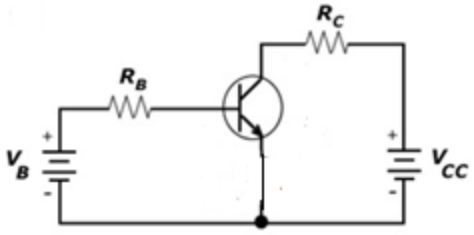
$V_{CC} – I_C R_C – V_{CE} - I_E R_E = 0$,

$V_{CC} – I_C R_C – V_{CE} – (\beta+1) I_B R_E = 0$, [since $I_E = I_C + I_B = \beta I_B + I_B = (\beta+1) I_B$]

$R_C = [V_{CC} – V_{CE} – (\beta+1) I_B R_E] / I_C = [12 – 5.5 - 101\times50\times0.00012]/ 12 = 491$ Ω

Example 5.7:

For following transistor CE circuit, V_{CC} = 10 V, R_B = 10 kΩ, R_C = 1 kΩ and β =100. The collector-emitter saturation voltage is $V_{CE,sat}$ = 0.1 V. Find the operating region when (a) V_B = 1.5 V and (b) V_B = 10.7 V.



Solution:

(a) When V_B = 1.5 V,

Applying KVL to base-emitter loop,

$I_B = (V_B - V_{BE})/ R_B = 0.08$ mA,

$I_C = \beta I_B = 8$ mA

Applying KVL to collector-emitter loop,

$V_{CE} = V_{CC} - I_C R_C = 2$ V

As $V_{CE} > V_{CE,sat}$, implies that the transistor is operating in active region.

(b) When $V_B = 10.7$ V,

$I_B = (V_B - V_{BE})/ R_B = 1$ mA

Then I_C would be, $I_C = \beta I_B = 100$ mA,

If $I_C = 100$ mA, then $V_{CE} = V_{CC} - I_C R_C = 10 - 100 = -90$ V

But V_{CE} cannot be greater than V_{CC} in magnitude; hence this is an impossible situation. We can conclude that transistor is working in saturation region and therefore $V_{CE} = V_{CE,sat} = 0.1$ V. Then collector current is

$I_C = (V_{CC} - V_{CE,sat})/R_C = (10 - 0.1)$ V/ 1 kΩ = 9.9 mA.

In saturation region, the current gain is $\beta = I_C / I_B = 9.9$ mA/ 1mA = 9.9

Clearly, the current gain β in saturation region is much smaller than the β in active region.

3.1 DIODE LOAD LINE ANALYSIS

A load line is used for the graphical analysis drawn on a current-voltage graph. This line represents the applied load and that is why called the load line. Consider a series diode circuit in forward biasing as shown in Fig. 5.18.

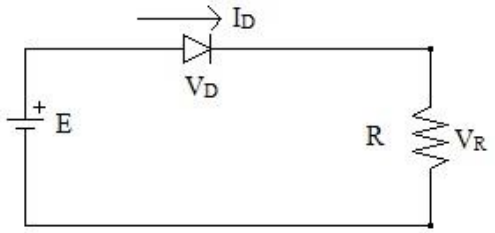


Figure 5.18: p-n diode forward bias series circuit.

Applying Kirchhoff 's voltage law (KVL) to above circuit,

$$E - V_D - I_D R = 0$$

$$E = V_D + I_D R \qquad (5.6)$$

To draw the load line on V-I graph, first put $V_D = 0$ in equation (5.6), to find a point on current axis (y-axis). So,

When $V_D = 0$, $I_D = E/R$, this I_D value is the first point on current axis.

Then set the value $I_D = 0$, which gives $V_D = E$, this V_D value is the second point on voltage axis (x-axis). Connect the two points by a straight line. This line is called the load line, as shown in Fig. 5.19.

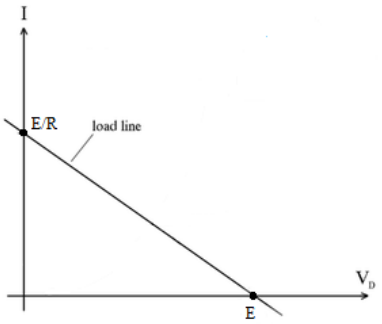


Figure 5.19: Load line

Q-point:

The intersection point between the load line and characteristics curve is called the Q-point or point of operation.

As we can see in Fig. 5.20, the load line and characteristics curve cut each other at a point, given by Q-point.

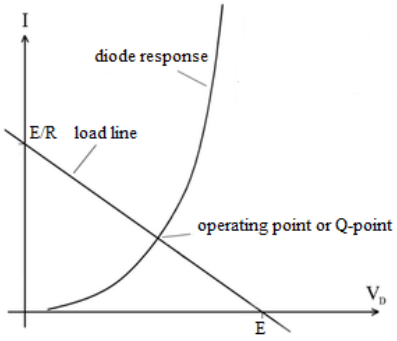


Figure 5.20: Load line and Q-point

The Q-point is important as it gives the values of operating voltage and current. The intersection of vertical line from Q-point to voltage axis gives you diode voltage. Similarly, the intersection of horizontal line from Q-point to current axis gives you the value of diode current.

5.7 TWO PORT NETWORKS

A network may be represented schematically by a rectangular box with input and output terminals. Two-input & two-output terminals make four terminal networks. In a four terminal network, a pair of terminals at which a signal enters or leaves is called port. In the Fig. 5.21, one pair of terminals or port is (1-1') and (2-2') is another port. So in this figure, two ports and four terminals (1, 1', 2, 2') are shown. For rest of our chapter, we will call port (1-1') as port-1 (or first port) and port (2-2') as port-2 (or second port). One pair of terminals or one port represents the input and another pair of terminals or another port represents the output. Transmission line, transformer and filter are the example of two port network.

Figure: 5.21

Voltage and current (variables) are assigned to each port. The voltage and current at the input terminal are V_1& I_1 and V_2& I_2 at the output terminal. Two of these four variables are independent and other two are independent. The choice of independent variables is arbitrary. The number of possible combinations generated by four variables, taken two at a time, is six. So, there are six possible sets of two equations.

5.7.1 Open Circuit Impedance Parameters

A general linear two-port network is shown in Fig. 5.22. In this network, I_1 and I_2 are independent variable while V_1 and V_2 are dependent variables.

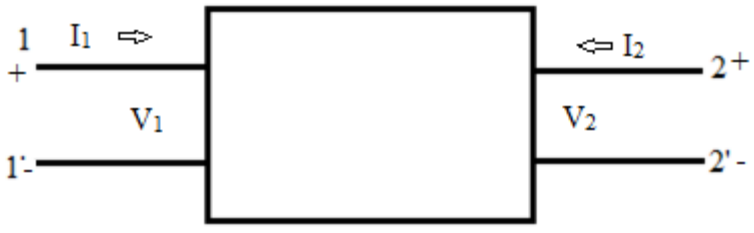


Figure: 5.22

The voltages V_1 and V_2 in terms of currents may be written as

$$V_1 = Z_{11} I_1 + Z_{12} I_2$$

$$V_2 = Z_{21} I_1 + Z_{22} I_2 \quad (5.7)$$

or in matrix form

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad (5.8)$$

Z_{11}, Z_{12}, Z_{21} and Z_{22} are called impedance (Z) parameters.

The individual Z parameters may be defined as follows:

Suppose port 2-2' is open circuited, then $I_2 = 0$

Thus (using equation 5.7), $Z_{11} = \frac{V_1}{I_1}$ with $I_2 = 0$

$$Z_{21} = \frac{V_2}{I_1} \text{ with } I_2 = 0$$

Z_{11} is driving-port impedance at port 1-1' with port 2-2' is open circuited. It is called open circuit input impedance. Z_{21} is the transfer impedance at port 1-1' with port 2-2' open circuited. It is called open circuit forward transfer impedance.

If port 1-1' is open circuited, then $I_1 = 0$

So,

$$Z_{12} = \frac{V_1}{I_2} \text{ with } I_1 = 0$$

$$Z_{22} = \frac{V_2}{I_2} \text{ with } I_1 = 0$$

Z_{12} is called open circuit reverse transfer impedance. Z_{22} is open circuit output impedance.

These Z parameters can be calculated for the two port network whose internal circuit is given. Since these equations involve open circuit at one of the ports, these parameters are known as open circuit parameters.

The equivalent circuit of two port network governed by open circuit impedance parameters is shown in Fig. 5.23.

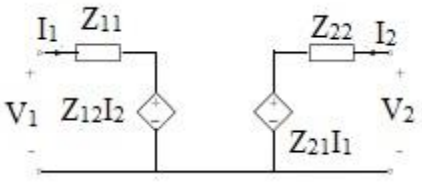


Figure: 5.23

The diamond shape symbol represents the dependent voltage sources. If the open circuit transfer impedances are equal, i.e. $Z_{12} = Z_{21}$, the two port network is called reciprocal network.

Example 5.8:

Find the Z-parameters for the circuit of Fig. 5.24.

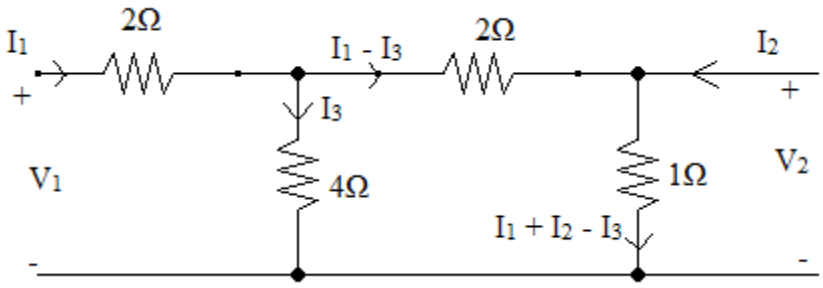


Figure: 5.24

Solution:

The Z-parameters can be determined using following procedure:

To find Z_{11}:

Let the current $I_2 = 0$, then impedance from port-1 end is Z_{11}.

So,

$$Z_{11} = [(2\Omega + 1\Omega) \parallel 4\Omega] + 2\Omega = 26/7\ \Omega$$

To find Z_{21}:

We know that Z_{21} is the ratio of V_2 and I_1 when $I_2 = 0$. So, we have to find V_2 which, in this circuit, is voltage across 1 Ω resistor when $I_2 = 0$.

Applying KVL to second loop,

$3(I_1 – I_3) – 4\ I_3 = 0$; which gives, $I_3 = 3/7\ I_1$

So, the voltage $V_2 = (I_1 – I_3) \times 1\Omega = (I_1 – 3/7\ I_1) \times 1 = 4/7\ I_1$

Hence, $Z_{21} = V_2/I_1 = 4/7\ \Omega$

To find Z_{22}:

Put current $I_1 = 0$. Find the impedance from prot-2 end.

So, $Z_{22} = (4\Omega + 2\Omega) \| 1\Omega = 6/7\ \Omega$

To find Z_{12}:

We know that Z_{12} is the ratio of V_1 and I_2 when $I_1 = 0$. So, we have to find V_1 which, in this circuit, is voltage across 4 Ω resistor when $I_1 = 0$.

Applying KVL to second loop,

$– 2I_3 + (I_1 – I_3) – 4\ I_3 = 0$; which gives, $I_3 = 1/7\ I_2$

So, the voltage $V_1 = 1/7\ I_2 \times 4\Omega = 4/7\ I_2$

Hence, $Z_{21} = V_1/I_2 = 4/7\ \Omega$

We can see the network is reciprocal as $Z_{21} = Z_{12}$.

Example 5.9:

Find the Z-parameters for the circuit given in Fig. 5.25

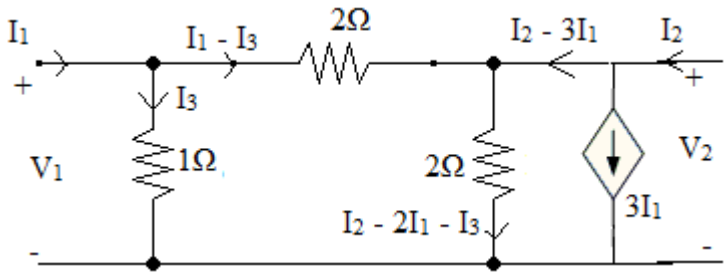


Figure: 5.25

Solution:

Another very simple way to find the Z-parameters is to obtain the equations in form of equation (2.1), after applying KVL to circuit.

Applying KVL to first loop,

$$V_1 = I_3 \tag{5.9}$$

V_2 is the voltage across 2Ω parallel connected resistor.

So, $V_2 = (I_2 - 2I_1 - I_3)\times 2 = 2I_2 - 4I_1 - 2I_3$

Now applying KVL to second loop,

$V_2 - I_3 + 2(I_1 - I_3) = 0$

Putting value of V_2,

$$2I_2 - 4I_1 - 2I_3 - I_3 + 2(I_1 - I_3) = 0 \tag{5.10}$$

Solving equations (5.9) & (5.10)

$V_1 = -0.4\, I_1 + 0.4\, I_2$&

$V_2 = -3.2\, I_1 + 1.2\, I_2$

Comparing the above equations with equation (5.7)

We have, Z11 = -0.4 Ω, $Z_{12} = 0.4\Omega$, $Z_{21} = -3.2\ \Omega$, $Z_{22} = 1.2\Omega$

5.7.2 Short Circuit Admittance (Y) Parameters

A general two port network is shown below again:

The Y-parameters may be defined by expressing the port currents I_1 and I_2 in terms of voltages V_1 and V_2. Here, I_1 and I_2 are dependent variables and V_1& V_2 are independent variables.

Therefore,

$I_1 = Y_{11}V_1 + Y_{12}\, V_2$

$$I_2 = Y_{21}\, V_1 + Y_{22}\, V_2 \tag{5.11}$$

Y_{11}, Y_{12}, Y_{21}& Y_{22} are called admittance parameters. The individual values of parameters can be obtained by short-circuiting the port voltages.

Short-circuit the second port, then $V_2 = 0$, and

$$Y_{11} = I_1/V_1 \& Y_{21} = I_2/V_1 \quad (5.12)$$

Y_{11} is short circuit input impedance and Y_{21} is short-circuit forward transfer admittance.

Short-circuit the first port, then $V_1 = 0$, and

$$Y_{12} = I_1/V_2 \& Y_{22} = I_2/V_2 \quad (5.13)$$

Y_{12} and Y_{22} are called short-circuit reverse transfer admittance and short circuit output impedance, respectively.

The equivalent circuit governed by equation (5.11) may be drawn as (Fig. 5.26)

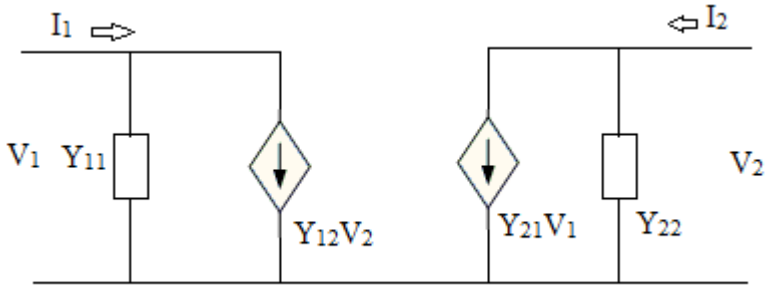


Figure: 5.26

Example 5.10:

Find the Y-parameters for the network shown in Fig. 5.27

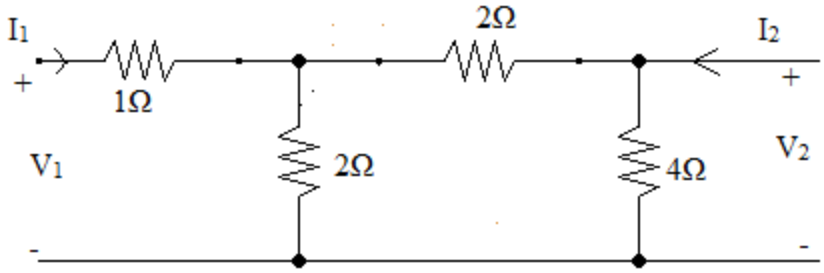


Figure: 5.27

Solution:

First Method:

Remove the source V_2 and short circuit port two. The 4 Ω resistor will become ineffective and the circuit will look like Fig. 5.28

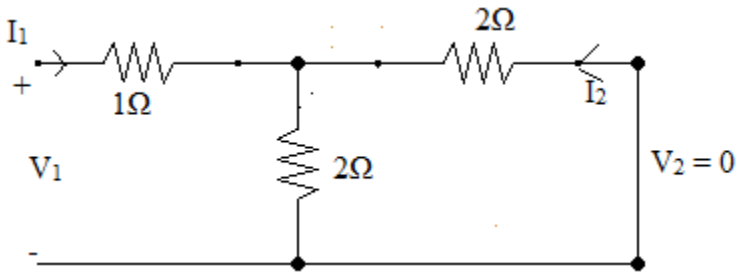


Figure: 5.28

The equivalent impedance from port-1 end is

$Z_{eq} = (2 \| 2)+1 = 2\Omega$

So, $V_1 = I_1 \times 2$

$Y_{11} = I_1/V_1 = 1/2$ ℧

Apply KVL to the right loop of Fig. 5.28

$2I_2 + 2 (I_1 + I_2) = 0$

$I_2 = - ½ I_1$

As $V_1 = 2 I_1$, Therefore $I_2 = - V_1/4$

And so, $Y_{21} = I_2/ V_1 = - ¼$ ℧

Similarly, to find other two parameters, short circuit the port-1 (Fig. 5.29), then the equivalent impedance from port-2 end is

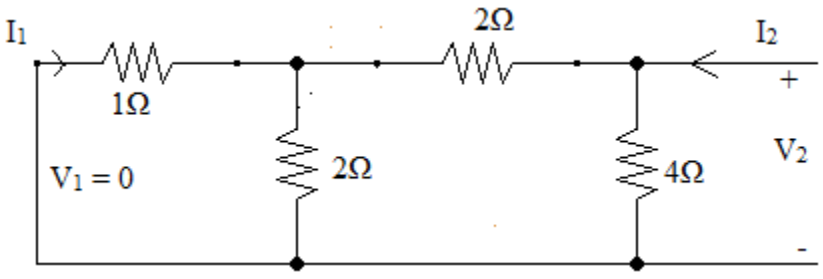


Figure: 5.29

$Z_{eq} = [(2 \| 1)+2] \| 4 = 8/5\ \Omega$

So, $V_2 = I_2 \times 8/5$

$Y_{22} = I_2/V_2 = 5/8\ \text{℧}$

Apply KVL on circuit of Fig. 5.28, we have

$I_1 = -2/5\ I_2$

As $V_2 = 8/5\ I_2$, Therefore $I_1 = -V_2/4$

And so, $Y_{12} = I_1/V_2 = -¼\ \text{℧}$

Second Method:

In another method to find Y-parameters, first find the Z-parameters of network.

Applying KVL to all three loops of the circuit of Fig. 5.27, we have after solving

$V_1 = 5/2\ I_1 + I_2$&

$V_2 = I_1 + 2\ I_2$

Comparing these equations with equation (5.7), we have

$Z_{11} = 5/2\ \Omega$, $Z_{12} = Z_{21} = 1\Omega$, $Z_{22} = 2\Omega$

So the Z-matrix may be written as

$$[Z] = \begin{bmatrix} 5/2 & 1 \\ 1 & 2 \end{bmatrix}$$

The magnitude of Z

$|Z| = 5/2 \times 2 - 1 \times 1 = 4$

Now the Y-matrix can be obtained using Z- parameters in following way.

$$[Y] = [Z]^{-1} = \begin{bmatrix} \frac{Z_{22}}{|Z|} & \frac{-Z_{12}}{|Z|} \\ \frac{-Z_{21}}{|Z|} & \frac{Z_{11}}{|Z|} \end{bmatrix}$$ (Remember this matrix)

Hence, $[Y] = \begin{bmatrix} \frac{2}{4} & \frac{-1}{4} \\ \frac{-1}{4} & \frac{5}{8} \end{bmatrix}$, Giving $Y_{11} = 1/2\ \text{℧}$, $Y_{12} = Y_{21} = -1/4\ \text{℧}$, $Y_{22} = 5/8\ \text{℧}$

This is the same answer as obtained using first method.

5.7.3 Transmission (ABCD) Parameters

These parameters are widely used in transmission line. Port-1 is called the sending end and port-2 is called receiving end. The transmission parameters provide a relationship between input and output and express the source variables V_1 and I_1 in terms of destination variables V_2 and I_2. The parameters are defined by the equations

$V_1 = A\ V_2 - B\ I_2$&

$$I_1 = C\ V_2 - D\ I_2 \qquad (5.14)$$

Here, minus sign in second term of right hand is used with I_2.

You have learned by now that how to find these parameters. Using equation (5.14),

$A = \left.\frac{V_1}{V_2}\right|_{I_2=0}$, $1/A = V_2/V_1$ is called open circuit voltage gain.

$-1/B = \left.\frac{I_2}{V_1}\right|_{V_2=0} = Y_{21}$, is short circuit transfer admittance

$1/C = \left.\frac{V_2}{I_1}\right|_{I_2=0} = Z_{21}$, is open circuit transfer impedance

$-1/D = \left.\frac{I_2}{I_1}\right|_{V_2=0} = \alpha$, is short circuit current gain.

Example 5.11:

Find the transmission or ABCD parameters for the circuit shown in Fig. 5.30

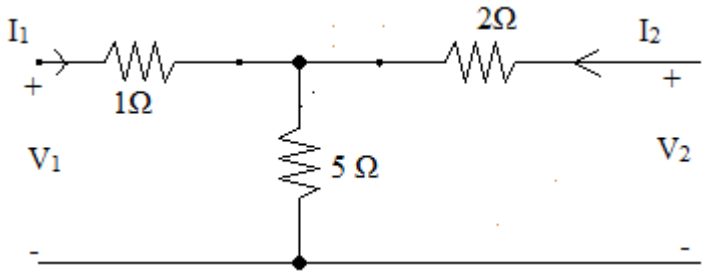


Figure: 5.30

Solution:

Open circuit port-2, so that $I_2 = 0$ (we can find A and C when $I_2 = 0$), then

$A = V_1 / V_2$,

Here, $V_1 = 6 I_1$ and $V_2 = 5 I_1$ and so

$A = 6/5$

$C = I_1 / V_2 = 1/5$ ℧

When port-2 is short circuited, then $V_2 = 0$ (we can find B and D when $V_2 = 0$), so

$B = -V_1 / I_2 = 17/5$ Ω (applying KVL when port-2 is short circuited)

& $D = -I_1/I_2 = 7/5$

5.7.4 Hybrid (h) Parameters

These parameters are useful in transistor circuits. Voltage at one port and current on other port are taken as independent variables. The dependent variables, V_1 and I_2 , are expressed in terms of independent variables V_2 and I_1 in following way:

$V_1 = h_{11} I_1 + h_{12} V_2$

$$I_2 = h_{21} I_1 + h_{22} V_2 \qquad (5.15)$$

When $V_2 = 0$, i.e. when port-2 (referring Fig. 5.22) is short circuited, then

$h_{11} = \left.\frac{V_1}{I_1}\right|_{V_2=0}$, short circuit input impedance

$h_{21} = \left.\frac{I_2}{I_1}\right|_{V_2=0}$, short circuit forward current gain

Similarly, by letting port-1 open, i.e. $I_1 = 0$

$h_{12} = \left.\frac{V_1}{V_2}\right|_{I_1=0}$, open circuit reverse voltage gain

$h_{22} = \left.\frac{I_2}{V_2}\right|_{I_1=0}$, open circuit output admittance

An equivalent circuit in terms of h-parameters is shown in Fig. 5.31.

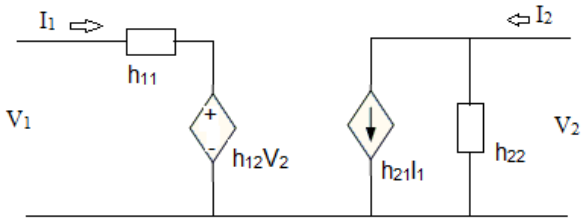


Figure: 5.31

Example 5.12:

Find the h-parameters of the network shown below [Fig. 5.32].

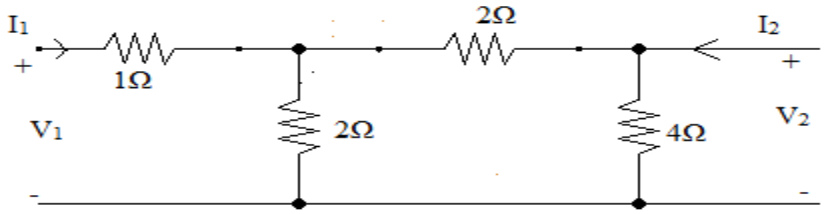


Figure: 5.32

Solution:

When port-2 is short circuited, $V_2 = 0$, and we may find h_{11} and h_{21}.

$h_{11} = V_1 / I_1$, here $V_1 = I_1 \times Z_{eq}$

$Z_{eq} = 2\ \Omega$, so $h_{11} = 2\ \Omega$

$h_{21} = I_2 / I_1$ when $V_2 = 0$; $I_2 = -I_1/2$

And so $h_{21} = -1/2$

Similarly, when port-1 is open, $I_1 = 0$, and we get

$h_{12} = 1/2$ and $h_{22} = ½\ \mho$

5.8 SUMMARY

In this chapter, we have studied a three terminal device, called transistor. The mode of operation and various configurations have been described. The input characteristics, output characteristics and current gain characteristics are studied for all three configurations. Transistors are mainly used in signal amplification circuits and switching circuits. The active region of operation is employed when transistor is to be used for amplification purpose. In digital electronics, transistors are mostly used as a switch. When working as switch, the saturation and cut-off regions are employed. The important transistor parameters like current gains α and β and their relations are discussed with solving some related problems.

5.9 GLOSSARY

Alpha (α)	Ratio of collector current to emitter current in a bipolar junction transistor (BJT).
Base	The region that lies between the emitter and collector of a bipolar junction transistor (BJT).
Beta (β)	The ratio of collector current to base current in a bipolar junction transistor (BJT).
Hybrid circuit	Circuit that combines two technologies
Admittance	Measure of how easily AC will flow through a circuit
Transistor	Semiconductor device that can be used as an amplifier or as an electronic switch.
Transmission line	Conducting line used to transmit signal energy between two points
Terminal	Point at which electrical connections are made.
Emitter	The region from which charge carriers are injected into the base of a bipolar junction transistor.
Collector	The region in a bipolar junction transistor through which a flow of charge carriers leaves the base region.

5.10 SUGGESTED READINGS

- Gupta Kumar Sharma-Hand book of electronics, Pragati Prakashan.
- V.K. Mehta & R. Mehta-Principles of Electronics- S. Chand
- Basic Electronics-B.L. Theraja-S. Chand

5.11 TERMINAL QUESTIONS

Long Answer type

1. What is transistor? Describe its construction and basic transistor operation.
2. What is the difference between npn and pnp transistors? Draw the symbols of both type of transistors showing the current directions.
3. What are current gains α and β? Derive the relation between α and β.
4. Describe the common-base transistor configuration. Plot and describe the input characteristics, output characteristics and current gain characteristics mentioning the various input and output terminal voltages and currents.
5. Describe the common-emitter transistor configuration. Plot and describe the input characteristics, output characteristics and current gain characteristics mentioning the various input and output terminal voltages and currents.
6. Describe the common-collector transistor configuration. Plot and describe the input characteristics, output characteristics and current gain characteristics mentioning the various input and output terminal voltages and currents.
7. Describe the transistor regions of operation by plotting all the regions.
8. What is avalanche breakdown? Why should we avoid this region of operation?

Numerical Answer type

1. If the emitter current I_E is 1.2 mA and β=60, find α, I_B and I_C.
Ans: 0.983, 19.6 μA, 1.176 mA

2. If the emitter current of a transistor is 8 mA and I_B is 1/100 of I_C, determine the levels of I_C and I_B.

Ans: I_C = 7.921 mA, I_B = 79.21 μA

3. Given an α_{dc} of 0.998, determine I_C if I_E = 4mA.

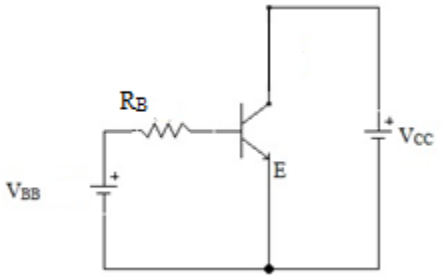
(a) Determine α_{dc} if I_E = 2.8 mA and I_B = 20 μA.
(b) Find I_E if I_B= 40 μA and α_{dc} is 0.98

Ans: (a) IC = 3.992 mA (b) 0.993 (c) 2 mA

4. Using the CB output characteristics of Fig. 4.7, determine the collector currents if I_E = 4.5 mA & V_{CB} = 4V and I_E = 4.5 mA & V_{CB} = 16 V. What is the change in level of collector current due to change in V_{CB}?

Ans: 4.5 mA (approx.), 4.5 mA (approx.), negligible

5. For the following transistor circuit, find I_B and R_B if V_{CC}= 9V, V_{BB} = 3 V, I_C = 2 mA, β = 50, V_{BE} = 0.7 V & V_{CE} = 4 V

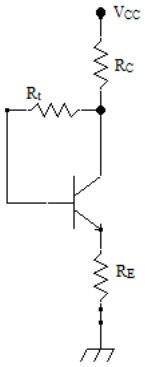


Ans: 0.4 mA, 5.75 kΩ

6. The base-emitter junction is forward biased with a supply voltage of 10 V through a base resistance of 500 kΩ. The collector junction is reverse biased with 15 V through 1.2 kΩ. Find out the base current, collector current and collector-emitter voltage V_{CE} if β = 100.

Ans: 20 μA, 2 mA, 13.6 V

7. For the following circuit, V_{CC} =12V, R_C =2.3 kΩ, R_E = 1 kΩ, V_{CE} = 5V, V_{BE} =0.7 V and β =50. Find I_B, I_C and R_t.



Ans: 41.59 μA, 2.121 mA, 103.39 kΩ

Objective Type Questions

1. A transistor has
   (a) One p-n junction
   (b) two p-n junction
   (c) three p-n junction
   (d) four p-n junction
2. The base of a transistor is …….. doped
   (a) heavily

(b) moderately
(c) lightly
(d) none of the above

3. Which are the minority carriers in npn transistor?
   (a) Free electrons
   (b) Holes
   (c) Donor ions
   (d) Acceptor ions
4. The emitter of a transistor is …… doped
   (a) Lightly
   (b) Heavily
   (c) Moderately
   (d) None of the above
5. The current relation for the transistor is
   (a) $I_C = I_E + I_B$
   (b) $I_E = 2 I_C + I_B$
   (c) $I_B = I_E + I_C$
   (d) $I_E = I_C + I_B$
6. The collector of a transistor is ……. doped
   (a) Lightly
   (b) Heavily
   (c) Moderately
   (d) None of the above
7. The value of current gain α of a transistor is
   (a) More than one
   (b) Less than one
   (c) 1
   (d) zero
8. $I_C = \alpha I_E +$ …….
   (a) I_B
   (b) I_{CEO}
   (c) I_{CBO}
   (d) βI_B
9. In a transistor, $I_C = 100$ mA and $I_E = 100.5$ mA. The value of β is
   (a) 1
   (b) 50
   (c) 100
   (d) 200
10. The relation between β and α is
   (a) $\beta = 1 / (1 - \alpha)$

(b) $\beta = (1 - \alpha) / \alpha$
(c) $\beta = \alpha / (1 - \alpha)$
(d) $\beta = \alpha / (1 + \alpha)$

11. The most commonly used transistor arrangement is
    (a) Common base
    (b) Common emitter
    (c) Common collector
    (d) None of the above
12. $I_C = [\alpha / (1 - \alpha)] I_B + [\ldots\ldots / (1 - \alpha)]$
    (a) I_{CBO}
    (b) I_{CEO}
    (c) I_B
    (d) I_E
13. If the value of α is 0.9, then value of β is
    (a) 0.9
    (b) 9
    (c) 9.9
    (d) 90
14. What is an essential possible condition of biasing for a transistor to operate in an active region?
    (a) Forward biasing of emitter base & collector base junctions
    (b) Reverse biasing of emitter-base & collector-base junctions
    (c) Forward biasing of emitter base & Reverse biasing of collector-base junctions
    (d) Reverse biasing of emitter-base & Forward biasing of collector-base junctions
15. Which operating region of BJT enables Emitter-base & Collector-base junctions to undergo perfect short-circuit configuration?
    (a) Active region
    (b) Saturation region
    (c) Cut-off region
    (d) None of the above

16. When transistors are used in digital circuits they usually operate in the:
    (a) Active region
    (b) Breakdown region
    (c) Saturation and cut-off regions
    (d) Linear region

Answers objective type questions

1.b 2.c 3.b 4. b 5.d 6.c 7.b 8.c 9.d 10.c 11.b 12. a 13.d 14.c 15.b 16.c

UNIT 6 JFETs and MOSFETs

Structure

6.1 INTRODUCTION

The field-effect transistor (FET), like bipolar junction transistor (BJT), is a three terminal semiconductor device. However, the FET operates under principles completely different from those of the BJT. A field effect transistor is called unipolar device because the current through it results from the flow of only one of the two kinds of charge carriers: holes or electrons. The name field effect is derived from the fact that the current flow is controlled by an electric field set up in the device by an externally applied voltage. FET exhibits a high input resistance, typically many mega-ohms and is less noisy than bipolar transistors. FET also exhibits no offset voltage at zero drain current, and hence makes an excellent signal chopper.

There are two main types of FETs: The junction field-effect transistor (JFET) and the metal-oxide-semiconductor FET (MOSFET) also known as insulated gate FET (IGFET). The MOSFET is the most important component in modern digital integrated circuits, such as microprocessors and computer memories.

6.2 OBJECTIVES

After studying this unit, you should be able to-

- know about FETs
- construct JFET and MOSFET
- define various biasing in JFET and MOSFET
- plot the characteristics of JFET and MOSFET

6.3 JUNCTION FIELD EFFECT TRANSISTOR (JFET)

6.3.1 Structure

Figure 1 shows a diagram of the structure of a JFET and identifies the three terminals to which external electrical connection are made. As shown in the figure, a bar of N-type material has regions of P material embedded in each side. The two P regions are joined electrically and the common connection between them is called the gate (G) terminal. A terminal at one end of the N-bar is called the source (S), and a terminal at the other end is called the drain (D). The region of N material between the two opposing P regions is called the channel. The transistor shown in the figure is therefore called an N-channel JFET, while a device constructed from a P-type bar with embedded N regions is called a P-channel JFET. If we compare JFET with bipolar junction transistor, the drain corresponds to collector of a BJT, the source corresponds to the emitter, and the gate corresponds to the base. As we shall see, the voltage applied to the gate controls the flow of the current between drain and source, just as the signal applied to the base of a BJT controls the flow of current between collector and emitter.

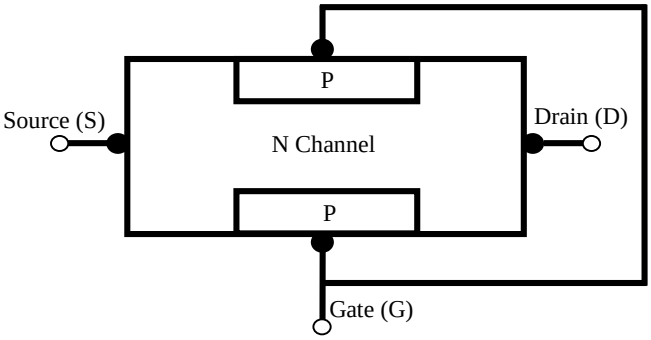


Figure 6.1: Structure of an N-channel JFET.

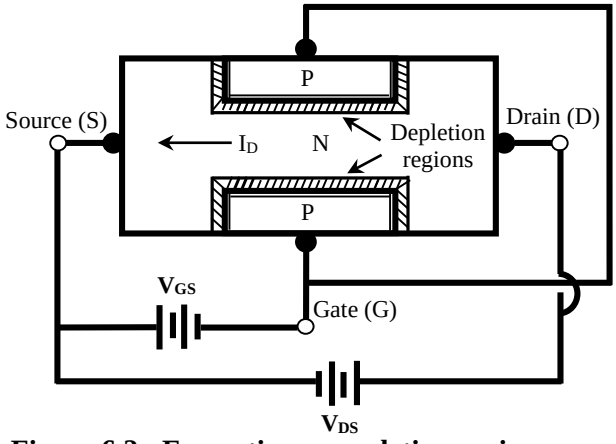
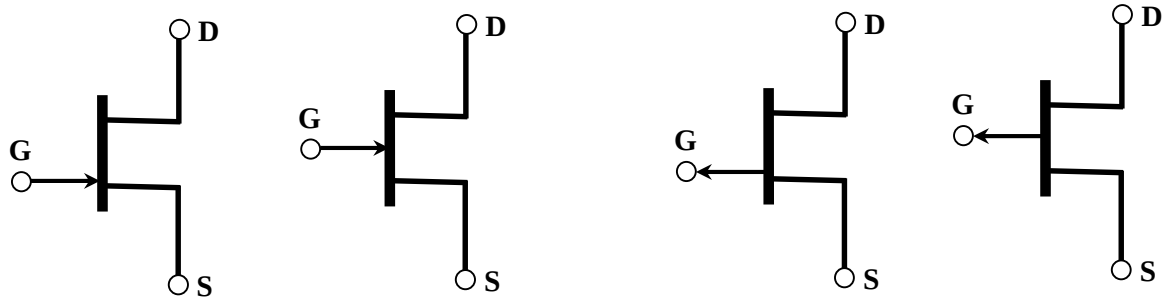


Figure 6.2: Formation of depletion regions.

6.3.2 Symbols

Figure 3 shows the schematic symbols used to represent N-channel and P-channel JFETs. Note that the arrowhead on the gate points into an N-channel JFET and outward for a P-channel device. The symbols showing the gate terminal off-centre are used as means of identifying the source: the source is the terminal drawn closet to the gate arrow. Some JFETs are manufactured so that the drain and source are interchangeable, and the symbols for these devices have the gate arrow drawn in the centre.



(a) N-Channel JFET (b) P-Channel JFET

Figure 6.3: Schematic Symbols of JFET.

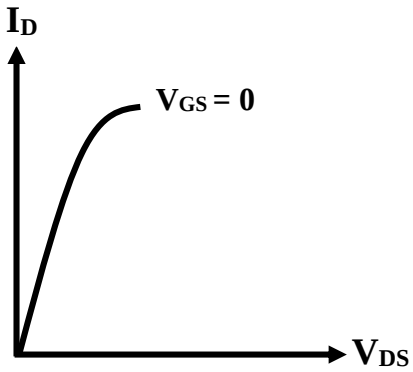
6.3.3 Working and Characteristics

When an external voltage is connected between the drain and the source of an N-channel JFET, so that the drain is positive with respect to the source, current is established by the flow of electrons through the N material from the source to the drain. (The source is so named because it is regarded as the origin of the electrons). Thus, conventional current flows from drain to source and is limited by the resistance of the N material. In normal operation, an external voltage is applied between the gate and the source so that the PN junctions on each side of the channel are reversed biased. Thus, the gate is made negative with respect to the source, as illustrated in fig. 2. Note in the figure that

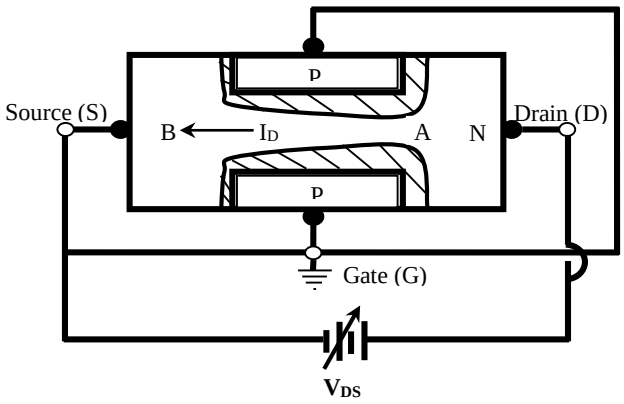
the reverse bias causes a pair of depletion regions form in the channel. The channel is more lightly doped than the gate, so depletion regions penetrate more deeply into the N-type channel than into the P material of the gate.

The width of the depletion regions in fig. 2 depends on the magnitude of the reverse-biasing voltage V_{GS}. As V_{GS} is made more negative, the depletion regions expand and the width of the channel decreases. The reduction in channel width increases the resistance of the channel and thus decreases the flow of current I_D from drain to source.

To investigate the effect of increasing V_{DS} on the drain current I_D, let us suppose for the moment that the gate is shorted to the source ($V_{GS} = 0$). As V_{DS} is increased slightly above 0, we find that the current I_D increases in direct proportion to it, fig. 4(a). This is as we would expect, since increasing the voltage across the fixed-resistance channel simply causes an Ohm's law increase in the current through it. As we continue to increase V_{DS}, we find that noticeable depletion regions begin to form in the channel fig. 4(b). Note that the depletion regions are broader near the drain end of the channel (in the vicinity of point A) than they are near the source end (point B). This is explained by the fact that current flowing through the channel creates a voltage drop along the length of the channel. Near the right of the channel, the channel voltage drops nearly equal to V_{DS}, so there is a large reverse-biasing voltage between the N channel and the P gate. As we proceed left to the channel, less voltage is available because of the drop that accumulates through the resistive N material. Consequently, the reverse-biasing potential between channel and gate becomes smaller and the depletion regions become narrower as we approach the source.



(a) Variation of drain current with V_{DS}

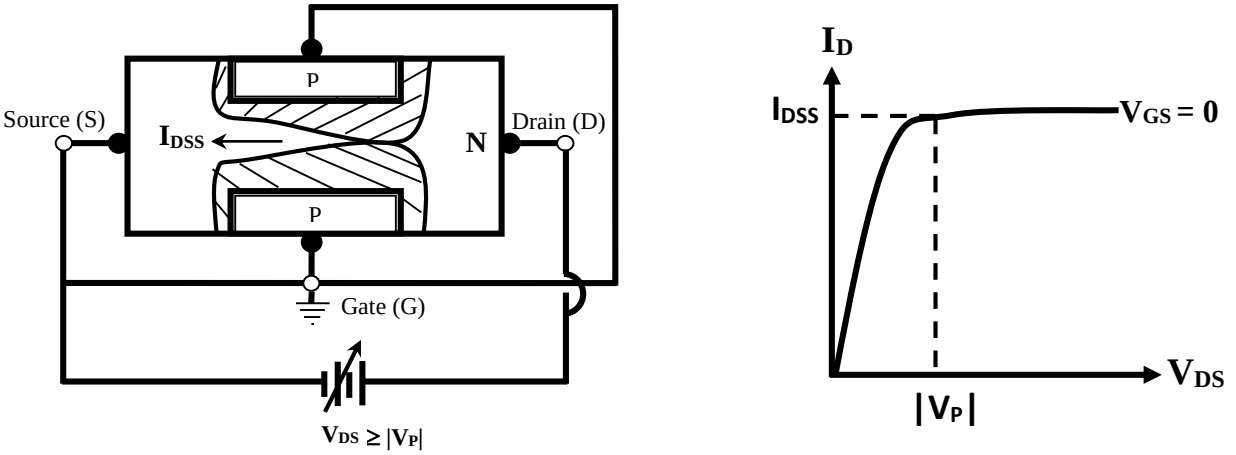


(b) Narrowing the channel width near the drain by increasing V_{DS}

Figure 6.4: Effects of increasing V_{DS} when $V_{GS} = 0$.

When V_{DS} is increased further, the depletion regions expand and the channel becomes very narrow in the vicinity of point A, causing the total resistance of the channel to increase. As a consequence, the rise in current is no longer directly proportional to V_{DS}. Instead, the current begins to level off, as shown by the curved portion of the plot in Fig. 4(a).

Figure 5(a) shows what happens when V_{DS} is increased to a value large enough to cause the depletion regions to meet at a point in the channel near the drain end. This condition is called pinch-off. At the point where pinch-off occurs, the gate-to-channel junction is reverse biased by the value of V_{DS}, so (the negative of) this value is called the pinch-off voltage, V_P. The pinch-off voltage is an important JFET parameter, whose value depends on the doping and geometry of the device. V_P is always a negative quantity for an N-channel JFET and a positive quantity for a P-channel JFET. Fig. 5(b) shows that the current reaches a maximum value at pinch-off and that it remains at that value as V_{DS} is increased beyond $|V_P|$. This current is called the saturation current and is designated I_{DSS} – the Drain-to-Source current with the gate shorted.



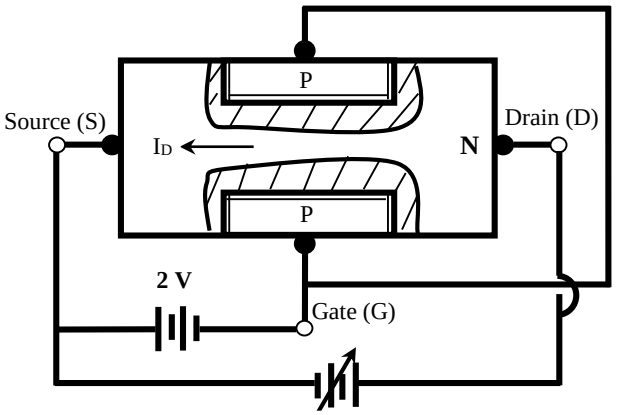
(a) When $V_{DS} \geq |V_P|$, the depletion regions meet in the channel (b) Beyond $|V_P|$, I_D remains at its saturation value I_{DSS}

Figure 6.5: N-Channel JFET at pinch-off.

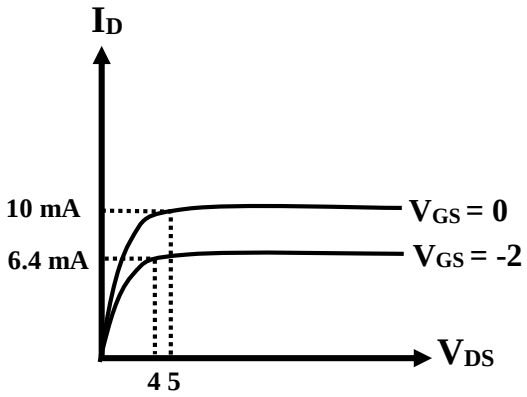
Despite the implication of the name 'pinch-off', note again that current continues to flow through the device when V_{DS} exceeds $|V_P|$. The value of the current remains constant at I_{DSS} because of a kind of self-regulating or equilibrium process that controls the current when V_{DS} exceeds $|V_P|$: Suppose that an increase in V_{DS} did cause I_D to increase; then there would be in the channel an increased voltage drop that would expand the depletion regions further and reduce the current to its original value. Of course, this change in current never actually occurs; I_D simply remains constant at I_{DSS}.

A typical set of values for V_P and I_{DSS} are -5 V and 10 mA, respectively. Suppose we connect a JFET having those parameter values in the circuit shown in fig. 6(a). Note that the gate is no longer shorted to the source, but a voltage V_{GS} = -2 V is connected to reverse bias the gate-to source junctions. The reverse bias causes the depletion region to penetrate the channel farther along the entire length of the channel than they did when V_{GS} was 0. If we now begin to increase V_{DS} above 0, we find that the current I_D once more begins to increase linearly, as shown in Fig. 6(b). Note that the slope of this line is not steep as that of the V_{GS} = 0 line, because the total resistance of the narrower channel is greater than before. As we continue to increase V_{DS}, we find that the depletion regions again approach each other in the vicinity of the drain. This further narrowing of the channel

increases its resistance and the current again begins to level off. Since there is already I-V reverse bias between the gate and the channel, the pinch-off condition, where the depletion regions meet, is now reached at V_{DS} = 4 V instead of 5 V ($V_{DS} = V_{GS} - V_P$). As shown in fig. 6(b), the current saturates at the lower value of 6.4 mA as V_{DS} is increased beyond 4 V.



(a) VGS = -2 V creates, along the length of the channel, a depletion region that is wider than when VGS = 0



(b) ID increases with increasing V_{DS}, until pinch-off occurs at V_{DS} = 4 V

Figure 6.6: Effect of increasing V_{DS} when V_{GS} = -2 V.

If the procedure we have just describes is repeated with V_{GS} set to -4 V instead of -2 V, we find that pinch-off is reached at V_{DS} = 3 V and that the current at I_D = 1.5 mA. It is clear that increasing the reverse-biasing value of V_{GS} (making V_{GS} more negative) causes the pinch-off condition to occur at smaller values of V_{DS} and that smaller saturation currents results. Figure 7 shows the family of characteristic curves, the drain characteristics, obtained when the procedure is performed for V_{GS} = 0, -1, -2, -3 and -4 V. The dashed line, which is parabolic, joins the points on each curve where pinch-off occurs. A value of $V_{DS(sat)}$ is the difference between V_{GS} and V_P: $V_{DS(sat)} = V_{GS} - V_P$, as we have already describe. The equation of parabola is

$$I_D = I_{DSS}\left(\frac{V_{DS(sat)}}{V_P}\right)^2 \qquad (1)$$

To illustrate, we have, in our example, V_P = -5 V and I_{DSS} = 10 mA; so at $V_{DS(sat)}$ = 4 V we find

$$I_D = (10\ mA)\left(\frac{4}{-5}\right)^2 = 6.4\ mA$$

Which is the saturation current at the V_{GS} = -1 V line (see Fig. 6(b)). Note in fig. 7 that the region to the right of the parabola is called the pinch-off region. This is the region in which the JFET is normally operated when used for small-signal amplification. It is also called the active region, or the saturation region. The region to the left of the parabola is called the voltage controlled resistance region, the Ohmic region, or the triode region. In this region, the resistance between

drain and source is controlled by V_{GS}, as we have previously discussed, and we can see that the lines become less steep (implying larger resistance) as V_{GS} becomes more negative. The device acts like a voltage-controlled resistor in this region, and there are some practical applications that exploit this characteristic.

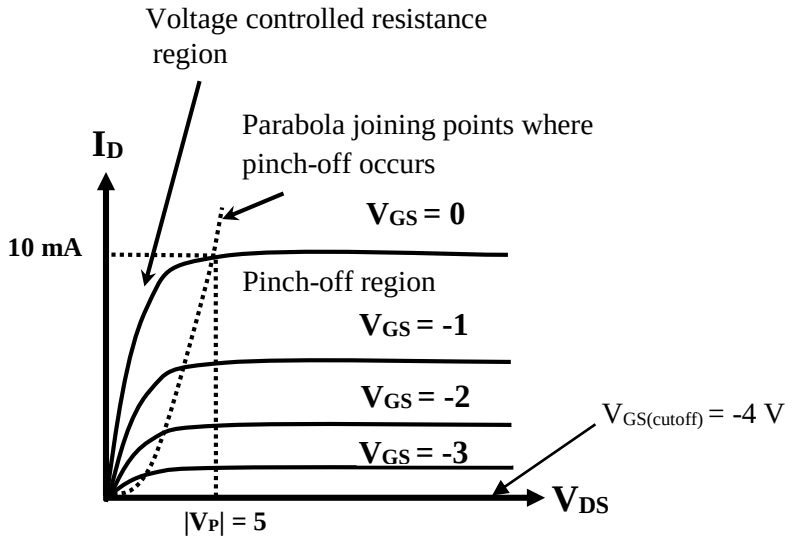


Figure 6.7: Drain characteristics of N-channel JFET.

The line drawn along the horizontal axis in fig. 7 shows that $I_D = 0$ when $V_{GS} = -4$ V, regardless of the value of V_{DS}. When V_{GS} reverse biases the gate-to-source junction by an amount equal to V_P, depletion regions meet along the entire length of the channel and the drain current is cut off. Since the value of V_{GS} at which the drain current is cut off is the same as V_P, the pinch-off voltage is also called the gate-to-source cutoff voltage. Thus, there are two ways to determine the value of V_P from a set of drain characteristics: It is the value of V_{DS} where I_D saturates when $V_{GS} = 0$; and it is the value of V_{GS} that causes all drain current to cease, i.e., $V_P = V_{GS(cutoff)}$.

One property of a field-effect transistor that makes it especially valuable as a voltage amplifier is the very high input resistance at its gate. Since the path from gate to source is a reverse-biased PN junction, the only current that flows into the gate is the very small leakage current associated with a reverse-biased junction. Therefore, very little current is drawn from a signal source driving the gate, and the FET input looks like a very large resistance. A dc input resistance of several hundred mega-ohms is not unusual. Although the gate of an N-channel JFET can be driven slightly positive, this action causes the input junction to be forward biased and radically decreases the gate-to-source resistance. In most practical application, the sudden and dramatic decrease in resistance when the gate is made positive would not be tolerable to a signal source driving a FET.

Similarly, structure and drain characteristics of a typical P-channel JFET can be drawn. Since the channel is P material, current is due to hole flow, rather than electron flow, between drain and

source. The gate material is of course N type. Note that all voltage polarities will be opposite those in the N-channel JFET.

Figure 8 shows the breakdown characteristics of an N-channel JFET.

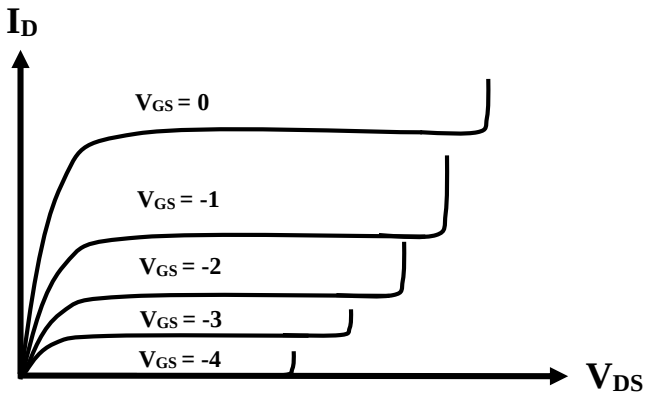


Figure 6.8: Breakdown characteristics of N-channel JFET.

Breakdown occurs at large values of V_{DS} and is caused by the avalanche mechanism described in connection with BJTs. Note that the larger the magnitude of V_{GS}, the smaller the value of V_{DS} at which breakdown occurs.

6.3.4 Transfer Characteristics of JFET

The transfer characteristic of a device is a plot of output current versus input voltage, for a fixed value of output voltage. When the input to a JFET is the gate-to-source voltage and the output current is drain current (common source configuration), the transfer characteristic can be derived from the drain characteristics. It is only necessary to construct a vertical line on the drain characteristics (a line of constant V_{DS}) and to note the value of I_D at each intersection of the line with a line of constant V_{GS}. The value of the I_D can then be plotted against the values of V_{GS} to construct the transfer characteristic. Figure 9 illustrates the process.

Figure 9 shows the transfer characteristics shown for V_{DS} = 10 V. As can be seen in the figure, this choice of V_{DS} means that all points are in the pinch-off region. For example, the point of intersection of the V_{DS} = 10 V line and the V_{GS} = 0 V line occurs at $I_D = I_{DSS}$ = 10 mA. At V_{DS} = 10 V and V_{GS} = -1 V, we find I_D = 6.4 mA. Plotting these combinations of I_D and V_{GS} produces the parabolic transfer characteristic shown. The nonlinear shape of the transfer characteristic can be anticipated by observing that equal increments in the values of V_{GS} on the drain characteristics (ΔV_{GS} = 1 V) do not produce equally spaced lines. Note that the intercepts of the transfer characteristic are I_{DSS} on the I_D-axis and V_P on the V_{GS}-axis.

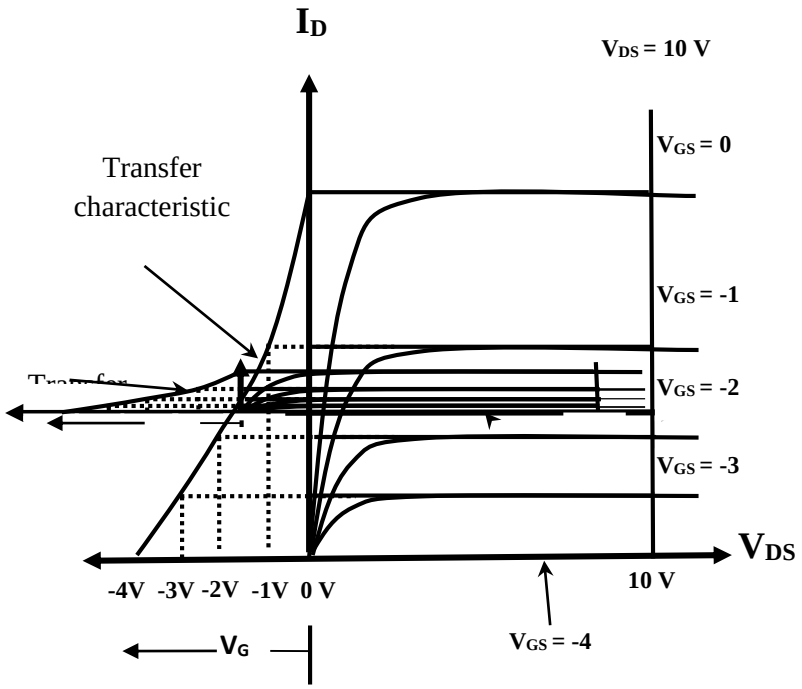


Figure 6.9: Formation of transfer characteristics of N-channel from the drain characteristics.

The equation for the transfer characteristic in the pinch-off region is, to a closed approximation,

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

Note that the above equation correctly predicts that $I_D = I_{DSS}$ when $V_{GS} = 0$ and that $I_D = 0$ when $V_{GS} = V_P$. The transfer characteristic is often called the square-law characteristic of a JFET and is used in some interesting applications to produce outputs that are nonlinear functions of inputs.

6.4 JFET Biasing

6.4.1 Fixed Bias Method

Like a bipolar transistor, a JFET used as an ac amplifier must be biased in order to create a dc output voltage around which ac variation can occur. When a JFET is connected in the common source configuration, the input voltage is V_{GS} and the output voltage is V_{DS}. Therefore, the bias circuit must set dc (quiescent) values for the drain to source voltage V_{DS} and drain current I_D. Figure 10 shows one method that can be used to bias N-channel JFET, similarly P-channel JFET can be biased.

Notice in fig. 10 that a dc supply voltage V_{DD} is connected to supply drain current to the JFET through resistor R_D, and that another dc voltage is used to set the gate to source voltage V_{GS}. This biasing method is called fixed bias because the gate to source voltage is fixed by the constant voltage applied across those terminals. Writing Kirchhoff's voltage law around the output loops in fig. 10, we find

$$V_{DS} = V_{DD} - I_D R_D \tag{3}$$

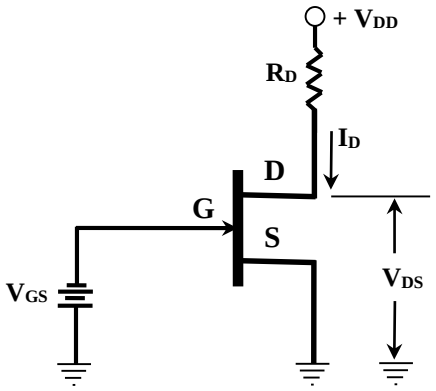


Figure 6.10: Fixed bias circuit for N-channel JFET.

When using this equation, always substitute a positive value for V_{DD} to ensure that the correct sign is obtained for V_{DS}. V_{DS} should always turn out to be a positive quantity in an N-channel JFET and a negative quantity in a P-channel JFET. Equation (3) can be written as –

$$I_D = -\left(\frac{1}{R_D}\right) V_{DS} + \frac{V_{DD}}{R_D} \tag{4}$$

Above equation is the equation of the dc load line for N-channel JFET and can be plotted on a set of drain characteristics to determine a Q-point. The load line intersects the V_{DS} axis at V_{DD} and the I_D axis at V_{DD}/R_D. The Q-point must be located in the pinch-off region for normal amplifier operation. To ensure that the Q-point is in the pinch-off region, the quiescent value of $|V_{DS}|$ must be greater than $|V_P| - |V_{GS}|$.

Of course the quiescent value of I_D can also be determined using the transfer characteristic of a JFET. Since the transfer characteristic is a plot of I_D versus V_{GS}, it is only necessary to locate the V_{GS} coordinate and read the corresponding value of I_D directly. The value of V_{DS} can then be determines using equation (3). While graphical techniques for locating the bias point are instructive, and provide insight to the way in which the circuit variables affect each other, the quiescent values of I_D and V_{DS} can be calculated using a straightforward computation, if the values of V_P and I_{DSS} are known.

6.4.2 Self Bias Method

Figure 11 shows a bias circuit that provides improved stability and requires only a single supply voltage. This bias method is called self bias, because the voltage drops across R_S due to the flow of quiescent current determines the quiescent value of V_{GS}. We can understand this fact by realizing that the current I_D in resistor R_S creates the voltage $V_S = I_D R_S$ at the source terminal, with respect to ground. For the N-channel JFET, this means that the source is positive with respect to the gate, since the gate is grounded. In other words, the gate is negative (by $I_D R_S$ volts) with respect to the source, as required for biasing an N-channel JFET; $V_{GS} = -I_D R_S$. Similarly, for the P-channel device, the gate is positive by $I_D R_S$ volts, with respect to the source: $V_{GS} = I_D R_S$. For N-channel the equation $V_{GS} = -I_D R_S$ describe straight lines when plotted on V_{GS} - I_D axis. Each line is called bias line. The quiescent value of I_D in the self bias circuit can be determined graphically by plotting the bias line on the same set of axis with the transfer characteristic. The intersection of the two locates the Q-point. The quiescent value of V_{DS} can be found by summing voltages (writing Kirchhoff's voltage law) around the output loops in Fig. 11 (here only N-channel is used for investigation).

$$V_{DS} = V_{DD} - I_D(R_D + R_S) \tag{5}$$

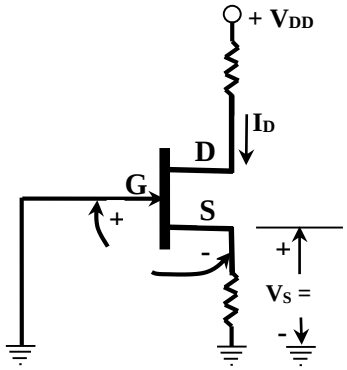


Figure 6.11: Self bias circuit for N-channel JFET.

6.4.3 Voltage Divider Bias Method

Figure 12 shows the voltage divider method for biasing a JFET which is identical to that used for a transistor. The resistors R_1 and R_2 form a voltage divider across drain supply V_{DD}. The voltage V_G across R_2 provides the necessary bias.

$$V_G = \left(\frac{R_2}{R_1 + R_2}\right) V_{DD} \tag{6}$$

The bias line equation for N-channel devices is

$$V_{GS} = V_G - I_D R_S \quad (7)$$

The circuit is so designed that $I_D R_S$ is larger than V_G so that V_{GS} is negative. This provides correct bias voltage. We can find operating point as under:

$$I_D = \frac{V_G - V_{GS}}{R_S} \quad (8)$$

and

$$V_{DS} = V_{DD} - I_D(R_D + R_S) \quad (9)$$

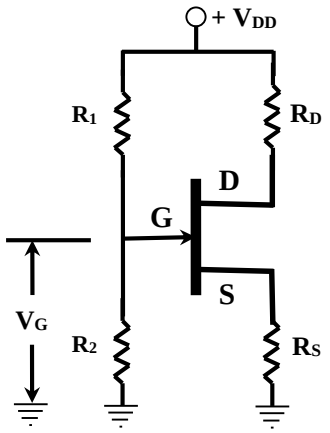


Figure 6.12: Voltage divider biasing circuit for N-channel JFET.

Although the circuit of voltage divider bias is little bit complex, but it provides good stability of the operating point.

6.5 Metal Oxide Semiconductor Field Effect Transistor (MOSFET)

The metal oxide semiconductor FET (MOSFET) is similar in many respect to its JFET counterpart, in that both have drain, gate and source terminals, and both are devices whose channel conductivity is controlled by a gate to source voltage. The principal feature that distinguishes a MOSFET from JFET is the fact that the gate terminal in a MOSFET is insulated from its channel region. For this reason, a MOSFET is often called an insulated gate FET, or IGFET. There are two kinds of MOSFETs: the depletion type and the enhancement types, also referred to as depletion mode and enhancement mode MOSFETs. These names are derived from the two different ways that the conductivity of the channel can be altered by variations in V_{GS}, as we shall see.

6.5.1 The Depletion type MOSFET

Figure 13 shows the structure of an N-channel, depletion type MOSFET. A block of high-resistance, P-type silicon forms a substrate, in which are embedded two heavily doped N-type wells, or pockets, labeled N^+. A thin layer of silicon dioxide (SiO_2), which is an insulating material,

is deposited along the surface. Metal contacts penetrate the silicon dioxide layer at the two N^+ wells and become the drain and source terminals. Between the two N^+ wells are a more lightly doped region of N material that forms the channel. Metal (aluminum) is deposited on the silicon dioxide opposite the channel and becomes the gate terminal. Note that the silicon dioxide insulates the gate from the channel. Going from gate to channel, we encounter metal, oxide and semiconductor, in that sequence, which accounts for the name MOSFET. Notice that there is no PN junction formed between gate and channel, as there is in a JFET.

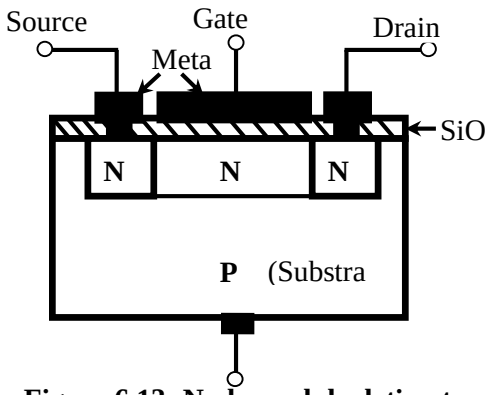


Figure 6.13: N-channel depletion type MOSFET.

Figure 14 shows the normal mode of operation of a depletion type, N-channel MOSFET. A voltage V_{DS} is connected between drain and source to make the drain positive with respect to the source. The substrate is usually connected to the source, as shown in figure. When the gate is made negative with respect to the source by V_{GS}, the electric field it produces in the channel drives electrons away from a portion of the channel near the SiO_2 layer. This portion is thus depleted of carriers and the channel width is effectively narrowed. The narrower the channel, the greater is resistance, and the smaller the current flow from drain to source. Thus the device behaves very much like an N-channel JFET, the principal difference being that the channel width is controlled by the action of the electric field rather than by the size of the depletion region of a PN junction. Since there is no PN junction, the voltage V_{GS} can be made positive without any concern for the consequences of forward biasing a junction. In fact, making V_{GS} positive attracts more electrons into the channel and increases, or enhances its conductivity. Thus, the gate voltage in a depletion type MOSFET can be varied through both positive and negative voltages and the device can operate in both depletion and enhancement modes. For this reason, the depletion type MOSFET is also called depletion-enhancement type.

Although there is a PN junction between the N material and the P substrate, this junction is always reverse biased and very little substrate current flows. The substrate has little bearing on the operation of the device and we will hereafter ignore its presence, other than to show its usual connection to the source. The resistance looking into the gate is extremely large. On the order of thousands of mega-ohms, because there is no PN junction and no path for current to flow through the insulating layer separating the gate and the channel.

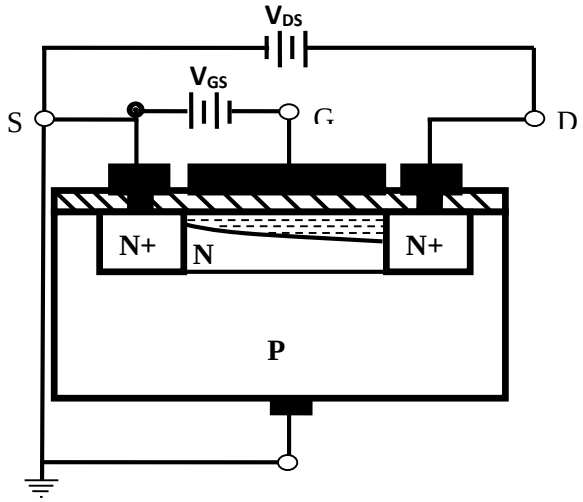


Figure 6.14 Functional diagram of N-channel, depletion type MOSFET.

6.5.2 Symbols of Depletion Type MOSFETs

Figure 15 shows the schematic symbols for N- and P-channel depletion type MOSFETs. The arrow on each symbol is drawn on the substrate terminal and its direction indicates whether the device is N-channel or P-channel. It points into the device for an N-channel and outward for a P-channel. Notice that the gate terminal is attached to a line that is separated from the rest of the symbol, to emphasize that the gate is insulated from the channel.

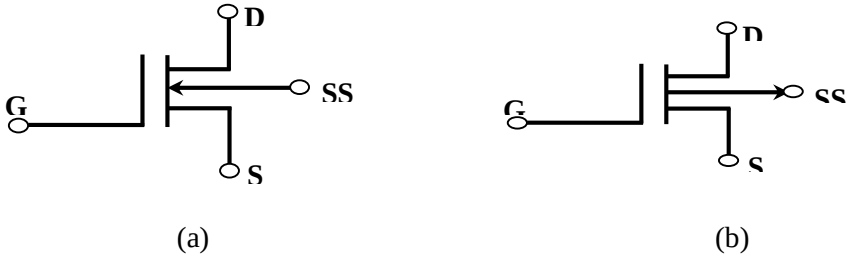


Figure 6.15 Symbols for (a) N-channel and (b) P-channel, depletion type MOSFET.

6.5.3 Characteristics of Depletion type MOSFET

Because of the similarity of a depletion type MOSFET to a JFET, we would expect it to have similar parameters and operating characteristics. This is indeed the case, as shown by the drain characteristics in Figure 16. Note that current increases linearly with increasing V_{DS} until a pinch-off condition is reached. Beyond pinch-off, the drain current remains constant at a saturation value depending on V_{GS}. More negative values of V_{GS} cause pinch-off to be reached sooner and result in smaller values of saturation current. If $V_{GS} = 0$, the drain current saturates at I_{DSS} when $V_{DS} = -V_P$ volts. If V_{GS} is made sufficiently negative to deplete the entire channel, the drain current is completely cut off. The value of V_{GS} at which this occurs is the gate to source cutoff voltage, $V_{GS(cutoff)} = V_P$. Note that the characteristics in fig. 16 also show operation in the enhancement mode, where V_{GS} is positive. Similarly, drain characteristics of a P-channel, depletion type

MOSFET can be drawn. Notice that depletion occurs in this device when V_{GS} is positive and enhancement when V_{GS} is negative.

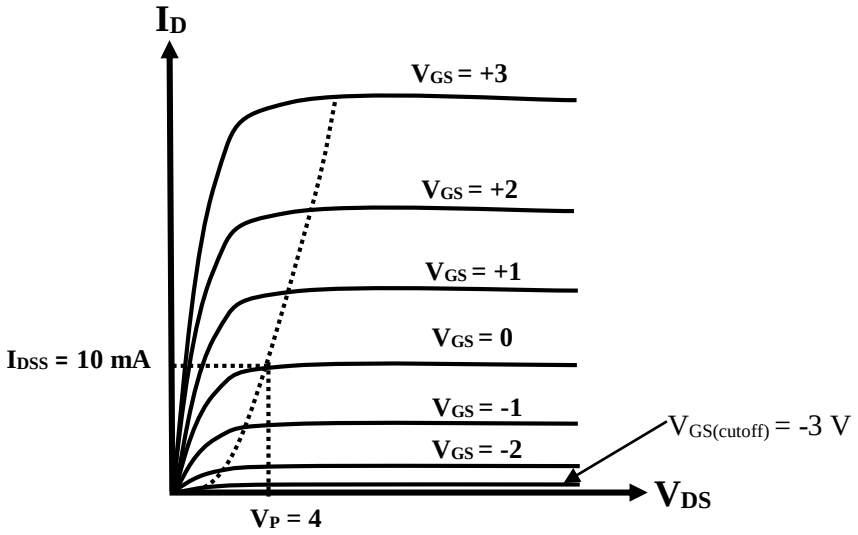


Figure 6.16 Drain characteristics of N-channel, depletion type MOSFET.

6.5.4 Transfer Characteristics of Depletion Type MOSFET

The square law equation for the transfer characteristic of a depletion type MOSFET is identical to that for a JFET:

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2 \tag{10}$$

This equation correctly predicts I_D when the depletion type MOSFET is operated in the enhancement mode. Note, for example, that when V_{GS} is positive in an N-channel device, $(1-V_{GS}/V_P) > 1$ (since V_P is negative), and therefore I_D is greater than I_{DSS}. Figure 17 shows transfer characteristics for N-channel device. Note that I_D exceeds I_{DSS} in the enhancement mode.

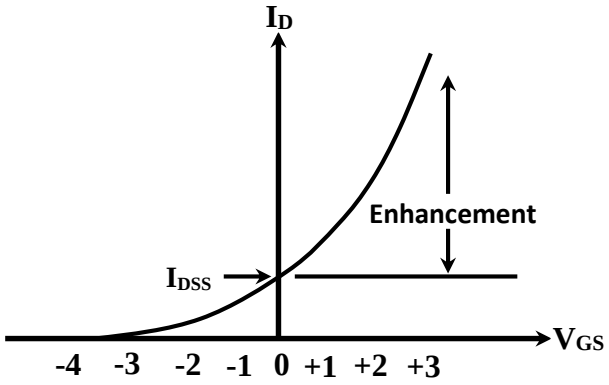


Figure 6.17: Transfer characteristics of N-channel, depletion type MOSFET.

The bias techniques we have already discussed for JFETs are wholly applicable to depletion type MOSFETs, since the characteristics of the two devices are so similar. We shall not repeat our discussion of those techniques.

6.5.5 Enhancement type MOSFETs

Recall that the channel of N-channel depletion MOSFET is the region of N material between the drain and the source (see Figure 13). In the enhancement MOSFET, there is no N-type material between the drain and the source; instead, the p type substrate extends all the way to the SiO_2 layer adjacent to the gate. This structure is shown in fig. 18. Apart from the absence on the N-type channel, the construction is the same as that of the depletion MOSFET.

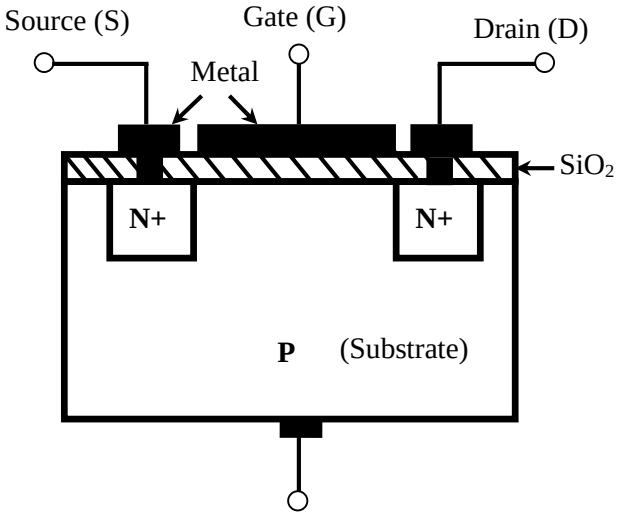


Figure 6.18 N-channel, enhancement type MOSFET.

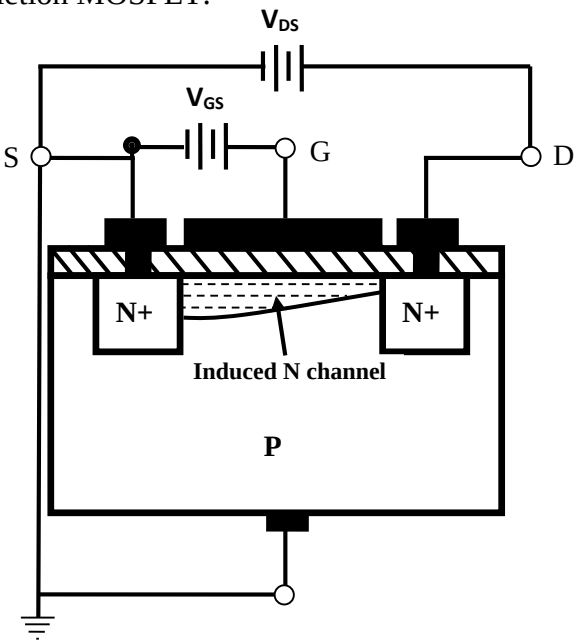


Figure 6.19 Functional diagram of N-channel, enhancement type MOSFET.

Figure 19 shows the normal electrical connections between drain, gate and source. As in the depletion MOSFET, the substrate is usually connected to the source. Notice that V_{GS} is connected so that the gate is positive with respect to the source. The positive gate voltage attracts electrons from the substrate to the region along the insulating layer opposite the gate. If the gate is made sufficiently positive, enough electrons will be drawn into that region to convert it to N-type material. Thus, an N-type channel will be formed between drain and source. The P material is said to have been inverted to form an N-type channel. If the gate is made still more positive, more electrons will be drawn into the region and the channel will widen, making it more conductive. In other words, making V_{GS} more positive enhance the conductivity of the channel and increases the flow of current from drain to source. Since electrons are induced into the channel to convert it to N-type material, the MOSFET shown in fig. 19 is often called an induced N channel, enhancement type MOSFET. When this device is referred to simply as an N-channel enhancement MOSFET, it is understood that the N-channel exists only when it is induced from the P substrate by a positive V_{GS}.

6.5.6 Symbols of Enhancement Type MOSFET

Figure 20 shows the schematic symbols used to represent N-channel and P-channel enhancement MOSFETs. As in previous FET symbols, the arrow is drawn pointing into the device for an N-channel FET and outward for a P-channel FET. The broken line symbolizes the fact that the channel is induced rather than being an inherent part of the structure.

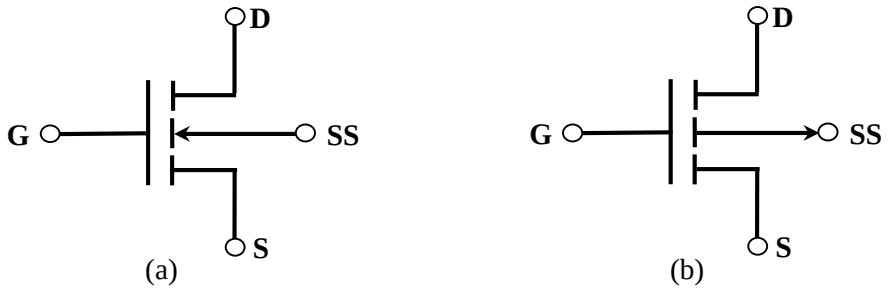


Figure 6.20: Symbols for (a) N-channel and (b) P-channel, enhancement type MOSFET.

6.5.7 Characteristics of Enhancement Type MOSFET

The induced N channel in fig. 21 does not become sufficiently conductive to allow drain current to flow until V_{GS} reaches a certain threshold voltage, V_T. In modern silicon MOSFETs, the value of V_T is typically in the range from 1 to 3 volts. Suppose that $V_T = 2$ V and that V_{GS} is set to some value greater than V_T, say 10 V. We will consider that happens when the drain to source voltage is gradually increased above 0 volts. As V_{DS} increases, the drain current increases because of

normal Ohm's law action. The current rises linearly with V_{DS}, as shown in fig. 21. As V_{DS} continues to increase, we find that the channel becomes narrower at the drain end, as illustrated in fig. 19. This narrowing occurs because the gate to drain voltage becomes smaller when V_{DS} becomes larger, thus reducing the positive field at the drain end. The positive gate to drain voltage decreases by the same amount that V_{DS} increases, so the electric field at the drain end is reduced and the channel is narrowed. As a consequence, the resistance of the channel begins to increase, and the drain current begins to level off. This leveling off can be seen in the curve of fig. 21. When V_{DS} reaches 8 V, then $V_{GD} = 10 - 8 = 2\ V = V_T$. That is, the positive voltage at the drain end reaches the threshold voltage and the channel width at that end shrinks to zero. Further increases in V_{DS} do not change the shape of the channel and the drain current does not increase any further; i.e., I_D saturates. This action is quite similar to the saturation that occurs at pinch-off in a JFET. When the process we have just described is repeated with V_{GS} fixed at 12 V, we find that saturation occurs at $V_{DS} = 12 - 2 = 10$ V. Letting $V_{DS(sat)}$ represent the voltage at which saturation occurs, we have, in the general case,

$$V_{DS(sat)} = V_{GS} - V_T \tag{11}$$

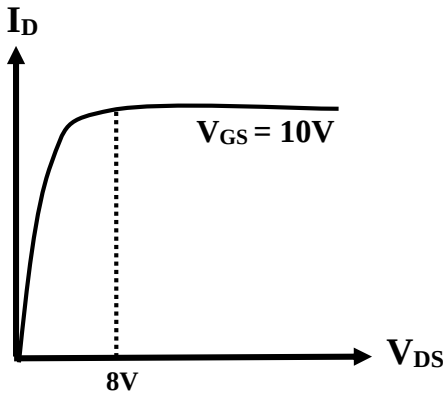


Figure 6.21: Variation of drain current with V_{DS} in N-channel, N-channel, enhancement type MOSFET.

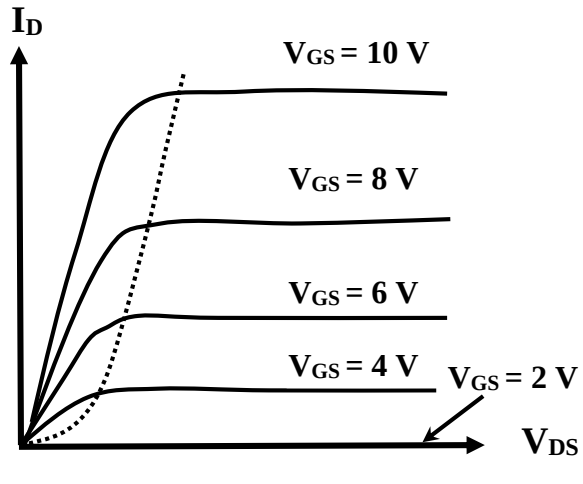


Figure 6.22: Drain characteristics of enhancement type MOSFET

Figure 22 shows a set of drain characteristics resulting from repetitions of the process we have describes, with V_{GS} set to different values of positive voltage. When V_{GS} is reduced to the threshold voltage $V_T = 2$ V, notice that I_D is reduced to 0 for all values of V_{DS}. The drain characteristics are similar to those of an N-channel JFET, except that all values of V_{GS} are positive in the case of the enhancement MOSFET. The enhancement MOSFET can be operated only in an enhancement mode, unlike the depletion MOSFET, which can be operated in both depletion and enhancement modes. The dashed, parabolic line shown on the characteristics in fig. 22 joins the saturation voltages, i.e., those satisfying equation 11. As in JFET characteristics, the region to the left of the parabola is called the voltage controlled resistance regions where the drain to source resistance

changes with V_{GS}. We will refer to the region to the right of the parabola as the active region. The device is normally operated in the active region for small-signal amplification.

Similarly the structure of a P-channel enhancement MOSFET and its electrical connections can be drawn. Note that the substrate is N-type material and that a P-type channel is induced by a negative V_{GS}. The field produced by V_{GS} drives electrons away from the region near the insulating layer and inverts it to P material. N-channel and P-channel MOSFETs are often called NMOS and PMOS devices for short.

6.5.8 Transfer Characteristics of Enhancement Type MOSFET

In the active region, the drain current and gate to source voltage are related by

$$I_D = 0.5\,\beta(V_{GS} - V_T)^2 \qquad V_{GS} \geq V_T \tag{12}$$

Where β is a constant, whose value depends on the geometry of the device, among other factors. A typical value of β is 0.5×10^{-3} and $V_T = 2$ V and corresponding transfer characteristic is shown in fig. 23.

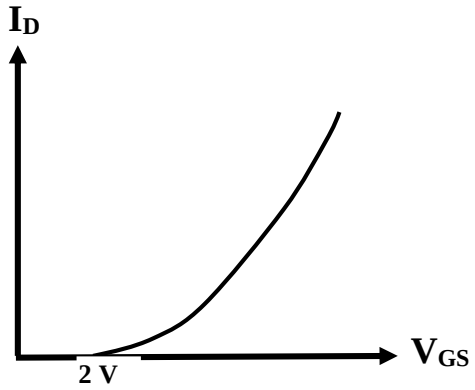


Figure 6.23: Transfer characteristic for N-channel, enhancement type MOSFET.

6.5.9 Enhancement MOSFET Bias Circuits

Although enhancement MOSFETs are most widely used in digital integrated circuits (and require no bias circuitry in those applications), they can, and occasionally do, find applications in small-signal amplifiers. Figure 24 shows one way to bias an enhancement NMOS for such application. This circuit appears to be the bias circuit we used for N-channel JFET (fig. 12) but it is quite different in principle. The resistor R_S does not provide self-bias as it does in the JFET circuit.

Self-bias is not possible with enhancement devices; it can occur only in depletion devices. In fig. 24, the resistor R_S is used to provide feedback for bias stabilization, in the same way that the emitter resistor does in a BJT bias circuit. The larger the value of R_S, the less sensitive the bias point is to changes in MOSFET parameters caused by temperature changes or by device

replacement. Figure 24 shows the voltage drops in the enhancement MOSFET bias circuit. R_1 and R_2 form a voltage divider that determines the gate to ground voltage V_G:

$$V_G = \left(\frac{R_2}{R_1 + R_2}\right) V_{DD} \tag{13}$$

The voltage divider is not loaded by the very large input resistance of the MOSFET, so the values of R_1 and R_2 are usually made very large to keep the ac input resistance of the stage large. Writing Kirchhoff's voltage law around the gate to source loop, we find

$$V_{GS} = V_G - I_D R_S \ (NMOS) \tag{14}$$

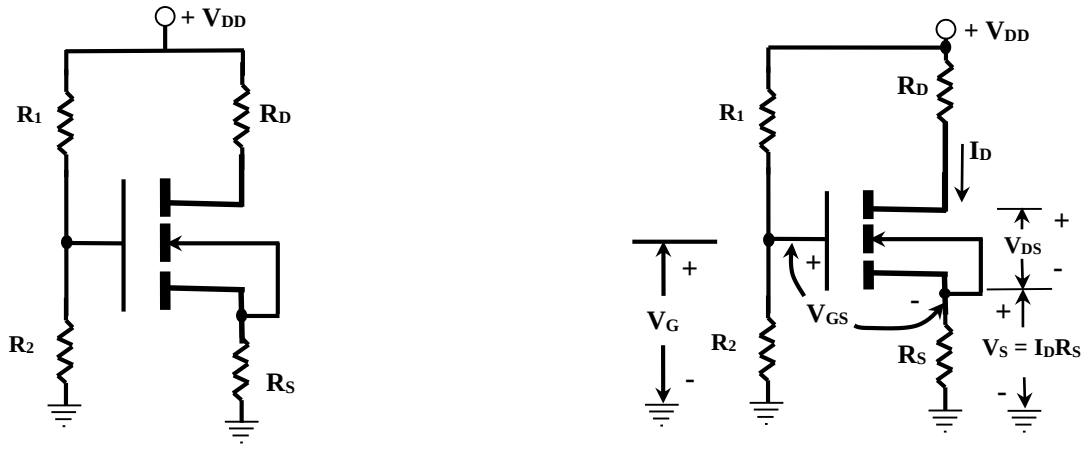


Figure 6.24: Voltage divider biasing circuit for N-channel, enhancement MOSFET.

For a PMOS device, V_G and V_{GS} are negative, so equation 14 would be written

$$V_{GS} = V_G + I_D R_S \ (PMOS) \tag{15}$$

(Note that I_D is considered positive in both equations). Writing Kirchhoff's voltage law around the drain to source loop, we find

$$V_{DS} = V_{DD} - I_D(R_D + R_S) \quad (NMOS) \tag{16}$$

Again regarding I_D as positive in both the NMOS and PMOS devices, the counterpart of equation 16 for PMOS device is

$$V_{DS} = -|V_{DD}| + I_D(R_D + R_S) \quad (PMOS) \tag{17}$$

V_{DS} is negative in a PMOS circuit; note that the absolute value of V_{DD} must be used in equation 17 to obtain the correct sign for V_{DS}.

Equation 14 can be rewritten in the form

$$I_D = -\left(\frac{1}{R_S}\right)V_{GS} + \frac{V_G}{R_S} \qquad (18)$$

Equation 18 is seen to be equation of a straight line on I_D-V_{GS} axis. It intercepts the I_D axis at V_G/R_S and the V_{GS} axis at V_G. The line can be plotted on the same set of axis as the transfer characteristics of the device, and the point of intersection locates the bias values of I_D and V_{GS}. The bias point can also be determined algebraically by solving equation 14 simultaneously with the transfer characteristic equation.

6.6 SUMMARY

The bipolar transistor relies on two types of charge: free electrons and holes. This is why they are called bipolar. This unit discusses another kind of transistor called the field effect transistor (FET). This type of device is unipolar because its operation depends on only one type of charge, either free electrons or holes. In other words, an FET has majority carriers but not minority carriers. There are two kinds of unipolar transistors: JFETs and MOSFETs. In this unit construction, working, characteristics and biasing of these FETs has been discussed.

6.7 GLOSSARY

Amplitude: The size of a signal, usually its peak value.

Bipolar Transistor: A transistor where both free electrons and hole are necessary for normal operation.

Depletion Layer: The region at the junction of p and n type semiconductors. Because of diffusion, free electrons and holes recombine at the junction. This creates pairs of oppositely charged ions on each side of the junction. This region is depleted of free electrons and holes.

Depletion Type MOSFET: A FET with an insulated gate that relies on the action of a depletion layer to control the drain current.

Enhancement Type MOSFET: A FET with an insulated gate that relies on an inversion layer to control its conductivity.

Gate: The terminal of a field-effect transistor that controls drain current.

Gate-Source Cut Off Voltage: The voltage between the gate and the source that reduces the drain current of a depletion type device to approximately zero.

Ohmic Region: The part of the drain curves that starts at the origin and end at the proportional pinch-off voltage.

Reverse Bias: Applying an external voltage across a diode to aid the barrier potential. The result is almost zero current.

Self Bias: The bias you get with a JFET because of the voltage produced across the source resistor.

Substrate: A region in the depletion mode MOSFET located opposite from the gate, forming a channel through which electrons flowing from source to drain must pass.

Unipolar Transistor: A transistor where either free electrons or hole is necessary for normal operation.

6.8 REFERENCES

1. 'Electronic Devices and Circuits', T.F. Bogart, Universal Book Stall, New Delhi
2. 'Electronic Principles', A.P. Malvino, TMC Publishing Company Ltd., New Delhi.

6.9 SUGGESTED READINGS

1. 'Basic Electronics and Linear Circuits', N.N. Bhargava, D.C. Kulshreshtha and S.C. Gupta, TMC Publishing Company Ltd., New Delhi.

2. 'Integrated Electronics: Analog and Digital Circuits and Systems', J. Millman and C.C. Halkias, TMC Publishing Company Ltd., New Delhi.

3. 'Principles of Electronics', V.K. Mehta & Rohit Mehta, S.Chand & Company, New Delhi.

6.10 TERMINAL QUESTIONS

1. The gate of a JFET is ……. Biased.

2. In a n-channel JFET, the charge carriers are …… .

3. When drain voltage equals the pinch-off voltage, the drain current ……. with the increase in drain voltage.

4. The input control parameter of a JFET is …… .

5. A MOSFET is sometime called …….. FET.

6. Explain the construction and working of a JFET.

7. Explain the construction and working of MOSFET.

8. Describe the biasing circuits used in JFET and MOSFET

6.11 ANSWERS

1. Reverse 2. Electrons 3. Remains constant 4. Gate voltage 5. Insulated gate

6. See the section 6.3.1 and 6.3.3.

7. See the section 6.5.1 and 6.5.5.

8. See the section 6.4 and 6.5.9.

UNIT 7 AMPLIFIERS

Structure

7.1 INTRODUCTION

An amplifier is a circuit that is used for amplifying a signal. Most of the times a signal may have the desired waveform and frequency but its amplitude is so small that it is not of any practical use. As we know that generally the input of electronic circuit or devices need definite levels of amplitude of signal for proper operation hence it needs to be strengthened or amplified. The input signal to an amplifier will be a current or voltage and the output will be an amplified version of the input signal. An amplifier circuit which is purely based on a transistor or transistors is called a transistor amplifier.

As already discussed, there are three transistor configurations that are used commonly i.e. common base (CB), common collector (CC) and common emitter (CE).

In common base configuration has a current gain less than unity however voltage gain is greater than unity. Common collector configuration (emitter follower) has a current gain greater than unity. However common emitter configuration has current as well as voltage gain and therefore power gain also greater than unity. Because of this reason we generally use common emitter configuration in amplifier. In this chapter we will discuss single stage transistor amplifier in detail.

A good transistor amplifier must have the following parameters; high input impedance, high band width, high gain, high slew rate, high linearity, high efficiency, high stability etc.

When only one transistor with associated circuit is used for increasing the strength of a weak signal, the circuit is known as single transistor amplifier; however, the gain of single stage amplifier is not sufficient practically we use a number of stage of amplification where output of first amplifier is used as input to the next amplifier. It is known as cascading of amplifier or multistage amplifier. In the next section we will discuss classification of amplifiers according to different criterion.

7.2 OBJECTIVE

After study of unit the student will be able

i) To understand the classification of amplifiers on different basis.
ii) The concept of single stage and multistage transistor amplifier.
iii) To explain different components and working of single stage common emitter amplifier and its mathematical analysis.
iv) To understand the circuit, working and frequency response of R-C coupled amplifier.
v) To understand the circuit, working and frequency response of transformer coupled amplifier.
vi) To understand different type of distortion in amplifier.

7.3 CLASSIFICATION OF AMPLIFIERS ON DIFFERENT BASIS

As we know an Amplifier circuit is one which strengthens the signal. Let us now try to understand the classification of amplifiers according to many considerations.

Based on its output

Depending upon the parameter that is amplified at the output, there are voltage and power amplifiers.

- **Current amplifiers** − The amplifier circuit that increases the current level of the input signal, is called as current amplifier.
- **Voltage amplifiers** − The amplifier circuit that increases the voltage level of the input signal, is called as Voltage amplifier.
- **Power amplifiers** − The amplifier circuit that increases the power level of the input signal, is called as Power amplifier.

Based on the Transistor Configuration

Depending upon the type of transistor configuration, there are CE CB and CC amplifiers.

- **CE amplifier** − The amplifier circuit that is formed using a CE configured transistor combination is called as CE amplifier.
- **CB amplifier** − The amplifier circuit that is formed using a CB configured transistor combination is called as CB amplifier.
- **CC amplifier** − The amplifier circuit that is formed using a CC configured transistor combination is called as CC amplifier.

Based on number of stages

Depending upon the number of stages of Amplification, there are Single-stage amplifiers and Multi-stage amplifiers.

- **Single-stage Amplifiers** − This has only one transistor circuit, which is a single stage amplification.
- **Multi-stage or cascade Amplifiers** − This has multiple transistor circuit, which provides multi-stage amplification.

Based on the input signals

Depending upon the magnitude of the input signal applied, they can be categorized as Small signal and large signal amplifiers.

- **Small signal Amplifiers** − When the input signal is so weak so as to produce small fluctuations in the collector current compared to its quiescent value, the amplifier is known as Small signal amplifier.
- **Large signal amplifiers** − When the fluctuations in collector current are large i.e. beyond the linear portion of the characteristics, the amplifier is known as large signal amplifier.

Based on the frequency range

Depending upon the frequency range of the signals being used, there are audio and radio amplifiers.

- **Audio Amplifiers** − The amplifier circuit that amplifies the signals that lie in the audio frequency range i.e. from 20Hz to 20 KHz frequency range, is called as audio amplifier.
- **Radio Amplifiers** − The amplifier circuit that amplifies the signals that lie in a very high frequency range (radio frequency range), is called as radio amplifier.
- **Video Amplifiers** − The amplifier circuit that amplifies the signals that lie in a very high frequency range (radio frequency range), is called as video amplifier.

Based on Biasing Conditions

Depending upon their mode of operation, there are class A, class B, class C and class AB amplifiers.

- **Class A amplifier** − The biasing conditions in class A power amplifier is such that the collector current flows for the entire AC signal applied.
- **Class B amplifier** − The biasing conditions in class B power amplifier are such that the collector current flows for half-cycle of input AC signal applied.
- **Class C amplifier** − The biasing conditions in class C power amplifier are such that the collector current flows for less than half cycle of input AC signal applied.
- **Class AB amplifier** − The class AB power amplifier is one which is created by combining both class A and class B in order to have all the advantages of both the classes and to minimize the problems they have.

Based on the Coupling method

Depending upon the method of coupling one stage to the other, there are RC coupled, Transformer coupled and direct coupled amplifier.

- **RC Coupled amplifier** − A Multi-stage amplifier circuit that is coupled to the next stage using resistor and capacitor (RC) combination can be called as a RC coupled amplifier.
- **Transformer Coupled amplifier** − A Multi-stage amplifier circuit that is coupled to the next stage, with the help of a transformer, can be called as a Transformer coupled amplifier.
- **LC Coupled amplifier**− A Multi-stage amplifier circuit that is coupled to the next stage using inductance and capacitor (RC) combination can be called as a LC coupled amplifier.

- **Direct Coupled amplifier** − A Multi-stage amplifier circuit that is coupled to the next stage directly, can be called as a direct coupled amplifier.

7.4 SINGLE STAGE COMMON EMITTER TRANSISTER AMPLIFIERS

Common emitter amplifier is the most commonly employed circuit as it gives the highest voltage gain, Fig 7.1 shows the circuit if common emitter NPN transistor in which emitter is common to both the input (emitter-base) and output (collector-emitter) circuits and is grounded. As necessary the emitter base junction is forward biased and collector is reverse biased.

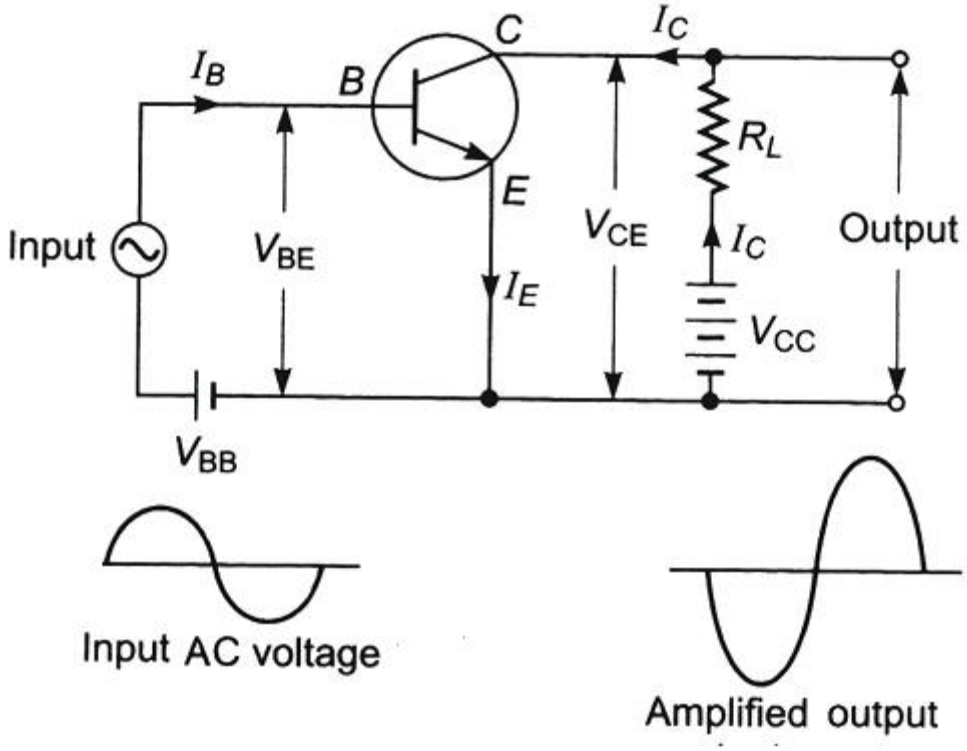


Figure 7.1: Single stage common emitter transistor amplifier.

Here input base current controls collector current and a very small change in base current corresponds to a lager change in collector current hence the current gain is also high. During the positive half cycle of input signal forward bias of emitter base junction is increased which causes an increase in base current hence also an increase in collector current because of its potential drop across collector decreases and collector voltage varies in opposite phase to input voltage through negative half cycle. Similarly, when input signal goes through negative cycle forward bias of emitter base junction is decreased which causes a decrease in base current hence also a decrease in collector current. It causes an increase in potential drop across collector and collector voltage

varies through negative half cycle opposite to input voltage Thus in common emitter amplifier output voltage is 180^0 out of phase with input signal. This is called phase reversal. Since the input section(base-emitter) is forward biased hence it has a low resistance (~ 100 Ω) and the output section (collector-emitter) is reverse biased (~ 10 KΩ), a large a.c. voltage will be developed across the load corresponding to a small change in the emitter voltage hence the input signal gets amplified.

Current gain: The ratio in change in output current to input current is defined as the current gain. In this case i_c is output current and i_b is input current. Hence current gain denoted by β is given by

$$\beta = \frac{\delta i_c}{\delta i_b}$$

Value of β >>1.hence the current gets amplified in common emitter configuration.

Voltage gain: Similarly, voltage gain of common emitter transistor amplifier is given by

$$A_v = \frac{\delta V_{out}}{\delta V_{in}}$$

Now change in input current (base current)$\delta i_b = \frac{\delta V_{in}}{R_{in}}$ wher ri is input resistance

And change in output collector current

$$\delta i_c = \beta\ \delta i_b = \beta \frac{\delta V_{in}}{R_{in}}$$

The change in output voltage (collect voltage) appearing across the load RL is given by

$$\delta V_{out} = \ \delta i_c \ \times \ R_L = \left(\beta \frac{\delta V_{in}}{R_{in}}\ \right) \times \ R_L$$

$$Hence\ A_v = \frac{\delta V_{out}}{\delta V_{in}} = \ \frac{\left(\beta \frac{\delta V_{in}}{R_{in}}\ \right) \times \ R_L}{\delta V_{in}} = \beta \frac{R_L}{R_{in}}$$

Which is also higher than that obtained in common base amplifier.

Power gain: Power gain of common emitter transistor amplifier is given by

Power gain = Current gain × Voltage gain

$$= \beta \times \beta \frac{R_L}{R_{in}} = \beta^2 \frac{R_L}{R_{in}}$$

Important: If we use PNP transistor in place of NPN transistor the polarities of batteries will be reversed as in this case base is made of N type semiconductor unlike P type in NPN transistor also emitter and collector are also changed but in order to make base forward biased and collector reverse biases hence polarities of battery need to be changed. The circuit connection will be as shown in figure.

Practical circuit of common emitter transistor amplifier: To obtain faithful amplification a transistor must be used with proper biasing and stabilizing circuit such that emitter-base junction always remains forward biased and collector base junction remains reverse biased. In practical

single stage common emitter amplifier, we use only one battery for the in the associated circuitry using suitable bias register for input and output (Fig. 7.2). Different circuit component and their functions are explained as follows.

i) **Biasing and stabilization circuit: P**roper biasing and stabilization is ensured using biasing resistances R_1 and R_2 and stabilizing circuit R_E and bypass capacitor C_E and battery source V_{CC} so that emitter base junction remains forward bias and collector base junction is reverse biased by establishing a proper operating point. The capacitor C_E connected across the emitter resistor R_E is of large value (~ 100μF) so that its reactance ωC_E is small for a.c. component thus bypasses the alternating component of emitter current and potential difference across Re is due to the d.c. component of emitter current. Otherwise amplified a.c. signal flowing through RE will cause an alternating voltage drop across emitter ad operating point will be no longer stable.

ii) **Input capacitor** Cin**:** the signal voltage is coupled to the base of transistor by input capacitor C_1 which is of the capacity of about 10 μF. This capacitor allows only ac signal to flow but isolates the signal source from R_2. If the input capacitor is not used the signal source resistance will come across the bias resistance R_2. and thus bias voltage of the base of the transistor will be changed.

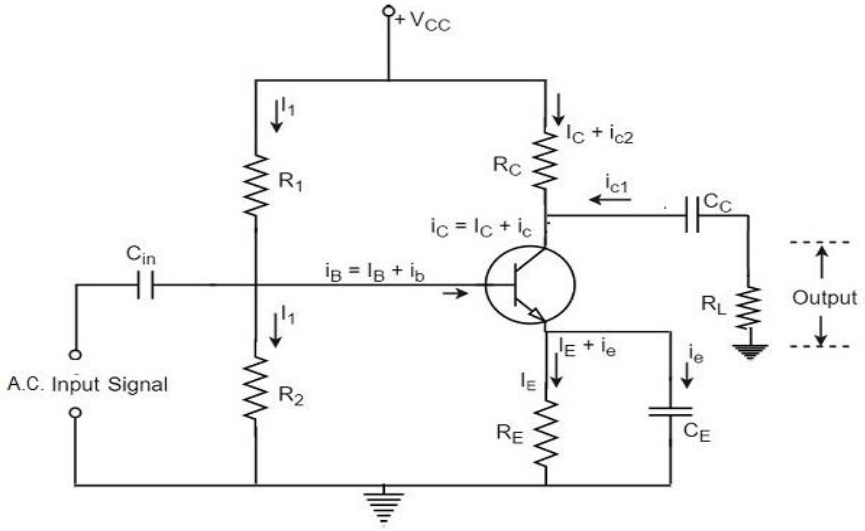


Figure 7.2: Practical circuit of common emitter transistor amplifier.

iii) **The Emitter bypass capacitor:** another capacitor C_E connected in parallel to emitter resistance R_E is of the capacity of about (~ 100 μF), called the emitter bypass capacitor. This capacitor is used to provide a low reactance path to the amplified ac signal. In the absence of emitter bypass capacitor, amplified ac signal flowing through Re will cause a voltage drop across it which will feedback the input side and reduce the output voltage.

iv) **Coupling or blocking capacitor C_C:** This capacitor is used to couple one stage of amplifier to the next stage and is of capacity of about 10 μF. It prevents the D.C. component of first stage from reaching the input of the second stage. In its absence R_C

will come in parallel with the biasing register R_L of the next stage and change the bias condition of next stage.

Various currents flowing in different branches of the amplifier circuit are indicated in fig. 7.2 With the application of input signal

Base current i_B=D.C. base current due to biasing current+A.C. base current due to A.C. input Signal

$=I_B+i_b$

Similarly, Collector current i_C=D.C. collector current due to biasing current + A.C. collector current due to A.C. input Signal

$=I_C+i_C$

And Emitter current i_E=D.C. Emitter current due to biasing current+ A.C. Emitter current due to A.C. input Signal

$=I_E+i_E$

Ex. 7.1 In a single stage common emitter transistor amplifier a change of 0.01V in the input signal causes a change of 5 μA in base current and 5 mA in collector current. Then calculate
i) Input impedance ii) effective a.c. loads iii) current gain, voltage gain and power gain.
Given collector load Rc =5kΩ and R_L=10 kΩ.

Solution. i) Input impedance = $R_{in} = \frac{\delta V_{BE}}{\delta i_b} = \frac{.01}{5\ \mu A} = 2000\ \Omega = 2\ \text{K}\Omega$

ii) Effective a.c. loads= $R_{Ac} = R_C \,||\, R_L = \frac{R_C \times R_L}{R_C + R_L} = \frac{5\times 10}{5+10}\text{K}\Omega = \frac{50}{15}\text{K}\Omega = 3.33\ \text{K}\Omega$

iii) Current gain $\beta = \frac{\delta i_c}{\delta i_b} = \frac{5\ mA}{5\ \mu A} = 1000$

Voltage gain

$$A_v = \beta \frac{R_{AC}}{R_{in}} = 1000\ \frac{3.33\ \text{K}\Omega}{2\text{K}\Omega} = 1665$$

Power Gain= voltage gain × Current gain = 1665 × 1000 = 1665000

Ex. 7.2 Find the output voltage of single stage common emitter amplifier if the input voltage is 1 mV. Given collector load R_c = 5 kΩ and R_L=10 kΩ , $\beta = 100$ and R_{in} = 2 KΩ.

Solution.

Effective a.c. load= $R_{Ac} = R_C \,||\, R_L = \frac{R_C \times R_L}{R_C + R_L} = \frac{5\times 10}{5+10}\text{K}\Omega = \frac{50}{15}\text{K}\Omega = 3.33\ \text{K}\Omega$

Voltage gain

$$A_v = \beta \frac{R_{AC}}{R_{in}} = 100\ \frac{3.33\ \text{K}\Omega}{2\text{K}\Omega} = 166.5$$

Also Voltage gain = $\frac{Output\ voltage}{Input\ voltage}$

hence Output voltage = Voltage gain × Input voltage.
= 166.5 × 1 mV= 166.5 mV.

7.5 MULTISTAGE AMPLIFIER: CASCADING OF AMPLIFIERS

In general, the current or voltage gain of a single stage amplifier is inadequate for most of the electronic applications, therefore more than one stage of amplification is used to achieve desired amplifications. For that the output of each amplifier stage is coupled with the help of some coupling method to the input of the next stage this is called cascading or multistage amplifier (Fig 7.3).

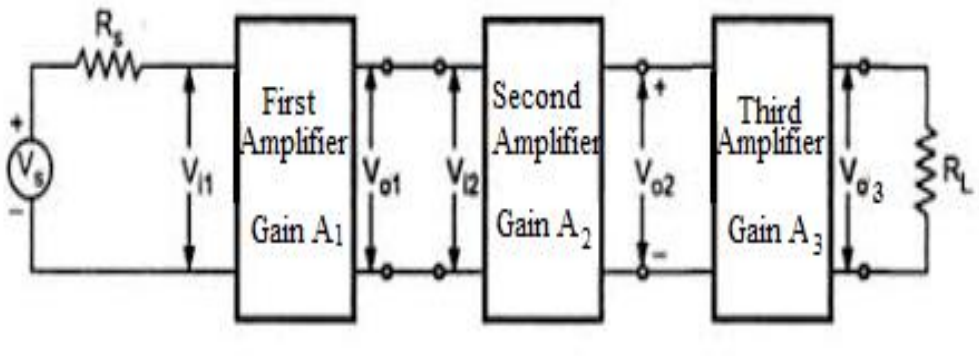


Figure 7.3: Block diagram of multistage amplifier.

Terminology frequently used in case of multistage amplifiers.

Gain: The overall gain of a multistage amplifier is equal to the product of the gains of the individual stages. If A_1, A_2, A_3 etc. are the voltage gain of individual amplifier used in multistage amplifier, then overall gain A of multistage amplifier is given by

$$A = A_1 \times A_2 \times A_3 \times \dots$$

This can be proved easily. As shown in Fig 7.3 overall gain A of multistage amplifier is

$$A = \frac{V_{o3}}{V_{i1}} = \frac{V_{o1}}{V_{i1}} \times \frac{V_{o2}}{V_{o1}} \times \frac{V_{o3}}{V_{o2}} = A_1 \times A_2 \times A_3$$

As $A_1 = \frac{V_{o1}}{V_{i1}}$ $A_2 = \frac{V_{o2}}{V_{o1}}$ $A_3 = \frac{V_{o3}}{V_{o2}}$

However, because of loading effect of the next stages overall gain is slightly less than the product of gain.

Gain in logarithmic scale: the gain of amplifier is sometimes expressed as logarithmic scale known as Bel or decibel which is of great importance particularly for audio amplifier as human ears response to sound is also logarithmic.

Bel power gain: the common logarithm (log to the collector 10) of power gain P_{out}/P_{in} is known as Bel power gain.

$$\text{i.e. Power gain} = \log_{10} \frac{P_{out}}{P_{in}}$$

In general decibel gain is used as bel is too large to most of the work.it is equivalent to one tenth of a bel.
So in decibel

$$\text{Power gain} = 10 \log_{10} \frac{P_{out}}{P_{in}} \text{ db} \qquad \text{(as 1 bel = 10 db)}$$

$$\text{voltage gain} = 20 \log_{10} \frac{P_{out}}{P_{in}} \qquad (\text{as } P \alpha V^2)$$

$$\text{and current gain} = 20 \log_{10} \frac{P_{out}}{P_{in}} \qquad (\text{as } P \alpha I^2)$$

However, if the gain is expressed in db the overall of a multistage amplifier is the sum of gains of individual stages in db. However absolute gain will be obtained by multiplying the gains of individual gain.

Frequency response curve: In general, the gain of an amplifier increase with frequency reaches to maximum for a particular frequency (resonant frequency) or frequency range (midband) and then becomes constant, therefor particular frequency as all the reactive component in the circuit have negligible effect and then decreases at higher frequency.
The graph drawn between voltage gain and signal frequency is called frequency response curve (Fig.7.4).

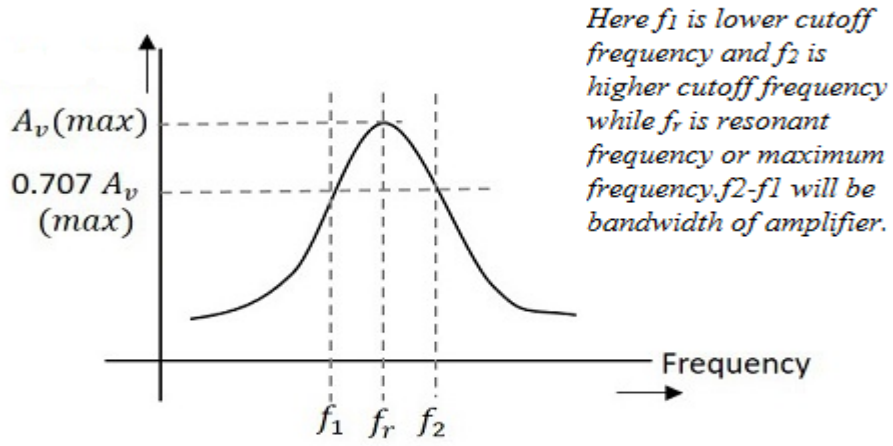


Figure 7.4: Frequency response curve and bandwidth.

Bandwidth: The frequency ranges over which the voltage gain is equal to or greater than $1/\sqrt{2}$ or 70.7 % of the maximum voltage gain or the power gain is ½ or 50% of the maximum power gain as shown in figure f_2-f_1 is bandwidth f_1 is called lower cut off frequency while f_2 is higher cut off frequency.f_1, f2 are also called half power points as power gain drops to ½ of maximum power gain. Hence bandwidth can also be defined as difference between two half power points.
Ex. 7.3 In a three stage common emitter amplifier the gain of individual amplifiers is 100,50,20 respectivly. find the total voltage gain. Also if the gain is 10db,20 db and 25 db then what will be overall gain.

Solution. Overall gain of multistage amplifier is given by $A= A_1 \times A_2 \times A_3$ Where A_1, A_2, A_3 are the voltage gain of individual amplifier used in multistage amplifier,

Hence here

Overall gain of multistage amplifier is given by $A= 100\times 50 \times 20=100000$

If the gain is represented in db, Overall gain of multistage amplifier is given by $A= A_1 + A_2 + A_3$

Hence A=10+20+25=55 db

This shows calculation in db scale is comparatively easier.

Ex. 7.4 In an amplifier the maximum gain is 2000 is at 10 Khz. What will be the value of its bandwidth if lower and upper cut off frequencies at 5 Khz and 15 Khz respectively. Also find gain at lower and upper cut off frequency.

Solution. Bandwidth of amplifier=upper cut off frequency-lower cutoff frequency

= 15 Khz-5Khz=10 Khz.

Gain at Lower cutoff frequency=.707 of maximum voltage gain=.707 × 2000=1414

Similarly gain at upper cutoff frequency is also =.707 of maximum voltage gain

=.707 × 2000 =1414

7.6 R-C COUPLED AMPLIFIER

7.6.1 Circuit Diagram

When the coupling from one stage to next is achieved by a coupling capacitor followed by a connection to a shunt resistor, such amplifiers are known as resistance-capacitance coupled amplifier or simply RC coupled amplifier.

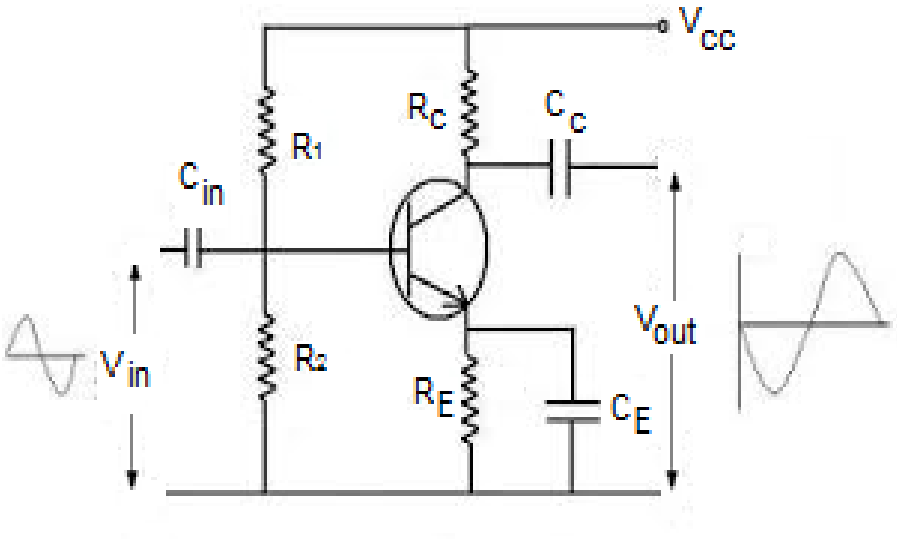


Figure 7.5: Single stage R-C coupled amplifier.

As you can see in the Fig 7.5, a coupling capacitor C_C is used to connect the output of first stage to the collector i.e. input of the second stage and this continues when more stages are connected. The resistances R_1, R_2 and R_E form the biasing and stabilization network. The emitter bypass capacitor offers a low resistance path to the signal. Without this capacitor the voltage gain of each stage would be lost. The coupling capacitor C_C pass a.c. signal but blocks d.c. This prevents d.c. interference between various stages and the shifting of operating point.

7.6.2 Operation

When a.c. signal is applied to the collector of the first transistor, it is amplified and appears across its collector load R_C. .Now the amplified signal developed across R_C is given to the collector of the next transistor through a coupling capacitor C_C .The second stage again amplifies this signal and the more amplified signal appears across the second stage collector resistance. In this way the cascaded stages amplify the signal and the overall gain is considerably increased.

7.6.3 Frequency Response

The figure below shows the frequency response of a typical RC coupled amplifier.

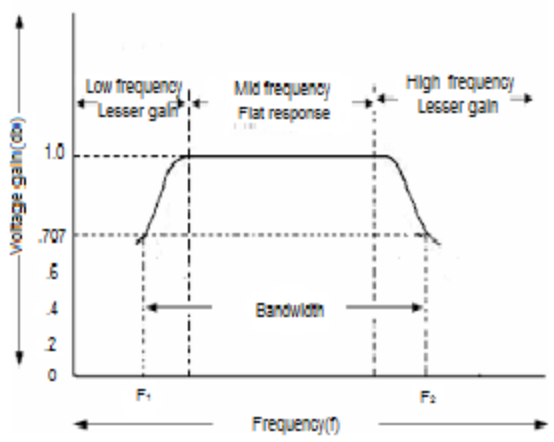


Figure 7.6: Frequency response curve of RC coupled amplifier.

You can notice from the above Fig. 7.6 that the voltage gain drops off at low (< 50 Hz) and high (> 20 KHz) frequencies. However, it is uniform over the mid-frequency range i.e. 50 Hz to 20 KHz. This behaviour of the amplifier can be explained as follows:

(i)At Low Frequencies

At low frequencies i.e. below 50 Hz, the reactance of coupling capacitor C_C is quite high and hence very small part of the signal will pass from one stage to the next stage. Again C_E cannot shunt the

emitter resistance R_E effectively because of its large reactance at low frequencies. These two factors cause the dropping of voltage gain at low frequencies.

(ii) At High Frequencies

At high frequencies i.e. above 20 KHz, the reactance of coupling capacitor C_C is quite small and hence it behaves as a short circuit. This increases the loading effect of next stage and results in decreased voltage gain. Again at high frequencies, capacitive reactance of collector-emitter junction is low which in result increases the collector current. This causes decrease in current amplification factor β. These two factors cause the dropping of voltage gain at high frequencies.

(iii) At Mid Frequencies

At mid frequencies i.e. between 50 Hz to 20 KHz , the voltage gain of the amplifier is constant. The effect of coupling capacitor in this frequency range is such that the voltage gain remains uniform. As the frequency increases in this range, reactance of C_C decreases which in result increases the gain. However, at the same time lower reactance means higher loading effect of first stage to the next one and hence gains decreases. Thus, these two factors almost cancel each other, resulting in a uniform gain at this mid frequency.

However, the total gain is less than the product of the gains of individual stages. It is because, when a second stage follows the first stage, the effective load resistance of first stage is reduced due to the shunting effect of the input resistance of second stage. This reduces the gain of the stage which is loaded by the next stage.

To explain it better, let us take an example of 2-stage amplifier (Fig.7.7). The gain of first stage will be reduced due to loading effect of the second stage. But the gain of the second stage which has no loading effect due to subsequent stage, remains unchanged.

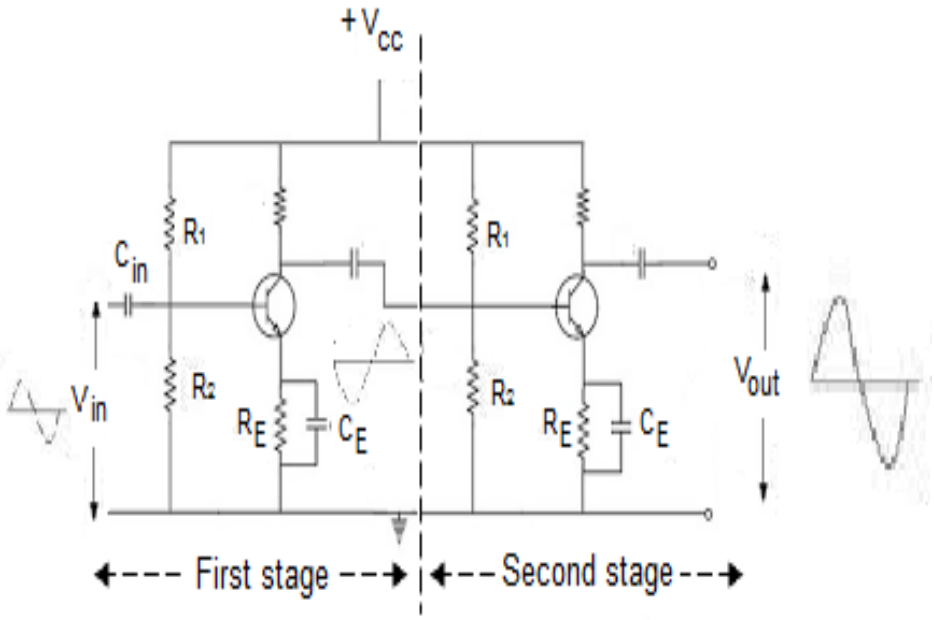


Figure 7.7: Two stage R-C coupled amplifier.

The overall gain is equal to the product of the gains of two stages.

7.6.4 Advantages, Disadvantages, Application of RC Coupled Amplifier

Advantages

i. It has a great frequency response. The gain is uniform over the audio frequency range which is important for speech, music etc.
ii. It employs only resistors and capacitors which are cheap, hence, it has low cost.
iii. The circuit is very compact due to the small size and light weight of resistors and capacitors.

Disadvantages

i. The RC coupled amplifiers have low voltage and power gain. Because, the low resistances presented by the input of each stage to the subsequent stage decreases the effective load resistance and hence decreases the gain.
ii. These amplifiers become noisy with age, particularly due to moist.
iii. Impedance matching is poor because the output impedance of RC coupled amplifier is several hundred ohms whereas the input impedance of a speaker is only few ohms.

Application

i. Used as voltage amplifiers for example in the initial stages of public address system.
ii. If other type of coupling such as transformer coupling is used in the initial stages, this results in frequency distortion which may be amplified in the next stage.
iii. But, due to its poor impedance matching, it is rarely used in the final stages.

7.7 TRANSFORMER COUPLED AMPLIFIER - OPERATION AND FREQUENCY RESPONSE

As we have already discussed in our previous section on RC coupled transistor amplifier the voltage and power gain are low since, the effective load resistance of each stage is decreased due to the low resistance presented by the input of each stage to the next stage.

However, the voltage and power gain could also be increased. If the effective load resistance of each stage could be increased, this can be achieved by transformer coupling. By using the impedance matching properties of transformer, the low resistance of one stage or load can be reflected as a high load resistance to the previous stage.

Transformer coupling is normally used when the load is small. It is mostly used for power amplification.

7.7.1 Circuit Diagram

The Fig. 7.8 below shows the circuit of a two stage transformer coupled amplifier.

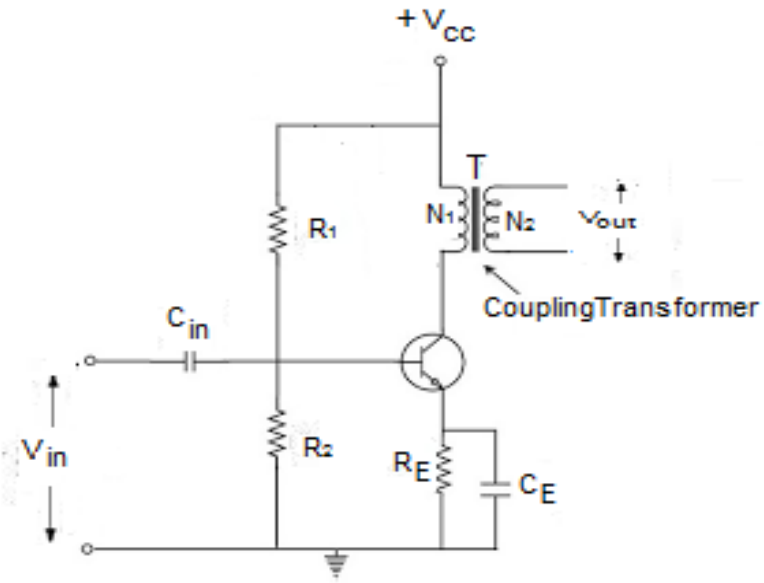


Figure 7.8: Single stage transformer coupled amplifier.

As you can see from the above Fig. 7.8 a coupling transformer is used to feed the output of one stage to the input of the next stage.

The primary P of this transformer is made the collector load and its secondary S supplies input to the next stage.

7.6.2 Operation

When an a.c. signal is applied to the collector of first transistor, it appears in the amplified form across the primary P of the coupling transformer.

Now the voltage developed across P is transferred to the input of the next stage by the transformer secondary S.

The second stage now performs the amplification in an exactly same manner.

7.6.3 Frequency Response

The frequency response of a transformer coupled amplifier is shown in the figure below.

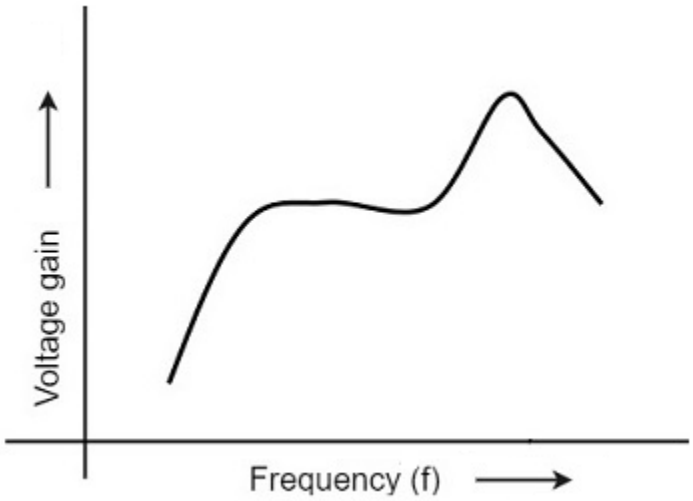


Figure 7.9: Frequency response curve of RC coupled amplifier.

It is clear from the above Fig. 7.9 that its frequency response is poor than the RC coupled amplifier.

The gain is constant only over a small range of frequency.

Since, the output voltage is equal to the collector current multiplied by reactance of primary, hence at low frequencies, as the reactance of primary begins to fall, the output voltage also decreases and hence the gain.

At high frequencies, the capacitance between turns of windings acts as a bypass condenser to reduce the output voltage and hence the gain.

Due to these two factors, there will be disproportionate amplification of frequencies in a complete signal such as music, speech etc.

Hence, transformer coupled amplifier introduces frequency distortion.

7.6.4 Advantages, Disadvantages, Application of Transformer Coupled Transistor Amplifier

Advantages

1. There is no loss of signal power in the collector or collector resistors.
2. An excellent impedance matching can be achieved in a transformer coupled amplifier.
3. Due to excellent impedance matching, transformer coupling provides higher gain. A properly designed single stage transformer coupling can provide the gain of two stages of RC coupling.

Disadvantages

1. It has a poor frequency response.
2. The coupling transformers are bulky and expensive at audio frequencies.
3. Frequency distortion is higher i.e. low frequency signals are less amplified as compared to the high frequency signals.
4. Transformer coupling introduces hum in the output.

Application

1. It is mostly used for impedance matching.
2. Normally the last stage of a multi stage amplifier is the power stage. Here it is required to transfer maximum power to the output device for example a loud speaker.
3. Usually the impedance of an output device is few ohms whereas the output impedance of the transistor is several hundred ohms.
4. In order to match the impedance, a step down transformer of proper turn ratio is used. The impedance of secondary of the transformer is made equal to the load impedance and primary impedance equal to the output impedance of the transistor.
5. Thus the load on the primary side is comparable to the output impedance of the transistor and this result in maximum power transfer from the transistor to the primary of transformer.

7.8 DISTORTION IN AMPLIFIERS

For an ideal amplifier the output waveform should be an exact replica of input waveform in amplified form however in practical amplifier whenever a signal is applied to the input of transistor, some change occurs in the output. Any change in the output wave shape from the input is called the distortion. It is not desirable as such change may cause the change in information carried out by signal.

Amplifier distortion can be classified into two broad categories named linear and nonlinear distortion, depending on whether the transistors acting on linear or non-linear region of its characteristics.

Nonlinear distortion: As transistor is nonlinear device, such type of distortion occurs when transistor works in nonlinear region of its output characterstics.it is generally occurs in case of large signal inputs. It is again classified as amplitude distortion, frequency distortion, Phase distortion and intermodulation distortion,

Amplitude distortion: When the signal is viewed in time domain, when there is additional frequency component are produced which are not present in the input, it is called amplitude

distortion these additional frequencies are generally harmonics of the input frequency. it is caused because of nonlinearity of trans conductance curve.

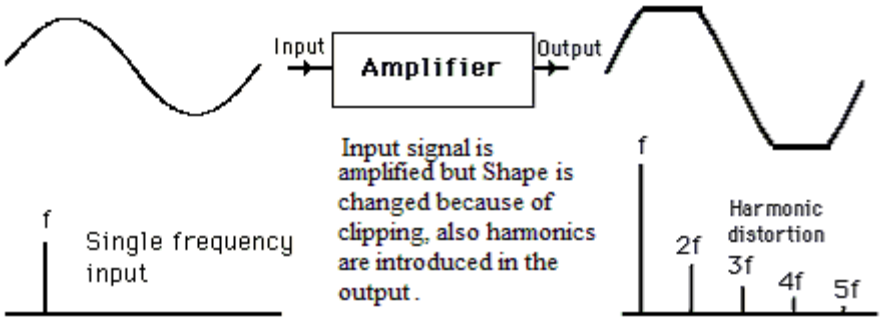


Figure 7.10: Harmonic distortion in amplifier.

Frequency distortion: It is caused in amplifiers when signal components of different frequencies are amplified differently. Frequency distortion is caused primarily either by the reactive element associated with the circuit or the interelectrode capacitance of the amplifying decvice. Generally, the gain of amplifier remains constant only in mid frequency region but falls off for low and high frequency therefore the quality of audio or music is affected by such signals. Such distortion occurs even with small signals.

Phase or delay distortion: When signals of different frequencies suffer unequal phase shift it is called phase or delay distortion It is caused by the electronic component used in the amplifier circuit. Such type of distortion is insignificant to audio amplifier as human ear is not sensitive to phase shift however in video signal is distorted as different frequency component of the signal are delayed by different amount of time if phase shift varies with frequency and it caused distortion in the picture.

Intermodulation Distortion:

The amplified output also contains components which are sum and difference of input frequencies (Fig 7.11). Non-linearity in amplifier components causes mixing of frequency components to form distorted output, this is called intermodulation distortion. It is particularly troublesome in the reproduction of music because generation of mixed frequency component in the output is easily noticeable. Harmonic distortion may also be serious, but at least the As musical sound probably already had harmonics present as part of the harmonic content of the sound, Harmonic distortion is bearable to a greater degree than intermodulation distortion.

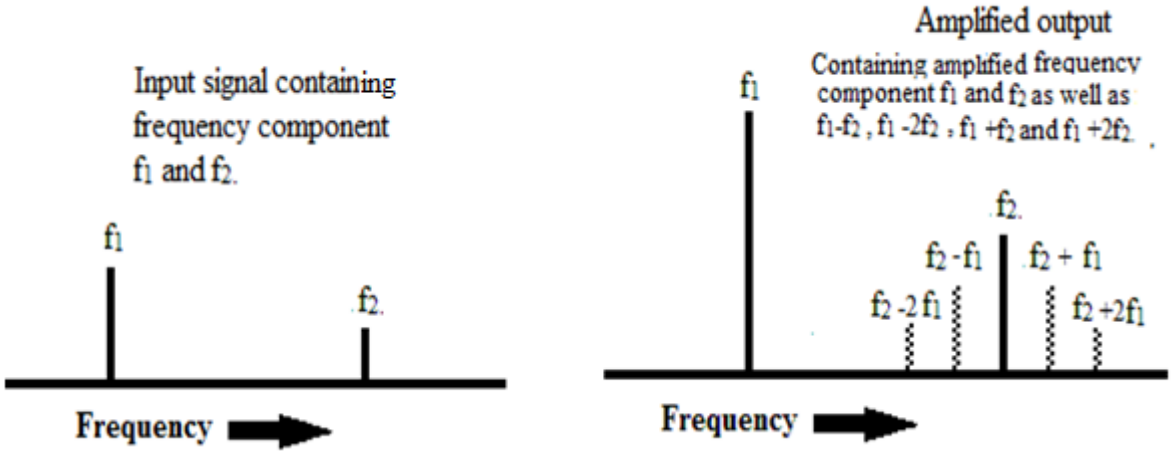


Figure 7.11: Intermodulation distortion in amplifier.

7.9 SUMMARY

An electronic circuit that is used for amplifying a signal is called an amplifier. A good transistor amplifier must have the following parameters; high input impedance, high band width, high gain, high slew rate, high linearity, high efficiency, high stability etc. When only one transistor with associated circuit is used for increasing the strength of a weak signal, the circuit is known as single transistor amplifier; however, the gain of single stage amplifier is not sufficient practically we use a number of stage of amplification where output of first amplifier is used as input to the next amplifier. It is known as cascading of amplifier or multistage amplifier. Based on the different considerations amplifier can be classified differently. When the coupling from one stage to next is achieved by a coupling capacitor followed by a connection to a shunt resistor, such amplifiers are known as resistance-capacitance coupled amplifier or simply RC coupled amplifier. Due to a smooth frequency response curve over a large range of middle frequency this amplifier is frequently used in audio amplifier. However, in RC coupled transistor amplifier the voltage and power gain are low since, the effective load resistance of each stage is decreased due to the low resistance presented by the input of each stage to the next stage. Voltage and power gain could also be increased, If the effective load resistance of each stage could be increased, this can be achieved by transformer coupling. By using the impedance matching properties of transformer, the low resistance of one stage or load can be reflected as a high load resistance to the previous stage.

As the frequency response curve of transformer is not as smooth as in the case of RC coupled amplifier, Transformer coupling is normally used when the load is small. It is mostly used for power amplification.

For an ideal amplifier the output waveform should be an exact replica of input waveform in amplified form. Any change in the output wave shape from the input is called the distortion.

Amplifier distortion can be classified into two broad categories named linear and nonlinear distortion, depending on whether the transistors acting on linear or non-linear region of its characteristics.

Nonlinear distortion is again classified as amplitude distortion, frequency distortion, and phase distortion

7.10 GLOSSARY

Amplifier: An amplifier is an electronic circuit capable of increasing either the current, voltage, power level, or all of these of an input electronic signal.

Cascading of amplifier: When output of first amplifier is given to the input of the next amplifier, in order to get the larger amplification. It is known as cascading of amplifier or multistage amplifier.

Distortion: Any change in the output wave shape from the input is called the distortion.

7.11 REFERENCES

1. Principles of electronics by V.K.Mehta

2. Basic Electronics by B.L. Thereja

3. Basic Electronics by D.C. Tayal

7.12 SUGGETSED READINGS

1. Electronic devices and circuits by Jacob Milliman and C.C. Halkais

2. The Feynman Lectures on Physics by Richard Feynman

3. Electronic devices and circuit theory by R.L. Boylestad and Louis N

7.13 TERMINAL QUESTIONS

7.13.1 Short Answer Type Questions

1. What do you understand by the term frequency response, decibel gain and cascading of amplifier?

2. Why is the function of coupling capacitor and bypass capacitor in amplifiers?

3. Why the gain of RC coupled amplifier becomes constant in mid frequency range?

4. Why coupling capacitors are not required in case of transformer coupled amplifier?

5. Why transformer coupled amplifier is generally used in the final stage of a multistage amplifier?

7.13.2 Long Answer Type Questions

1. Explain the working of single stage common emitter amplifier with the help of a circuit diagram? Derive the expression for current gain, voltage gain, power gain, input impedance and output impedance.

2. What do you understand by multistage transistor amplifiers? Discuss different methods of coupling amplifiers and their relative merits and demerits. Show that overall voltage gain in multistage amplifiers is equal to the product of the gains of individual stages.

3. Discuss the circuit diagram, working, frequency response curve, advantages and disadvantages of RC coupled amplifiers.

4. Explain transformer coupled amplifier in detail and compare it frequency response curve with RC coupled amplifiers.

5. What is distortion in amplifiers? Discuss different type of distortion their causes and how they can be limited in amplifiers.

7.13.3 Numerical Questions

1. In a CE transistor amplifier the collector and collector current is changed by 5μA and 0.5 mA respectively if the input signal is changed by 0.05 V. Calculate the value of current gain, input impedance, a.c. load, voltage gain and power gain of amplifier. Given that R_c=5 KΩ and R_L =5KΩ (Ans. current gain=100, input impedance=10 kΩ , a.c. load=2.5 kΩ, voltage gain=25 and power gain=2500 .)

2. The voltage gain and current gain of an amplifier is 75 and 100 when load resistance is 5 KΩ

(Ans. 3146 Ω)

3. Calculate the frequency of oscillation of Hartley oscillator having the parameters L_1=400μH

L_2=5000μH **,**M=500μH and C=100 pF. (Ans. 199.04 KHz)

4. In a crystal oscillator calculate the value of inductance offered by the crystal at oscillation frequency of 1 Mhz. Ans. (0.2533 H)

5. In a symmetrical collector coupled astable multivibrator having collector resistors of value 2.5 KΩ collector resistors are of value 200 KΩ and coupling capacitors of value .04 μF. calculate the value of time period and frequency of circuit oscillations also calculate the minimum value of transistor β. Ans. (.14 ms,7.2 kHz,12.5)

UNIT 8 FEEDBACK PRINCIPLE AND OSCILLATORS

Structure

8.1 INTRODUCTION

The process of injecting a fraction of output energy of some device back to the input is known as Feedback. It has been found that feedback is very useful in reducing noise and making the amplifier operation stable. An amplifier circuit simply increases the signal strength. But while amplifying, it also tends to amplify noise and distortion along with signal in its output, which is very undesirable.

These can be considerably reduced by using negative feedback done by injecting a fraction of output in phase opposition to the input signal.

Types of Feedbacks

Depending upon whether the feedback signal strengthen or weakens the input signal, there are two types of feedbacks used.

8.1.1 Positive Feedback

The feedback in which the feedback energy i.e., either voltage or current is in phase with the input signal and thus aids it, is called Positive feedback. Though the positive feedback increases the gain of the amplifier, it has the disadvantages such as increased distortion, Instability, sensitivity, increased noise etc.because of these disadvantages the positive feedback is seldom used in amplifiers. However, If the positive feedback is sufficiently large (barkhausan criterion), it leads to oscillations, therefore positive feedback is mostly used in oscillator circuits.

8.1.2 Negative Feedback

The feedback, in which the feedback voltage or current is out of phase with the input and thus opposes it, is called as negative feedback. In negative feedback, the amplifier introduces a phase shift of 180^o into the circuit while the feedback network is so designed that it produces no phase shift or zero phase shift. Thus the resultant feedback voltage V_f is 180^o out of phase with the input signal V_{in}. Though the negative feedback decreases the gain of the amplifier, it has the other advantages such as increased bandwidth, decreased distortion, stability, sensitivity, decreased noise etc.because of these advantages it is frequently used in amplifiers.

8.2 OBJECTIVES

After study of unit the student will be able

i) To understand the concept of feedback in amplifiers, their types and advantages and disadvantages.
ii) To classify different types of oscillators and their fundamentals.
iii) To explain different types of LC and RC oscillators, their construction and their working.
iv) To understand working and uses of crystal oscillator.
v) To describe relaxation oscillators.
vi) To explain construction and working of different multivibrators

8.3 PRINCIPLE OF FEEDBACK AMPLIFIER

A feedback amplifier generally consists of two parts, an amplifier and the feedback circuit. The feedback circuit usually consists of resistors. The principle of feedback amplifier can be understood from the following figure.

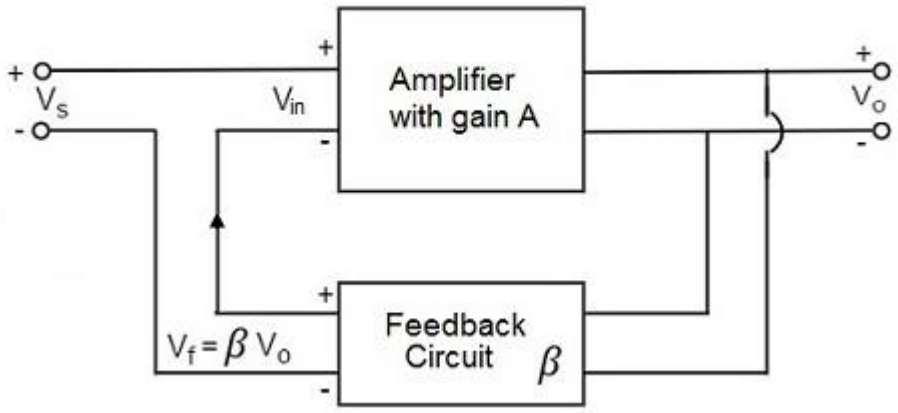


Figure 8.1: Principle of feedback amplifier.

Let the gain of the amplifier without feedback is A. as the gain of the amplifier is the ratio of output voltage V_o to the input voltage V_{in}.

$$A = \frac{V_o}{V_{in}}$$

Let β be the fraction of output given back to input known as feedback ratio or feedback fraction.

Hence $\beta = V_f/V_o$.

The feedback voltage $V_f = \beta V_o$ will be given back to input from the output V_o of the amplifier.

Now,

$$V_{in}=V_s \pm V_f = V_s \pm \beta V_o$$

Feedback voltage is added for positive feedback and subtracted for negative feedback, from the signal voltage V_s.

Let us consider the case of negative feedback. The output V_o must be equal to the input voltage (V_s - βV_o) multiplied by the gain A of the amplifier.

Hence,

$$(V_s - \beta V_o)A=V_o$$

Or

$$AV_s - A\beta V_o=V_o$$

Or

$$AV_s=V_o(1+A\beta)$$

Therefore,

$$V_o / V_s=A /(1+A\beta)$$

Let A_f be the overall gain (gain with the feedback) called closed loop gain of the amplifier. This is defined as the ratio of output voltage V_o to the applied signal voltage V_s, i.e.,

$$A_f = \text{Output voltage/ Input signal voltage} = V_o / V_s$$

So, from the above two equations, we can understand that,

The equation of gain of the feedback amplifier, with negative feedback is given by

$$A_f = A / (1+A\beta)$$

This is clear from above equation that gain of the amplifier is decreased by a factor $1/(1+A\beta)$ by the introduction of negative feedback. However, at the same time it also increases bandwidth, stability and reduce distortion and noise hence frequently used in amplifier.

The term β is called feedback ratio, βA is called feedback factor and $(1\pm A\beta)$ is called loop gain.

Similarly, the equation of gain of the feedback amplifier, with positive feedback is given by

$$A_f = A / (1-A\beta)$$

In general, $$A_f = A / (1\pm A\beta)$$

where + sign is taken for negative feedback and- sign for positive Feedback.

This is the standard equations to calculate the gain of feedback amplifiers.

8.4 ADVANTAGES OF NEGATIVE FEEDBACK

Though the gain of amplifier using negative feedback is reduced, there are many advantages of negative feedback such as

- Stability of gain is improved
- Bandwidth is Increased
- Distortion is reduced
- Noise is reduced
- Input impedance is increased
- Output impedance is increased

Let us discuss these advantages in detail

Improved stability of gain: gain of the amplifier, with negative feedback is given by

$$A_f = A / (1+A\beta)$$

If $A\beta >> 1$, then $A_f \backsimeq 1/\beta$. It means the gain becomes independent of A. Therefore, any change in supply voltage, transistor parameter etc. have almost no effect with negative feedback. Hence the gain is stabilized. If the feedback circuit does not contain any reactive component or β is independent of frequency the overall gain of negative feedback amplifier also become independent of frequency.

I**ncrease in bandwidth:** For an amplifier gain-bandwidth product remain same, hence

$$A\times \text{B.W.} = A_f \times (\text{B.W.})_f$$

As the gain of the amplifier in negative feedback is reduced by a factor $1/ (1+A\beta)$ hence bandwidth will be increased by factor $(1+A\beta)$ to make gain bandwidth product constant.

Reduction in distortion: With the introduction of negative feedback, any harmonic present in the output will also be feed back to the input and appear in amplified form in the output in opposite

phase with those due to the distortion of original applied signal thus the nonlinear distortion in output will be reduced.

Reduction of output Noise: Negative feedback reduces the level of noise generated within the amplifier only. If a noise voltage appears just at the input signal it also gets amplified by amplifier however when it is negatively feedback to the input total noise decreases considerably.

Increase in input impedance: Introduction of negative feedback increases the input impedance of the amplifier thus does not load the preceding stage or input voltage source. With simple calculation it can be shown that

$$R_{if}=R_i(1+A\beta)$$

Thus Input impedance increases by a factor 1+Aβ.

Decrease in output impedance Similarly the expression for output impedence with feedback can be proved as

$$R_{of} = R_o/(1+A\beta)$$

Thus Output impedance increases by factor (1+Aβ).

It is because of these advantages negative feedback is frequently employed in amplifiers.

Example 8.1. In a negative feedback amplifier, input voltage is 10mV feedback ration is 0.1 and the gain without feedback is 40. Calculate

i) Feedback factor ii) gain with feedback

iii) Feedback voltage

Solution: Given A=40 V_i=10mV β=0.1

i) Feedback factor βA= 0.1×40=4

ii) The equation of gain of the feedback amplifier, with negative feedback is given by

$$A_f = \frac{A}{1 + A\beta}$$

Hence gain with feedback $A_f = \frac{50}{1+40\times0.1} = 10$

iii) In order to calculate feedback voltage, we first have to calculate output voltage

Hence output voltage $V_o' = A_f \times V_i = 10 \times 10\ \text{mV} = 100\ \text{mV}$

Therefore feedback voltage = $\beta V_o' = 4 \times 100\ \text{mV} = 400\ \text{mV}$ or .4 volt.

Example 8.2. An amplifier has a bandwidth of 500 KHz and a voltage gains of 100 dB.What will be its bandwidth if 4 % negative feedback is introduced.

Solution: Given gain without feedback A=100 dB, Feedback ratio β = 4% $= \frac{4}{100} = 0.04$

Hence gain with negative feedback $A_f = \frac{A}{1+A\beta} = \frac{100}{1+100\times0.04} = 20\ \text{dB}$

Now band width without feedback= 500 KHz.

Since product of gain and bandwidth remains constant therefore

$$A \times BW = A_f \times BW' \text{ or } BW' = \frac{A\times BW}{A_f} = \frac{100\times500}{20} = 2500\ KHz.$$

Hence new bandwidth=2500 KHz

8.5 OSCILLATYORS FUNDAMENTALS

An oscillator is an electronic circuit used to produce oscillations. It is an essential part in electrical and electronic instruments and generates alternating voltages or currents of known frequency and amplitude.

The basic difference between an amplifier and oscillator is that in oscillator the frequency, waveform and magnitude of the a.c. power generated is controlled by the circuit itself and it does not require an input signal either to start or maintain oscillations. However, it is not creating any energy but merely acts as a energy converter and converts dc power from dc source we of supply into ac power of required frequency depending upon circuit components. It keeps generating an ac signal as long as the dc power source is connected. While we use an input signal which gets amplified using an amplifier (Fig.8.2).

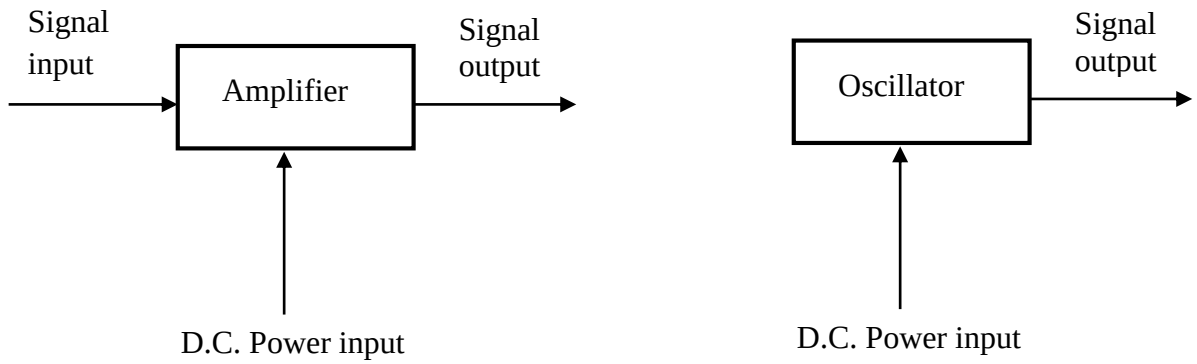


Figure 8.2: Basic difference between an amplifier and oscillator.

Essentials of oscillators: As shown in the following Fig. 8.3 essential components of oscillator are oscillatory circuit or frequency determining element, transistor amplifier and positive feedback circuit.

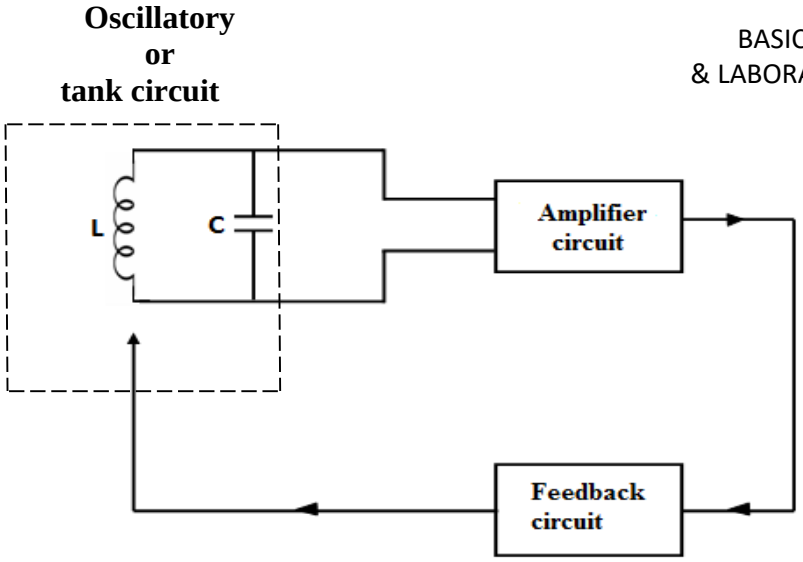


Figure 8.3: Essential components of oscillator.

1) **Oscillatory circuit or frequency determining element:** It can be either an inductance-capacitance network (L-C tank circuit) or resistance –capacitance (R-C network) or a quartz crystal depending upon the frequency and waveform desired. These components determine the frequency of oscillation and the circuit.
2) **Transistor amplifier:** It changes the d.c. power received from the battery into a.c. power and supplies it to the tank circuit to meet the losses of oscillation produced in tank circuit.
3) **Feedback circuit:** It supplies a part of the output power to the frequency determining element in correct phase to aid the oscillation i.e. it insures the positive feedback.

8.6 CLASSIFICATION OF OSCILLATOR

All oscillators having different names are used to perform similar action i.e. to produce continuous undammed oscillations, however depending on the energy supplying method to the tank circuit or, frequency determining network, shape of waveform and frequency range they can be broadly classified as under.

a) Depending upon the output waveform shape: According to shape of output waveform the oscillators are classified as

1. Sinusoidal oscillators: The waveform of these oscillators is similar to sine wave hence the name. These can be again classified as LC oscillator, RC oscillator.

LC oscillator: The feedback circuit comprises of inductance and capacitance.

1. Tuned base oscillator, tuned emitter oscillator, tuned collector oscillator.
2. Hartley oscillator
3. Colpitt's oscillator
4. Clapp's oscillator
5. Crystal oscillator

RC oscillator: The feedback circuit comprises of resistance and capacitance.

1. Wein bridge oscillator
2. Phase shift oscillator

2. Non-sinusoidal or relaxation oscillators: The output waveform is either square, rectangular, triangular or pulse type.i.e. non sinusoidal. Examples of relaxation oscillator are: 1. Multivibrator
2. Blocking oscillator

b) **Depending upon the method of producing oscillations:** According to this method, oscillators are subdivided as:

1. Feedback oscillators: In feedback oscillators, a part of output is feedback positively in the input circuit i.e. feedback voltage is in same phase to the input. Under proper condition (Barkhausen criterion) we get an output without any input so we can assume the oscillator as a positive feedback amplifier with infinite gain.

2.Negative resistance oscillators: If in a circuit, the current decrease with an increase in voltage the circuit is said to have a negative résistance. A resistance introduced in a circuit takes power from it so a negative resistance delivers power to the circuit. Oscillators based on this principle are called negative resistance oscillator or dynatron oscillator.

c) **Depending upon the frequency of produced oscillators:** According to this method, oscillators are again subdivided as:

1. Audio frequency or AF oscillators: These oscillators produce voltages of frequency from about 15 Hz to 20 KHz.

2. Radio frequency or RF oscillators: These oscillators produce voltages of frequency from about 10 KHz to 30 MHz.

3. Ultra high frequency or UHF oscillators: These oscillators produce voltages of frequency above 30 MHz also known as microwave oscillators.e.g. Reflex Klystron, Magnetron.

8.7 BARKHAUSEN CRITERION FOR OSCILLATION

Barkhausen criterion gives us the condition for sustained oscillations used in case of positive feedback; Let us consider a positive feedback network consisting a feedback amplifier, whose voltage gain without feedback is A and a feedback network with feedback fraction β (Fig.8.4). Let v_{in} be the input voltage and v_0 be the output voltage.

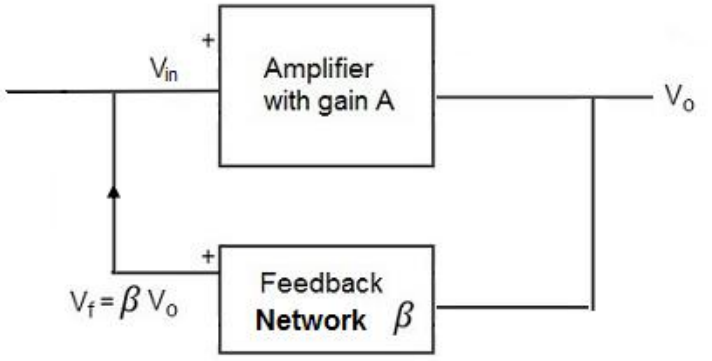


Figure 8.4: Barkhausen criterion for oscillation.

With positive feedback the effective input of the amplifier is $v_{in} + \beta v_0$ which when amplified A times by the amplifier give the output v_0 .Thus

$$(v_{in} + \beta v_0)A = v_0$$

$$\text{or} \quad v_{in} = \frac{v_o(1-\beta A)}{A}$$

If $\beta A=1$ Or $\beta=1/A$ then $v_{in} = 0$,i.e. we obtain an output voltage without any input signal. And amplifier functions as an oscillator.Therfore the condition for the maintenance of oscillation is $\beta A=1$ Or $\beta=1/A$. This equation is known as Barkhausen criterion for sustained oscillations.

Under the condition, the voltage gain after feedback

$$A_f = \frac{v_0}{v_{in}} = \frac{A}{(1-\beta A)} = \infty$$

It means that there is an output without input i.e. the amplifier becomes an oscillator.

8.8 HARTLEY OSCILLATOR

It is a very popular oscillator and widely used as a local oscillator in radio receiver.

Circuit diagram: As shown in Fig. 8.5, the frequency determining network consists of two inductors L_1 and L_2 and a capacitor C is connected in parallel to these. The output of the amplifier is applied across inductor L_1 and the voltage across L_2 forms the feedback voltage. The coil L_1 is inductively coupled to coil L_2 and the combination function as auto transformer. The capacitor C_c blocks d.c. and provides an a.c. path from the collector to the tank circuit. Radio frequency choke RFC provides d.c. load for the collector and also prevents a.c. from reaching the d.c. supply V_{cc}.capacitor C_{in} and resistor R_2 produce self bias while the combination R_e and C_e for stabilization of operating point.

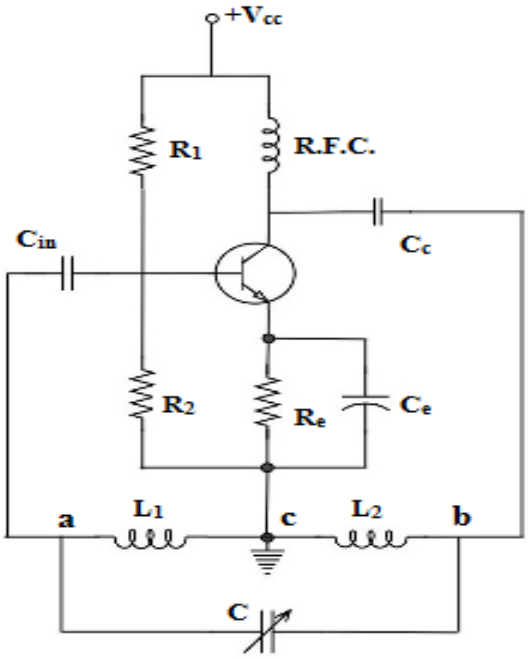


Figure 8.5: Hartley oscillator.

Working: When the collector supply voltage V_{cc} is switched on by closing the switch s, collector current starts rising and charges the capacitor C.When this capacitor is fully charged, it discharges through coils L_1 and l_2.setting up damped harmonic oscillations in the tank circuit. The oscillatory circuit in the tank circuit produces an ac voltage across L which is applied to the base emitter junction of the transistor. The amplified output is feedback to input through autotransformer action (i.e. by mutual inductance between L_1 and L_2.The energy supplied to tank circuit is in phase with the generated ososcillations.the ends a and b of the auto transformer L_1-L_2 are out of phase by180^0.point c is grounded i.e. at zero potential. Now with respect to point c, if point a is positive; point b will be at a negative potential. So that a phase difference of 80^0 is always maintained across L_1 and L_2.transistor amplifier also introduces phase difference of 80^0 between input and output voltages thus the total phase difference becomes 360^0 or 0^0.hence continuous oscillations are produced by oscillator using positive or regenerative feedback.

Frequency of oscillation: In Hartley oscillator the tank circuit is made up of L_1 and L_2 and C .The frequency of oscillation is given by: $f = \frac{1}{2\pi\sqrt{(L_1 + L_2 + 2M)C}}$ .Where M =mutual inductance between L_1 and L_2.

Feedback fraction: In Hartley oscillator the feedback voltage is across L_2 and output voltage is across L_1.hence feedback fraction is = $\beta = \frac{L_2}{L_1}$

8.9 COLPITTS OSCILLATOR

It is almost same as Hartley oscillator except it uses a tapped capacitance in place of tapped inductance in Hartley oscillator.

Circuit diagram: As shown in figure 8.6, the frequency determining network consists of two capacitor C_1 and C_2 and a inductance L is connected in parallel to these and the centre of the two capacitors is tapped. The output of the amplifier is applied across capacitance C_1 and the voltage across C_2 forms the feedback voltage. The capacitor C_c blocks d.c. and provides an a.c. path from the collector to the tank circuit. Radio frequency choke RFC provides d.c. load for the collector and also prevents a.c. from reaching the d.c. supply V_{cc}.capacitor C_{in} and resistor R_2 produce self bias while the combination R_e and C_e for stabilization of operating point.

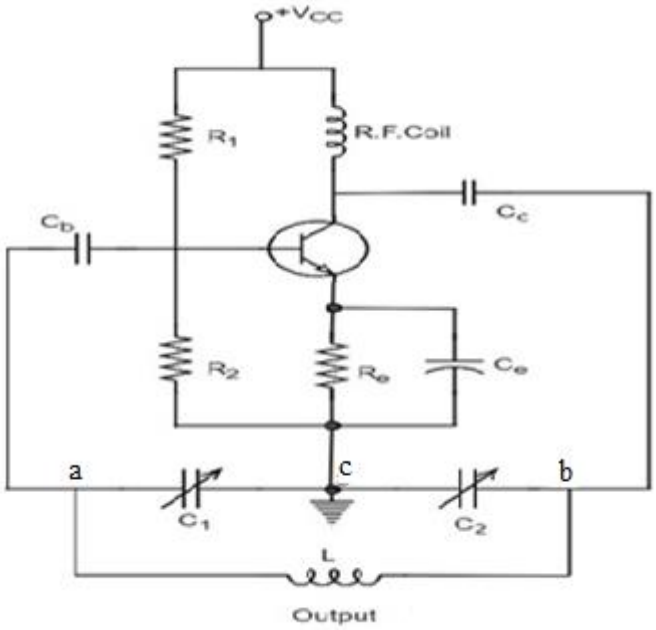


Figure 8.6: Colpitts oscillator.

Working: When the collector supply voltage V_{cc} is switched on by closing the switch, collector current starts rising and charges the capacitor C1 and C2.When this capacitor is fully charged, it discharges through coils L setting up damped harmonic oscillations in the tank circuit. The oscillatory circuit in the tank circuit produces an ac voltage across C1 which is applied to the base emitter junction of the transistor. The amplified output is in the collector circuit overcomes the losses occurring in the Tank circuit.

The energy supplied to tank circuit is in phase with the generated oscillations as follows. As point c is grounded if point a is at positive potential, b will be at negative potential with respect to c at that instant hence a phase difference of 180^0 is always maintained point a and point b.transistor amplifier also introduces phase difference of 80^0 between input and output voltages thus the total

phase difference becomes 360^0 or 0^0.hence continuous oscillations are produced by oscillator using positive or regenerative feedback

Frequency of oscillation: In Colpitt oscillator the tank circuit is made up of C_1 and C_2 and L. The frequency of oscillation is given by: $f = \frac{1}{2\pi\sqrt{LC}}$ .Where C = $\frac{C_1C_2}{C_1+C_2}$

Feedback fraction: In Colpitt oscillator the feedback voltage is across C_2 and output voltage is across L_1.hence feedback fraction is = $\beta = \frac{C_2}{C_1}$

Example 8.3. Find the frequency of oscillations of a Hartley oscillator if L_1=50 μH, L_2=1.5 mH,mutual inductance between them ia M=25 μH and C=100 pF.

Solution: Given L_1=50 μH, L_2=1.5 mH,M=25 μH and C=100 pF.

Frequency of oscillations of Hartley oscillator is given by=

$$f = \frac{1}{2\pi\sqrt{(L_1 + L_2 + 2M)C}}$$

$$= \frac{1}{2 \times 3.14 \times \sqrt{(50 + 1500 + 2 \times 25) \times 10^{-6} \times (100 \times 10^{-12})}}$$

$$= 398.09\ KHz.$$

Example 8.4. In colpitt's oscillator if L=10 mH,C_2=0.1μF and resonant frequency f=50000 Hz calculate the value of C_1.

Solution: frequency of oscillation of colpitt's oscillator is given by

$$f = \frac{1}{2\pi\sqrt{LC}} \qquad \text{where C} = \frac{C_1C_2}{C_1 + C_2}$$

Given L=10 mH,C_2=0.1μF and f=50000 Hz Hence

$$50000 = \frac{1}{2\pi\sqrt{10 \times 10^{-3} \times C}}$$

$$\mathbf{C = 3.19 \times 10^{-6}\ F}$$

Now $C = \frac{C_1C_2}{C_1 + C_2}$ so $3.19 \times 10^{-6} = \frac{C_1 \times .1 \times 10^{-6}}{C_1 + .1 \times 10^{-6}}$

$$(C_1 + .1 \times 10^{-6})3.19 \times 10^{-6} = (C_1 \times .1 \times 10^{-6})$$

Hence $C_1 = (.35 \times 10^{-6})$= .35 μF

8.10 WEIN BRIDGE OSCILLATOR

It is the standard oscillator circuit for all low frequencies in the range of 5 Hz to 1 MHz. It is generally used in commercially audio frequency oscillators and other low frequency applications. It is also a low distortion amplifier as the distortion is less than one percent.

This type of RC oscillators consists of two stage amplifier with RC lead –lag network. The first transistor serves as an oscillator and the second transistor as inverter. Thus the total phase shift between the input to the first stage and output of the second stage is 360^0.

Wein bridge network: The feedback RC network consists of two arms as shown in fig. 8.7.The first arm comprises a resistor R_1 and a capacitor C_1 in series. A parallel combination of R_2 and C_2 forms the second arm. The phase shift and attenuation by the network can be calculated as follows:

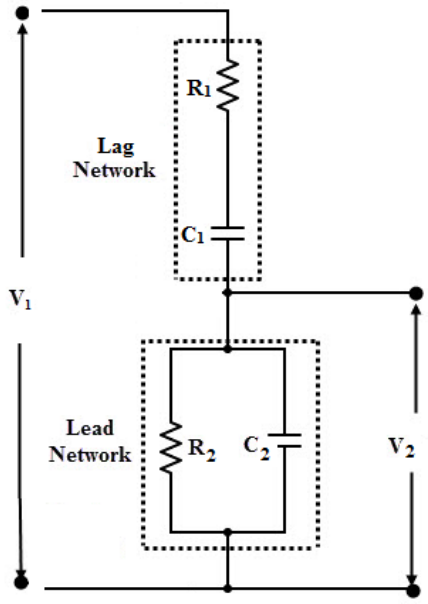


Figure 8.7: Wein bridge network.

$$Z_{parallel} = \frac{-R_2 \times \frac{j}{\omega C_2}}{R_2 - \frac{j}{\omega C_2}}$$

$$Z_{series} = R_1 - \frac{j}{\omega C_1}$$

Hence $$\frac{v_2}{v_1} = \frac{Impedence\ of\ parallel\ combination}{Total\ impedence}$$

$$\frac{v_2}{v_1} = \frac{\dfrac{-R_2 \times \dfrac{j}{\omega C_2}}{R_2 - \dfrac{j}{\omega C_2}}}{R_1 - \dfrac{j}{\omega C_1} - \dfrac{R_2 \times \dfrac{j}{\omega C_2}}{R_2 - \dfrac{j}{\omega C_2}}}$$

$$= \frac{\dfrac{R_2}{1+j\omega C_2 R_2}}{R_1 + \dfrac{j}{\omega C_1} + R_2 \left[1+j\omega C_2 R_2\right]}$$

$$= \frac{R_2}{R_1 + R_2 + \dfrac{C_2 R_2}{C_1} + \left[j\omega C_2 R_1 R_2 - \dfrac{1}{\omega C_1}\right]} \qquad \text{-------- (8.1)}$$

For the network to have phase shift of 0^0 or 80^0,imaginary part must be zero.
Therefore

$$\omega C_2 R_1 R_2 = \frac{1}{\omega C_1}$$

Or

$$\omega = \frac{1}{\sqrt{C_1 C_2 R_1 R_2}}$$

Hence frequency of oscillation

$$f_0 = \frac{1}{2\pi\sqrt{C_1 C_2 R_1 R_2}}$$

If the circuit component are such that $R_1 = R_2 = R$ and $C_1 = C_2 = C$ We get

$$f_0 = \frac{1}{2\pi RC}$$

Then

$$\frac{v_2}{v_1} = \frac{R_2}{R_1 + R_2 + \dfrac{C_2 R_2}{R_1}}$$

Again for $R_1 = R_2 = R$ and $C_1 = C_2 = C$ We get

$$\frac{v_2}{v_1} = \frac{1}{3}$$

This is the condition for sustained oscillations i.e. the feedback factor should be $\frac{1}{3}$ in case of wein bridge oscillator for the case $R_1 = R_2 = R$ and $C_1 = C_2 = C$

Equation 8.1shows that at very low frequencies (f < f0) the phase shift (i.e. phase angle of the output voltage v_2 with respect to the input voltage v_1) is positive as and at very high frequencies (f > f0) phase shift is negative. Thus the network circuit acts as a lead network at low frequency and as lag network at high frequencies. The phase shift is zero at frequency f=f0 known as resonant

frequency. Equation 3 shows that for sustained oscillation βAv>>1 or Av>>0.Since each common emitter amplifier provides a high voltage gain thus a negative feedback is required to control the total voltage gain. Therefore, a weinbridge oscillator uses positive and negative feedback. The positive feedback is through lead lag network of the non-inverting input and negative feedback is through the voltage divider to the inverting input (Fig.8.8). The two arms of the lead lag network and the two arms of voltage divider form a bridge known as Wein Bridge.

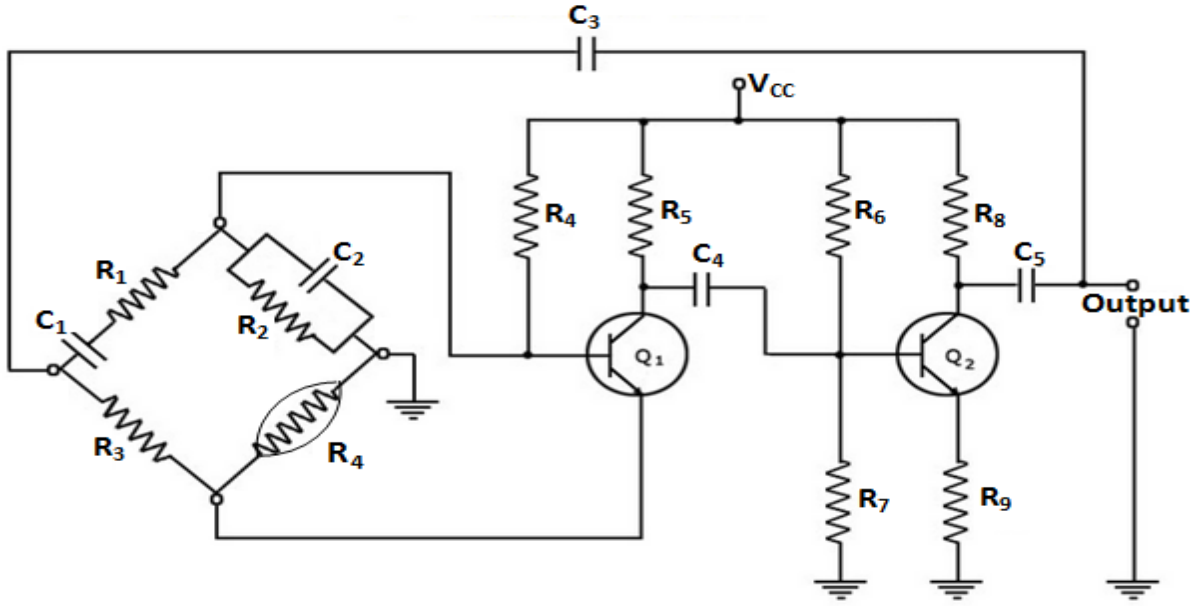


Figure 8.8: Wein bridge oscillator.

8.11 PHASE SHIFT OSCILLATOR

A phase shift oscillator essentially consists of RC circuit which provides an additional phase shift of 180^0 at same frequency as common emitter amplifier also produces 180^0 phase shift to the input signal thus the amplifier and the feedback circuit loop provides total phase shift of 360^0 at this frequency. This particular frequency will be the frequency at which the circuit will oscillate.

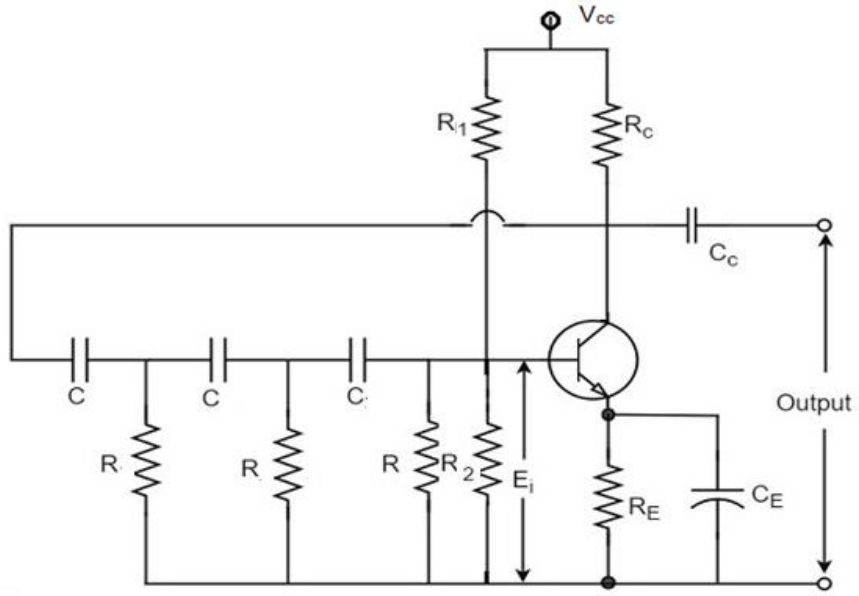


Figure 8.9: Phase shift oscillator.

The alternating current I so does v across r will lead the applied voltage by the angle

$$\tan\varphi = \frac{1}{\omega CR} = \frac{1}{2\pi fCR}$$

The values of C and R may be chosen such that for the frequency f phase shift $\varphi = 60^0$.

Since phase shift varies with frequency the frequency at which the RC circuit produces phase shift of 60^0 is called critical frequency.

Mathematical analysis: Since the bias and stabilization resistances R_1,R_2,R_E have sufficiently large value they are ignored in a.c equivalent circuit (Fig. 8.10)

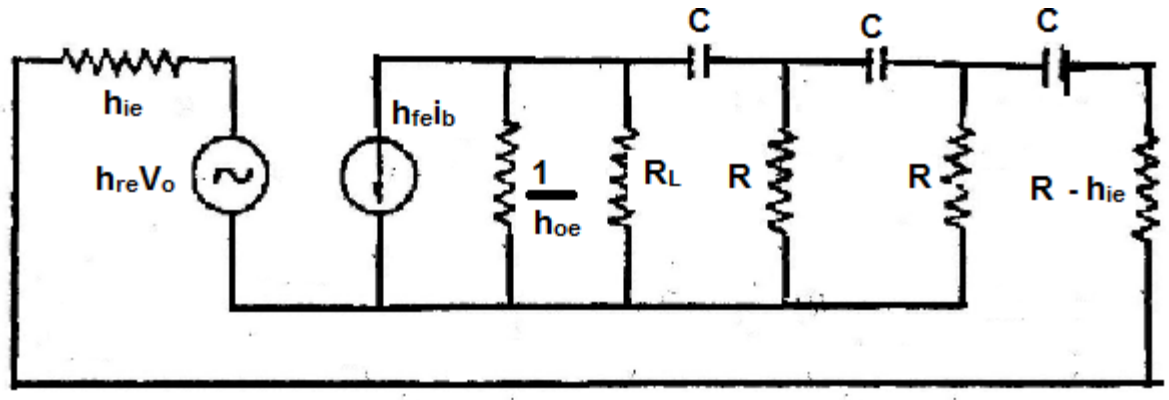


Figure 8.10: A.C. equivalent circuit of Phase shift oscillator.

To make mathematical analysis simpler $h_{re}V_o$ can be neglected as H_{re} of transistor is very small. We can also omit $\frac{1}{h_{oe}}$ as it is much larger than R_L. Also the current source $h_{fe}i_b$ has been replaces by an equivalent thevenin source voltage $h_{fe}i_bR_L$ with impedance R_L in series.Hence equivalent circuit may be reduces as shown in Fig. 8.11,

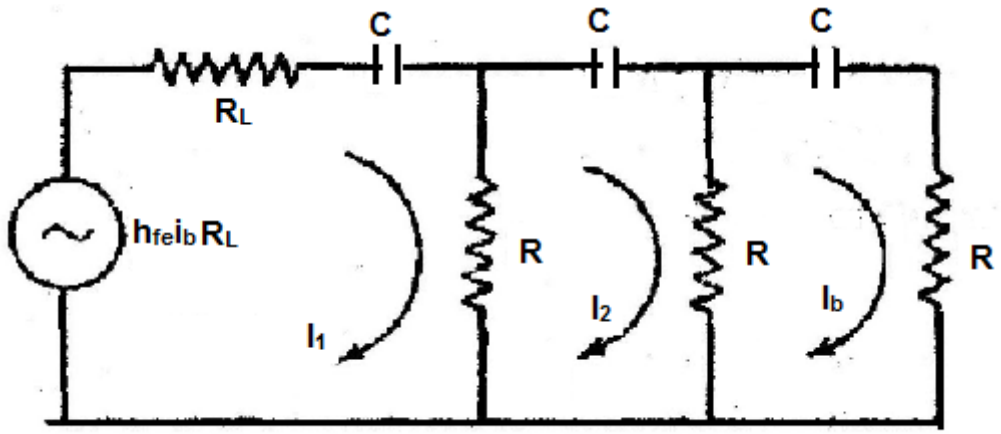


Figure 8.11: Simplified equivalent circuit of phase shift oscillator.

Applying Kirchoff's law for 3 loops in Fig. 8.11

$$\left(R + R_L + \frac{1}{j\omega C}\right) I_1 - R\, I_2 + h_{fe} R_L I_b = 0$$

$$-RI_1 + \left(2R + \frac{1}{j\omega C}\right) I_2 - RI_b = 0$$

$$-RI_2 + \left(2R + \frac{1}{j\omega C}\right) I_b = 0$$

It's determinant is given by

$$\begin{vmatrix} \left(R + R_L + \frac{1}{j\omega C}\right) & -R & h_{fe} R_L \\ -R & \left(2R + \frac{1}{j\omega C}\right) & -R \\ -R & -R & \left(2R + \frac{1}{j\omega C}\right) \end{vmatrix}$$

Or

$$\left(R + R_L + \frac{1}{j\omega C}\right)\left(3R^2 + \frac{4R}{j\omega C} - \frac{1}{\omega^2 C^2}\right) - R^2 \left(2R + \frac{1}{j\omega C}\right) + h_{fe} R_L R^2 = 0$$

Or

$$(R + R_L)\left(3R^2 - \frac{1}{\omega^2 C^2}\right) - (R + R_L)\left(\frac{j4R}{\omega C}\right) - j\left(\frac{3R^2}{\omega C} - \frac{1}{\omega^3 C^3}\right) - \frac{4R}{\omega^2 C^2} - 2R^3 + \frac{jR^2}{\omega C} + h_{fe} R_L R^2 = 0$$

To find the frequency of oscillation equating the imaginary part of equation to zero

$$-(R + R_L)\left(\frac{4R}{\omega C}\right) - \left(\frac{3R^2}{\omega C} - \frac{1}{\omega^3 C^3}\right) + \frac{R^2}{\omega C} = 0$$

$$-(R + R_L)4R - 3R^2 + \frac{1}{\omega^2 C^2} + R^2 = 0$$

$$-4R^2 - 4RR_L - -2R^2 + \frac{1}{\omega^2 C^2} = 0$$

$$6R^2 + 4RR_L - \frac{1}{\omega^2 C^2} = 0$$

$$0r \qquad \omega^2 = \frac{1}{C^2(6R^2+4RR_L)}$$

Therefore frequency of oscillation is $f = \frac{1}{C\sqrt{(6R^2+4RR_L)}}$

If R_L=R then frequency of oscillation will be $f = \frac{1}{RC\sqrt{(10)}}$

To find the condition for sustained oscillation equating the real part to zero

$$(R + R_L)\left(3R^2 - \frac{1}{\omega^2C^2}\right) - \frac{4R}{\omega^2C^2} - 2R^3 + h_{fe}R_LR^2 = 0$$

$$3R^3 - \frac{R}{\omega^2C^2} + 3R^2R_L - \frac{R_L}{\omega^2C^2} - \frac{4R}{\omega^2C^2} - 2R^3 + h_{fe}R_LR^2 = 0$$

Putting $\omega^2 = \frac{1}{C^2(6R^2+4RR_L)}$,the equation becomes

$$h_{fe}R_LR^2 = 29R^3 + 23\,R^2R_L + 4RR_L{}^2$$

Or

$$h_{fe} = 29\,\frac{R}{R_L} + 23\ + 4\frac{R_L}{R}$$

This is the required condition for sustained oscillation.

If R_L=R then frequency of oscillation will be $f = \frac{1}{RC\sqrt{(10)}}$ and $h_{fe} = 56$

Example 8.5. Find the frequency of oscillations of a Wein bridge oscillator with R= 10 kΩ and C=100 pF.

Solution: Frequency of oscillations of a Wein bridge oscillator is given by=

$$f = \frac{1}{2\pi RC} = \frac{1}{2\ \times 3.14\ \times 10\ \times 10^3\ \times 1000\ \times 10^{-12}} = 15.92\ KHz$$

Example 8.6. A phase shift oscillator uses three identical RC sections with R=50 KΩ and C=0.05 μF.Calculate the frequency of oscillation of phase shift oscillator.

Solution: For the identical section frequency of oscillation of phase shift oscillator is given by

$$f = \frac{1}{2\pi RC\sqrt{10}}$$

Given R=50 KΩ =5×10^4 Ω and C=0.05 μF = 5×10^{-8} F.

Hence

$$f = \frac{1}{2\pi\ \times 5 \times 10^4\ \times 5 \times 10^{-8}\sqrt{10}}$$

= 49.6 Hz

8.12 STABILITY OF AN OSCILLATOR-CRYSTAL OSCILLATOR

For the oscillator frequency of an oscillator must remain constant. The stability of an oscillator is a measure of its ability to maintain a frequency over a long interval of time. Some of the prominent factor causing the change in frequency of an oscillator are variation of supply voltage, variation in load, variation in the reactive component due to change in temperature with time, variation in

transistor elements due to thermal changes etc. Different methods are adopted to improve the frequency stability. Crystal oscillator is one such method in which the usual electrically tune circuit is replaced by properly cut quartz crystal. Crystal oscillators are oscillators where the primary frequency determining element is a quartz crystal. Because of the inherent characteristics of the quartz crystal the crystal oscillator may be held to extreme accuracy of frequency stability. Temperature compensation may be applied to crystal oscillators to improve thermal stability of the crystal oscillator.

Piezoelectric effect in quartz crystal: Quartz crystal is a very common form of silica and generally found in the form of a hexagonal prism with hexagonal pyramids at both the ends The line joining the two pointed ends of the crystal is called optical or z axis. X axis are known as electrical axis and Y as mechanical axis "(Fig. 8.12).

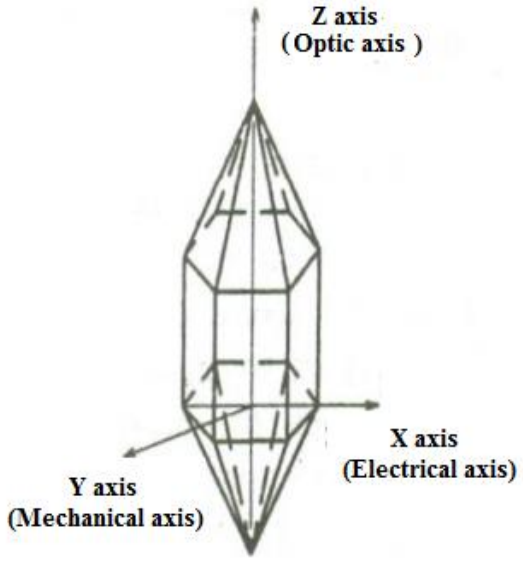


Figure 8.12: Quartz crystal.

In certain crystal like quartz, tourmaline etc. electrical and mechanical properties are interconnected in a crystal it is a called piezoelectric effect according to which When mechanical stress is applied along y axis of quartz crystal electrical charges appear on the faces of the crystal perpendicular to Y axis and vice versa. Now if we apply a.c. voltage across a quartz crystal it expands and contracts alternately,i.e.set into vibration at a frequency equal to applied alternating voltage.

Circuit diagram of crystal oscillator: Also known as pierce oscillator; as shown in Fig 8.13the resistance R_1, R_2 and R_E form the biasing and stabilization network same as in any transistor amplifier. Capacitor C_e by passes ac component and isolates R_e to avoid degeneration. Radio frequency choke RFC provides dc collector load and prevent any ac signal from entering the power supply. Quartz crystal is connected as a series element in the feedback path from collector to base. The crystal is excited in the series resonance mode. In series resonance mode crystal provides minimum impedance hence amount of positive feedback is maximum. The voltage

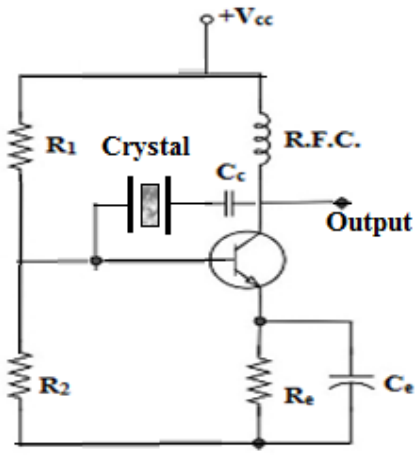


Figure 8.13: Crystal oscillator.

feedback signal is maximum for minimum crystal impedance. Coupling capacitor C_c has negligible impedance at the circuit operating frequency but high impedance for any dc component between collector and base therefore block dc component. The crystal is excited by the feedback energy from the output to input through crystal.

The oscillation frequency is equal to the series resonance frequency of the crystal and is given by

$$f_0 = \frac{1}{2\pi\sqrt{LC_1}}$$

Advantages and disadvantages of crystal oscillators are:

Advantages:

1. Very simple, needs no tank circuit. the frequency of oscillation is determined by the crystal which is independent of supply voltage and transistor parameters.
2. Oscillation frequencies can be changed by simply replacing one crystal with another.
3. The quality factor is as high as 10^4 to 10^6 in compassion with LC circuit having quality factor of the order of 100.
4. Very small frequency drift at room temperature which also can improved by putting crystal in crystal oven in which temperature is thermostatically controlled.

Disadvantages:

1. Generally used for frequencies exceeding 100 kHz. Have a very limited tuning range or no tuning at all.
2. Crystals are fragile therefore only used in low power circuits.

Example 8.7. In a crystal oscillator frequency of oscillation=10MHz and C=100 PF.Calculate the value of inductance offered by crystal.

Solution: frequency of oscillation of crystal oscillator is given by

$$f = \frac{1}{2\pi\sqrt{LC}}$$

Given f=10 MHz = 10×10^6 Hz and C=100 PF = 100×10^{-12} F.

$$10 \times 10^{6} = \frac{1}{2\pi\sqrt{\mathrm{L} \times 100 \times 10^{-12}}}$$

$$L = \frac{1}{4\pi^{2} \times 100 \times 10^{12} \times 10^{-10}}$$

= 2.54 H

8.13 RELAXATION OSCILLATORS

The output waveform is either square, rectangular, triangular or pulse type.i.e. non sinusoidal. Examples of relaxation oscillator are:

1. Multivibrator
2. Blocking oscillator

8.14 MULTIVIBRATORS

Multivibrator is a relaxation oscillator which consists a two stage R-C coupled amplifier such that output of one amplifier is positively feedback with input of the other. The circuit operates such that at a time if one transistor is conducting (on) and other cut off (on) controlled by circuit condition. There are only two possible state of a multivibrator as follows:

First state: Transistor Q_1 on and transistor Q_2 off.

Second state: Transistor Q_1 off and transistor Q_2 on.

The condition in which the multivibrator may remain indefinitely until the circuit is triggered by some external signal is known as stable state while if it changes automatically to other state after remaining in one state for finite time determined by circuit component is called quasi stable state. Depending upon the type of coupling and no. of stable states the multivibrators can be of three types.

1 Astable multivbrator
2. Monostable multivbrator
3.Bistable multivbrator

8.14.1 Astable Multivibrator or Free Running Multivibrator

An astable or free running multivibrator has no stable states but two quasi stable state. Once the Multivibrator is ON, it just changes its states on its own after a certain time period which is determined by the R_C time constants. A dc power supply or V_{cc} is given to the circuit for its operation.

Construction of Astable Multivibrator

Astble multivibrator consists of Two identical transistors named Q_1 and Q_2 connected to one another using positive feedback.,The collector of transistor Q_1 is connected to the base of transistor Q_2 through the capacitor C_1 and vice versa. The emitters of both the transistors are connected to the ground. The collector load resistors R_{L1} and R_{L2} and the biasing resistors R_1 and R_1 are of equal values. The capacitors C_1 and C_2 are of equal values.

The following Fig. 8.14 shows the circuit diagram for Astable Multivibrator.

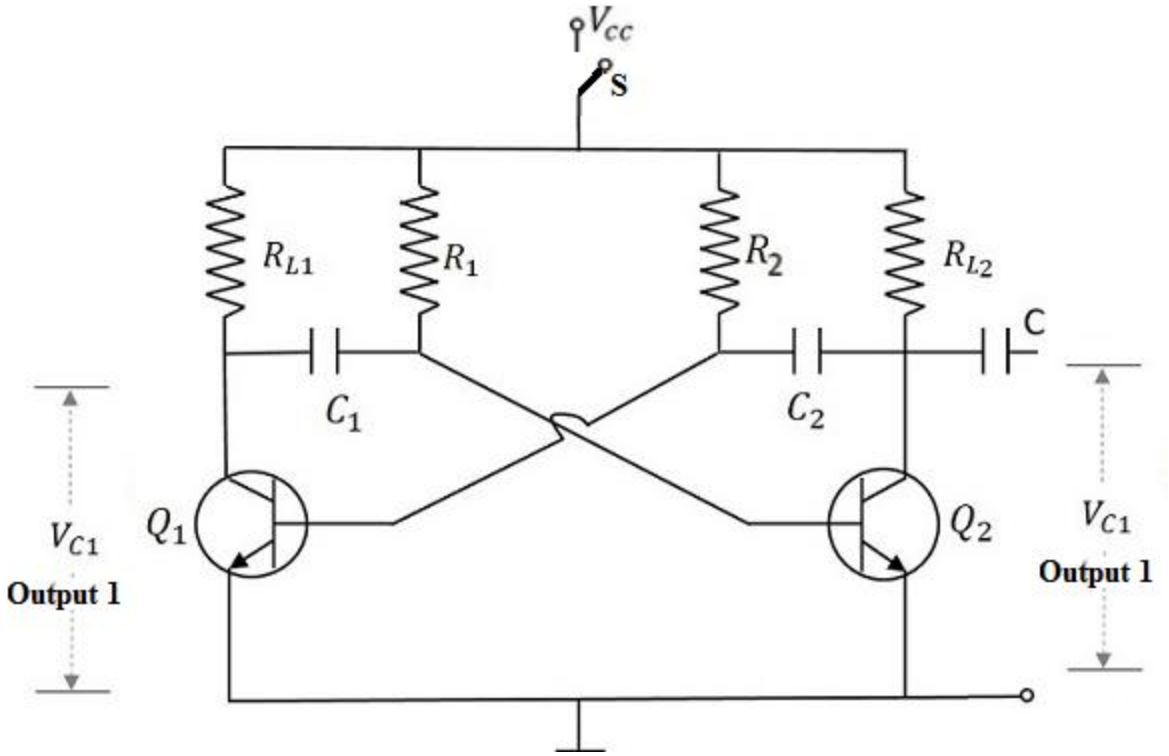


Figure 8.14: Astable or free running multivibrator.

Operation of Astable Multivibrator

When V_{cc} is applied, the collector current of the transistors increase. The collector current depends upon the base current,

$$Ic=\beta I_b$$

As no transistor characteristics are alike, one of the two transistors say Q_1 has its collector current increase and thus conducts. The collector of Q_1 is applied to the base of Q_2 through C_1. This connection lets the increased negative voltage at the collector of Q_1 to get applied at the base of Q_2 and its collector current decreases. This continuous action makes the collector current of Q_2 to decrease further. This current when applied to the base of Q_1 makes it more negative and with the cumulative actions Q_1 gets into saturation and Q_2 to cut off. Thus the output voltage of Q_1 will be $V_{CE\ (sat)}$ and Q_2 will be equal to V_{CC}.

The capacitor C_1 charges through R_1 and when the voltage across C_1 reaches 0.7V, this is enough to turn the transistor Q_2 to saturation. As this voltage is applied to the base of Q_2, it gets into

saturation, decreasing its collector current. This reduction of voltage at point B is applied to the base of transistor Q_1 through C_2 which makes the Q_1 reverse bias. A series of these actions turn the transistor Q_1 to cut off and transistor Q_2 to saturation. Now point A has the potential V_{CC}. The capacitor C_2 charges through R_2. The voltage across this capacitor C_2 when gets to 0.7V, turns on the transistor Q_1 to saturation.

Hence the output voltage and the output waveform are formed by the alternate switching of the transistors Q_1 and Q_2. The time period of these ON/OFF states depends upon the values of biasing resistors and capacitors used, i.e., on the R_C values used. As both the transistors are operated alternately, the output is a square waveform, with the peak amplitude of V_{CC}.

Output Waveforms

The output waveforms of astable multivibrator at the collectors of Q_1 and Q_2 are shown in the following Fig. 8.15.

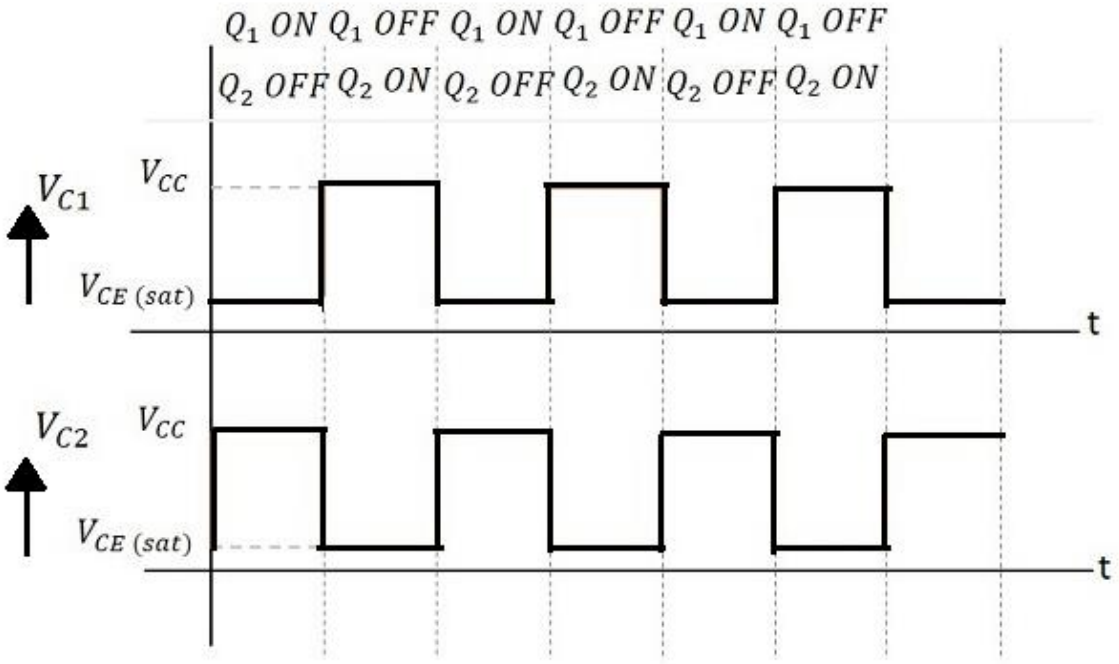


Figure 8.15 Output waveform of astable or free running multivibrator.

Frequency of Oscillations

The ON time of transistor Q_1 or the OFF time of transistor Q_2 is given by

$$t_1 = 0.69R_1C_1$$

Similarly, the OFF time of transistor Q_1 or ON time of transistor Q_2 is given by

$$t_2 = 0.69R_2C_2$$

Hence, total time period of square wave

$$t = t_1 + t_2 = 0.69(R_1C_1 + R_2C_2)$$

As $R_1 = R_2 = R$ and $C_1 = C_2 = C$, the total time period of square wave

$$t = 1.38\ RC$$

frequency of square wave will be

$$f = 1/t = 1/(1.38RC) = 0.7/RC$$

8.14.2 Monostable Multivibrator or Single Shot Multivibrator

A monostable multivibrator, as the name implies, has only one stable state and one quasi stable state. When one transistor conducts (i.e. on state), the other remains in non-conducting state (i.e. off state).As we already know that a stable state is such a state where the transistor remains without being altered, unless disturbed by some external trigger pulse. On applying external triggering pulse it switch to quasi stable state and returns to the original stable state after a certain time delay determined by the value of circuit components and remained there until another trigger pulse is applied. Hence it is also called delay or single shot multivibrator.

Thus a monostable multivibrator supplies a single output wave of a desired duration for every input trigger pulse unlike astable multivibrator which generates continuous square wave without the need of triggering.

Construction of Monostable Multivibrator

As shown in fig. 8.16 , monostable multivibrator consists two similar transistors Q_1 and Q_2 , connected to one another using positive feedback. The collector of transistor Q_1 is connected to the base of transistor Q_2 through the capacitor C_1. The base Q_1 is connected to the collector of Q_2 through the resistor R_2 and capacitor C. Another dc supply voltage $-V_{BB}$ is given to the base of transistor Q_1 through the resistor R_3. The trigger pulse is given to the base of Q_1 through the capacitor C_2 to change its state. R_{L1} and R_{L2} are the load resistors of Q_1 and Q_2.

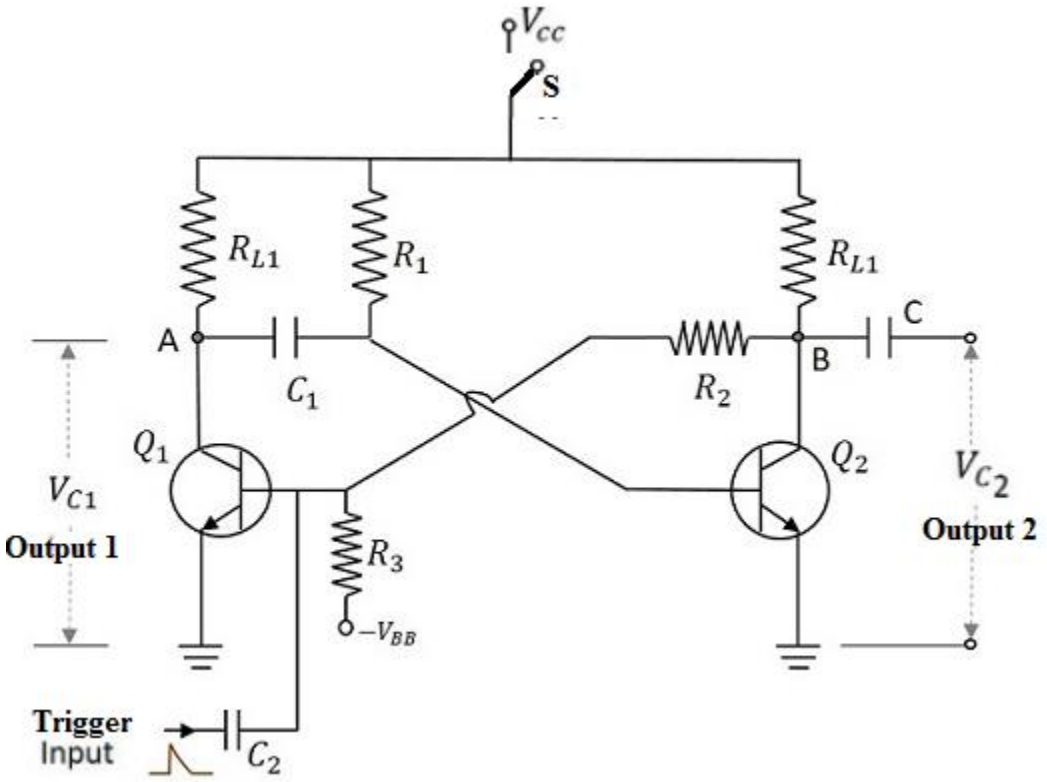


Figure 8.16: Monostable or single shot multivibrator.

One of the transistors, when gets into a stable state, an external trigger pulse is given to change its state. After changing its state, the transistor remains in this quasi-stable state or Meta-stable state for a specific time period, which is determined by the values of RC time constants and gets back to the previous stable state.

Operation of Monostable Multivibrator

Firstly, when the circuit is switched ON by closing the switch S and no triggering pulse is applied at C2, transistor Q_1 will be in OFF state and Q_2 will be in ON state. This is the stable state. As Q_1 is OFF, the collector voltage will be V_{CC} at point A and hence C_1 gets charged. Now if we apply A positive trigger pulse at the base of the transistor Q_1 through C2 ,It gives forward bias to transistor Q1 hence Q1 starts conducting. then the collector current of Q1 will rise and causing a decrease in its collector voltage. The resulting negative voltage is applied to the base of Q2 through C1this decreases the forward bias on Q2 and causes decrease in its collector current. Consequently, potential of point B increase due to lesser drop over RL2.It further increase the forward bias on Q1and the cumulative action soon Q1 is driven to saturation and Q2 to cut off. But This is the quasi-stable state or Meta-stable state.

The transistor Q_2 remains in OFF state, until the capacitor C_1 discharges completely. After this, the transistor Q_2 turns ON with the voltage applied through the capacitor discharge. This turn ON the transistor Q_1, which is the previous stable state.

Output Waveforms

The output waveforms of monostable multivibrator at the collectors of Q_1 and Q_2 along with the trigger input given at the base of Q_1 are shown in the following Fig 8.17.

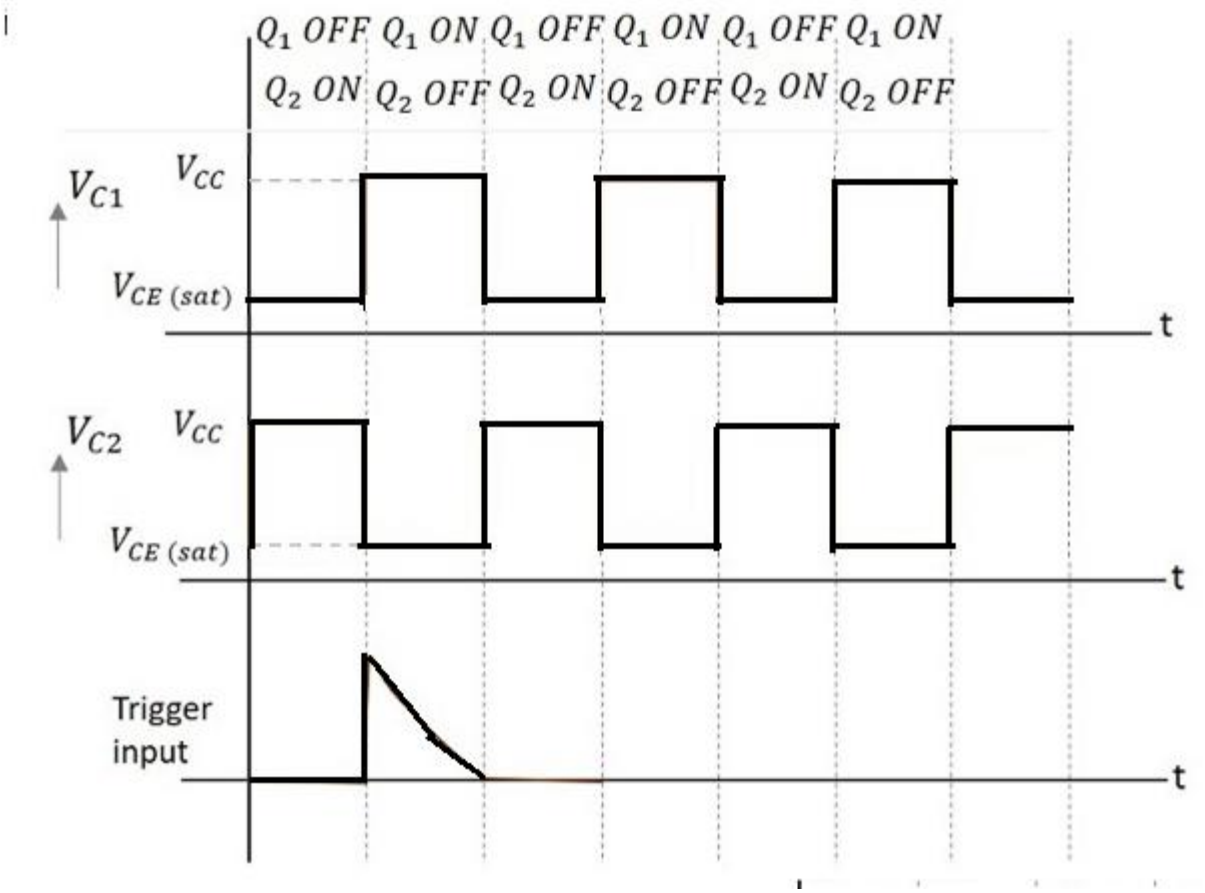


Figure 8.17: Output waveform of monostable or single shot multivibrator.

The width of this output pulse depends upon the RC time constant. Hence it depends on the values of R_1C_1. The duration of pulse is given by

$$T=0.69R1C1$$

The trigger input given will be of very short duration, just to initiate the action. This triggers the circuit to change its state from Stable state to Quasi-stable or Meta-stable or Semi-stable state, in which the circuit remains for a short duration. There will be one output pulse for one trigger pulse.

8.14.3 Bistable Multivibrator or Flip Flop

Bistable Multivibrators have two stable states ("Bi" means two) and maintain a given output state indefinitely unless an external trigger is applied forcing it to change state. To change from one state to the other, the bistable circuit requires a suitable trigger pulse for each stage. It is also known

as “flip-flop” as it “flips” into one logic state, remains there and then changes or “flops” back into its first original state.

Construction

Bistable multivibrator consists two identical common emitter transistors Q_1 and Q_2 with collector resistance R_{C1} and R_{C2} are connected in positive feedback to one another. The base resistors R_3 and R_4 are joined to a common source $-V_{BB}$. The feedback resistors R_1 and R_2 are shunted by capacitors C_1 and C_2 known as Commutating or speed up Capacitors as they reduce the transition time (time taken for the transfer of conduction from one transistor to the other). Triggering of transistor Q_1 and Q_2 is done by giving triggering pulse at the base of respective transistor through the capacitor C_3 and C_4.Outpur can be taken across collector emitter circuit of either transistor. The following Fig. 8.18 shows the circuit diagram of a self-biased bistable multivibrator.

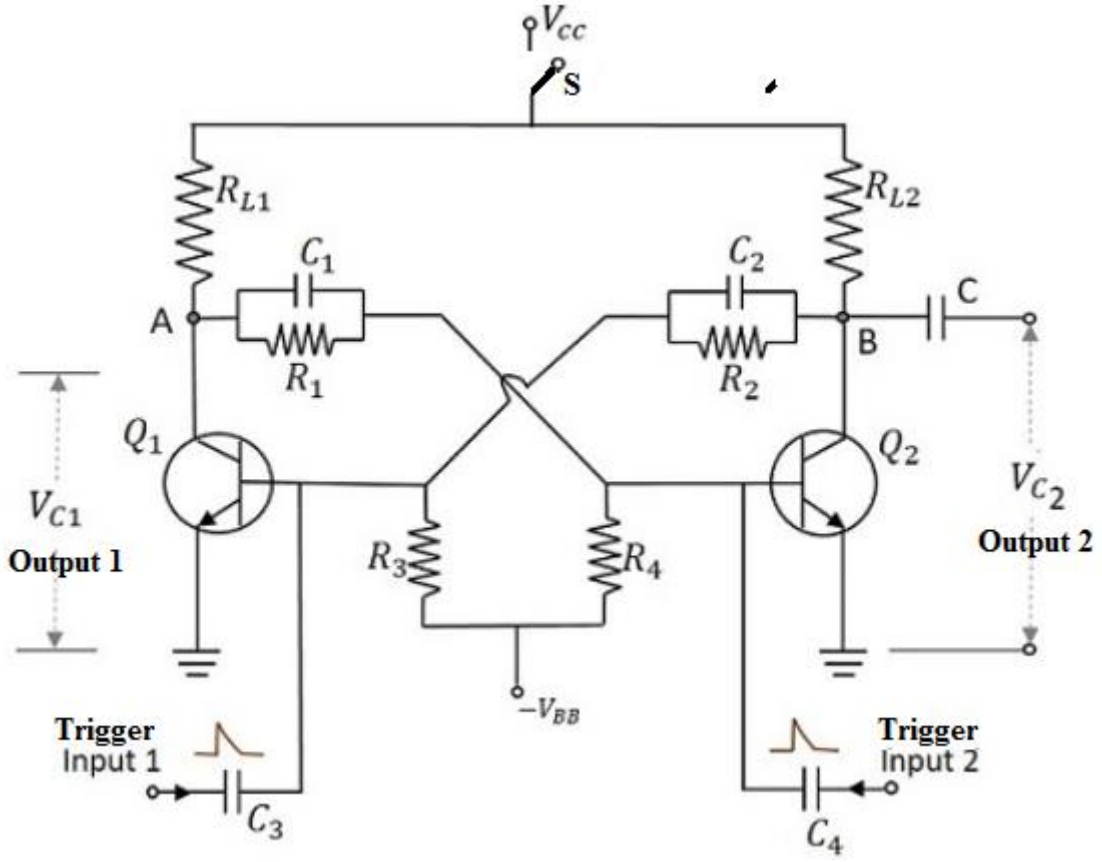


Figure 8.18: Bistable multivibrator or flip flop.

Working of Bistable Multivibrator

When the circuit is switched ON by closing switch S, due to some circuit imbalances as no two transistors can be exactly similar, one of the transistors, say Q_1 gets will conduct more than other. then the collector current of Q1 will rise at a faster rate causing a decrease in its collector voltage. The resulting negative signal is applied to the base of Q2 through C1 and consequently the collector voltage of Q2(positive going signal) is fed to the base of transistor Q1causes further decrease in collector current of Q1 and is applied to base of Q2.This regenerative action soon make transistor

Q1 switched ON, while the transistor Q_2 gets switched OFF. This is a stable state of the Bistable Multivibrator and it will stay indefinitely in this condition. For switching the multivibrator to other state, an external trigger pulse must be applied at the base 0f transistor.

now by applying a negative trigger at the base of transistor Q_1 or a positive trigger pulse at the base of transistor Q_2, we can change this stable state condition to other stable state i.e. transistor Q1 off and Transistor Q2 off. So, to understand this let a negative pulse is applied at the base of transistor Q_1. As a result, the collector voltage increases, which forward biases the transistor Q_2. The collector current of Q_2 as applied at the base of Q_1, reverse biases Q_1 and this cumulative action, makes the transistor Q_1 OFF and transistor Q_2 ON. This is another stable state of the Multivibrator. The same purpose is also served by applying a positive trigger pulse at the base of transistor Q_2

Now, if this stable state has to be changed again, then either a negative trigger pulse at transistor Q_2 or a positive trigger pulse at transistor Q_1 is applied.

Output Waveforms

The output waveforms of bistable multivibrator at the collectors of Q_1 and Q_2 along with the trigger inputs given at the bases of Q_W and Q_2 are shown in the following figures.

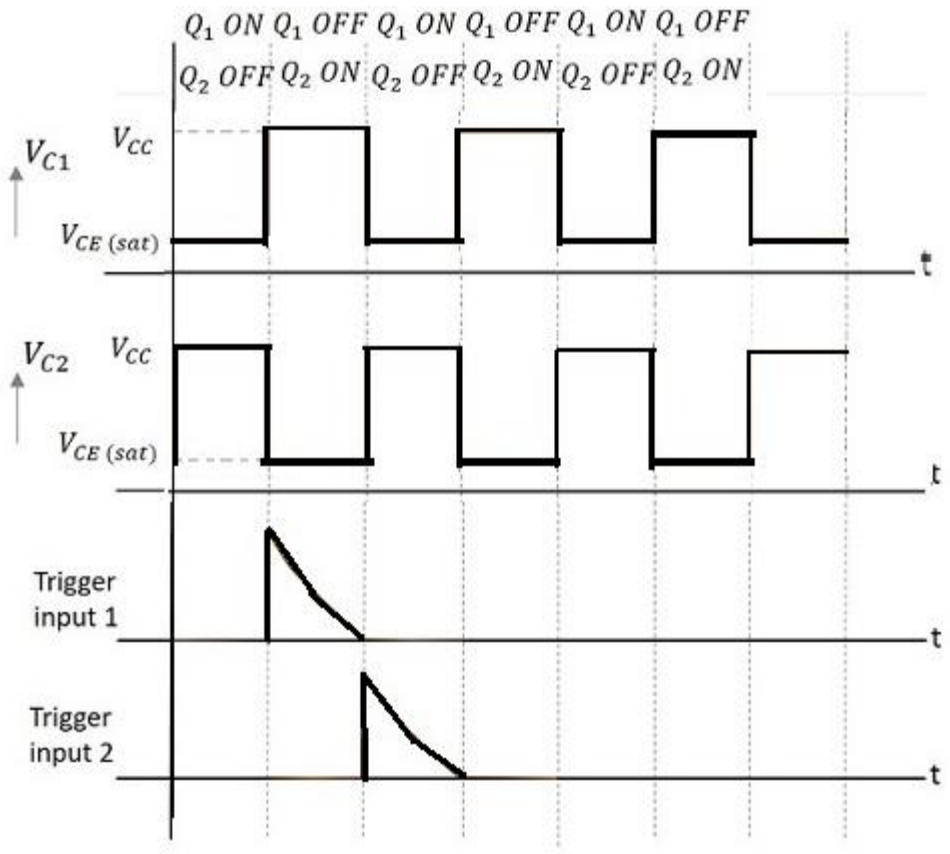


Figure: 8.19 Output waveform of bistable multivibrator.

Example 8.8. In a free running multivibrator, $R_1=R_2=20$ KΩ , $C_1=C_2=100$pF and $R_{L1}=R_{L2}=1$ KΩ.Calculate the frequency of oscillation and minimum value of transistor β.

Solution: As per given values it is clear that both stages are identical hence

$R= R_1=R_2=20$ KΩ$=20\times10^3\Omega$, $C= C_1=C_2=100$pF $=100\times10^{-12}$

and $R_L =R_{L1}=R_{L2}=1\times10^3\Omega$

frequency of oscillation of Astable multivibrator is given by

$$f = \frac{0.7}{RC}$$

$$f = \frac{0.7}{20 \times 10^3 \times 100 \times 10^{-12}}$$

$$f = 350\ KHz$$

To ensure oscillation transistor must saturate for which minimum value of β is given by

$$\beta_{min} = \frac{R}{R_L} = \frac{20\times10^3}{1\times10^3} = 20$$

8.15 SUMMARY

Feedback is a process in which a part or fraction of output is injected back to the input, it is called. In phase feedback with input is known as positive feedback while opposite phase feedback is called negative feedback. Positive feedback though increases the gain of the amplifier but also causes increased instability and distortion hence seldom used in amplifiers but used in oscillators. While negative feedback has many advantages except reduces gain such as gain stability, reduced distortion as well as noise, increased bandwidth, increase in input impedance and decrease in output impedance hence frequently used in amplifiers. Oscillator is an electronic device capable to generate a periodic a.c. signal without input. Amplifiers with positive feedback such that Aβ≥1 satisfies the breakhausan criterion of sustained oscillation. An oscillatory circuit or frequency determining elemental amplifier and positive feedback network are the three essential requirements for an oscillator. Oscillators which use inductance-capacitance (L-C) circuits as their tank circuits are called L-C oscillators. They are generally used as radio frequency oscillators. Examples are tuned collector, tuned base, Colpitt's, Hartley, etc. on the other hand crystal oscillators are very stable in which usual L-C circuit is replaced by a piezoelectric crystal. If resistance and capacitance are the frequency determining element(R-C) such oscillators are called R-C Oscillators. Wien Bridge and phase shift oscillator are examples of R-C oscillators and frequently used as audio oscillator. Another type of sinusoidal oscillator is negative resistance oscillators which make use of negative resistance elements such as dynatron and tunnel diode. The frequency range, power or voltage requirements, accuracy, stability, distortion and output impedance of an oscillator determines oscillator's suitability for a particular application. multivibrators are regenerative circuit mainly used in timing and switching applications. Astable or free running multivibrator have no stable state and two quasi stable state and changes from one quasi stable to other without any external triggering hence used as square wave generator.

Monostable multivibrator has one stable and one quasi stable stable state while bistable multivibrator has two stable state known also known as flip flop.

8.16 GLOSSARY

Feedback: The process of injecting a part of the output back to input is called feedback.

Positive feedback: If the feedback voltage is in same phase with input thus aids it and gain of the amplifier increases, it is called positive feedback. Generally, employs in oscillators.

Negative feedback: If the feedback voltage is 180^0 out of phase with input thus opposes it and gain of the amplifier decreases, it is called negative feedback. Despite of decreased gain because of other advantages such as increased bandwidth, increased stability, low noise and distortion etc. it is frequently used in oscillator.

Oscillator: An electronic device used to produce oscillation.

Stable and quasi-stable state: The condition in which the multivibrator may remain indefinitely until the circuit is triggered by some external signal is known as stable state while if it changes automatically to other state after remaining in one state for finite time determined by circuit component is called quasi stable state.

Relaxation oscillator: An oscillator which produces non sinusoidal oscillations is called relaxation oscillator.e.g. multivibrator.

8.17 REFERENCE BOOKS

1. Principles of electronics by V.K.Mehta
2. Basic Electronics by B.L.Thereja
3. Basic Electronics by D.C.Tayal

8.18 SUGGESTED READINGS

1. Electronic devices and circuits by Jacob Milliman and C.C.Halkais
2. The Feynman Lectures on Physics by Richard Feynman
3. Electronic devices and circuit theory by R.L.Boylestad and Louis N

8.19 TERMINAL QUESTIONS

8.19.1 Short Answer Type Questions

1. Why negative feedback is generally employed in amplifiers?
2. What is an oscillator? How does it differ from an amplifier?
3. What is barkhausan criterion for the feedback amplifier?
4. How does the colpitt's oscillator differ from Hartley's oscillator?
5. Why the bistable multivibrator is also called the flip flop multivibrator?

8.19.2 Long Answer Type Questions

1. Explain principle of feedback. Discuss the advantages of negative feedback in detail?
2. Draw the circuit diagram of colpitt's oscillator. Explain its working and establish an expression for its frequency of oscillation and condition for sustained oscillation.
3. Draw the circuit diagram of wein bridge oscillator. Explain its operation and derive an expression for its frequency of oscillation and condition for amplitude stabilization..
4. What are different types of multivibrators? Discribe with circuit diagram ,working and waveforms of astable multivibrator.
5. Write short notes on,

a) Phase shift oscillator
b) stable and quasi stable state
c) Crystal oscillator
d) Bistable multivibrator

8.19.3 Numerical Questions

1. A feedback amplifier has a voltage gain of 500 without feedback. If the feedback ratio is 0.1.Determine the voltage gain if negative feedback is used**.** (Ans. 9.8)
2. For a Hartley Oscillator it is given that If L_1=1 mH L_2=.1 mH and M=20μH .Calculate the operating frequency and feedback fraction for such Hartley oscillator.
(Ans. Frequency=1052KHz, feedback fraction=0.1)
3. In the phase shift oscillator if $R_1 = R_2 = R_3$ =1MΩ and $C_1 = C_2 = C_3$=100 pF. What will be its frequency of oscillation? (Ans. 650 Hz)
4. In a Quartz crystal oscillator circuit the parameters of its equivalent circuit are L= 0.5 H,C=0.05 pF,Cs=6 pF Cp=1 pF. calculate series and parallel resonant frequency.
(Ans. Series Frequency=1.03 MHz, parallel Frequency=1.0065 MHz)
5. In an astable multivibrator, the base resistors are of value 10 KΩ and the capacitors are of 0.01μF.Determine the pulse repeating rate. (Ans. 7000 Hz)

UNIT 9 NUMBER SYSTEMS AND BINARY CODES

Structure

9.1 INTRODUCTION

Digital electronics deals with the electronic circuits in which they perform operations by counting digits. A digital circuit is quite different from an analog circuit. In a digital systems one deal with only two possible states ‘zero’ and ‘one’, they basically represent ‘off and ‘on’ within a circuit. Digital Electronics has a number of applications in so many different fields. Ranging from something as simple as the Display of your digital watch or the Count-down timer which you see at traffic signals to something as complex as rocket science, quantum computing and many more. Now a days the applications of digital electronics are everywhere. Thus, the need of study to digital electronics is quite necessary for us.
In this chapter, we will discuss the basic concept of the binary numbers and their interconversion. We will also study the different types of available binary code and their interconversion.

9.2 OBJECTIVES

- Understand the concept of 'number system'
- Introduce basics of numbering systems
- Use decimal, octal, binary and hexadecimal
- Convert one number system to other
- Add octal, binary & hexadecimal numbers
- Understand the concept of 1’s and 2’s Complements
- Explain the common binary codes and their uses

9.3 NUMBER SYSTEM

A number system is system used for representing numbers in a certain form. The most common number system used in daily life is decimal number system (0, 1, 2, 3, 4, 5, 6, 7, 8, 9). Humans invented the decimal number system a few hundred years ago. This decimal number solved many purposes but by the development of machines and systems this number system is unable to solve the complex functions. So mathematician developed a number of other number systems to solve the above-mentioned problem. There are several other number systems also available. Such as binary, hexadecimal, octal, etc. In the following section, we will discuss the different types of available number system in detail:

9.3.1 Binary Numbers

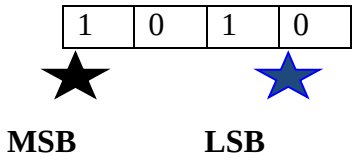
The simplest number system is binary number system. The concept of binary system was first introduced in the 1930s by George Boole. He was an English mathematician, logician and educator. Binary number consists of only two digits i.e., 0 (zero), 1 (one). As the binary system has only two digits, hence it has a base of two. The concept of binary number system was refined by Gottfried Wilhelm Leibniz and established that the binary system joined the principles of

arithmetic and logic. The position of 1, 0 in a binary number represent its weight within he number. In a binary number weight of each successive higher position is defined as the binary power of two. Binary number system follows the same rules as the frequently used decimal (base-10) number system. So as a substitute of $\mathbf{10^n}$ (powers of ten) in decimal system, $\mathbf{2^n}$ (powers of two) used binary system.

9.3.2 Decimal-Binary Conversion

For all computer and digital systems, the conversion of binary to decimal (base-2 to base-10) numbers and decimal to binary (base-10 to base-2) is an important concept.
In binary number the right hand most bit is known as the “Least Significant Bit” or “LSB”, and the left hand known as the Most Significant Bit or MSB.

Example 9.1: In a binary number 1010, 0 is LSB and 1 is MSB.



The conversion of decimal to binary is obtained by dividing the decimal number by 2 frequently until quotient 0 is observed.

Table 9.1 Converting a decimal number to a binary.

0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
10	1010

Steps to convert decimal to binary:

- Divide the decimal number by 2.
- Keep the remainder in an array.
- The above two steps have been repeated until the number is greater than zero.
- Write the array in reverse order.

Example 9.2: Conversion of 26_{10} to a hexadecimal number.

Division	Reminder
26/2=13	0
13/2=6	1
6/2=3	0
3/2=1	1
1/2=0	1

Put the remainder in reverse order: 11010_2

Result: $26_{10} = 11010_2$

For fractional decimal number to binary number conversion following steps are followed:

- Take the fractional part of the decimal number as multiplicand.
- Multiple fractional part of the decimal number by 2.
- Put the value of integer part of result in an array.
- Repeat the above-mentioned step until the number became zero.
- Print the array.

Example 9.3: Conversion of decimal fractional number 0.250 into binary number.

Multiplication	Carry
0.250 x 2=0.500	0
0.500 x 2=1.00	1
0.00 x 2=0.00	0

Put these resultant reminders, which is 0.010. Hence, the binary fractional number equivalent of decimal fractional 0.250 will be 0.010.

For conversion of binary to decimal the LSB position has been multiplied by 2^0 (1). Other higher significant position is multiplied by the ascending power of 2 (2^1, 2^2, 2^3, etc.).

Example 9.4: Conversion of 101101 binary number system into decimal number system.

$$
\begin{aligned}
101101_2 &= 1 \quad 0 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1 \\
&= 1 \times 2^6 \quad 0 \times 2^5 \quad 1 \times 2^4 \quad 1 \times 2^3 \quad 1 \times 2^2 \quad 0 \times 2^1 \quad 1 \times 2^0 \\
&= 1 \times 64 \quad 0 \times 32 \quad 1 \times 16 \quad 1 \times 8 \quad 1 \times 4 \quad 0 \times 2 \quad 1 \times 1 \\
&= 1 \times 64 \quad 0\ 32 \quad 1 \times 16 \quad 1 \times 8 \quad 1 \times 4 \quad 0 \times 2 \quad 1 \times 1 \\
&= 64+16+8+4+1 \\
&= 93
\end{aligned}
$$

Result: $101101_2 = 93_{10}$

Conversions of fractional binary to decimal have been done in a similar manner.

Example 9.5: Conversion of 0.1101 binary number into decimal number.

$$
\begin{aligned}
0.1101_2 &= 1 \quad 1 \quad 0 \quad 1 \\
&= 1 \times 2^{-1} \quad 1 \times 2^{-2} \quad 0 \times 2^{-3} \quad 1 \times 2^{-4} \\
&= 0.5000+0.2500+0.0000+0.0625 \\
&= 0.8125
\end{aligned}
$$

Result: $0.1101_2 = 0.8125_{10}$

9.3.3 Octal Numbers

The digits 0, 1, 2, 3, 4, 5, 6 and 7 have been used in the octal system. The base of the octal system is 8. The LSB position has a weight of 8^0 (1). Other higher significant positions are having weight as the ascending power of 8 (8^1, 8^2, 8^3, etc.).

Table 9.2 Converting a binary number to an octal number system.

Binary	Octal
000	0
001	1
010	2
011	3
100	4
101	5
110	6
111	7

The conversion of decimal to octal is obtained by dividing the decimal number by 8 frequently until quotient 0 is observed.
Steps to convert decimal to octal:

- Divide the decimal number by 8.
- Keep the remainder in an array.
- The above two steps have been repeated until the number is greater than zero.
- Write the array in reverse order.

Example 9.6: Conversion of 210_{10} to an octal number.

Division	Reminder
210/8=26	2
26/8=3	2
3/8=0	3

Put the remainder in reverse order: 322
Which is equivalent octal number of decimal integer 210.

For fractional decimal number conversion following steps are followed:

- Take the fractional part of the decimal number as multiplicand.
- Multiple fractional part of the decimal number by 8.
- Put the value of integer part of result in an array.
- Repeat the above-mentioned step until the number became zero.
- Print the array.

Example 9.7: Conversion of decimal fractional number 0.140869140625 into octal number.

Multiplication	Carry
0.140869140625 x 8=0.12695313	1
0.12695313 x 8=0.01562504	1
0.01562504 x 8=0.12500032	0
0.12500032 x 8=0.00000256	1
0.00000256 x 8=0.000020544	0

Put these resultant reminder, which is approximately 0.11010. Hence, the octal fractional number equivalent of decimal fractional 0.140869140625 will be 0.11010.

9.3.4 Octal-Binary Conversion

Conversion of octal to binary number system or binary to octal number system is quite easy task. For this type of conversion, we have to replace each significant digit by its 3 bit binary equivalent.

Example 9.8: Conversion of 256 octal number system into binary number system.

$$256_8 = 2 \quad 5 \quad 6$$
$$= 010 \quad 101 \quad 110$$

Result: $256_8 = 010101110_2$

For conversion of binary to octal the reverse procedure has been used. Starting from the LSB, We make group of 3bits and then covert their decimal equivalent.

Example 9.9: Conversion of 110101111 binary number system into octal number system.

$$110101111_2 = 110 \quad 101 \quad 111$$
$$= 6 \quad 5 \quad 7$$

Result: $10101111_2 = 657_8$

9.3.5 Hexadecimal Numbers

Hexadecimal number systems have a base of 16 and uses 16 symbols namely, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F. Decimal 10, 11, 12 , 13, 14, 15 are represented by A, B, C, D, E, F in hexadecimal number systems. The LSB position has a weight of 16^0 (1). Other higher significant position are having weight as the ascending power of 16 (16^1, 16^2, 16^3, etc.).

Table 8.3 Converting a decimal number to a hexadecimal number system.

Decimal	Binary	Hexadecimal
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7

8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F

The conversion of decimal to hexadecimal is obtained by dividing the decimal number by 16 frequently until quotient 0 is observed.

Steps to convert decimal to hexadecimal:

- Divide the decimal number by 16.
- Keep the remainder in an array.
- The above two steps have been repeated until the number is greater than zero.
- Write the array in reverse order.

Example 9.10: Conversion of 112_{10} to a hexadecimal number.

Divison	Reminder
112/16=7	0
7/16=0	7

Put the remainder in reverse order: 70_{16}

Result: $112_{10} = 70_{16}$

For fractional decimal number conversion following steps are followed:

- Take the fractional part of the decimal number as multiplicand.
- Multiple fractional part of the decimal number by 16.
- Put the value of integer part of result in an array.
- Repeat the above-mentioned step until the number became zero.
- Print the array.

9.3.6 Hexadecimal-Binary Conversion

Conversion of hexadecimal to binary number system or binary to hexadecimal number system is quite easy task. For this type of conversion, we have to replace each significant digit by its 4 bit binary equivalent.

Example 9.11: Conversion of $2B8_{16}$ hexadecimal number system into binary number system.

$$2B8_{16} = 2 \quad B \quad 8$$
$$= 0010 \quad 1011 \quad 1000$$

Result: $2B8_{16} = 001010111000_2$

For conversion of binary to hexadecimal the reverse procedure has been used. Starting from the LSB, we make group of 4 bits and then covert their decimal equivalent.

Example 9.12: Conversion of 110101111 binary number system into octal number system.

$$111011011111_2 = 1110 \quad 1101 \quad 1111$$
$$= E \quad D \quad F$$

Result: $10101111_2 = EDF_{16}$

9.4 BINARY ARITHMETIC

Binary arithmetic is necessary part of all the digital computers and digital system. Arithmetic rule for addition, subtraction, multiplication and division of binary numbers are the key for binary system. There are four rules of binary addition, subtraction, multiplication and division of binary numbers:

Table 9.4 Arithmetic rules for addition, subtraction, multiplication and division.

S. No.	Addition (A+B)	Subtraction	Multiplication	Division
1	0+0=0	0-0=0	0 X 0=0	0 ÷ 0= not allowed
2	0+1=1	1-0=1	0 X 1=0	0 ÷ 1=0
3	1+0=1	1-1=1	1 X 0=0	1 ÷ 0= not allowed
4	1+1=10	10-1=1	1 X 1=1	1 ÷ 1=1

9.4.1 Binary Addition

Two binary numbers can be added in a similar way as decimal number can be added. Addition of a binary numbers started with the LSB and then proceeded to MSB. Every time carry resulting from the LSB is added to the next bit.

Example 9.13: Binary addition of 1110 to 1111.

In 1110, 0 is LSB and 1 is MSB.

Steps for addition

1. Firstly, LSB of the binary numbers added, i.e. 1 + 0 = 1.
2. There is no carry. So next significant bits are added, i.e. 1 + 1 = 0 with 1 carry.
3. Carry obtained in the previous step is added to the next significant bits, i.e. 1 + 1 +1 = 1 with 1 carry.
4. Again, the carry obtained in the previous step is added to the next significant bits, i.e. 1 + 1 +1 = 1 with 1 carry.
5. Thus, the sum of the two binary digits is 11101.

9.4.2 Binary Subtraction

Binary subtraction is performed is done in a similar manner as binary addition. Subtraction is carried out with LSB and proceeded to next higher significant digits. If 1 is subtracted from 0 then 1 is borrowed from the higher significant digit.

Example 9.14: Binary subtraction of 1100 to 1010.

In 1110, 0 is LSB and 1 is MSB.

Steps for subtraction

1. Firstly, LSB of the binary numbers subtraction, i.e. 0 - 0 = 0.
2. Next significant bits are subtracted, i.e. 0 - 1. So we have to borrow a bit from the next higher significant bit. Which results 1 as difference and 1 borrow.
3. Move to the next significant bit. As 1 is borrowed, so 1 is changed to 0 here. Then performed the subtraction i.e. 0-0 = 0.
4. Again, move to the next significant bit and perform the subtraction. i.e. 1 - 1 = 0.
5. Thus, the subtraction of the two binary digits is 0010.

9.4.3 Binary Multiplication

The procedure of binary multiplication is quite similar as digital multiplication.

Steps for multiplication

1. At first LSB of binary bits are multiplied. If the multiplicand or multiplier is 0 the result will be 0. If the multiplier is 1 the result will be multiplicand itself.
2. Next significant bits are multiplied and product is written with a shift.
3. The process is repeated until we get the MSB.
4. Finally, all the product terms are added.

Example 9.15: Binary multiplication of 1100 to 1010.

$$\begin{array}{r} 1\ 1\ 0\ 0 \\ \times\ 1\ 0\ 1\ 0 \\ \hline 0\ 0\ 0\ 0 \\ 1\ 1\ 0\ 0 \\ 0\ 0\ 0\ 0 \\ 1\ 1\ 0\ 0 \\ \hline 1\ 1\ 1\ 1\ 0\ 0\ 0 \end{array}$$

9.4.3 Division

Binary division follows the same method as digital division. In case of binary division, the division by 0 is not allowed.

Example 9.16: Platform the following division.

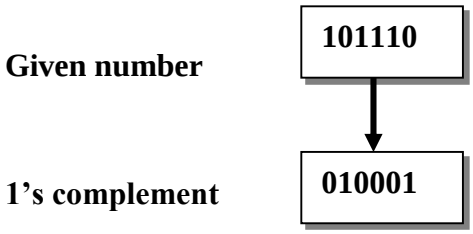
11110 ÷ 110

$$\begin{array}{r} 1\ 0\ 1 \\ \hline 110\overline{)\ 1\ 1\ 1\ 1\ 0} \\ 1\ 1\ 0 \\ \hline 1\ 1\ 0 \\ 1\ 1\ 0 \\ \hline 0\ 0\ 0 \end{array}$$

Therefore, the result is 101.

9.5 1'S AND 2'S COMPLEMENTS

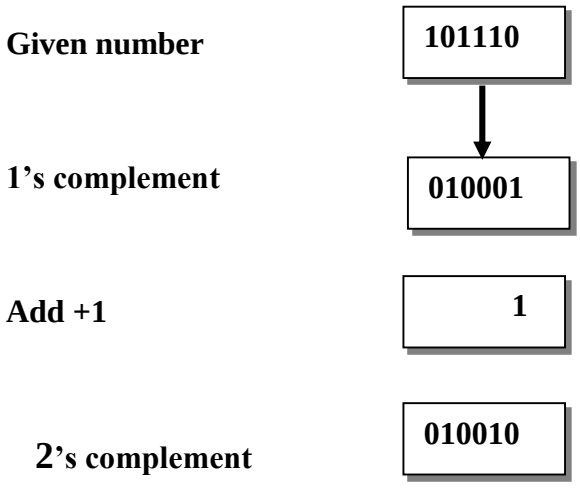
In the digital computers, to simplify the subtraction operation and for the logical manipulations complements are used. Basically, 1's complement of a binary number is obtained by changing all the bits in it, that is transforming the 1 bit to 0 bit and 0 bit to 1 bit.
For example: 1's complement of a binary number "101110" will be "010001".



The 2's complement of a binary number is obtained by adding 1 to the 1's complement of a binary number.

Example 9.17: 2's complement of a binary number "101110".

1. First calculate the 1's complement a binary number "101110" → 010001
2. Add 1 in the 1's complement → 010001+1=010010



9.5.1 1's Complement Subtraction

1's complement of a binary number is used for subtraction of binary numbers. This method allows us to subtract using addition.

For subtracting a smaller number form a larger number the following procedure follows:

1. Calculate the 1's complement of a smaller number
2. Then add the 1's complement to the largest number
3. Check the result of addition. If they have carry also known as "end around carry". Remove this carry and add this to the result.
4. If there is no end around carry, then final result will be the 1's complement of the result of addition and it is negative sign.

Example 9.18: Subtract 110101 – 10011 by 1's complement method.

Steps

1. 1's complement of 10011 is 011010.
2. 011010+110101=**1** 0 0 1 1 1 1
3. There is carry over, so result will be 0 0 1 1 1 1 + 1 =0 1 0 0 0 0

Example 9.19: Subtract 1010 – 1000 by 1's complement method.

Steps

1. 1's complement of 1010 is 0101.
2. 1000+0101=1 1 0 1
3. No carry over. Hence the difference is negative and is obtained by writing the 1's complement of 1 1 0 1. Hence, the difference is –0010.

9.5.2 2's Complement Subtraction

12's complement of a binary number is also used for subtraction of binary numbers. Similar to above mentioned method this method also allows us to subtract using addition.

For subtracting a smaller number, form a larger number the following procedure follows:

1. Calculate the 2's complement of a smaller number
2. Then add the 2's complement to the largest number
3. Check the result of addition. If the there is a carry. Discard the carry.
4. If there is no carry, the result is in 2's complement form with negative sign.

Example 9.20: Subtract 1010 – 1111 by 2's complement method.

Steps

1. 2's complement of 1010 is (1010 + 1) i.e. 0110.
2. 0110+1111= **1**0101
3. There is carry over, the result of subtraction is 0101.

Example 9.21: Subtract 10110 – 11010 by 2's complement method.
Steps

1. 2's complement of 11010 is (00101 + 1) = 00110.
2. 10110+00110= 1 1 1 0 0
3. There is no carry over, the result of subtraction is negative and is obtained by writing the 2's complement of 11100 i.e. (00011 + 1) or 00100. Hence the difference is – 100.

9.6 BINARY CODE

In the coding, "numbers, letters or words are represented by a specific group of symbols". These specific groups of symbols are called as a **binary code**. These binary codes are represented by the different number as well as alphanumeric letter. There are a number of advantages of binary codes, some of them are given below

Advantages of binary code:

There are a number of advantages of binary codes. Some advantages offered by binary codes are listed below.

1. We used binary codes for the computer application.
2. Binary codes are used for the analysis and designing computer digital circuit.
3. As in binary codes we use only 0, 1. So the performance becomes so easy.
4. Binary codes are suitable for digital communication.

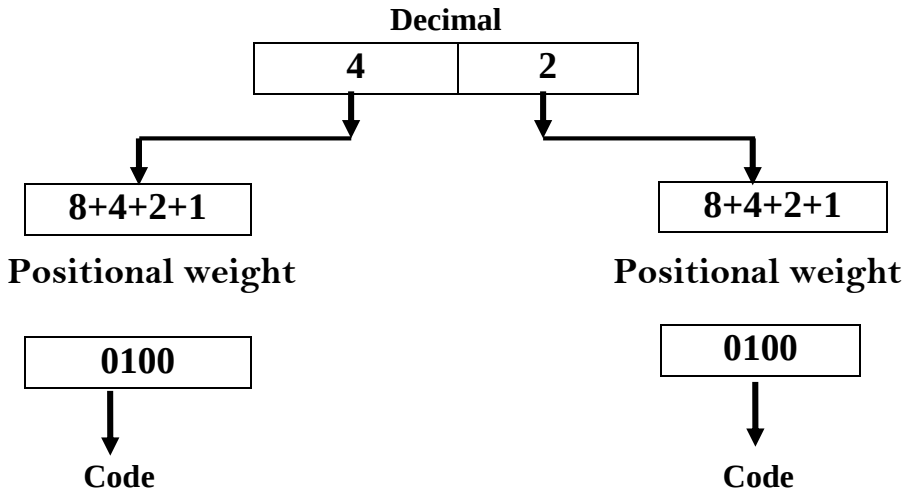
The binary codes are broadly

- Weighted Codes
- Non-Weighted Codes
- Binary Coded Decimal Code
- Alphanumeric Codes
- Error Detecting Codes
- Error Correcting Codes

In the following sections we have discuss only three types of binary codes.

9.6.1 Weighted Code

In the case of weighted code each digit of a code has its positional weight. A weighted code contains three parts in each code. The positional digit, Base or Radix(r), Exponents i.e. the positional weightage.



9.6.2 Non Weighted Code

In this type of code, each digit of the code does not assign any positional weight as in case of weighted code. There are a number of available non weighted codes, such as Excess3 code, Gray code etc. In the following section, we will discuss the Excess3 code, Gray code in detail.

9.6.2.1 Excess-3 Code

Excess-3 code is also known as XS-3 code. It is a non weighted code and used to express decimal number system. The Excess-3 code words are obtained from 8421 BCD code words by adding $(0011)_2$ or $(3)_{10}$. Table represent decimal number system their BCD and excess-3 code.

Decimal Number ⟶ 8421 BCD $\xrightarrow{\text{Add 0011}}$ Excess-3

Table 9.5 Converting BCD code to Excess-3 code.

Decimal	BCD 8 4 2 1	Excess-3 BCD+0011
0	0 0 0 0	0 0 1 1
1	0 0 0 1	0 1 0 0
2	0 0 1 0	0 1 0 1

3	0 0 1 1	0 1 1 0
4	0 1 0 0	0 1 1 1
5	0 1 0 1	1 0 0 0
6	0 1 1 0	1 0 0 1
7	0 1 1 1	1 0 1 0
8	1 0 0 0	1 0 1 1
9	1 0 0 1	1 1 0 0

9.6.2.2 Gray Code

Gray code is also a non-weighted code. The specific feature of that code is that only one bit will change each time the decimal number is incremented. Gray code is also known as unit distance code, as only one bit changes at a time as shown in table. It cannot be used for arithmetic operation. Gray code is used in shaft position encoders as it follows cyclic code. Cyclic code means one digit changes at one time. "A shaft position encoder produces a code word, which represents the angular position of the shaft".

Table 9.6 Converting a BCD to Gray code.

Decimal	BCD 8 4 2 1	Gray code
0	0 0 0 0	0 0 0 0
1	0 0 0 1	0 0 0 1
2	0 0 1 0	0 0 1 1
3	0 0 1 1	0 0 1 0
4	0 1 0 0	0 1 1 0
5	0 1 0 1	0 1 1 1
6	0 1 1 0	0 1 0 1
7	0 1 1 1	0 1 0 0
8	10 0 0	1 1 0 0
9	1 0 0 1	1 1 0 1

9.6.3 Binary Coded Decimal (BCD) code

Binary Coded Decimal is also known as BCD code. In BCD code, a 4-bit binary number represents each decimal digit. In the BCD, with four bits we can represent sixteen numbers (0000 to 1111). However, in BCD code only first ten of these are used (0000 to 1001). The remaining six code combinations i.e. 1010 to 1111 are invalid in BCD.

Table 9.7 Converting a decimal number to a BCD code.

Decimal	BCD 8 4 2 1
0	0 0 0 0
1	0 0 0 1
2	0 0 1 0
3	0 0 1 1
4	0 1 0 0
5	0 1 0 1
6	0 1 1 0
7	0 1 1 1
8	10 0 0
9	1 0 0 1

The main advantage of BCD codes is that it is very similar to decimal system. Therefore, we need to remember binary equivalent of decimal numbers 0 to 9 only. BCD Codes have their own set of disadvantages also. The BCD arithmetic is little more complicated. As BCD needs more number of bits than binary to represent the decimal number. So BCD is less efficient than binary.

9.7 CODE CONVERSION

A number of techniques can be used to convert a code from to another format. We will show here the following methods:

- Binary to BCD Conversion
- BCD to Binary Conversion
- BCD to Excess-3
- Excess-3 to BCD

1. Binary to BCD Conversion

Binary to BCD conversion includes only two steps:

Step 1 – Conversion of binary to decimal.
Step 2 – Conversion of decimal to BCD.

Example 9.22: convert $(11101)_2$ to BCD.

Step 1 – Conversion of $(11101)_2$ binary to decimal.
$(11101)_2 = ((1 \times 2^4) + (1 \times 2^3) + (1 \times 2^2) + (0 \times 2^1) + (1 \times 2^0))_{10}$
Binary − 11101_2 = Decimal − 29_{10}

Step 2 – Conversion of 29_{10} decimal to BCD.
$29_{10} = 0010_2\ 1001_2$
$29_{10} = 00101001_{BCD}$

Result: $(11101)_2 = 00101001_{BCD}$

2. BCD to Binary Conversion

BCD to Binary conversion includes only two steps:

Step 1 – Conversion of BCD to decimal.
Step 2 – Conversion of decimal to binary.

Example 9.23: convert $(00101001)_{BCD}$ to binary.

Step 1- Convert to $(00101001)_{BCD}$ BCD to binary
Convert a group of four digits and obtain decimal equivalent for each group.
$(00101001)_{BCD} = 0010_2\ 1001_2$
$(00101001)_{BCD} = 2_{10}\ 9_{10}$
$(00101001)_{BCD} = 29_{10}$
BCD − $(00101001)_{BCD}$ = Decimal − 29_{10}

Step 2: Convert decimal to binary.
Long division method has been used for decimal to binary conversion.
Decimal − 29_{10} = Binary − 11101_2

Result: $(00101001)_{BCD} = (11101)_2$

3. BCD to Excess-3 Conversion

BCD to Excess-3 conversion includes the following steps

- Step 1 – Conversion of BCD to decimal.
- Step 2 – Add $(3)_{10}$ to converted decimal.
- Step 3 – Finally binary is converted to get excess-3 code.

Example 9.24: Conversion of $(1001)_{BCD}$ to Excess-3.

Step 1 − Convert to decimal
$(1001)_{BCD} = 9_{10}$

Step 2 − Add 3 to decimal
$(9)_{10} + (3)_{10} = (12)_{10}$

Step 3 − Convert to Excess-3
$(12)_{10} = (1100)_2$

Result: $(1001)_{BCD} = (1100)_{XS-3}$

4. Excess-3 to BCD Conversion

Excess-3 to BCD conversion include only one step

- Step 1 -- Subtract $(0011)_2$ from each 4 bit of excess-3 digit to obtain the corresponding BCD code.

Example 9.25: convert $(10011010)_{XS-3}$ to BCD.

Given XS-3 number = 1 0 0 1 1 0 1 0
Subtract $(0011)_2$ = 1 0 0 1 0 1 1 1

BCD = 0 1 1 0 0 1 1 1

Result: $(10011010)_{XS-3} = (01100111)_{BCD}$

9.8 SUMMARY

In this unit, you have studied about different types of number systems, binary arithmetic, binary codes and their interconversion. To present the clear understanding of number systems, some commonly used number systems like decimal, binary, octal, hexadecimal number system have been discussed. You have studied that the base of decimal is 10, binary is 2, octal is 8 and hexadecimal is 16. You have also studied the interconversion of different number systems. For a binary system different arithmetic rules of addition, subtraction, multiplication, division have been

discussed in detail. The concept of binary codes and their interconversion have been elaborated. Basically, different types of available numbers, letters or words are represented by a specific group of symbols are known as codes. Among the available codes we have discussed weightage, non weightage and BCD code in detail. For understanding of different terminology glossary have been discussed. To make the concepts clear numerous solved examples are given in the unit. To check your progress, different terminal questions are given at the end of the unit.

9.9 GLOSSARY

Binary: Numerical value represented in the Binary or Base-2 number system.

Binary Coded Decimal (BCD): Four bit code used to port every digit of a decimal number by its 4 bit binary equivalent.

Digital State: It represent voltage condition existing. Normally interpret as "high" or "low".

Hexadecimal: Number system, which has a base of 16. Hexadecimal number can express as digits 0 through 9 plus letters A through F.

High: Logic level 1.

Low: Logic level 0.

Logic: The decision making capability of gate circuits in digital electronics.

Octal: Number system having a base value 8; digits from 0 to 7 are used to express an octal number.

9.10 REFERENCES

1. Digital Electronics (Oxford Higher Education) by G.K. Kharate
2. Digital Electronics, IGNOU, New Delhi
3. Digital Electronics by D.C. Green, Pearson
4. Digital Electronics: Principles and Applications by S K Mandal, McGraw Hill Education; 1 edition (1 July 2017)
5. Digital Electronics: Principles and Integrated Circuits by Anil K. Maini, Wiley; 1 edition (2007) A Textbook of Digital Electronic by S.S. Bhatti, Rahul Malhotra, I K International Publishing House Pvt. Ltd; 1st Edition 2011 edition (30 November 2011)

6. Electronic devices and circuits by S Salivahanan and Suresh Kumar, McGraw Hill Education

9.11 SUGGESTED READINGS

1. Digital Electronics by D.C. Green, Pearson
2. Electronic devices and circuits by S Salivahanan and Suresh Kumar, McGraw Hill Education
3. Digital Electronics by Morris Mano, Pearson Education

9.12 TERMINAL QUESTIONS

(Should be divided into Short Answer type, Long Answer type, Numerical type)

9.12.1 Short Answer type

1. What is the base of binary, hexa and octal number system?
2. What is meant by 1's and 2's Complements of a binary number?
3. What is a binary coded decimal?
4. Explain how BCD addition can be carried out?
5. Convert each decimal number to octal:

   (i) 16 (ii) 45 (iii) 270

9.12.2 Long Answer Type

1. What do you mean by binary code in digital electronic? Explain the different types of binary codes in detail.
2. Explain the concept of binary arithmetic in detail.

9.12.3 Numerical Answer Type

1. Convert Decimal Number 29_{10} to Binary Number.
2. Convert Binary Number 11101_2 to Decimal Number.
3. Convert Octal Number 25_8 to Decimal Number.

4. Convert Octal Number to 25_8 = Binary Number.
5. Convert Decimal Number to 21_{10} = Binary Number.
6. Convert Octal Number to 25_8 = Binary Number.
7. Convert Binary Number to 10101_2 = Hexadecimal Number.
8. Convert Hexadecimal Number to 15_{16} = Binary Number.
9. Convert $(11101)_2$ to BCD.
10. Convert $(1001)_{BCD}$ to Excess-3.
11. Convert $(10011010)_{XS-3}$ to BCD.

9.13 ANSWERS

Answer numerical type question

Ans 1: 11101_2 , 2: 29_{10} ,3: 21_{10}, 4: 10101_2, 5: 10101_2, 6 : 10101_2, 7: 15_{16}, 8: 10101_2, 9: $(00101001)_{BCD}$, 10: $(1100)_{XS-3,}$ 11: $(01100111)_{BCD}$

UNIT 10 BOOLEAN ALGEBRA AND LOGIC GATES

Structure

10.1 INTRODUCTION

In this chapter, we will discuss about logic gates, De Morgan's theorem, Demorganisation, Combinational circuits such as adder substractor. A logic gate is a building block of a digital circuit. Most logic gates have two inputs and one output and are based on Boolean algebra. In electronics, a logic gate is an idealized or physical device implementing a Boolean function ; that is, it performs a logical operation on one or more binary inputs and produces a single binary output.

10.2 OBJECTIVES

- To learn about logic gates
- To learn about Half-adder and Full-adder
- To learn about Half-adder and Full-adder
- To discuss Boolean laws
- To discuss De Morgan's theorem
- To Disorganisation

10.3 LOGIC GATS

A logic gate is a building block of a digital circuit. Most logic gates have two inputs and one output and are based on Boolean algebra. At any given moment, every terminal is in one of the two binary conditions *false* (high) or *true* (low). False represents 0, and true represents 1. Depending on the type of logic gate being used and the combination of inputs, the binary output will differ. A logic gate can be thought of like a light switch, wherein one position the output is off 0, and in another, it is on 1. Logic gates are commonly used in integrated circuits (IC).

In electronics, a logic gate is an idealized or physical device implementing a Boolean function; that is, it performs a logical operation on one or more binary inputs and produces a single binary output. Depending on the context, the term may refer to an ideal logic gate, one that has for instance zero rise time and unlimited, or it may refer to a non-ideal physical device.

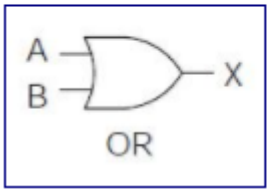
Logic gates are primarily implemented using diodes or transistor acting as electronic switches, but can also be constructed using vacuum tubes, electromagnetic relay (relay logic), fluidic logic, pneumatic logic, optics, molecules or even mechanical elements. With amplification, logic gates can be cascaded in the same way that Boolean functions can be composed, allowing the construction of a physical model of all of Boolean logic, and therefore, all of the algorithms and mathematics that can be described with Boolean logic.

Logic circuits include such devices as multiplexers, registers, arithmetic logic units (ALUs), and computer memory, all the way up through complete microprocessor, which may contain more than 100 million gates. In modern practice, most gates are made from MOSFETs (metal–oxide–semiconductor field effect transistor).

Compound logic gates AND-OR - Invent (AOI) and OR-AND-Invert (OAI) are often employed in circuit design because their construction using MOSFETs is simpler and more efficient than the sum of the individual gates.

10.3.1 OR Gate

The OR gate is an electronic circuit that gives a high output (1) if one or more of its inputs are high. A plus (+) is used to show the OR operation.



OR gate

Input A	Input B	Output
0	0	0
1	0	1
0	1	1
1	1	1

The output, Q of a "Logic OR Gate" only returns "LOW" again when ALL of its inputs are at a logic level "0". In other words, for a logic OR gate, any "HIGH" input will give a "HIGH", logic level "1" output.

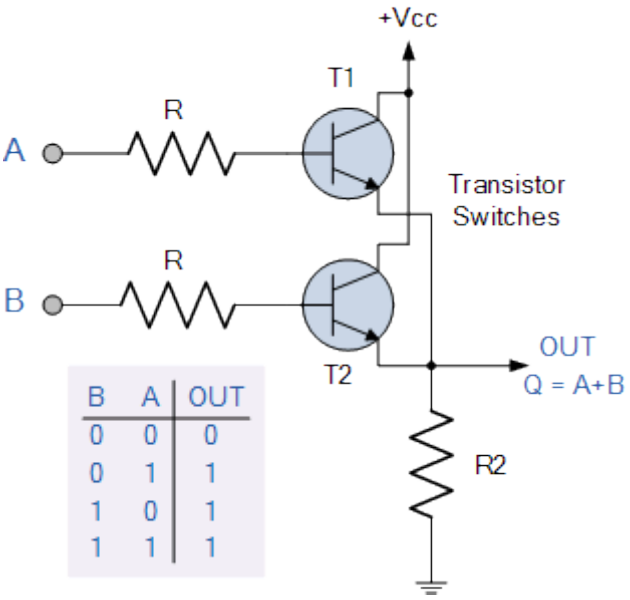
The logic or Boolean expression given for a digital logic OR gate is that for Logical Addition which is denoted by a plus sign, (+) giving us the Boolean expression of: A+B = Q.

Thus a logic OR gate can be correctly described as an "Inclusive OR gate" because the output is true when both of its inputs are true (HIGH). Then we can define the operation of a 2-input logic OR gate as being:

"If either A or B is true, then Q is true"

10.3.1.a 2-input Transistor OR Gate

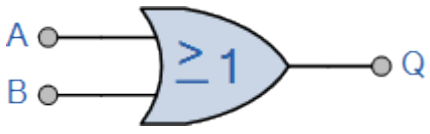
A simple 2-input inclusive OR gate can be constructed using RTL Resistor-transistor switches connected together as shown below with the inputs connected directly to the transistor bases. Either transistor must be saturated "ON" for an output at Q.



Logic OR Gates are available using digital circuits to produce the desired logical function and is given a symbol whose shape represents the logical operation of the OR gate.

10.3.1.b The 2-input Logic OR Gate

Symbol



2-input OR Gate

Truth Table

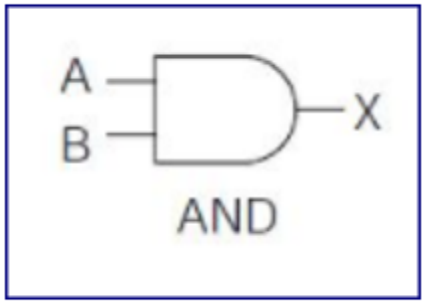
B	A	Q
0	0	0
0	1	1
1	0	1
1	1	1

Boolean Expression Q = A+B

Read as A OR B gives Q

10.3.2 AND Gate

The AND gate is an electronic circuit that gives a **high** output (1) only if all its inputs are high. A dot (.) is used to show the AND operation i.e. A.B. Bear in mind that this dot is sometimes omitted i.e. AB.



AND gate

Input A	Input B	Output
0	0	0
1	0	0
0	1	0
1	1	1

The output state of a digital logic AND gate only returns "LOW" again when ANY of its inputs are at a logic level "0". In other words for a logic AND gate, any LOW input will give a LOW output.

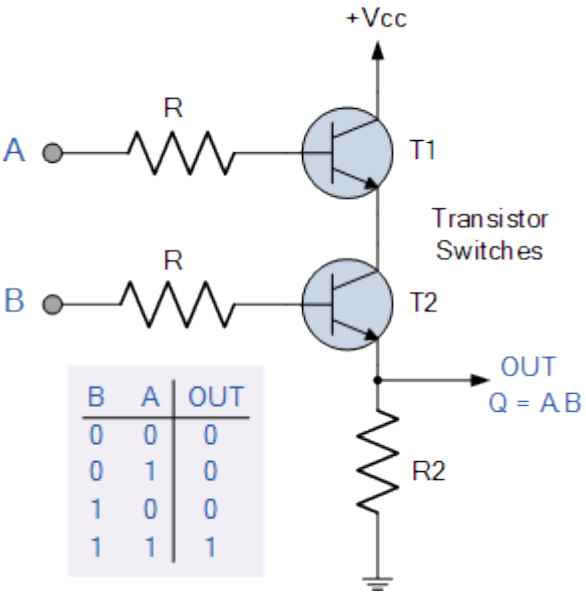
The logic or Boolean expression given for a digital logic AND gate is that for *Logical Multiplication* which is denoted by a single dot or full stop symbol, (.) giving us the Boolean expression of: A.B = Q.

Then we can define the operation of a digital 2-input logic AND gate as being:

"If both A and B are true, then Q is true"

10.3.2.a 2-input Transistor AND Gate

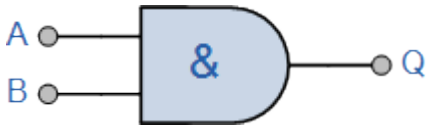
A simple 2-input logic AND gate can be constructed using RTL Resistor-transistor switches connected together as shown below with the inputs connected directly to the transistor bases. Both transistors must be saturated "ON" for an output at Q.



Logic AND Gates are available using digital circuits to produce the desired logical function and is given a symbol whose shape represents the logical operation of the AND gate.

10.3.2.b The 2-input Logic AND Gate

Symbol



2-input AND Gate

Truth Table

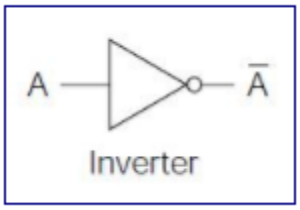
B	A	Q
0	0	0
0	1	0
1	0	0
1	1	1

Boolean Expression Q = A.B **Read as A AND B gives Q**

10.3.3 NOT Gate

The NOT gate is an electronic circuit that produces an inverted version of the input at its output. It is also known as an inverter. If the input variable is A, the inverted output is known as NOT A. This is also shown as A', or A with a bar over the top, as shown at the outputs. The diagrams below

show two ways that the NAND logic gate can be configured to produce a NOT gate. It can also be done using NOR logic gates in the same way.



Input	Output
0	1
1	0

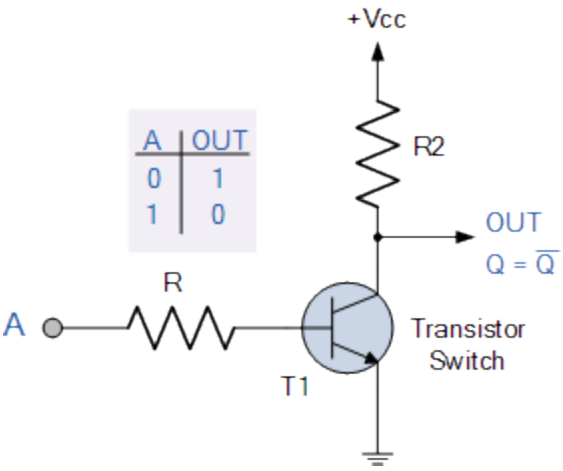
Inverting NOT gates are single input devices which have an output level that is normally at logic level "1" and goes "LOW" to a logic level "0" when its single input is at logic level "1", in other words it "inverts" (complements) its input signal. The output from a NOT gate only returns "HIGH" again when its input is at logic level "0" giving us the Boolean expression of: A = Q.

Then we can define the operation of a single input digital logic NOT gate as being:

"If A is NOT true, then Q is true"

10.3.3.a Transistor NOT Gate

A simple 2-input logic NOT gate can be constructed using a RTL Resistor-transistor switches as shown below with the input connected directly to the transistor base. The transistor must be saturated "ON" for an inverted output "OFF" at Q.

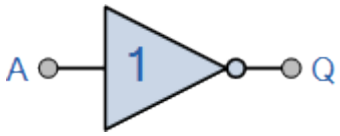


Logic NOT Gates are available using digital circuits to produce the desired logical function. The standard NOT gate is given a symbol whose shape is of a triangle pointing to the right with a circle at its end. This circle is known as an "inversion bubble" and is used in NOT, NAND and NOR

symbols at their output to represent the logical operation of the NOT function. This bubble denotes a signal inversion (complementation) of the signal and can be present on either or both the output and/or the input terminals.

The Logic NOT Gate Truth Table

Symbol



Inverter or NOT Gate

Boolean Expression Q = not A or $\overline{A}$

Truth Table

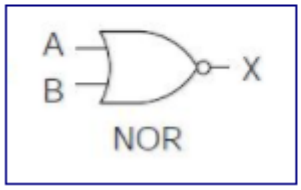
A	**Q**
0	**1**
1	**0**

Read as inverse of A gives Q

Logic NOT gates provide the complement of their input signal and are so called because when their input signal is "HIGH" their output state will NOT be "HIGH". Likewise, when their input signal is "LOW" their output state will NOT be "LOW". As they are single input devices, logic NOT gates are not normally classed as "decision" making devices or even as a gate, such as the AND or OR gates which have two or more logic inputs. Commercial available NOT gates IC's are available in either 4 or 6 individual gates within a single IC package.

10.3.4 NOR Gate

This is a NOT-OR gate which is equal to an OR gate followed by a NOT gate. The outputs of all NOR gates are low if any of the inputs are high.
The symbol is an OR gate with a small circle on the output. The small circle represents inversion.

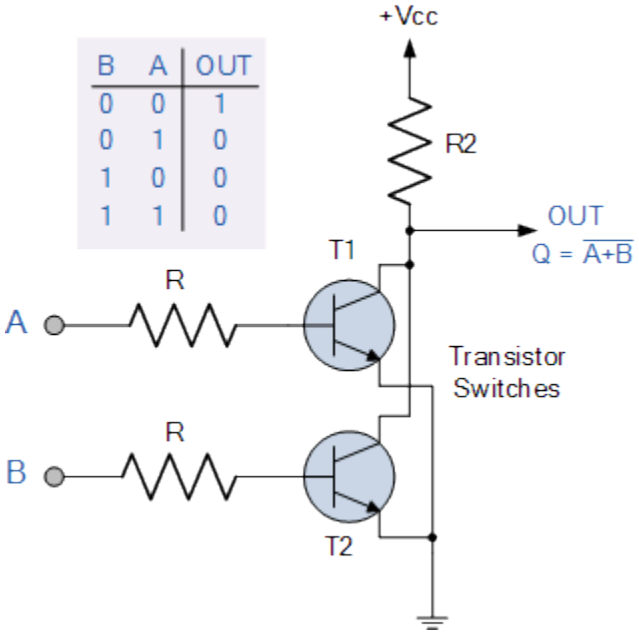


NOR gate

Input A	Input B	Output
0	0	1
1	0	0
0	1	0
1	1	0

10.3.4.a Transistor NOR Gate

A simple 2-input logic NOR gate can be constructed using RTL Resistor-transistor switches connected together as shown below with the inputs connected directly to the transistor bases. Both transistors must be cut-off "OFF" for an output at Q.

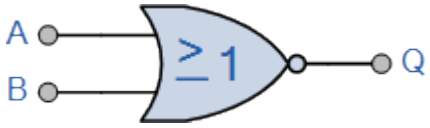


Logic NOR Gates are available using digital circuits to produce the desired logical function and is given a symbol whose shape is that of a standard OR gate with a circle, sometimes called an "inversion bubble" at its output to represent the NOT gate symbol with the logical operation of the NOR gate given as.

The Digital Logic "NOR" Gate

2- input NOR Gate

Symbol



2-input NOR Gate

Truth Table

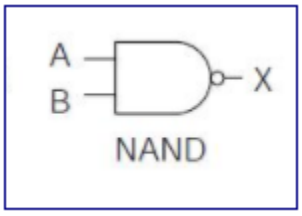
B	A	Q
0	0	1
0	1	0
1	0	0
1	1	0

1	0	0
1	1	0

Boolean Expression $Q = \overline{A+B}$ Read as A OR B gives NOT Q

10.3.5 NAND Gate

This is a NOT-AND gate which is equal to an AND gate followed by a NOT gate. The outputs of all NAND gates are high if any of the inputs are low. The symbol is an AND gate with a small circle on the output. The small circle represents inversion.



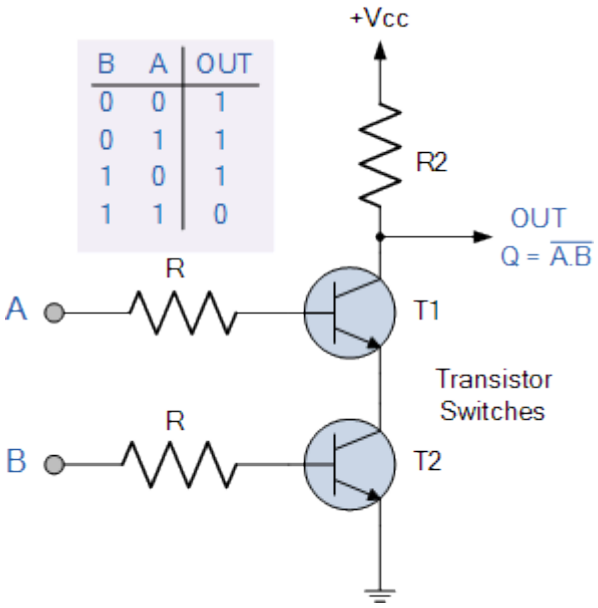
NAND gate

Input A	Input B	Output
0	0	1
1	0	1
0	1	1
1	1	0

The NAND (Not – AND) gate has an output that is normally at logic level "1" and only goes "LOW" to logic level "0" when ALL of its inputs are at logic level "1". The **Logic NAND Gate** is the reverse or "*Complementary*" form of the AND gate we have seen previously.

10.3.5.a Transistor NAND Gate

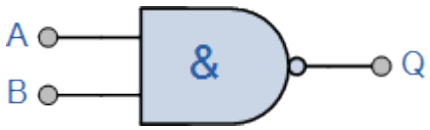
A simple 2-input logic NAND gate can be constructed using RTL Resistor-transistor switches connected together as shown below with the inputs connected directly to the transistor bases. Either transistor must be cut-off "OFF" for an output at Q.



Logic NAND Gates are available using digital circuits to produce the desired logical function and is given a symbol whose shape is that of a standard AND gate with a circle, sometimes called an "inversion bubble" at its output to represent the NOT gate symbol with the logical operation of the NAND gate given as.

10.3.5.b 2-input Logic NAND Gate

Symbol



2-input NAND Gate

Truth Table

B	A	Q
0	0	1
0	1	1
1	0	1
1	1	0

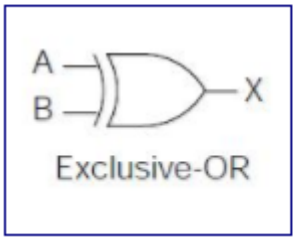
Boolean Expression $Q = \overline{A.B}$

Read as A AND B gives NOT Q

10.3.6 Exclusive-OR Gate

The exclusive-OR gate has a graphic symbol similar to the OR gate except for the additional curved line on the input side.

The output of the gate is 1 if any input is 1 but excludes the combination when both inputs are 1. It is similar to an odd function; that is, its output is 1 if an odd number of inputs are 1.



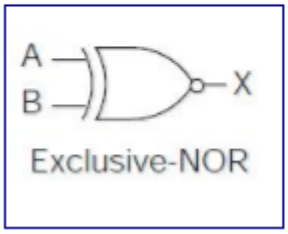
EX-OR gate

Input A	Input B	Output
0	0	0
1	0	1
0	1	1
1	1	0

10.3.7 Exclusive-NOR Gate

The exclusive-NOR is the complement of the exclusive-OR, as indicated by the small circle in the graphic symbol.

The output of this gate is 1 only if both the inputs are equal to 1 or both inputs are equal to 0.



EX-NOR gate

Input A	Input B	Output
0	0	1
1	0	0
0	1	0
1	1	1

10.4 BOOLEAN ALGEBRA

Boolean Algebra is used to analyze and simplify the digital (logic) circuits. It uses only the binary numbers i.e. 0 and 1. It is also called as Binary Algebra or logical Algebra. Boolean algebra was invented by George Boole in 1854.

10.4.1 Rule in Boolean Algebra

Following are the important rules used in Boolean algebra.

- Variable used can have only two values. Binary 1 for HIGH and Binary 0 for LOW.
- Complement of a variable is represented by an overbar (-). Thus, complement of variable B is represented as $\overline{B}$. Thus if B = 0 then $\overline{B}$ = 1 and B = 1 then $\overline{B}$ = 0.
- ORing of the variables is represented by a plus (+) sign between them. For example ORing of A, B, C is represented as A + B + C.
- Logical ANDing of the two or more variable is represented by writing a dot between them such as A.B.C. Sometime the dot may be omitted like ABC.

10.4.2 Boolean Laws

There are six types of Boolean Laws.

10.4.2.a Commutative law

Any binary operation which satisfies the following expression is referred to as commutative operation.

(i) $A.B = B.A$ (ii) $A + B = B + A$

Commutative law states that changing the sequence of the variables does not have any effect on the output of a logic circuit.

10.4.2.b Associative law

This law states that the order in which the logic operations are performed is irrelevant as their effect is the same.

(i) $(A.B).C = A.(B.C)$ (ii) $(A + B) + C = A + (B + C)$

10.4.3.c Distributive law

Distributive law states the following condition.

$A.(B + C) = A.B + A.C$

10.4.4.c AND law

These laws use the AND operation. Therefore, they are called as **AND** laws.

(i) $A.0 = 0$ (ii) $A.1 = A$

(iii) $A.A = A$ (iv) $A.\overline{A} = 0$

10.4.5.d OR law

These laws use the OR operation. Therefore, they are called as **OR** laws.

(i) $A + 0 = A$ (ii) $A + 1 = 1$

(iii) $A + A = A$ (iv) $A + \overline{A} = 1$

10.4.5.e INVERSION law

This law uses the NOT operation. The inversion law states that double inversion of a variable results in the original variable itself.

$\overline{\overline{A}} = A$

Self-assessment questions(SAQ)

1. Simplify the expression, using laws of Boolean algebra

   (AB+C) (AB+ D)

2. Simplify the expression, using laws of Boolean algebra
   AC + ABC =AC

10.5 DE MORGAN'S THEOREM

A mathematician named DE Morgan developed a pair of important rules regarding group complementation in Boolean algebra. By *group* complementation,

Theorem 1. The complement of the sum of two (or more) variables is equal to the product of the complements of the variables, i.e,

$$\overline{A + B} = \bar{A}.\bar{B}$$

Theorem 2. The complement of the product of two (or more) variables is equal the sum of complements of the variables i.e,

$$\overline{A.B} = \bar{A} + \bar{B}$$

Proof.

1. **To prove that $\overline{A+B} = \bar{A}.\bar{B}$**

Case 1. When $A = 0, B = 0$
Then $\overline{A+B} = \overline{0+0} = \bar{0} = 1$
And $\bar{A}.\bar{B} = \bar{0}.\bar{0} = 1.1 = 1$
Case 2. When $A = 0, B = 1$
Then $\overline{A+B} = \overline{0+1} = \bar{1} = 0$
And $\bar{A}.\bar{B} = \bar{0}.\bar{1} = 1.0 = 0$
Case 3. When $A = 1, B = 0$
Then $\overline{A+B} = \overline{1+0} = \bar{1} = 0$
And $\bar{A}.\bar{B} = \bar{1}.\bar{0} = 0.1 = 0$
Case 4. When $A = 1, B = 01$
Then $\overline{A+B} = \overline{1+1} = \bar{1} = 0$
And $\bar{A}.\bar{B} = \bar{1}.\bar{1} = 0.0 = 0$

We can see that in each case, the left hand side is equal to the right hand side. Hence the theorem is verified. Hence we conclude that $\overline{A+B} = \bar{A}.\bar{B}$

2. **To prove that $\overline{A.B} = \bar{A} + \bar{B}$**

Case 1. When $A = 0, B = 0$
Then $\overline{A.B} = \overline{0.0} = \bar{0} = 1$
And $\bar{A} + \bar{B} = \bar{0} + \bar{0} = 1 + 1 = 1$

Case 2. When $A = 1, B = 0$
Then $\overline{A.B} = \overline{1.0} = \bar{0} = 1$
And $\bar{A} + \bar{B} = \bar{1} + \bar{0} = 0 + 1 = 1$
Case 3. When $A = 0, B = 1$
Then $\overline{A.B} = \overline{0.1} = \bar{0} = 1$
And $\bar{A} + \bar{B} = \bar{0} + \bar{1} = 1 + 0 = 1$
Case 4. When $A = 1,\ B = 1$
Then $\overline{A.B} = \overline{1.1} = \bar{1} = 0$
And $\bar{A} + \bar{B} = \bar{1} + \bar{1} = 0 + 0 = 0$

We can see that in each case, the left hand side is equal to the right hand side. Hence the theorem is verified. Hence we conclude that $\overline{A.B} = \bar{A} + \bar{B}$

10.5 .a Demorganisation

Step 1: Remove overall NOT sign

Step 2: Change all Ors to ANDs and ANDs to ORs

Step 3: Complement all individual variables.

$$\overline{A + BC} = \bar{A}.\overline{BC}$$

$$= \bar{A}.(\bar{B} + \bar{C})$$

$$= \bar{A}.\bar{B} + \bar{A}.\bar{C}$$

Example 1.

consider

$$\overline{(A + B).(C + D)} = (\overline{A + B}) + (\overline{C + D})$$

$$= (\bar{A}.\bar{B}) + \bar{C}.\bar{D})$$

$$= \bar{A}.\bar{B} + \bar{C}.\bar{D}$$

For expression under the NOT sign, the opposite procedure would be followed. For example,

$$\overline{A + (B + C)} = \bar{A}.\overline{(B + C)}$$

$$= \bar{A}.(\bar{B}.\bar{C})$$

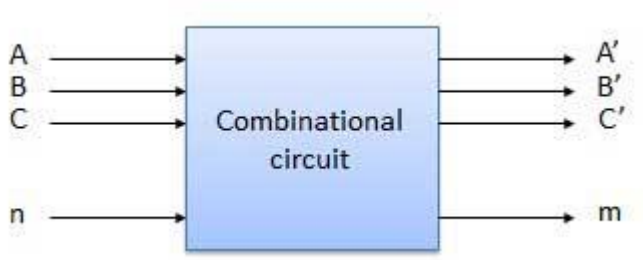
$$= \bar{A}.\bar{B}.\bar{C}$$

10.6 COMBINATIONAL CIRCUIT

Combinational circuit is a circuit in which we combine the different gates in the circuit, for example encoder, decoder, multiplexer and demultiplexer. Some of the characteristics of combinational circuits are following −

- The output of combinational circuit at any instant of time, depends only on the levels present at input terminals.
- The combinational circuit do not use any memory. The previous state of input does not have any effect on the present state of the circuit.
- A combinational circuit can have an n number of inputs and m number of outputs.

Block diagram

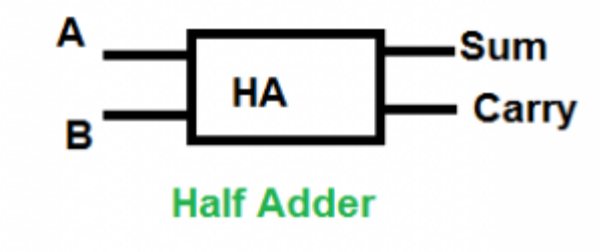


We're going to elaborate few important combinational circuits as follows.

10.6.a Half Adder

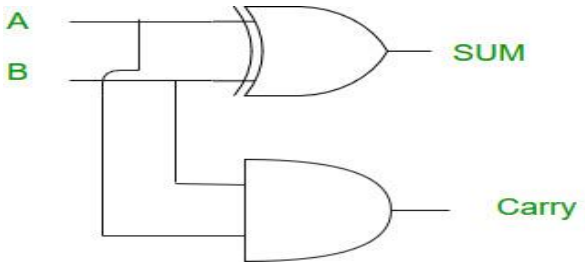
Half adder is a combinational logic circuit with two inputs and two outputs. The half adder circuit is designed to add two single bit binary number A and B. It is the basic building block for addition of two **single** bit numbers. This circuit has two outputs **carry** and **sum**.

Block diagram



\

Circuit Diagram



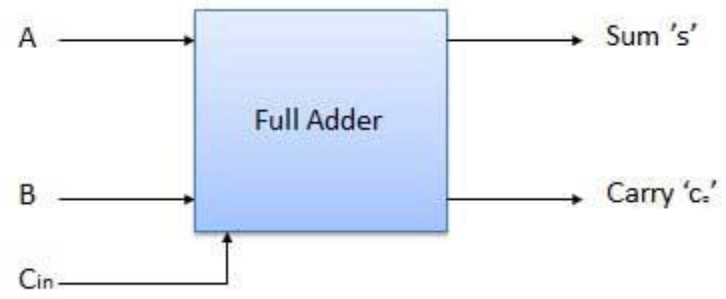
Truth Table

Input		Output
A	B	S C
0	0	0 0
0	1	1 0
1	0	1 0
1	1	1 1

10.6.b Full Adder

Full adder is developed to overcome the drawback of Half Adder circuit. It can add two one-bit numbers A and B, and carry c. The full adder is a three input and two output combinational circuit.

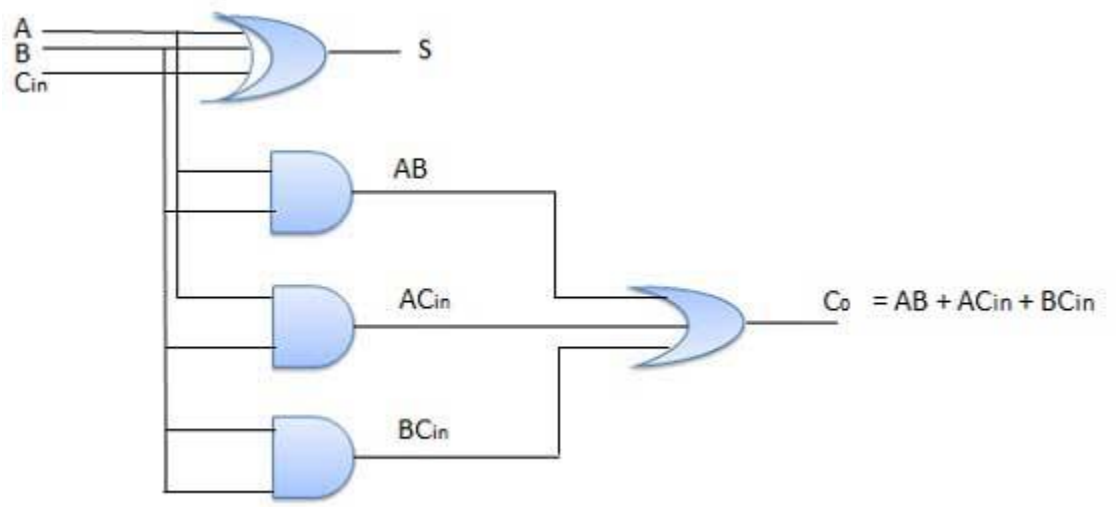
Block diagram



Truth Table

Inputs			Output
A	B	C_{in}	S C_O
0	0	0	0 0
0	0	1	1 0
0	1	0	1 0
0	1	1	0 1
1	0	0	1 0
1	0	1	0 1
1	1	0	0 1
1	1	1	1 1

Circuit Diagram



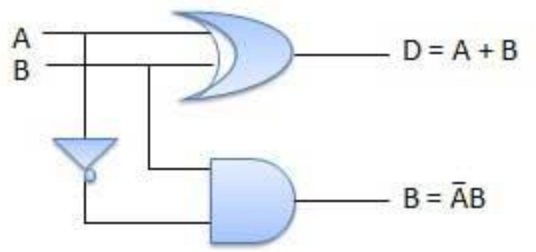
10.6.c Half Subtractors

Half subtractor is a combination circuit with two inputs and two outputs (difference and borrow). It produces the difference between the two binary bits at the input and also produces an output (Borrow) to indicate if a 1 has been borrowed. In the subtraction (A-B), A is called as Minuend bit and B is called as Subtrahend bit.

Truth Table

Inputs		Outputs	
A	B	(A-B)	Borrow
0	0	0	0
0	1	1	1
1	0	1	0
1	1	0	0

Circuit Diagram



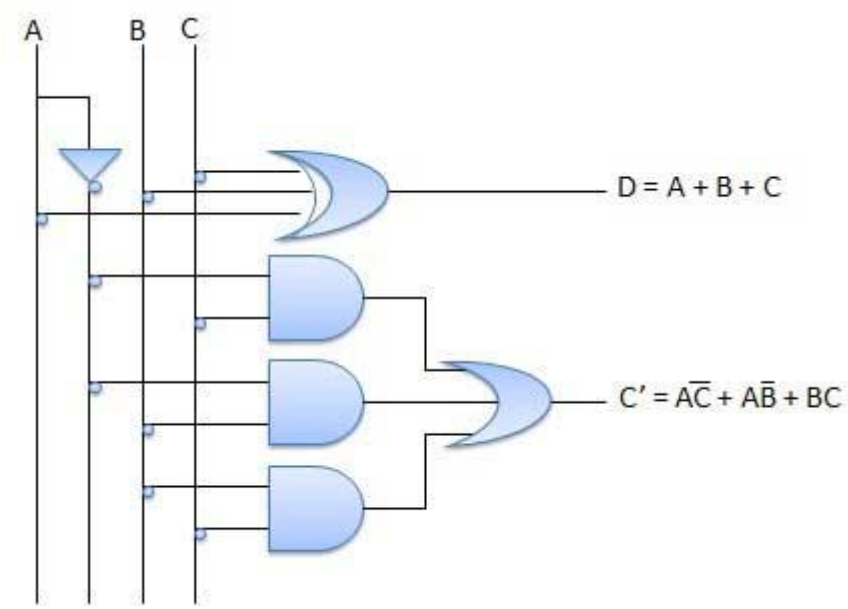
10.6.d Full Subtractors

The disadvantage of a half subtractor is overcome by full subtractor. The full subtractor is a combinational circuit with three inputs A,B,C and two output D and C'. A is the 'minuend', B is 'subtrahend', C is the 'borrow' produced by the previous stage, D is the difference output and C' is the borrow output.

Truth Table

Inputs			Outputs	
A	**B**	**C**	**A-B-C**	**C'**
0	**0**	**0**	**0**	**0**
0	**0**	**1**	**1**	**1**
0	**1**	**0**	**1**	**1**
0	**1**	**1**	**0**	**1**
1	**0**	**0**	**1**	**0**
1	**0**	**1**	**0**	**0**
1	**1**	**0**	**0**	**0**
1	**1**	**1**	**1**	**1**

Circuit Diagram



10.7 SUMMARY

In this chapter, we have learnt about logic gates devices. The logic gates become quite useful in digital operation. The Boolean expressions can be simplified using De Morgan's theorem.

10.8 TERMINAL QUESTIONS

10.8.1 Short Answer Type Questions

1. Prove the expression, using laws of Boolean algebra
   A+AB=A
2. Prove the expression, using laws of Boolean algebra
   $(A + B)(A + \bar{B})(\bar{A}$ +C) = AC
3. State and prove De Morgan's theorem?
4. What is Boolean algebra? Explain laws of Boolean algebra with suitable examples?

10.8.2 Objective Type Questions

1. When an input signal 1 is applied to a NOT gate, its output is

(a) 1
(b) 0
(c) Either 0 or 1
(d) Any positive value

2. The value of 1+1 is
   (a) 2
   (b) 0
   (c) 1
   (d) 10
3. A binary half adder is
   (a) Is a 2 bit adder
   (b) Is a 1 bit adder
   (c) Performs arithmetic addition
   (d) Has one output
4. Boolean expression for OR gate
   (e) A+ B
   (f) A-B
   (g) AB
   (h) None of these

Answer Objective type questions

1.b 2.c 3.a 4.a

LABORATORY COURSE

PHY (N)-350L

Experiment 1: To plot the characteristics of Zener diode and study of Zener diode as a voltage regulator

1.1 Object:

After performing this experiment, you should be able to

- What are semiconductor device
- Formation of Zener diode
- Characteristics of Zener diode
- Power regulation of Zener diode

1.2 Introduction:

Semiconductors are the material whose conductivity lies between conductor and insulator. Atoms of semiconductors are attached by covalent bonds. At absolute zero temperature semiconductor is insulator as no free electron to conduct and at higher temperatures semiconductor behave as conductor due free electrons which are produced by brakeage of some bonds. By Adding impurity of trivalent and pentavalent impurity, the pure semiconductor is made P and N type semiconductor respectively. In P type semiconductor there are holes as majority charge carrier and electrons in minority and in N type semiconductor electrons are in majority and holes are in minority charge carrier.

Zener diode is a heavily doped PN junction used in reverse bias. In Zener diode the current (reverse) remains constant over long voltage range until avalanche breakdown occurs. Due to this reason, Zener diode is used in voltage regulation

1.3 Apparatus Used:

Zener diode, microammeter, rheostat (100Ω), voltmeter, battery and connection wires.

1.4 Theory and Formula Used:

Zener diode is PN junction whose both P and N regions are heavily doped. Due to heavy doping depletion layer is narrow. When reverse voltage become more than Zener voltage and less than avalanche breakdown voltage, the current increases abruptly due to sudden increase of hole-electron pairs. The electron having energy less potential barrier has finite probability to penetrate barrier. It is quantum mechanical phenomena and potential barrier penetration of PN junction is explained by tunneling effect.

At Zener breakdown, large and saturated current flows in diode. It happens nearly at 20V and called Zener voltage. If the reverse voltage is increased further the current will not change because Zener voltage is sufficient to attract all the electrons produce at Zener breakdown. This property Zener diode can be used it in voltage regulator.

1.5 About apparatus:

The symbol of Zener diode is presented in figure1. The apparatus of Zener diode to study its characteristics is a circuit in which PN junction is revere biased and a resistance R_1 is connected to series with Zener diode. The input voltage is varied with the help of rheostat. The circuit diagram is presented in figure 2.

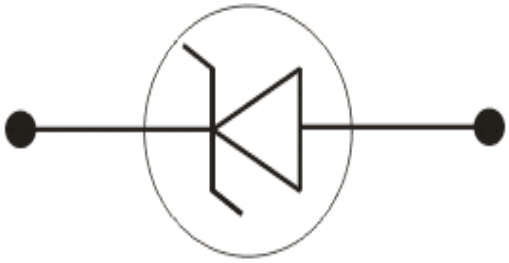


Figure 1

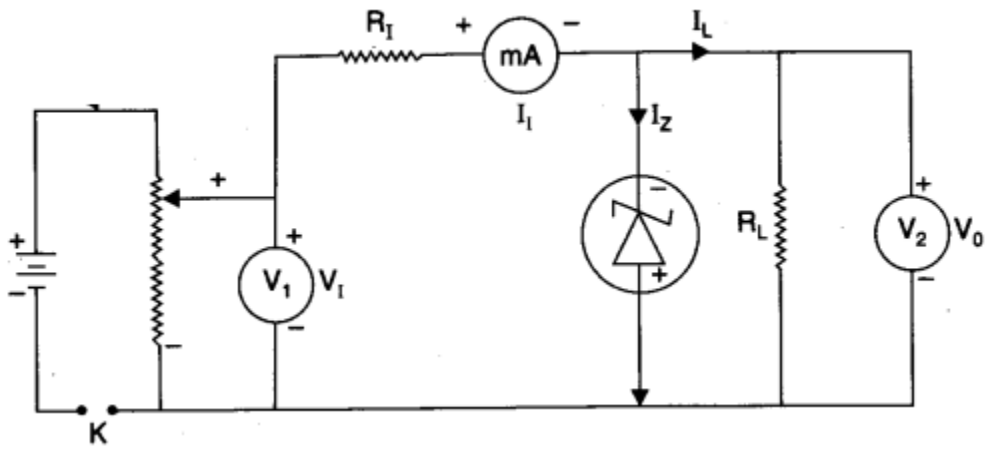


Figure 2 Reverse bias

1.6 Procedure:

For Zener characteristics

1. Make connection as shown in circuit diagram (figure 2).
2. With the help of rheostat vary the applied voltage in step of 1V.

3. Note down the ammeter reading in micrometer up to zener breakdown.
4. Draw the graph between voltage V and current I.

For Zener diode as voltage regulator (varying inputs)

1. Make connection as shown in circuit diagram (figure 2).
2. For a given input voltage, measure the output voltage across Zener diode.
3. Measure the output voltage across Zener diode on changing input voltage.
4. Draw the graph between output voltage V_O and input voltage V_I.

For Zener diode as voltage regulator (varying loads)

1. Make connection as shown in circuit diagram (figure 3).
2. For a given input voltage, measure the output voltage across Zener diode.
3. Measure the output voltage across Zener diode on changing load in output circuit R_L.
4. Draw the graph between load R_L and output voltage.

1.7 Observation:

Table1. For Zener characteristics

S.N.	Voltage V (In V)	Current I (In mA)
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

Table2. For Zener diode as voltage regulator (varying inputs)

S.N.	Input voltage V (In V)	Output voltage V (In V)
1		
2		
3		

4		
5		
6		

Table3. For Zener diode as voltage regulator (varying loads)

S.N.	Loads (In KΩ)	Output voltage V (In V)
1		
2		
3		
4		
5		
6		

1.8 Results:

1. The characteristics of Zener diode is presented the graph and Zener breakdown occurs at……volts.
2. The output voltage across Zener is constant on varying input voltage for the input range ……..
3. The output voltage across Zener in constant on varying loads in output circuit for load range…….

1.9 Precaution and source of error:

1. Sensitive voltmeter and sensitive ammeter sould be used.
2. The direction about maximum reverse voltage given by manufacturer should be strictly followed.
3. The graphs should drawn smothly.
4. There should not be any fluctuation on the power.
5. To avoid over heating of junction, current should not passed for long time.

1.10 Summary:

1. Zener diode is a heavily doped PN junction used in reverse bias.
2. Zener breakdown occurs before avalanche breakdown due to thin junction.
3. The voltage across Zener is constant on varying input voltage.
4. The voltage across Zener is constant on varying loads in output circuit.

1.11 Glossary

` Semiconductor: Materials that conduct to electricity between conductor and insulator.

Extrinsic semiconductor: Impure semiconductor for increasing conductivity.

Doping: Adding impurity on pure semiconductor.

Tunnel effect: It is quantum mechanical phenomena for penetration through potential barrier. The electron having energy less potential barrier has finite probability to penetrate the barrier.

Voltage regulation: The constant output for varying inputs or loads.

Loads: High resistance in output circuit.

1.12 References:

1. C. L. Arora, *B.Sc. Practical Physics*, S. Chand publication, Delhi.
2. C.L. Arora and P.S. Hemne, *Physics for Degree students (BSc III year)*, S. Chand publication, Delhi.
3. S.L.Gupta, V.Kumar *A Hand book of electronics* Pragati prakashan, Meerut.
4. B.L.Thereja *A Hand book of electronics* S. Chand publication, Delhi.
5. V.K.Mehta and Rohit Mehta,*Principle of electronics* S. Chand publication, Delhi.
6. S.L.Gupta, V.Kumar, *Practical Physics*, Pragati prakashan, Meerut.
7. *https://en.wikipedia.org.*

1.13 Viva-voce questions:

Question1. What is meant Zener breakdown?

Answer: A little current flows in Zener diode on applying voltage. At a voltage that is less than avalanche breakdown voltage, the current increases abruptly. This is called Zener breakdown

Question2. Is there any difference between Zener diode and ordinary PN junction diode?

Answer: Yes, Zener diode is heavily doped to achieve sharp breakdown before avalanche breakdown.

Question3.What is voltage regulation?

Answer: Voltage regulation ability of instrument to maintain constant output for varying load or input.

Experiment 2: To draw output characteristics of NPN transistor in common emitter configuration

2.1 Object:

After performing this experiment, you should be able to

- What is NPN transistor?
- How output characteristics of NPN transistor is drawn.
- What is use of NPN transistor?

2.2 Introduction:

With the advance of semiconductor the NPN and PNP transistor are most important device, which are used in amplification. As we know that P and N type semiconductor are made by doping trivalent and pentavalent impurity in intrinsic semiconductor respectively. The NPN transistor is made by sandwiching P type semiconductor by two N type semiconductor on either side. The middle portion of NPN transistor is called base and either portion is called emitter and collector. The base of NPN transistor is thin and lightly doped, while emitter is highly doped and collector is lightly doped and have wider region. As requirements the NPN transistor can be use in common emitter, common base and common collector configuration. NPN and PNP transistors are called bipolar devices as both hole and electron are responsible for conduction inside transistor.

2.3 Apparatus Used: NPN transistor, battery, DC Voltmeter (0-10V), DC ammeter (0-50 mA) and rheostat (100Ω).

2.4 Theory and Formula Used:

Common emitter configuration is most used configuration in amplification as its amplification factor (β) is very high. The circuit diagram of NPN transistor is presented in figure. In circuit the base-emitter junction is kept forward bias while emitter-collector junction is kept reverse bias.

Two N regions have mobile electrons (negative) as majority and hole as minority, while in middle section p region have mobile holes (positive) in majority and electrons as minority. As shown in figure 2, the N-P junction (emitter base junction) is made small forward bias and P-N junction (base collector junction) is made reverse bias. Under the forward bias the electrons in emitter move toward base, while holes move toward base o emitter. Since base is thin, most of electrons pass on collector and few of them combine with hole constitute base current from base to emitter. Electrons reaching collector is attracted by positive voltage given to collector and thus constituting collector current. There are following two characteristics of NPN transistor in common emitter configuration.

(a) Input characteristics: To draw the input characteristics, the collector voltage V_c is made zero. The emitter voltage V_e is increased from zero onward and corresponding emitter current I_e is noted. Graph is plotted between emitter voltage V_e and emitter current I_e. Another graph is plotted for constant collector voltage V_c say for 25 volt. These curves show **input characteristics** of NPN transistor.

(b) Output characteristics: These characteristics are obtained by plotting collector current I_C versus collector-emitter voltage V_C at a fixed value of base current I_B. The base current is changed to some other fixed value and the observations of I_C versus V_{CE} are repeated .All the represents the output characteristics of a common-emitter circuit.

2.5 About apparatus:

The apparatus in the experiment is a NPN transistor as shown in figure 1. The base-emitter junction is forward bias by connecting 6V battery (i.e. P of junction is given +ve voltage and N is given –ve voltage) and bias while emitter-collector junction is kept reverse bias (i.e. P of junction is given -ve voltage and N is given +ve voltage). For the measurement of base current (input current) and collector current (output current) the microammeter and milliammeter are used in circuit.

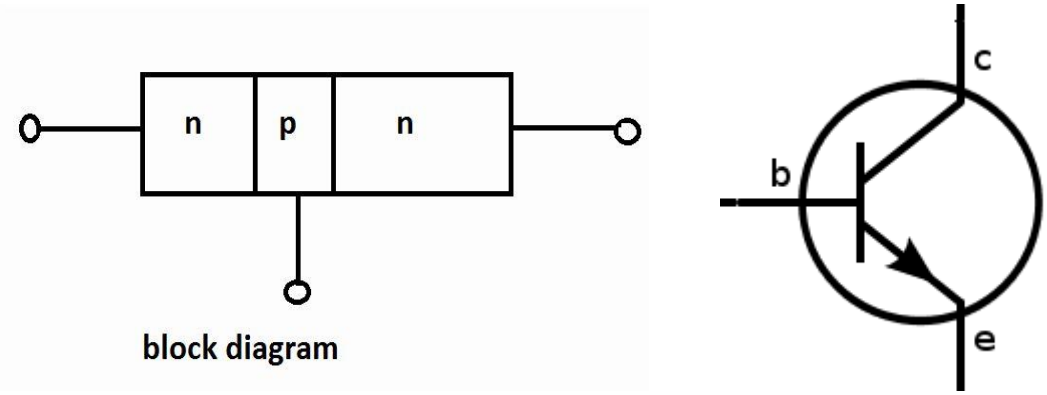


Figure 1

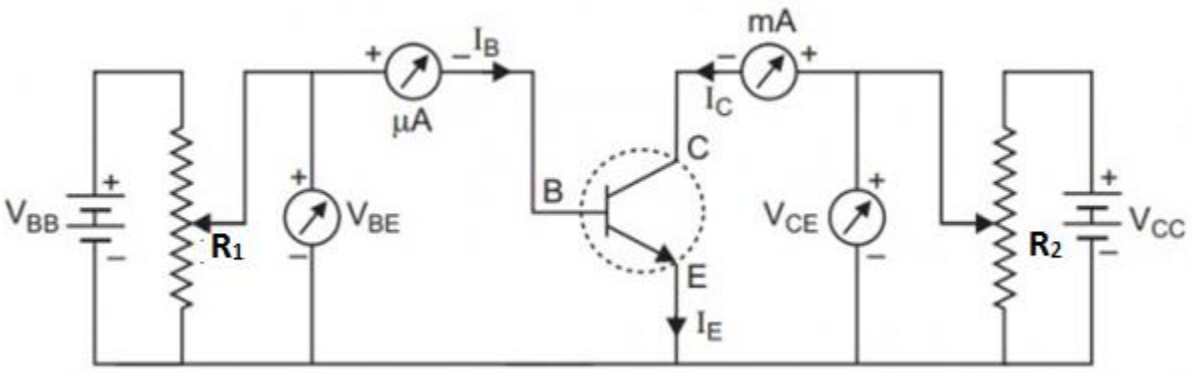


Figure 2

2.6 Procedure:

Let us perform the experiment in following steps.

1. Make connection as shown in figure 2.
2. Adjust base current say at 50 µA by means of R_1.
3. Increase collector voltage in step of 1 volt by changing R_2 and note the corresponding collector current.
4. Now note collector current with changing collector voltage for constant base current say for 75 µA, 100 µA, and 125 µA.
5. Plot curve between collector voltage and collector current at particular constant base current.

2.7 Observation:

Table for Output characteristics (I_C/V_C)

S.N.	Collector Voltage V_C (In volt)	Collector current I_C in mA for base current			
		I_b=50 µA	I_b=75 µA	I_b=100 µA	I_b=125 µA
1					
2					
3					
4					
5					
6					
7					
8					

2.8 Result: The output characteristics plotted between collector voltage V_C and collector I_C at constant base current I_b are plotted and presented in the graph .

2.9 Precaution and source of error:

1. Sensitive voltmeter and sensitive ammeter sould be used.
2. Specification about apparatus as given by manufacturer should be take care.
3. There should be proer biasing while performing experiment.
4. There should be not be fluctuation of power

2.10 Summary:

1.In Commen emitter transistor, the emitter section is used in both input circuit and output circuit.

2.Emitter-base junction is made forword bias.

3. Emitter-collector junction is made reverse bias.

4.The collector current increases on increase base current.

5. Base current of order 100μA and collector current mA flows in transistor

2.11 Glossary

Semiconductor: Material whose conductivity lies between conductor and insulator.

Intrinsic semiconductor: Si and Ge in their pure form.

Current amplification factor: It is the ratio of output current (collector current) and input current (base current).

Amplifier: A device which raises the strength of weak signal.

2.12 References:

1. C. L. Arora, *B.Sc. Practical Physics*, S. Chand publication, Delhi.
2. C.L. Arora and P.S. Hemne, *Physics for Degree students (BSc III year)*, S. Chand publication, Delhi.
3. S.L.Gupta, V.Kumar *A Hand book of electronics* Pragati prakashan, Meerut.
4. B.L.Thereja *A Hand book of electronics* S. Chand publication, Delhi.
5. V.K.Mehta and Rohit Mehta,*Principle of electronics* S. Chand publication, Delhi.
6. S.L.Gupta, V.Kumar, *Practical Physics*, Pragati prakashan, Meerut.
7. *https://en.wikipedia.org.*

2.13 Viva-voce questions:

Question1.What is a semiconductor transistor?

Answer: It is a semiconductor device having three sections namely emitter, base and collector.

Question2. What is thickness of base section and what is region of it?

Answer: The thickness of base region is of order 10μ (10^{-5}M). It is made small so that minimum (2 to 3%) recombination of hole electron pair take place.

Question3.In how many ways a transistor can be used?

Answer. The transistor can be used in following three ways

(i) Common base configuration (CB)
(ii) Common emitter configuration(CE)
(iii) Common collector configuration(CC)

Question4. What is current gain in CE transistor?

Answer: It the ratio of collector current to base current. Its value is about 50.

Question5. Why CE configuration is preferred than CB configuration?

Answer: Due to its large current gain in CE than CB configuration

Question6. What is order of current in CE transistor?

Answer: The base current is order of 100μA, emitter & collector current is of order mA .

.

Experiment 3: To draw output characteristics of PNP transistor in common base configuration

3.1 Object:

After performing this experiment, you should be able to

- What is PNP transistor?
- How common base connection is made.
- How output characteristics of PNP transistor is drawn.
- What is use of PNP transistor?

3.2 Introduction:

With the advance of semiconductor the NPN and PNP transistor are most important device, which are used in amplification. As we know that P and N type semiconductor are made by doping trivalent and pentavalent impurity in intrinsic semiconductor respectively. The PNP transistor consists of a very thin slice N type semiconductor diffused by two by two type semiconductor on either side. The middle portion of PNP transistor is called base and either portion is called emitter and collector. The base of PNP transistor is thin and lightly doped, while emitter is highly doped and collector is lightly doped and have wider region. As requirements the NPN transistor can be use in common emitter, common base and common collector configuration. NPN and PNP transistors are called bipolar devices as both hole and electron are responsible for conduction inside transistor. The block diagram and symbols of PNP transistor is shown in figure1.

3.3 Apparatus Used:

PNP transistor, battery, DC Voltmeter (0-10V), DC ammeter (0-50 mA) and Rheostat (100Ω).

3.4 Theory and Formula Used:

The curves representing the variation of current with voltage in PNP transistor is called transistor characteristics. In Common base configuration the base is made common to both input and output. The circuit diagram of PNP transistor is presented in figure 2. In circuit the base-emitter junction (input circuit) is kept forward bias while base-collector (output circuit) junction is kept reverse bias. The voltage gain in this configuration is low and less than unity while voltage gain and power gain is high.

Two P regions have mobile holes (positive) as majority and electrons (negative) as minority, while in middle section N region have mobile electrons (negative) in majority and holes (positive) as minority. As shown in figure 2, the P-N junction (emitter base junction) is made small forward bias and N-P junction (base collector junction) is made reverse bias. Under the forward bias the

holes in emitter move toward base, while electrons move toward base to emitter. Since base is thin, most of holes pass on collector and few of them combine with electrons constitute base current from emitter to base. Holes reaching collector is attracted by negative voltage given to collector and thus constituting collector current. There are two characteristics of PNP transistor as following

(a) Input characteristics: To draw the input characteristics, the collector voltage V_c is made zero. The base voltage V_b is increased from zero onward and corresponding emitter current I_e is noted. Graph is plotted between base voltage V_b and emitter current I_e. Another graph is plotted for constant collector voltage V_c say for 25 volt. These curves show input characteristics of PNP transistor.

(b) Output characteristics: These characteristics are obtained by plotting collector current I_C versus collector voltage V_C at a fixed value of emitter current Ie. The base current is changed to some other fixed value and the observations of I_C versus V_{CB} are repeated represents the output characteristics of a common-emitter circuit.

3.5 About apparatus:

The apparatus in the experiment is a PNP transistor as shown in figure 1. The base-emitter junction is forward bias by connecting 6V battery (i.e. P of junction is given +ve voltage and N is given –ve voltage) and bias while base-collector junction is kept reverse bias (i.e. P of junction is given -ve voltage and N is given +ve voltage). For the measurement of base current (input current) and collector current (output current) the micro ammeter and milliammeter are used in circuit.

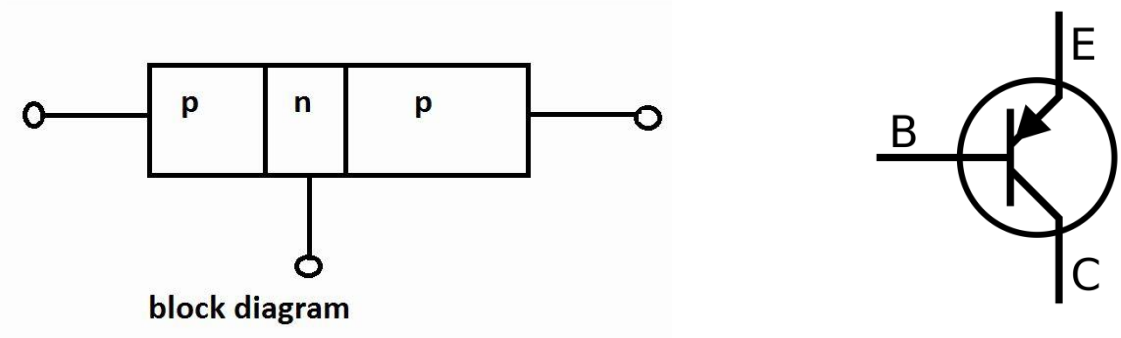


Figure 1

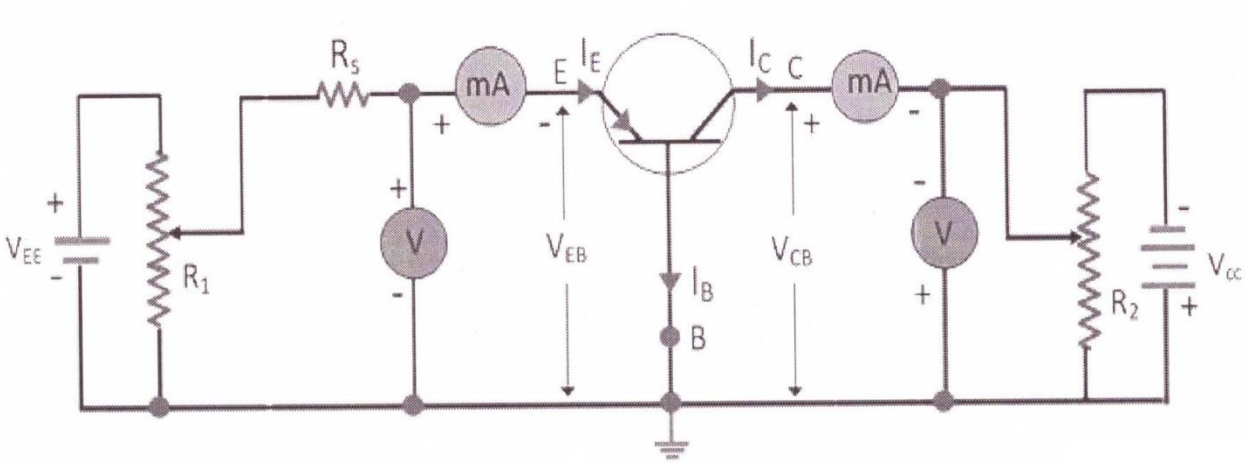


Figure 2

3.6 Procedure:

Let us perform the experiment in following steps.

1. Make connection as shown in figure 2.
2. Adjust emitter current say at 2 mA by means of R_1.
3. Increase collector voltage in step of 1 volt by changing R_2 and note the corresponding collector current.
4. Now note collector current with changing collector voltage for constant emitter current say for 4 mA, 6 mA, and 8 mA.
5. Plot curve between collector voltage and collector current at particular constant base current.

3.7 Observation:

Table for Output characteristics (I_C/V_C)

S.N.	Collector Voltage V_C (In volt)	Collector current I_C in mA for emitter current			
		I_e=2 mA	I_e=4 mA	I_e=6 mA	I_e=8 mA
1					
2					
3					
4					
5					
6					
7					
8					

3.8 Result:

The output characteristics plotted between collector voltage V_C and collector I_C at constant emitter current I_e are plotted and presented in the graph .

3.9 Precaution and source of error:

1. Sensitive voltmeter and sensitive ammeter sould be used.
2. Specification about apparatus as given by manufacturer should be take care.
3. There should be proer biasing while performing experiment.
4. There should be not be fluctuation of power

3.10 Summary:

1. In Commen baser transistor, the base section is used in both input circuit and output circuit.
2.Emitter-base junction is made forword bias.
3. Base-collector junction is made reverse bias.
4.The collector current increases on increase emitter current.
5. Base current of order 100μA and emitter & collector current mA flows in transistor.
6. The current gain is low and less than unity in common base configuration.
7. The voltage gain and power gain is high in common base configuration.

3.11 Glossary

Semiconductor: Material whose conductivity lies between conductor and insulator.

Intrinsic semiconductor: Si and Ge in their pure form.

Current gainr: It is the ratio of output current (collector current) and input current (eitter current).

3.12 References:

1. C. L. Arora, *B.Sc. Practical Physics*, S. Chand publication, Delhi.
2. C.L. Arora and P.S. Hemne, *Physics for Degree students (BSc III year)*, S. Chand publication, Delhi.
3. S.L.Gupta, V.Kumar *A Hand book of electronics* Pragati prakashan, Meerut.
4. B.L.Thereja *A Hand book of electronics* S. Chand publication, Delhi.
5. V.K.Mehta and Rohit Mehta,*Principle of electronics* S. Chand publication, Delhi.
6. S.L.Gupta, V.Kumar, *Practical Physics*, Pragati prakashan, Meerut.

3.13 Viva-voce questions:

Question1.What is a semiconductor transistor?

Answer: It is a semiconductor device having three sections namely emitter, base and collector.

Question2. What is thickness of base section and what is region of it?

Answer: The thickness of base region is of order 10μ (10^{-5}M). It is made small so that minimum (2 to 3%) recombination of holes and electrons take place.

Question3.In how many ways a transistor can be used?

Answer. The transistor can be used in following three ways

(iv) Common base configuration (CB)
(v) Common emitter configuration(CE)
(vi) Common collector configuration(CC)

Question4. What is current gain in CB transistor?

Answer: It the ratio of collector current to emitter current and is low (less than unity) in CB configuration.

Question4. What is relation of current gain in CB and CE configuration of transistor?

Answer: $\beta = \frac{\alpha}{1-\alpha}$, symbols are in usual meaning.

Experiment 4: To measure the hybrid parameters of PNP transistor in common emitter configuration

4.1 Object:

After performing this experiment, you should be able to

- What is common emitter configuration
- What are hybrid parameters
- What are unit of hybrid parameters

4.2 Introduction:

Every linear circuit having input and output terminals can be analyzed by four parameter (one measured in ohm, one in mho and two dimensionless) called hybrid parameter. Hybrid means mixed. Since these parameters have mixed dimensions so called hybrid parameters

Let consider a linear circuit shown in figure1 has input voltage (V_1) and current (I_1) and output voltage (V_2) and current (I_2)

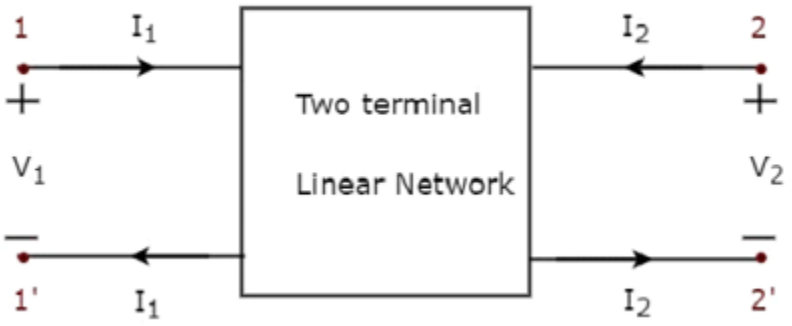


Figure1

Let $$V_1 = h_{11}I_1 + h_{12}V_2$$

$$I_2 = h_{21}I_1 + h_{22}V_2$$

In above equation h's are called hybrid parameter and it is fixed for given circuit. All the hybrid parameters can be determined by

(i) By putting $V_2 = 0$ (output circuit is short circuited)

$h_{11} = \frac{V_1}{I_1}$, Input impedance

$h_{21} = \frac{I_2}{I_1}$, Forward current gain

(ii) By putting $I_1 = 0$ (input circuit is open circuited)

$h_{12} = \frac{V_1}{V_2}$, Reverse voltage gain

$h_{21} = \frac{I_2}{I_1}$, output admittance

Two of above parameters h_{12} and h_{21} are dimensionless. Other two h_{11} and h_{22} have dimension of impedance and admittance.

4.3 Apparatus Used:

Transistor AC 126, two inductances, two resistances, capacitor, milliammeter (0-10mA), VTVM, AF oscillator, two batteries and connection wires.

4.4 Theory and Formula Used:

In case of transistor subscript b, e and c is written for common base, common emitter and common collector respectively. Hence the parameters h_{ie}, h_{re}, h_{fe} and h_{oe} stand for input impedance, reverse voltage ration, forward current gain and output admittance for common emitter configuration. Lets us design an amplifying circuit of PNP transistor in common base configuration as in figure 2.

When S in ON in above circuit and low signal (1K Hz) is applied between terminal 1 & 2 the current in input circuit I_b will be V_{31}/R_1 (As L_1 offered low impedance) and base-emitter voltage V_b isV_{23}. The current in output circuit will be $I_c = (V_{46}-V_{56})/R_2$

$$h_{ie} = \frac{V_b}{I_b} = R_1 \frac{V_{23}}{V_{31}} \qquad h_{fe} = \frac{I_c}{I_b} = R_1 \frac{V_{46} - V_{56}}{R_2 V_{31}}$$

S in OFF in above circuit and low signal (1K Hz) is applied between terminal 5 & 6. So I_b will be zero, base-emitter voltage V_b is V'_{32} and $I_c = (V'_{56}- V'_{46})/R_2$

$$h_{re} = \frac{V_b}{V_c} = \frac{V'32}{V'46}, \qquad h_{oe} = \frac{I_c}{V_c} = \frac{V'56 - V'46}{R_2 V'46}$$

4.5 About apparatus:

The apparatus is a circuit of PNP transistor (AC 125) as shown in figure 2. The circuit have biasing voltage V_{BB} and V_{CC}, inductance L_1 and L_2 and capacitor C. 1 KHz signal will provided by AF oscillator. The voltage between ant two terminals will be measured by VTVM. The apparatus is amplifying PNP transistor in common emitter configuration.

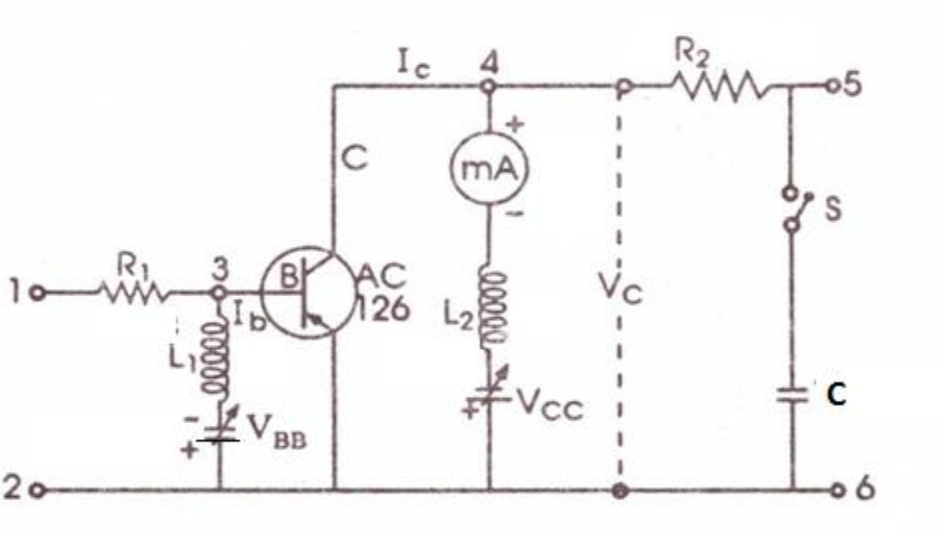


Figure 2

4.6 Procedure: let us perform following steps

1. For biasing of circuit there V_{BB} and V_{CC} adjusted such that I_C 2mA and V_{CE} is 5 volt.
2. Switch on (S) so $V_c = 0$
   Apply 1 KHz to input by AF oscillator (At 1and 2) set output voltage say 10 mV
   Measure V_{23}, V_{31}, V_{46} and V_{56}
3. Switch off (S) so $I_b = 0$
   Apply 1 KHz to input by AF oscillator (At 5and 6), set output voltage say 200mA or 300mA. And measure V'_{32}, V'_{46} and V'_{56}

4.7 Observation:

R_1 = ·····Ω R_2 =·····Ω

S.N.	$V_c = 0$ (Switch S is ON)	$I_b = 0$ (Switch S is OFF)
1	V_{23} =·····volt	V'_{32} =····· volt

2	V_{31} =..... Volt	V'_{46} =..... volt
3	V_{46} =..... volt	V'_{56} =..... Volt

4.8 Calculation: The h parameters of transistor will be calculated by following observations

$$h_{ie} = R_1 \frac{V_{23}}{V_{31}}$$

$$h_{fe} = R_1 \frac{V_{46} - V_{56}}{R_2 V_{31}}$$

$$h_{re} = \frac{V'32}{V'46}$$

$$h_{oe} = \frac{V'56 - V'46}{R_2 V'46}$$

4.9 Results: The h parameters for the given transistor at 1 KHz are

h_{ie} = ... ohm
h_{re} = ...
h_{fe} = ...
h_{oe} = ... ohm

4.10 Precaution and source of error:

1. Sensitive voltmeter and sensitive ammeter sould be used.
2. The direction about maximum reverse voltage given by manufacturer should be strictly followed.
3. The frequency should be adjusted carefully with oscillator.
4. There should not be any fluctuation on the power.
5. To avoid over heating of junction, current should not passed for long time.

4.11 Summary:

1. The 'h' parameters of circuit are mixed parameter as have mixed units.
2. Two parameters h_{re} and h_{fe} are dimensionless.
3. The 'h' parameters changes with temperature and operating point.

4.12 Glossary

` Semiconductor: Materials that conduct to electricity between conductor and insulator.

Common emitter configuration: Emitter of transistor is common to both input and output.

AF oscillator: Audio frequency generator of signal.

VTVM: Voltage tube voltmeter used to measure voltage (dC & AC), resistance etc.

4.13 References:

1. 1C. L. Arora, *B.Sc. Practical Physics*, S. Chand publication, Delhi.
2. C.L. Arora and P.S. Hemne, *Physics for Degree students (BSc III year)*, S. Chand publication, Delhi.
3. S.L.Gupta, V.Kumar *A Hand book of electronics* Pragati prakashan, Meerut.
4. B.L.Thereja *A Hand book of electronics* S. Chand publication, Delhi.
5. V.K.Mehta and Rohit Mehta,*Principle of electronics* S. Chand publication, Delhi.
6. S.L.Gupta, V.Kumar, *Practical Physics*, Pragati prakashan, Meerut.
7. *https://en.wikipedia.org.*

4.14 Viva-voce questions:

Question1. What are hybrid ‘h’ parameters?

Answer: Mixed parameter since have mixed dimensions to describe the constant of four terminal (two input and two output) are called h parameters.

Question2. What is effect of temperature on ‘h’ parameters?

Answer: ‘h’ parameters change with temperature.

Question3. Is h’ parameters also changes with operating points of transistor?

Answer: yes.

Question4. For what value ‘h’ parameters of transistor can be found correct?

Answer: ‘‘h’ parameters can found correct for small signal.

Question5. What is dimension of h_{ie}?

Answer: Ohm.

Experiment 5: To plot output characteristics of FET and measure pinch off voltage

5.1 Object:

After performing this experiment, you should be able to

- What is FET?
- What are bipolar and unipolar device.
- How the operation of FET is different from NPN and PNP transistor.
- What is use of FET?
- What is pinch off voltage?

5.2 Introduction:

After the advance of semiconductor technology, because of small size, low operation voltage and low noise many semiconductor devices are manufactured and used. Among these all, FET have important and used in amplification. FET is a field effect transistor in which electric field controls the flow of current. As we know that P and N type semiconductor are made by doping trivalent and pentavalent impurity in intrinsic semiconductor respectively. FET is a three terminals device and is a semiconductor bar (either P or N type semiconductor) on which two PN junction is made on opposite sides in middle parts. The terminals from either side of bar along the bar are called drain and source terminals and terminal from PN junction is called gate. The FET is called unipolar device as current inside the bar is due to movements of only one type of charge carrier (i.e. holes or electrons). The symbolic diagram is presented in figure 1.

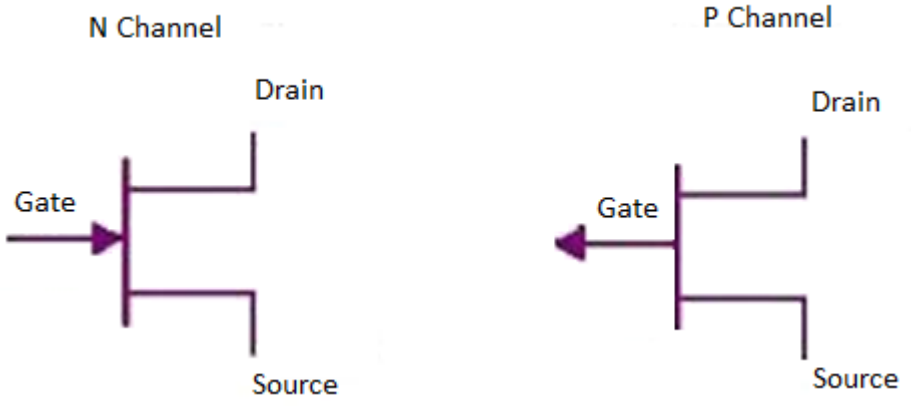


Figure 1

5.3 Apparatus Used:

FET, DC power supplies for drain source and gate source biasing, DC Voltmeters (0-10V & 0-30V), DC ammeter (0-50 mA) and rheostat (100Ω).

5.4 Theory :

The two PN junctions at the two sides of bar of conduction channel form depletion layers. The current through N channel bar is by electrons and in P channel bar is by holes. The width and hence resistance of channel can be controlled by changing the input voltage between gate and source V_{GS}. The greater reverse voltage V_{GS} will narrow the channel as the depletion layer of PN junction become wider.

The FET operate on the principle that the drain current through channel can be change by changing reverse gate source voltage V_{GS}. The working of FET is as

(1) When is drain kept at positive voltage (V_{DS}) with respect to source at zero gate potential, the drain current flows in channel of the FET. On increasing V_{DS} drain current increases and becomes saturated after a fixed voltage called pinch of voltage (V_P).
(2) When a reverse voltage V_{GS} is applied between gate and source, the depletion layers on the two sides of bar become wider and hence drain current decreases. The drain current I_p decreases on increasing reverse voltage V_{GS} and at some fixed reverse voltage the channel become completely cut off i.e. the drain current become zero. This reverse gate source voltage is called cut off voltage.

5.5 About apparatus:

The apparatus in the experiment is a field effect transistor (FET) as shown in figure 2. The drain is kept at positive voltage with respect to source by the battery (0-30V) and gate is kept at negative voltage with the source by the battery (0-10V). For the measurement of drain current milliammeter is used in circuit.

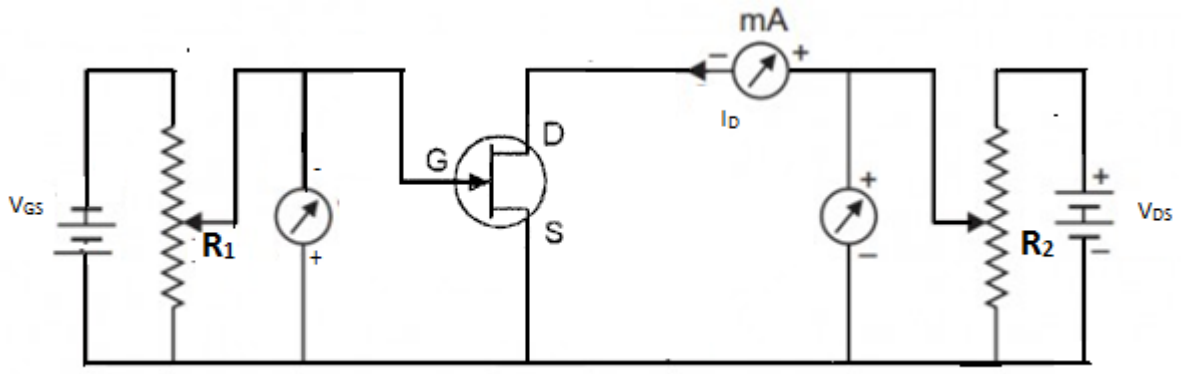


Figure 2

5.6 Procedure:

Let us perform the experiment in following steps.

1. Make connection as shown in figure 2.
2. Adjust reverse gate voltage say at 1V by means of R_1.
3. Increase drain voltage in step of 1 volt by changing R_2 and note the corresponding drain current.
4. Now note drain current with changing drain voltage for constant reverse gate voltage say for 1V, 2V, 3V and 4Vetc.
5. Plot curve between drain voltage and drain current for constant reverse gate voltage.
6. Obtain the pinch off voltage from the graph between drain voltage and drain current.
7. Repeat the observation of drain current with drain voltage for different V_{GS} say1V, 2V, 3V and 4V and find the pinch off voltage.

5.7 Observation:

Table for Output characteristics (I_d/V_{DS})

S.N.	Drain Voltage V_{DS} (In volt)	Drain current I_d (mA) for base current			
		V_{GS}=1V	V_{GS}=2V	V_{GS}=3V	V_{GS}=4V
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

5.8 Result:

The graph between drain voltage with drain current at constant gate source reverse voltage V_{GS} are plotted in the graphs and pinch off voltage V_P are ……....

5.9 Precaution and source of error:

FET should handle carefully.

1. Sensitive voltmeter and sensitive ammeter sould be used.
2. There should be proper biasing i.e. gate is netaive and drain is positive with respect ro source.
3. Voltge should not exceeds the rated value of FET.
4. There should be not be fluctuation of power.

5.10 Summary:

1. FET is a three terminal unipolar device.

2. Three terminals in FET are called source, gate and drain.

3. Drain is always kept positive with respect to source, while gate is negative with respect to source.

4. Drain and source terminals are interchangeable.

5. The drain voltage for which drain current become saturate is called pinch off voltage.

5.11 Glossary:

Semiconductor: Material whose conductivity lies between conductor and insulator.

Intrinsic semiconductor: Si and Ge in their pure form.

Amplifier: A device which raises the strength of weak signal.

Unipolar device: The semiconductor inside which the conduction is by either (holes or electrons) charge carriers

Pinch off voltage: The drain voltage for the drain current becomes saturated is called pinch off voltage.

5.12 References:

1. C. L. Arora, *B.Sc. Practical Physics*, S. Chand publication, Delhi

2. C.L. Arora and P.S. Hemne, *Physics for Degree students (BSc Second year)*, S. Chand publication, Delhi

3. Indu Prakash, Ram Krishna, A K Jha, *A text Book of Practical Physics*, Kitab Mahal Publication Delhi.

5. S.L.Gupta, V.Kumar, *Practical Physics*, Pragati prakashan, Meerut.

4. *https://en.wikipedia.org.*

5.13 Viva-voce questions:

Question1.How a FET is different from npn/pnp transistor?

Answer: In FET current conduction is either by holes or electrons and is controlled by means by electric field while in npn/pnp transistor input current controls output current and conduction is due to both holes & electrons.

Question2. Define pinch off voltage?

Answer: It is drain-source voltage at which the drain current becomes constant.

Question3.What is use of FET?

Answer. FET is used as amplifier

Question4. What is biasing of gate in FET?

Answer: Reverse bias.

Experiment 6: To draw characteristics of PN junction diode

6.1 Object:

After performing this experiment, you should be able to

- What are semiconductor device
- Formation of PN junction diode
- Characteristics of PN junction diode.
- Use of PN junction diode.

6.2 Introduction:

Semiconductors are the material whose conductivity lies between conductor and insulator. Atoms of semiconductors are attached by covalent bonds. At absolute zero temperature semiconductor is insulator as no free electron to conduct and at higher temperatures semiconductor behave as conductor due free electrons which are produced by brakeage of some bonds. By Adding impurity of trivalent and pentavaleny impurity, the pure semiconductor is made P and N type semiconductor respectively. In P type semiconductor there are holes as majority charge carrier and electrons in minority and in N type semiconductor electrons are in majority and holes are in minority charge carrier.

A P-N junction is made by diffusing N type semiconductor with P type semiconductor.

6.3 Apparatus Used:

P-N junction diode, milliammeter and microammeter, Rheostat (100Ω), voltmeter, battery and connection wires.

6.4 Theory and Formula Used:

A PN junction diode is a semiconductor crystal having acceptor impurities (P type) in one region and donor impurities (N type) in another region. The boundary between two regions is called PN junction. The characteristics property of PN junction is that current can be pass through it much more easily in one direction than in other.

As shown in figure 1, the P region has holes (positive) as majority charge carriers and equal numbers of fixed negatively charge acceptor ions (The material as hole is neutral). Similarly, N type region has electrons (negative) as majority and equal number of fixed positively charged donor ions. In addition to majority charge carriers, there are few minority charge carrier in each region. The P region contains few electrons and N region contains few holes.

When the P region is diffused over N region, there forms a depletion region of order 10 µM and a junction (potential barrier of 0.1 to 0.5V) is formed. In the absence of any external voltage applied across the PN junction there is no current in diode. Under this condition few minority charge carriers have sufficient energy to cross over the junction. This constitutes equal and opposite current and hence balances and net current is zero. By applying external voltage, the junction is made conductive in following two ways

(I) **Forward Bias:** When +ve end of battery is connected to P region and -ve end of battery connected to N region of PN junction, A electric field is set up directed from P to N inside the semiconductor. When the external voltage exceeds the junction barrier potential, the holes move from P to N region and electrons move from N to P region. The moving majority charge carriers constitute current which increases on increasing external voltage and become saturated after some vale of external voltage. This biasing is called forward bias as shown in Figure 2.

(II) **Reverse Bias:** When -ve end of battery is connected to P region and +ve end of battery connected to N region of PN junction, A electric field is set up directed from N to P inside the semiconductor. This biasing is called reverse bias a**s** in Figure 3. In this condition the majority holes in P region and electrons in N region will not cross the junction, but few minority charge carrier will cross the barrier. The current constituted by minority charge carries will form small reverse current. When reverse voltage is increased, the fast moving minority electron will break the covalent bond, which will available the charge carrier. In this condition a large amount of reverse current will flow. This condition is called avalanche breakdown.

6.5 About apparatus:

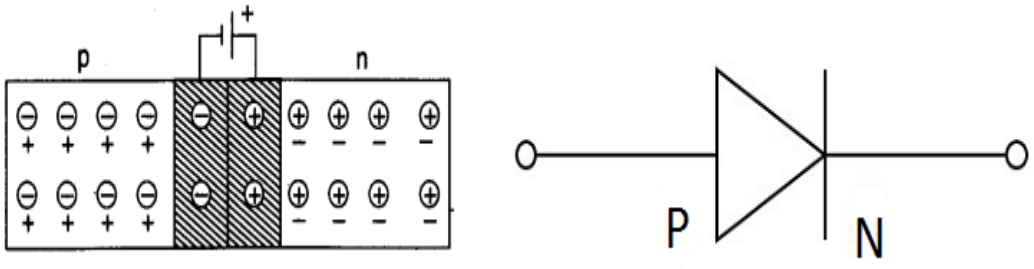


Figure 1

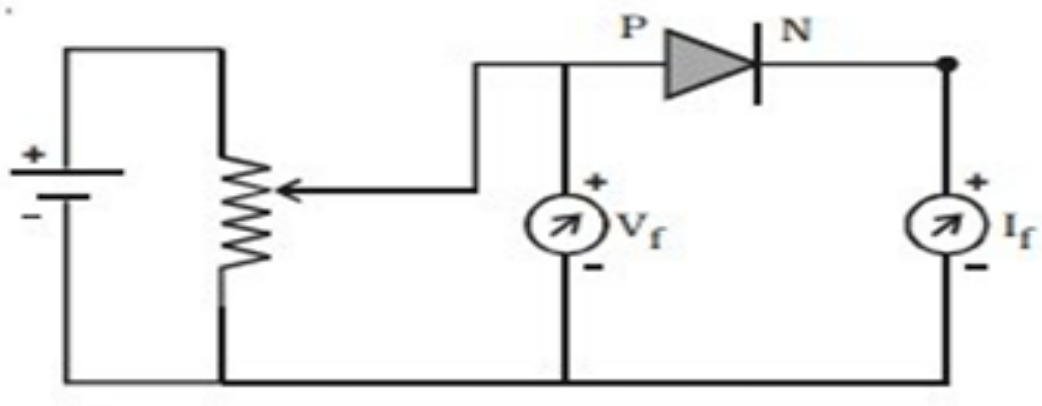


Figure 2 Forward bias

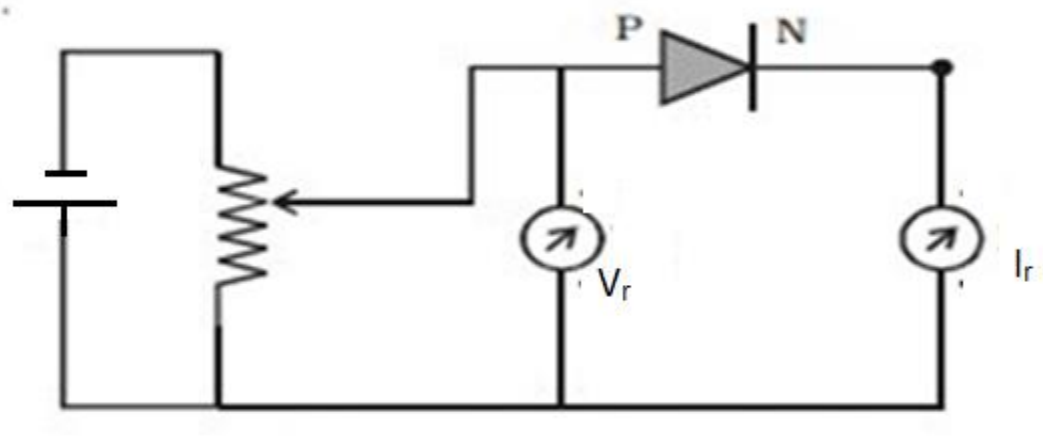


Figure 3 Reverse bias

6.6 Procedure:

For Forward bias

6. Make connection as shown in circuit diagram (figure 2).
7. With the help of rheostat vary the applied voltage in step of 0.1V.
8. Note down the ammeter reading in milliamter.
9. Draw the graph between voltage V_f and current I_f.

For Reverse bias

5. Make connection as shown in circuit diagram (figure 3).
6. With the help of rheostat vary the applied voltage in step of 0.5V.
7. Note down the ammeter reading in micrometer.
8. Draw the graph between voltage V_r and current I_r.

6.7 Observation:

Table1.For plate characteristics

S.N.	Forward bias		Reverse bias	
	Voltage V (In V)	Current I (In mA)	Voltage V (In V)	Current I (In mA)
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

6.8 Result: The forward and reverse characteristics of PN junction diode is presented the graph.

6.9 Precaution and source of error:

1. Sensitive voltmeter and sensitive ammeter sould be used.
2. The direction about maximum plate voltage given by manufacturer should be strictly followed.
3. The graphs should drawn smothly.
4. There should not be any fluctuation on the power.
5. To avoid over heating of PN junction, current should not passed for long time.

6.10 Summary:

1. By addition of trivalent and pentavalent impurity pure Ge or Si become P and N type semiconductor respectively.
2. A junction is formed when N type semiconductor is diffuse over N type semiconductor. This is called PN junction.
3. The PN junction conduct when forward bias. So can be used as rectifier.
4. The current in PN junction is carried by electrons and holes.

6.11 Glossary

` Semiconductor: Materials that conduct to electricity between conductor and insulator.

Extrinsic semiconductor: Impure semiconductor for increasing conductivity.

Forward bias: When P is given +ve and N is given –ve voltage.

Reverse bias: When P is given -ve and N is given +ve voltage.

Rheostat: It is used to vary potential

Rectifier: The device which chanhes AC voltage to DC voltage.

Resistance: Obstruction to the flow of current.

6.12 References:

1. C. L. Arora, *B.Sc. Practical Physics*, S. Chand publication, Delhi.
2. C.L. Arora and P.S. Hemne, *Physics for Degree students (BSc III year)*, S. Chand publication, Delhi.
3. S.L.Gupta, V.Kumar *A Hand book of electronics* Pragati prakashan, Meerut.
4. B.L.Thereja *A Hand book of electronics* S. Chand publication, Delhi.
5. V.K.Mehta and Rohit Mehta,*Principle of electronics* S. Chand publication, Delhi.

6.13 Viva-voce questions:

Question1. What is meant by P type and N type semiconducor?

Answer: If we add III group impurity as boron (B), aluminum (Al) to pure semiconductor (Ge or Si), then the charge carriers are holes. This resultant material is called P type semiconductor.

On other 1.hand If we add V group impurity as arsenic (As), antimony (Sb) to pure semiconductor (Ge or Si), then the charge carriers are electrons. This resultant material is called N type semiconductor.

Question2. How PN junction is formed?

Answer: When the P type semiconductor is diffused over N type semiconductor the movement of hole (from P type semiconductor) and electrons (from N type semiconductor) leaves their parent atoms as ions. These ions restrict further movement of charge carriers hence form a junction called PN junction.

Question3. What will happen if high voltage is applied to PN junction?

Answer. In this case high current will flow and will damage the junction.

Question4.What is order of current in junction.

Answer: In forward bias, current is in milliampere while in reverse current is in microampere.