

# ECE 1051 : BASIC ELECTRONICS

## TABLE OF CONTENTS

### Part - I : Analog Electronics

---

#### Chapter – 1 : Diodes and Applications

##### **Module-1 : PN Junction Diodes**

1.1.1	Introduction	5
1.1.2	Concept of PN junction	5
1.1.3	PN junction under bias	6
1.1.4	V-I Characteristics of diode	9
1.1.5	Static and dynamic Resistance of Diode	13
1.1.6	Ideal and Practical diodes	14
1.1.7	Equivalent circuit of diode	15
1.1.8	Break down phenomenon in diodes	18
1.1.9	Zener diode characteristics	19
1.1.10	Diode as a capacitor	20

##### **Module-2 : Applications of Diodes**

1.2.1	Introduction	23
1.2.2	Basic Block diagram of DC power supply	24
1.2.3	Half wave rectifier	28
1.2.4	Full Wave Rectifier	31
1.2.5	Full wave bridge rectifier	34
1.2.6	Comparison of Rectifiers	37
1.2.7	Rectifier with filter	38

##### **Module-3 : Voltage Regulators**

1.3.1	Zener Voltage Regulator	44
1.3.2	IC Voltage Regulators	51

##### **Module-4 : Special purpose diodes**

1.4.1	Light Emitting Diodes	53
1.4.2	Photo diodes and applications	54
1.4.3	Optocoupler	55
1.4.4	Solar Cell	56

---

#### Chapter – 2 : BJT and Applications

##### **Module-1 : BJT Characteristics**

2.1.1	BJT construction and operation	61
2.1.2	BJT configurations	62

2.1.3 BJT current components 63

2.1.4 BJT characteristics 65

#### **Module-2 : BJT Biasing**

2.2.1 DC load line and Ned for Biasing 68

2.2.2 Fixed biasing 70

2.2.3 Voltage divider biasing or self –biasing 73

#### **Module-3 : Transistor as an amplifier**

2.3.1 RC coupled Amplifier 77

2.3.2 Frequency Response of an Amplifier 78

2.3.3 Multistage amplifier 79

#### **Module-4 : Transistor as a Switch**

2.4.1 Introduction 81

2.4.2 Transistor as LED driver 82

2.4.3 Transistor as Inverter 82

### **Chapter – 3 : Operational amplifier and applications:**

---

#### **Module-1 : Operational Amplifier**

3.1.1 Introduction 84

3.1.2 Internal Block diagram of Op-amp 84

3.1.3 Op-Amp Characteristics 88

#### **Module-2 : Linear applications of Operational Amplifier**

3.2.1 Inverting amplifier 90

3.2.2 Non-inverting Amplifier 91

3.2.3 Voltage follower 92

3.2.4 Inverting Summing Amplifier 92

3.2.5 Difference amplifier 93

3.2.6 Integrator 94

3.2.7 Differentiator 94

#### **Module-3 : Non-linear applications of Operational Amplifier**

3.3.1 Voltage Comparator 96

3.3.2 Square wave generator 96

## **Part - II : Digital Electronics**

### **Chapter – 4 : Number systems & Codes**

---

#### **Module-1 : Number Systems**

4.1.1 Decimal number system 100

4.1.2 Binary number system 101

4.1.3 Octal number system 101

4.1.4 Hexadecimal number systems 102

4.1.5	Number system – Arithmetic	103
4.1.6	Complementation of Numbers	105

**Module-2 : Codes**

4.2.1	Weighted binary codes	108
4.2.2	Non-Weighted codes	110
4.2.3	Self-complimenting codes	110
4.2.4	Error detecting and correcting codes	111

**Chapter – 5 : Boolean Algebra & Logic gates**

---

**Module-1 : Introduction to Boolean Algebra**

5.1.1	Boolean algebraic theorems	115
5.1.2	Simplification of Boolean algebraic expressions	118

**Module-2 : Logic gates**

5.2.1	Logic gates and operation	121
5.2.2	Concept of Universal Logic	126
5.2.3	Classification of digital circuits	128
5.2.4	Implementation of Combinational Circuits	129

**Module-3 : Simplification of Boolean functions using K - map**

5.3.1	Standard form of Boolean expressions	135
5.3.2	Introduction to K-map	137
5.3.3	Simplification of Boolean expressions using K-map	138

**Chapter – 6 : Sequential circuits**

---

**Module-1 Flip-flops and applications**

6.1.1	Introduction to Flip Flops	142
6.1.2	Counters	147
6.1.3	Shift Registers	151

**Part - III : Principles of Electronic Communication****Chapter – 7 : Fundamentals of Analog communication**

---

**Module-1 : Amplitude Modulation**

7.1.1	Principle of Electronic Communication	156
7.1.2	Need for modulation	158
7.1.3	Amplitude modulation	159
7.1.4	Time domain analysis	160
7.1.5	Spectrum of AM signal	162
7.1.6	Power content of AM	162
7.1.7	Different types of AM signals	165

7.1.8	Detection of AM signal	166
7.1.9	Super heterodyne principle of AM reception	167

**Module-2 : Frequency Modulation**

7.2.1	Frequency Modulation	170
7.2.2	Time domain analysis of FM signal	171
7.2.3	Comparison of AM and FM	174

**Chapter – 8 : Introduction to digital communication**

---

**Module-1 : Digitization of Analog signals**

8.1.1	Basic principle of Sampling	176
8.1.2	Pulse Amplitude Modulation	179
8.1.3	Pulse Width Modulation	180
8.1.4	Pulse Position Modulation	180
8.1.5	Digital communication system	180
8.1.6	Digital Modulation Techniques	183

**Chapter – 9 : Communication Networks**

---

**Module-1 : Introduction to Communication networks**

9.1.1	Introduction to Data communication	186
9.1.2	Introduction to communication networks	187
9.1.3	Types of communication networks	188
9.1.4	Communication Network topology	189
9.1.5	Network protocols and Reference models	190

**Chapter – 10 : Introduction to Mobile Communication**

---

**Module-1 : Introduction to Mobile Communication**

10.1.1	Principle of Cellular communication	195
10.1.2	Multiple access Technology	196
10.1.3	Global System for Mobile Communications (GSM)	198

**Part -I**  
**ANALOG ELECTRONICS**  
**Chapter -1: Diodes and Applications**

The term diode is used to represent a device or element which has two electrodes. These devices are characterized by the fact that they allow electric current to flow in one direction, and block flow of current in the opposite direction. This unilateral behaviour is predominantly used in switching and rectification. Diode is in fact, the very first electronic device invented. Initially, for several years vacuum tube version was in use, and were bulky in size, required higher power for operation, and were slow. Today, we have semiconductor diodes, which are very small in size, requires relatively low power and operates at higher speeds. Semiconductor diodes are available in various forms and are used in wide variety of applications. In this unit, we shall look at the operating behavior and characteristics of semiconductor diodes along with their typical applications such as rectifiers, voltage regulators and some special purpose applications.

### **Module – 1 : Diodes**

#### **Learning Outcomes:**

At the end of this module, students will be able to:

1. Explain the operation of PN junction diode under different biasing conditions.
2. Plot the I-V characteristics of the diode.
3. Define static and dynamic resistance of the diode.
4. Explain the breakdown phenomenon observed in diodes.
5. Describe the working of Zener diode and plot its I-V characteristics.
6. Explain the operation of diode as a capacitor.

#### **1.1.1 Introduction**

Materials are broadly classified as metals, insulators and semiconductors. A semiconductor like Germanium or Silicon has electrical conductivity lying between conductor and insulator. Semiconductors are the basic materials used in modern electronics. For example, Diodes, Transistors, Solar cells, Light-emitting diodes (LEDs), and integrated circuits.

#### *Self Reading:*

1. *Crystal structure of Germanium and Silicon*
2. *Intrinsic and extrinsic semiconductors*
3. *N-type and P-type semiconductors and concept of minority and majority carriers*
4. *Diffusion and drift currents*

#### **1.1.2 Concept of PN Junction**

As you know, P-type semiconductor has large number of holes while, N-type semiconductor has large number of free electrons. When P-type and N-type materials are joined together, a gradient of charge carriers densities is created at the junction. This will

cause the electrons to move from N-type material to the P-type material and holes to move from N-type material to the P-type material. This process of movement of charge carriers form the region of higher concentration to a lower concentration in the absence of external electric field is called diffusion. Diffusion of charge carriers across the junction will continue until the equilibrium condition is established. Also at the junction, N-type material will have positively charged immobile ions and P-type material have negatively charged immobile ions. Thus the regions on either sides of p-n interface lose their charge neutrality and become charged. For this reason it is called space charge region. As the region is devoid or depleted of mobile charge carriers it is also called depletion region and is as shown in Figure. 1.1.1.

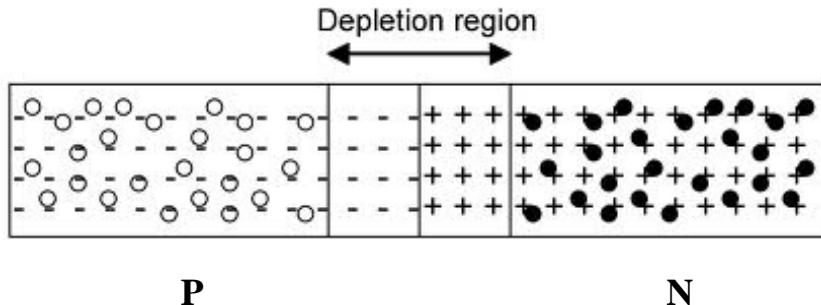


Figure 1.1.1. Schematic of PN junction

The space charge on either sides of the junction causes a potential difference across the P-N junction and it is called the barrier potential. This is the minimum amount of voltage required to initiate flow of charge carriers across the junction. Doped germanium has a barrier potential of about 0.3 volts whereas, doped silicon has a barrier voltage of about 0.7 volts.

### 1.1.3 P-N junction under bias

Application of external voltage across the diode is called biasing. Depending upon the polarity and magnitude of voltage applied, we can have three biasing conditions, as listed below.

- a) No bias or Zero bias
- b) Forward bias
- c) Reverse bias

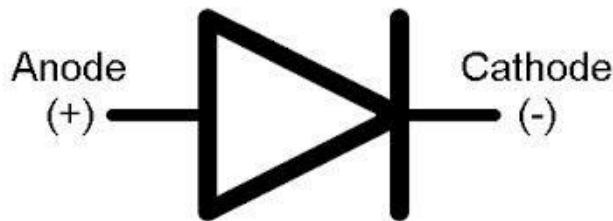


Figure 1.1.2. Circuit symbol of a diode

- a) **Zero Bias:** In the absence of any bias voltage, the net flow of charge carriers in any one direction for a semiconductor diode is zero. This occurs because minority carriers (holes) in the N-type material will encounter barrier in the depletion region to cross the junction and move to the P-type region. Same is the case for electrons in P-type material. This results in depletion region with high impedance, and hence no current flows through the diode. The built-in potential varies from 0.3 to 0.7 eV depending upon the type of semiconductor material. A diode operated without any biasing is shown in Figure 1.1.3.

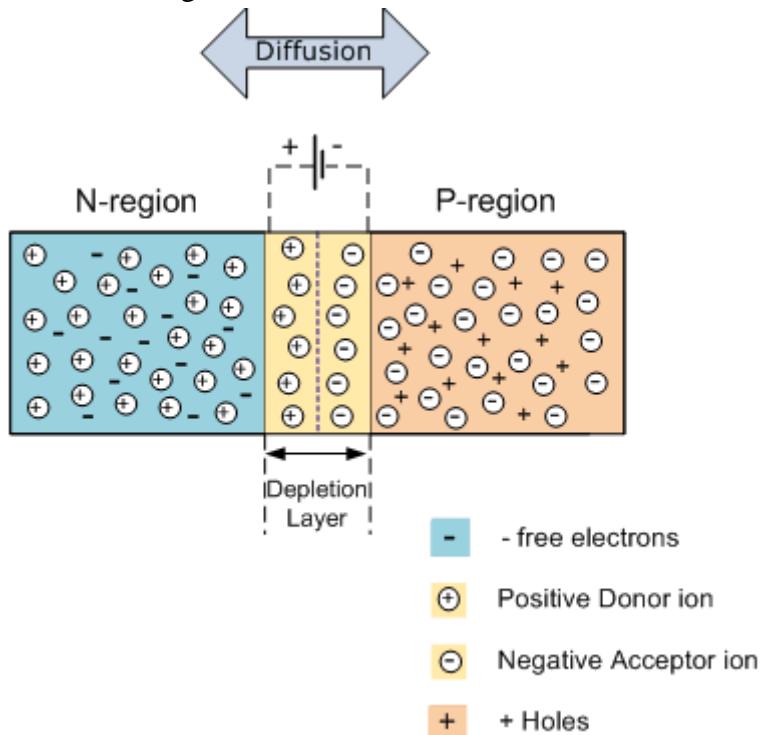


Figure. 1.1.3. P-N diode with zero biase.

- b) **Forward Bias:** When a negative voltage is applied to the N-type material and a positive voltage is applied to the P-type material, the diode is said to be in a Forward Bias condition. Figure 1.1.4. shows the diode with forward biase. If the external voltage applied is greater than the value of the barrier potential, the carriers start crossing eth junction and hence there will be a forward current. The diode is said to be in the ON condition.

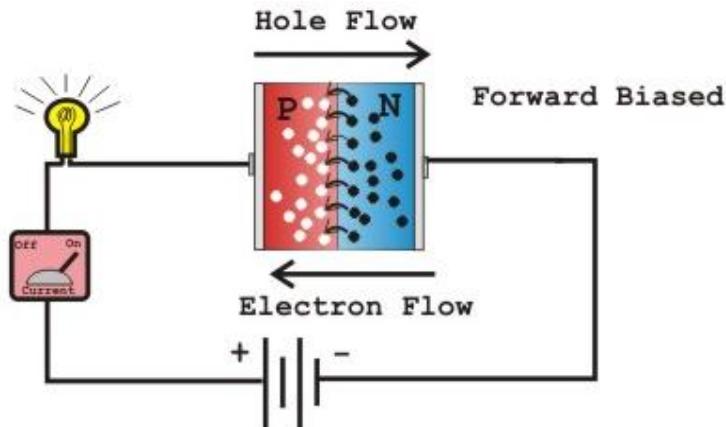


Figure 1.1.4. Forward biasing of P-N junction diode.

[<http://www.imagesco.com/articles/photovoltaic/photovoltaic-pg3.html>].

- c) **Reverse Bias:** When a positive voltage is applied to the N-type material and a negative voltage is applied to the P-type material , the diode is said to be in a reverse biased condition, as shown in Figure. 1.1.5. The positive voltage applied to the N-type semiconductor attracts electrons towards the positive electrode and hence away from the junction. At the same time, the holes in the P-type semiconductor are attracted towards the negative electrode. This results in widening of depletion layer due to a lack of electrons and holes near the junction and presents a high impedance path for the majority carriers. The height of potential barrier is increased, which prevents the flow of forward current through the diode. However, the applied potential favors the movement of minority carriers across the junction causing flow of current in the reverse direction. This current is called reverse saturation current and is represented by  $I_0$  or  $I_S$ .

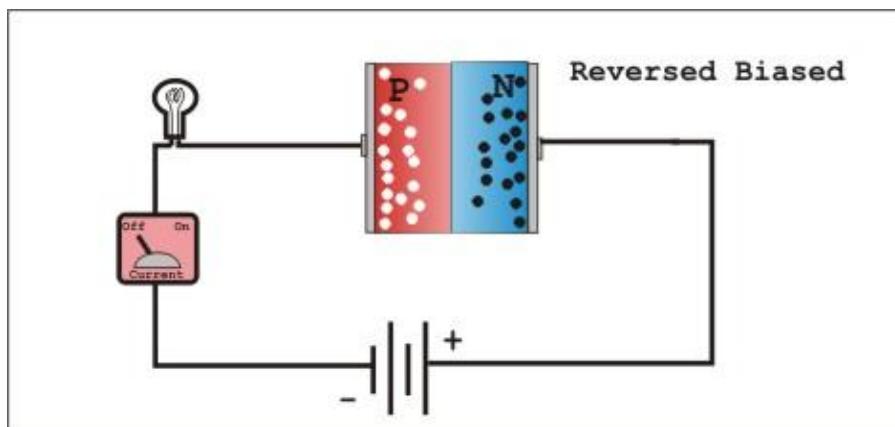


Figure. 1.1.5. Reverse biasing of P-N junction diode.

[<http://www.imagesco.com/articles/photovoltaic/photovoltaic-pg3.html>].

**Self test:**

1. The arrow direction in the diode symbol indicates
  - a. Direction of electron flow.
  - b. Direction of hole flow (Direction of conventional current)
  - c. Opposite to the direction of hole flow
  - d. None of the above
  
2. When the diode is forward biased, it is equivalent to
 

a. An OFF switch	b. An On switch
c. A high resistance	d. None of the above
  
3. The barrier potential voltage of Si diode is
 

a. 0.2 V	b. 0.7 V	c. 0.8 V	d. 1.0 V
----------	----------	----------	----------
  
5. List the methods available for testing the diode. How cut-in voltage of diode is measured in practice?

**1.1.4 I-V characteristics of diode**

I-V characteristics of practical diode is shown in Figure 1.1.6. When the forward biased voltage is applied to diode, current is initially zero and then increases sharply after crossing the cut-in voltage. In this case, the diode behaves like a closed switch. Similarly, in reversed biased condition, the diode behaves like an open switch and very small current flows due to minority charge carriers, which is known as reverse saturation current. In the reverse biased condition, beyond a particular reverse voltage, a sudden rise of current will be observed and this voltage is called breakdown voltage.

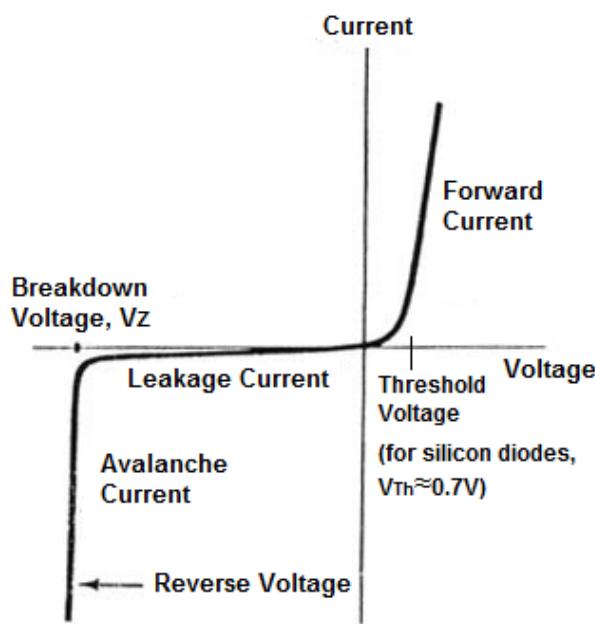


Figure. 1.1.6. I-V characteristics of practical diode  
[\[http://www.learningaboutelectronics.com/Articles/Ideal-diode.php\]](http://www.learningaboutelectronics.com/Articles/Ideal-diode.php)

**a) Forward biased Characteristic:**

The application of a forward bias voltage to the junction diode results in the depletion layer becoming very thin and narrow which represents a low resistance path through the junction thereby aiding flow of current through the diode. The point at which this sudden increase in current takes place is called “knee” point and is represented on the static I-V characteristics as shown in Figure 1.1.7.

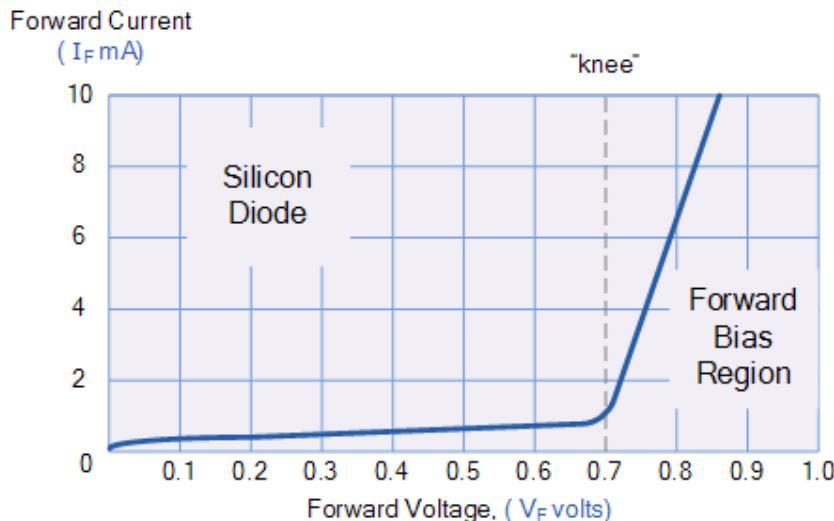


Figure 1.1.7. I-V characteristics of P-N junction diode under forward biased condition.

[[http://www.electronics-tutorials.ws/diode/diode\\_3.html](http://www.electronics-tutorials.ws/diode/diode_3.html)]

**b) Reverse biased Characteristic:**

In this case, PN junction offers high resistance value and practically zero current flows through the junction diode with an increase in bias voltage. However, a very small leakage current flows through the junction which is in the order of microamperes ( $\mu A$ ) for ordinary rectifier diodes.

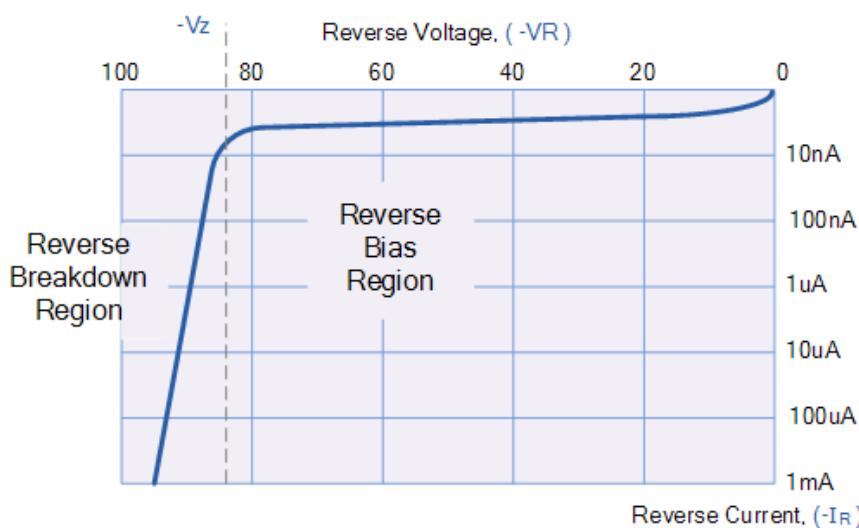


Figure. 1.1.8. I-V characteristics of P-N junction diode under reversed biased condition.

[ [http://www.electronics-tutorials.ws/diode/diode\\_3.html](http://www.electronics-tutorials.ws/diode/diode_3.html) ]

If the reverse bias voltage  $V_R$  applied to the diode is increased beyond certain limits there will be a large current due to avalanche effect and cause breakdown. This is shown in figure 1.1.8.

**c) The Diode Current:**

The current flowing through the diode is given by the following equation.

$$I_D = I_0 \left( e^{\frac{V_D}{\eta V_T}} - 1 \right) \quad (1.1.1)$$

$I_D$  = Diode current

$I_0$  = Reverse saturation current.

$V_D$  = Applied bias voltage (Positive for forward and negative for reverse bias)

$V_T = \frac{T}{11600}$  Volt equivalent of temperature (T is in degree Kelvin)

for Germanium  $\eta = 1$  and for Silicon  $\eta = 2$

For large forward bias, equation 1.1.1 approximates to

$$I_D = I_0 (e^{\frac{V_D}{\eta V_T}}) \quad (1.1.2)$$

For large reverse bias, equation 1.1.1 approximates to

$$I_D = -I_0 \quad (1.1.3)$$

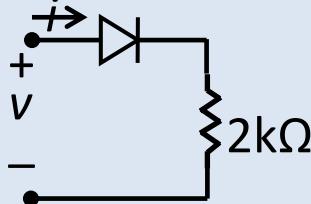
**d) Effect of Temperature on the Reverse current:**

The reverse saturation current is a temperature dependent parameter. It doubles for every  $10^\circ C$  rise in temperature. Let  $I_{01}$  be the reverse saturation current at temperature  $T_1$  and  $I_{02}$  be the reverse saturation current at temperature  $T_2$ , where  $T_2 > T_1$ . Thus the rise in reverse saturation current can be modelled as

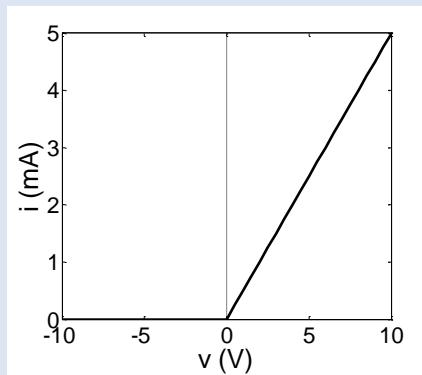
$$I_{02} = I_{01} 2^{(T_2 - T_1)/10} \quad (1.1.4)$$

**Exercise Problem 1:**

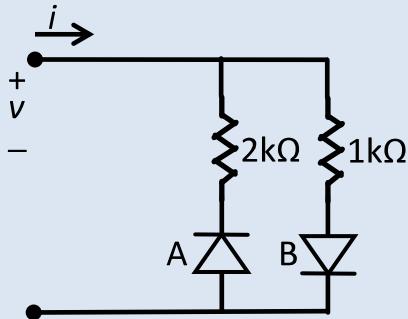
1. Sketch  $I$  versus  $V$  to scale for each of the circuits shown below. Assume that the diodes are ideal and allow  $V$  to range from  $-10 \text{ V}$  to  $+10 \text{ V}$ .



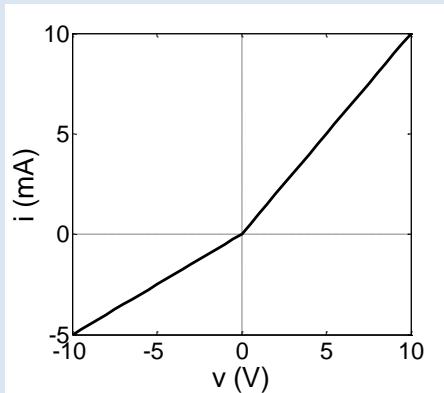
**Solution:** Diode is ON for  $v > 0$  ;



2. Sketch  $I$  versus  $V$  to scale for each of the circuits shown below. Assume that the diodes are ideal and allow  $V$  to range from  $-10 \text{ V}$  to  $+10 \text{ V}$ .



**Solution:** Diode B is ON for  $v > 0$  and  $R=1\text{k}\Omega$ . Diode A is on for  $v < 0$  and  $R=2\text{k}\Omega$ .



### 1.1.5 Static and Dynamic Resistance of Diode

There are two types of diode resistance namely, DC and AC resistance.

#### a) DC or Static Resistance

The application of a dc voltage to a circuit containing a semiconductor diode will result in an operating point on the characteristic curve that will not change with time. The resistance of the diode at the operating point can be found by the corresponding values of  $V_D$  and  $I_D$  as shown in Figure 1.1.9. The equation for the DC resistance can then be written as,

$$R_D = \frac{V_D}{I_D} \quad (1.1.5)$$

Note that the DC or static resistance of a diode does not depend on the curve shape, it depends only on the operating point or the values of diode voltage and current.

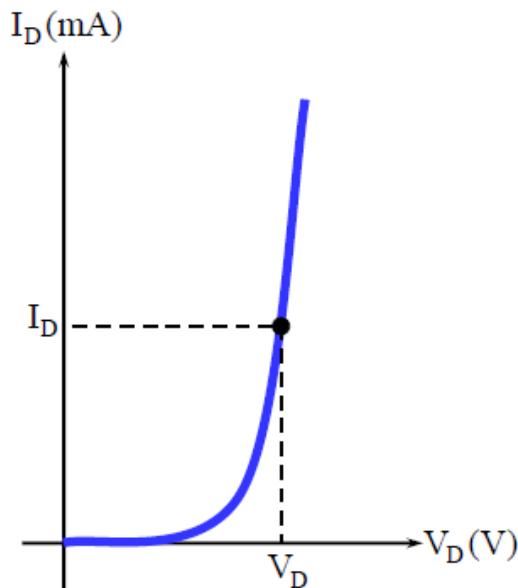


Figure. 1.1.9. Static resistance of a diode [http://www.freewebs.com]

#### b) AC or Dynamic Resistance:

To determine the dynamic resistance of a diode, a tangent is drawn to the curve through the operating point as shown in Figure 1.1.10.

The dynamic resistance of the diode is found using the following equation:

$$r_d = \frac{1}{slope} = \frac{\Delta V_d}{\Delta I_d} \quad (1.1.6)$$

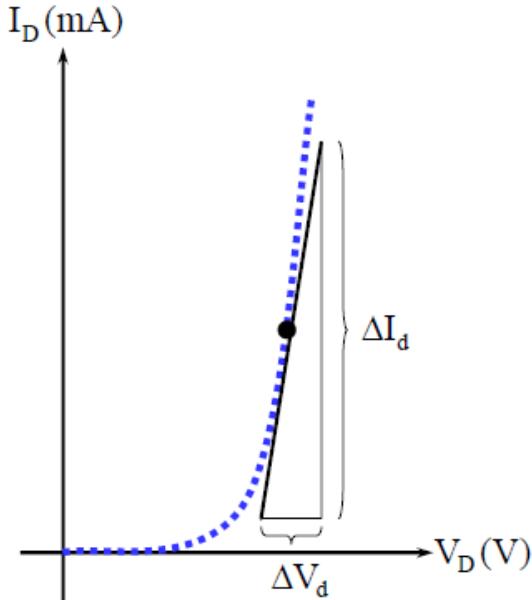


Figure 1.1.10. Dynamic resistance [http://www.freewebs.com]

Also, dynamic resistance is found by the derivative of the diode equation; where:

$$I_D = I_s \left[ \exp\left(\frac{V_D}{\eta V_T}\right) - 1 \right], \quad (1.1.7)$$

$$\frac{dI_D}{dV_D} = \frac{d}{dV_D} \left( I_s \left[ \exp\left(\frac{V_D}{\eta V_T}\right) - 1 \right] \right) = \left( \frac{I_s}{\eta V_T} \exp\left(\frac{V_D}{\eta V_T}\right) \right) \quad (1.1.8)$$

$$\frac{dV_D}{dI_D} = \left( \frac{\eta V_T}{I_D + I_s} \right) \quad (1.1.9)$$

Since  $I_D \gg I_s$ , then

$$r_d = \frac{dV_D}{dI_D} = \frac{\eta V_T}{I_D} \quad (1.1.10)$$

$$\text{For Ge , } \eta=1 ; \quad V_T \approx \frac{T}{11600} \quad (1.1.11)$$

$$\text{And at room temperature, } T=300 \text{ K, } \quad r_d = \frac{dV_D}{dI_D} = \frac{300}{11600 \times I_D} \quad (1.1.12)$$

### 1.1.6 Ideal and Practical diode

In an ideal diode, current flows freely through the device when forward biased, offering no resistance. An ideal diode is simply a P-N junction where the change from P-type to N-type material is assumed to occur instantaneously, also referred to as an abrupt junction.

The simplified diode model ignores the effect of diode resistance in comparison with values of other elements of the circuit. The voltage drop across the diode is zero.

A practical diode does offer some resistance to current flow when forward biased. Since there is some resistance (built-in potential), there will be some power dissipated when current flows through a forward biased diode. Therefore, there is a practical limit to the amount of current a diode can conduct without damage. A reverse biased diode has very high resistance and excessive reverse bias can cause the diode to damage.

### 1.1.7 Equivalent circuit of diode:

Diode is often replaced by its equivalent circuit during circuit analysis and design. For DC diode model, characteristics of an ideal diode and the modifications that were required due to practical considerations has been considered for following cases:

- (i) Ideal Diode
- (ii) Second approximation of diode
- (iii) Practical diode

(i) For an Ideal diode  $V_\gamma = 0$ ,  $R_R = \infty$  and  $R_F = 0$  as shown in Figure 1.1.14. In other words, the ideal diode is a short in the forward bias region and an open in the reverse bias region.

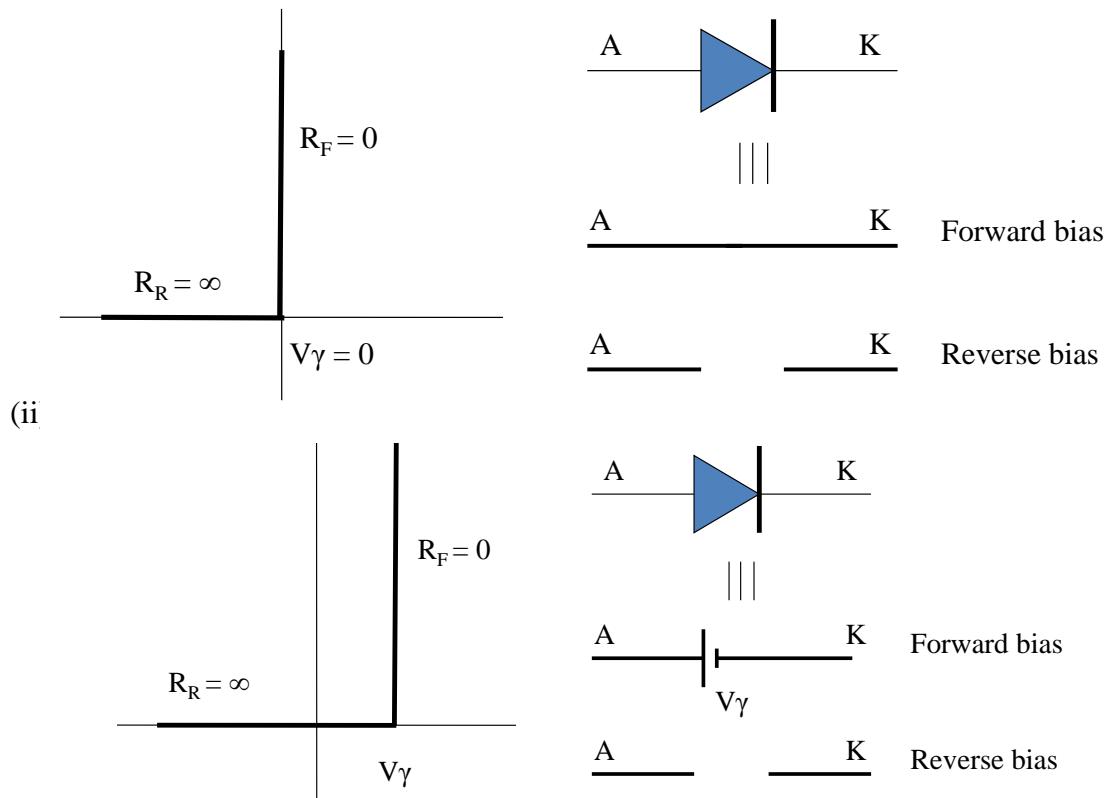


Figure 1.1.11 Equivalent circuit of diode for second approximation.

(iii) In Practical diode (silicon)  $V_\gamma = 0.7$  V,  $R_R < \infty$  (typically several MΩ),  $R_F \approx r_d$  (typically  $< 50 \Omega$ ) as shown in Figure 1.1.12.

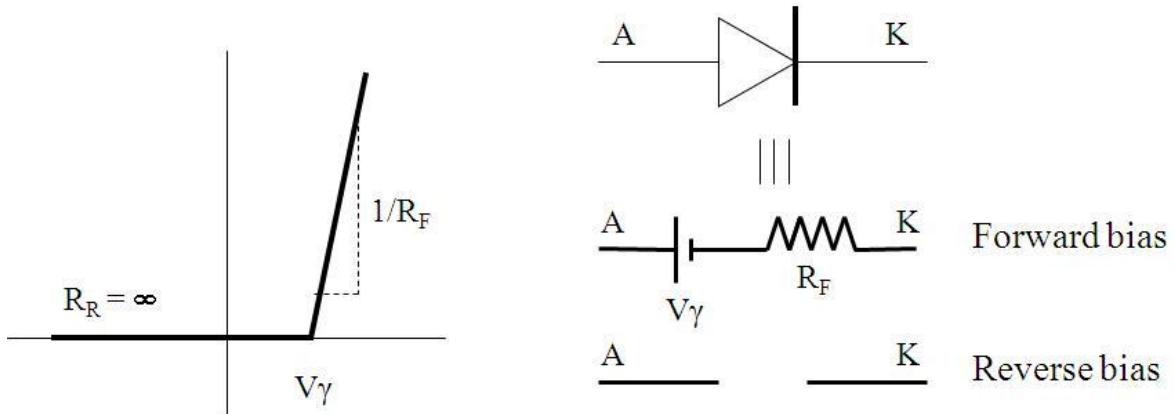


Figure 1.1.12 Equivalent circuit of practical diode.

**Exercise Problem 1:**

1. A Silicon diode has a saturation current of  $1\text{pA}$  at  $20^\circ\text{C}$ . Determine (a) Diode bias voltage when diode current is  $3\text{mA}$  (b) Diode bias current when the temperature changes to  $100^\circ\text{C}$ , for the same bias voltage.

Ans. Given: The diode current  $I_D=3\text{mA}$ ,

Reverse saturation current  $I_0=1\times 10^{-12} \text{ A}$ , Temperature  $T=20^\circ\text{C}=273+20=293\text{K}$

The diode is silicon  $\eta=2$

The equation for the diode current  $I_D$  is given by

$$(a) \text{The diode bias voltage } I_D = I_0 \left( e^{\frac{V_D}{\eta V_T}} - 1 \right) \text{ and } V_T = \frac{T}{11600} = \frac{293}{11600} = 25.25 \text{ mV}$$

$$V_D = \eta V_T \ln \left( 1 + \frac{I_D}{I_0} \right) = 1.103 \text{ V}$$

(b) The diode current when the temperature is  $1000^\circ\text{C}$

$$V_T = \frac{T}{11600} = \frac{393}{11600} = 32.15 \text{ mV}$$

The temperature is raised to  $100^\circ\text{C}$

(So the reverse saturation current  $I_0$  changes)

$$I_{02} = I_{01} 2^{(T_2 - T_1)/10} = 10^{-12} \left( 2^{\frac{(100-20)}{10}} \right) = 256 \text{ pA}$$

$$I_D = 256 \times 10^{-12} \left( e^{\frac{1.103}{(2 \times 32.15 \times 10^{-3}) - 1}} \right) = 7.21 \text{ mA}$$

2. Find the static and dynamic resistance of a P-N junction germanium diode if

the temperature is  $27^{\circ}\text{C}$  and  $I_0=1\mu\text{A}$  for an applied forward bias of  $0.2\text{V}$ .

**Solution:**

Given: Applied forward voltage =  $0.2\text{ V}$ .

Reverse saturation current  $I_0=1\times 10^{-6}\text{ A}$ .

Temperature  $T=27^{\circ}\text{C} = 273+27 = 300^{\circ}\text{K}$ .

The diode is Ge,  $\eta=1$ .

$$V_T = \frac{T}{11600} = \frac{300}{11600} = 25.86\text{ mV}$$

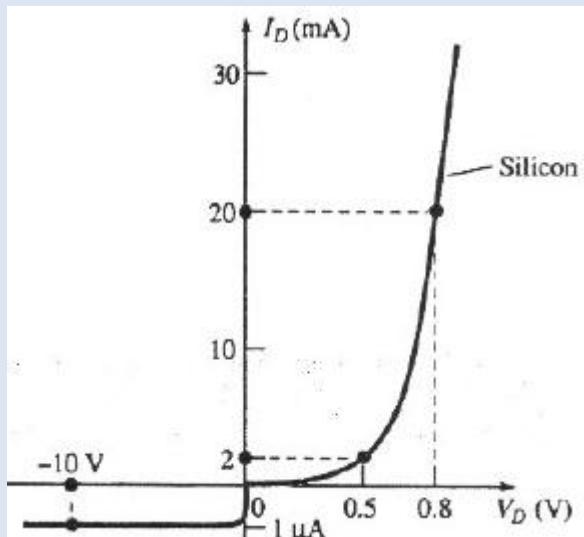
$$I_D = 1\times 10^{-6} \left( e^{\frac{0.2}{(1\times 25.86\times 10^{-3}-1)}} \right) = 2.28\text{ mA}$$

Static Resistance:  $\frac{V}{I} = 0.087\text{ K}\Omega$

$$\text{Dynamic resistance} = \frac{\eta V_T}{I_D + I_0} = \frac{1\times 25.86\times 10^{-3}}{2.28\times 10^{-3} + 1\times 10^{-6}} = 11.33\Omega$$

3. Determine the dc resistance levels for the diode at

- (a)  $I_D=2\text{ mA}$ , (b)  $I_D=20\text{ mA}$ , (c)  $V_D=-10\text{ V}$



**Solution:** From the graph, find corresponding voltage for the mentioned current values.

- (a)  $R_D = V_D/I_D = 0.5\text{V}/2\text{mA} = 250\Omega$ .  
 (b)  $R_D = V_D/I_D = 0.8\text{V}/20\text{mA} = 40\Omega$ .  
 (c) At  $V_D = -10\text{V}$ ,  $I_D = -Is = -1\mu\text{A}$  (from the curve) and  
 $R_D = V_D/I_D = 10\text{V}/1\mu\text{A} = 40\Omega$

### 1.1.8 Break down phenomenon in diodes

Breakdown voltage is the largest reverse voltage that can be applied without causing an exponential increase in the current in the diode. As long as the current is limited, exceeding the breakdown voltage of a diode does no harm to the diode.

Two breakdown mechanisms exist in diode. They are

- a) Zener breakdown
- b) Avalanche breakdown

#### a) Zener breakdown:

In Zener breakdown, the electric field established due to the reverse voltage capable of getting the electrons out of their covalent bonds and away from their parent atoms as shown in Figure 1.1.13. Electrons are transferred from the valence to the conduction band. In this situation, the current can still be limited by the limited number of free electrons produced by the applied voltage so it is possible to cause Zener breakdown without damaging the semiconductor.

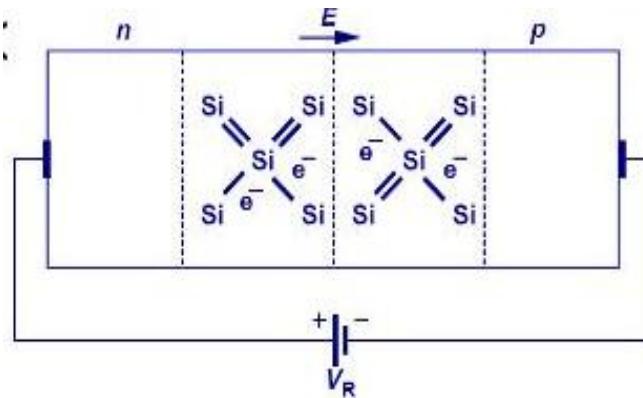


Figure. 1.1.13 Schematic of the zener breakdown mechanisms

<http://shrdocs.com/presentations/12656/index.html>

#### b) Avalanche Breakdown:

Avalanche breakdown occurs when the applied voltage is so large that electrons achieve kinetic energy sufficiently high and collide with the silicon atoms and knock off more electrons. These electrons are then also accelerated and subsequently collide with other atoms. Each collision produces more electrons which leads to more collisions etc as shown in Figure. 1.1.14. The current in the semiconductor rapidly increases and the material can quickly be destroyed.

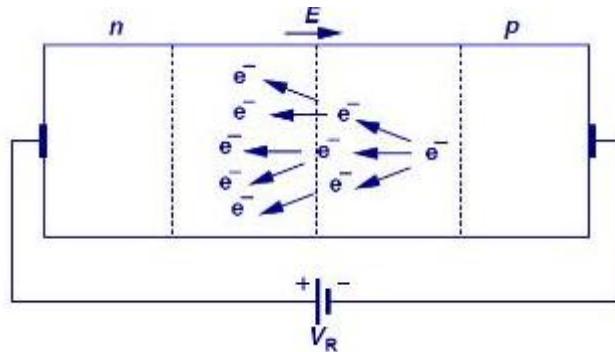


Figure. 1.1.14. Avalanche breakdown phenomenon  
<http://shrdocs.com/presentations/12656/index.html>

### 1.1.9 Zener diode characteristics:

Zener diode (also referred as regulated diode) is a two terminal device that is widely used in voltage regulators. When the reverse bias, applied to the semiconductor, has reached to zener voltage  $V_z$ , the current will be dramatically increased while the voltage keeps constant. Circuit symbol of Zener diode is shown in Figure 1.1.15. It is special kind of diode that is heavily doped during manufacturing. This results in a high number of free carriers. These additional current carriers permit reverse current flow when a specific reverse bias voltage is reached. This voltage level is referred to as the avalanche point or Zener point. When forward biased, the Zener diode acts like a regular diode. Zener diodes are rated according to the voltage at which they will "turn ON", or begin to conduct reverse bias current. The regulated values of the zener diode are thus distributed in the range from 3V to several hundreds of volts, whereas the power range is distributed from 200mW to 100W.

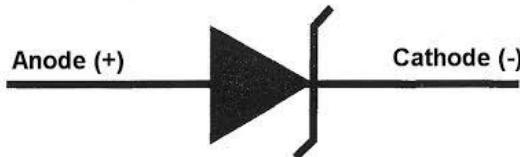


Figure. 1.1.15. Circuit symbol of Zener diode.

Zener Diodes are used in the "REVERSE" bias mode, i.e. the anode is connected to the negative supply. From its I-V characteristics curve as shown in Figure. 1.1.16, it can be said that Zener diode has a region in its reverse bias characteristics of almost a constant voltage regardless of the current flowing through the diode. This voltage across the diode (Zener Voltage,  $V_z$ ) remains nearly constant even with large changes in current through the diode caused by variations in the supply voltage or load. This ability to control itself can be used to great effect to regulate or stabilise a voltage source against supply or load variations. The diode will continue to regulate until the diode current falls below the minimum  $I_z$  value in the reverse breakdown region.

When the Zener diode is forward biased it behaves like ordinary diode. In the reverse biased condition, as the reverse voltage is increased beyond the break down voltage of the diode, the current rises sharply with the applied voltage but the voltage across the diode

remains constant. This behavior of Zener diode is used to provide a constant reference voltage such as in the case of voltage regulation.

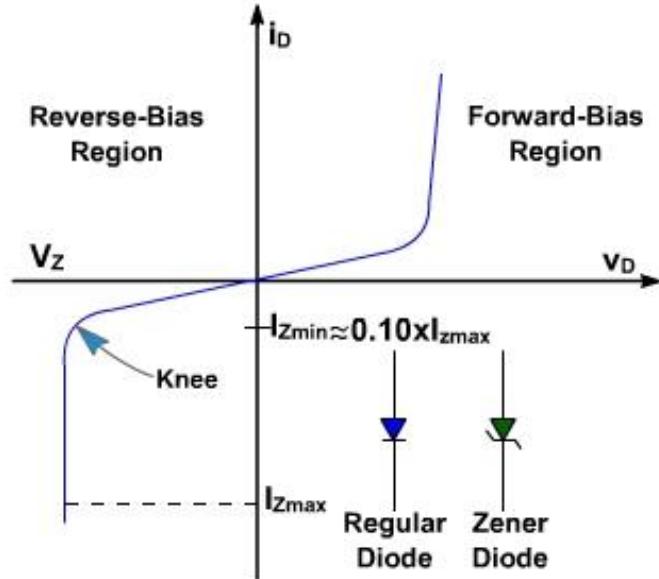


Figure 1.1.16. I-V characteristic of Zener diode  
[<http://www.nptel.ac.in>].

$I_{ZK}$  or  $I_{Zmin}$  – Minimum current necessary to maintain breakdown.

$I_{ZM}$  or  $I_{ZMax}$  – Maximum current that can be safely passed through the zener diode

$P_{ZM}$  or  $P_{ZMax}$  – Maximum power dissipation across zener diode

$$P_{ZM} = V_Z \cdot I_{ZM}$$

Equivalent circuit of zener diode in different conditions is shown in Figure 1.1.17.

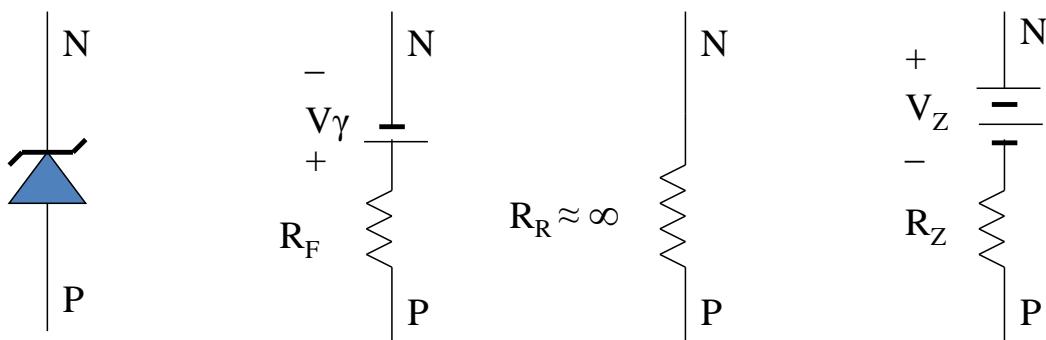


Figure. 1.1.17 (a) Zener diode symbol (b) equivalent circuit in forward biased condition (c) equivalent circuit in reversed biased condition (d) equivalent circuit in breakdown condition

### 1.1.10 Diode as a capacitor

Varactor diode is a special purpose diode that acts as a capacitor and always operates in reverse-bias. It is doped to maximize the inherent capacitance of the depletion region. The

depletion region acts as a capacitor dielectric because of its nonconductive characteristic. The p and n regions are conductive and acts as the capacitor plates, as shown in Figure 1.1.18.

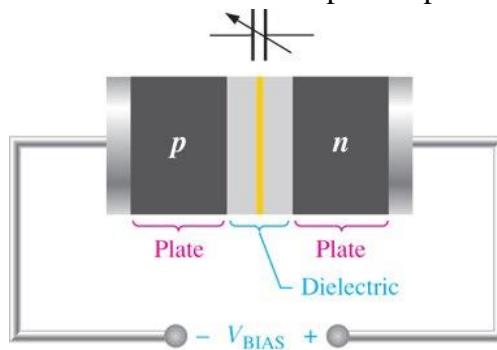


Figure. 1.1.18. Schematic of varactor diode

Capacitance is determined by the parameters of plate area (A), dielectric constant ( $\epsilon$ ) , and plate separation (d) and is given by

$$C = \frac{Ax\epsilon}{d} \quad (1.1.13)$$

As the reverse-bias voltage increases the depletion region widens, effectively increasing the plate separation, thus decreasing the capacitance.

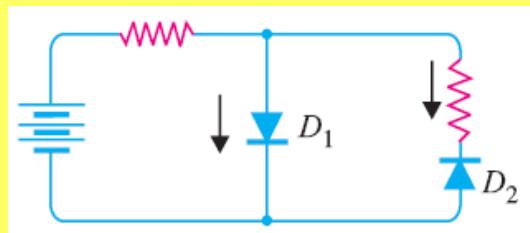
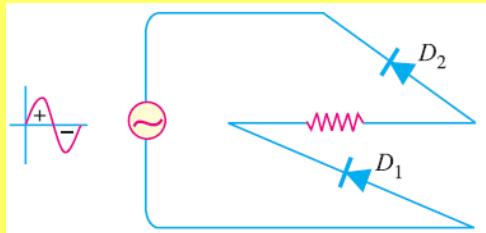
Major application of varactors is in tuning circuits. For example very high frequency (VHF), ultra high frequency (UHF), and satellite receivers employ varactors. When used in a parallel resonant circuit, the varactor acts as a variable capacitor. Thus, allowing the resonant frequency to be adjusted by a variable voltage level. They are also used in Frequency Modulation circuits.

### Summary

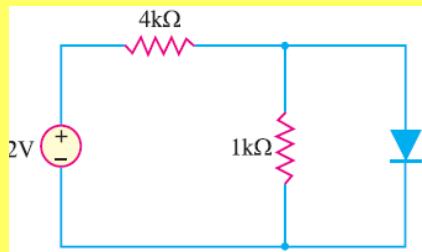
1. PN junction diode in forward biased condition behaves like a closed switch and in reversed biased condition behaves like an open switch.
2. The reverse saturation current of a diode doubles for every  $10^0$  rise in temperature.
3. There are two breakdown mechanisms in a zener diode: avalanche breakdown and zener breakdown.
4. The zener diode is generally used in reverse breakdown region.
5. A zener diode maintains a nearly constant voltage across its terminals over a specified range of zener currents.
6. Varactor diode operates in reverse biased condition and acts as a variable capacitance.

**Exercise Problems:**

1. In each diode circuit shown below, find whether the diodes are forward or reverse biased.



2. Determine the state of diode for the circuit shown below and find  $I_D$  and  $V_D$ . Assume practical model for the diode.



3. Calculate the dynamic forward and reverse resistance of a PN junction diode, when the applied voltage is 0.25V for Germanium Diode.  $I_0 = 1\mu A$  and  $T = 300$  K.

(Ans:  $r_f=1.734\Omega$ ;  $r_r=390M\Omega$ )

4. A germanium diode has reverse saturation current of  $0.19\mu A$ . Assuming  $\eta = 1$ , find the current in the diode when it is forward biased with 0.3 V at  $27^\circ C$ .

(Ans: 19.5mA)

5. The forward current in a Si diode is 15 mA at  $27^\circ C$ . If reverse saturation current is  $0.24nA$ , what is the forward bias voltage? (Ans: 0.93V)

6. A germanium diode carries a current of 10mA when it is forward biased with 0.2V at  $27^\circ C$ . (a) Find reverse saturation current. (b) Find the bias voltage required to get a current of 100mA. (Ans:  $4.42\mu A$ ,  $0.259V$ )

## Module 2: Application of Diodes

In the previous module we have discussed the behaviour and V-I characteristics of PN junction diode. Diode conducts when it is forward biased and behaves like a closed switch. Whereas, during the reverse bias, it goes off and behaves like an open switch. This unilateral behaviour depending on the polarity of external voltage applied to the diode is used in many circuit applications. One such application is conversion of AC voltage to DC, called rectification. In this module we will study various forms of rectifier circuits and their analysis in detail.

### Learning Outcomes:

At the end of this module, students will be able to:

1. Explain the need for AC to DC conversion
2. Draw and explain the block diagram of a basic dc power supply unit.
3. Discuss the importance of the various components used in the rectifier circuits.
4. Discuss the working of a half wave and full wave rectifier circuits.
5. Analyse the performance of rectifier circuits and compare them.
6. Explain the working of rectifier circuits with capacitor filter.

### 1.2.1. Introduction

Today, we cannot imagine life without electronic products like cell phones, computers, laptops, music systems etc., in our daily lives. Some of these electronic systems work on a constant DC voltage derived from the AC mains, while others use internal batteries, which requires regular charging. In any case, some device which at one end receives AC voltage, which is 230V sinusoidal signal of 50Hz as input and produces a constant DC voltage at its output is required. This module aims at providing an insight into some of the basic circuits which converts AC mains in to a constant DC voltage. Such circuits or devices are called regulated power supplies.

A signal obtained from main AC power supply is purely sinusoidal that can be defined in terms of Peak amplitude and Frequency.

- Peak amplitude: The maximum amplitude of an alternating signal on either sides measured from its zero value.
- Frequency: Number of cycles that passes a given point per second. It is equal to reciprocal of time taken to complete one full cycle.

It is mathematically expressed as  $V(t) = A \sin(\omega t)$  and plotted as shown in Figure 1.2.1, where Peak amplitude  $A = 230 \times \sqrt{2}$  V,  $\omega = 2\pi f$  radians/sec, frequency( $f$ ) = 50Hz = 1/20ms.

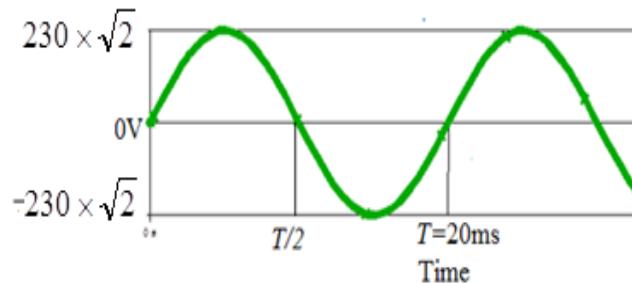


Figure 1.2.1: AC sinusoidal signal with amplitude  $A = 230 \times \sqrt{2}$  V,  $f=50$  Hz signal

*Average or DC value:* The dc value of a signal  $V(t)$  is the average value of that signal. It is mathematically evaluated as:

$$V_{av} = V_{dc} = \frac{1}{\text{Time Period}} \left[ \int_0^T V(t) dt \right]$$

$$V_{av} = V_{dc} = \frac{1}{2\pi} \left[ \int_0^{2\pi} V(t) dt \right] \quad (1.2.1)$$

The root mean square(RMS) value of the signal  $V(t)$  mathematically evaluated as:

$$V_{rms} = \sqrt{\frac{1}{\text{Time Period}} \left[ \int_0^T V^2(t) dt \right]} \quad \text{or} \quad V_{rms} = \sqrt{\frac{1}{2\pi} \left[ \int_0^{2\pi} V^2(t) dt \right]} \quad (1.2.2)$$

Note 1 : A pure sinusoidal signal has an average value equal to zero. It means the dc value of this signal is zero.

### 1.2.2. Basic Block Diagram of DC power supply

The DC power supply converts the AC sinusoidal signal to a DC signal. The block diagram of a basic DC power supply unit is as shown in Figure 1.2.2.

A DC regulated power supply unit consists of the following key components.

- a) Step down transformer
- b) Rectifier circuits
- c) Filter circuit
- d) Regulator

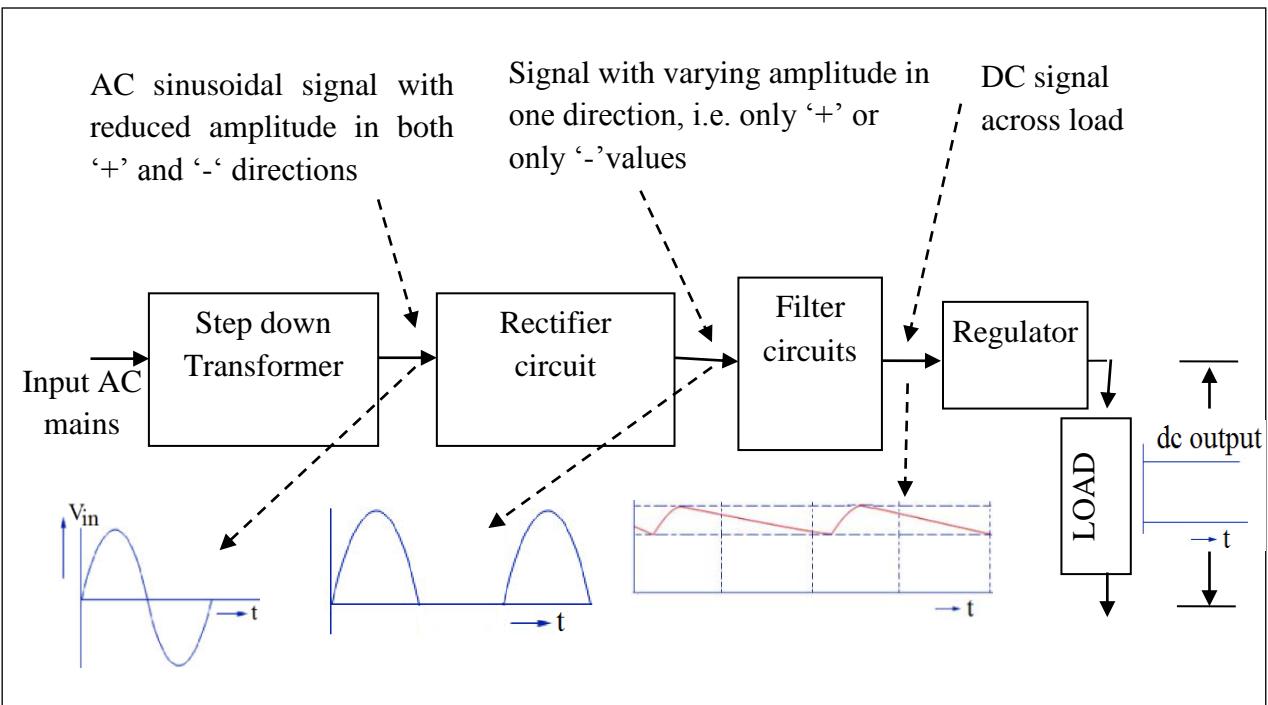


Figure 1.2.2: Basic block diagram of DC power supply

**a) Step down Transformer:**

A transformer is used to bring voltage up or down in an AC electrical circuit. A step down transformer consists of two coils of wire called primary and secondary winding placed such that they are not in contact with each other as shown in Figure 1.2.3a. The symbolic representation of the transformer is shown in Figure 1.2.3b.

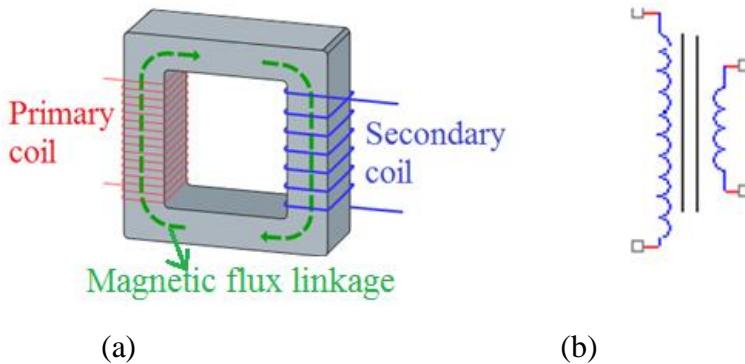


Figure 1.2.3: (a) Transformer core with primary and secondary windings (b) circuit representation of a step down transformer.

A step down transformer converts a high voltage low current power to a low voltage high current power. A step down transformer has large number of turns in primary coil compared to the number of turns in the secondary coil. Hence in this case the secondary voltage is less than the primary voltage and equivalently there is rise in secondary current. The ratio of primary voltage to secondary voltage is proportional to the ratio of number of turns in the primary to the number of turns in the secondary. The 155V AC mains applied at the primary of a step down transformer is stepped down by ten times at the secondary as shown in Figure 1.2.4.

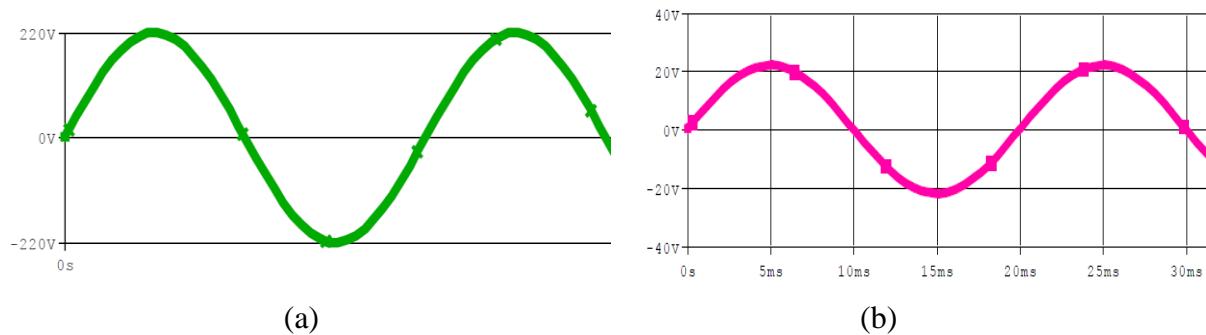


Figure 1.2.4: (a) Primary voltage waveform with peak amplitude of 220V  
 (b) Stepped down version is available at the secondary with peak amplitude of 22V

Transformer used in the DC power supply units plays two important roles. First role is, as discussed earlier, it steps down AC voltage to a suitable value and secondly it provides electrical isolation to the low voltage low power components on the secondary side of the transformer. The second equally important role is to provide electrical isolation to the voltage low power components on the other side of the transformer. This also provides some amount of safety to the equipments using such DC supplies.

*Note 2: The frequency of the primary voltage is equal to the frequency of the secondary voltage of the transformer.*

### b) Rectifier circuit:

A rectifier circuit is the heart of a DC power supply. It converts an AC sinusoidal signal that is bidirectional (with both positive and negative amplitudes) into a signal which is unidirectional (either only positive or only negative). Thus rectifier circuit forces the current through the load to flow in only one direction. The rectified output is usually called a pulsating DC voltage. Thus in general the process of converting an AC signal into pulsating DC signal is called *rectification* and the circuit is called *rectifier*. Rectification is commonly performed using semiconductor diodes because of its inherent unidirectional conduction property.

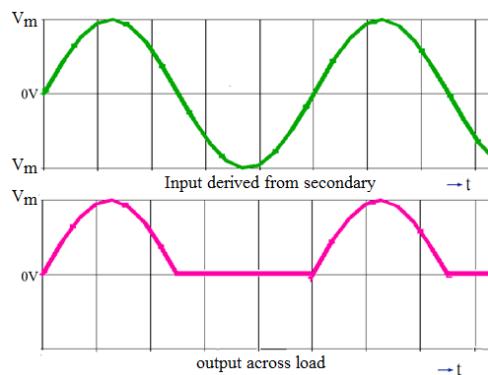


Figure 1.2.5: Secondary transformer voltage and rectified voltage waveform

The rectified output voltage with respect to the input voltage using a single diode is as shown in Figure 1.2.5. The rectifier passes positive half cycle to the output and blocks the negative half cycle. In the case of a full wave rectifier both half cycles will be rectified and available as unidirectional pulses at the output.

Note 3: The AC component of a rectified output signal is not equal to zero. That is the rectified pulsating DC signal has both DC and AC components.

**c) Filter circuit:**

The pulsating DC signal is not suitable for appliances that require pure DC voltage. Filters can be used to minimize (smooth out) the pulsations or eliminate the AC content from the rectified signal to achieve approximately constant valued (or a DC) signal. A filter contains a capacitor, an energy storing component that can hold the voltage to the peak value of the rectified pulsating DC and then dissipate energy to load when the pulsating DC drops as shown in Figure 1.2.6. Essentially, the filter minimises the ac component present in the output of the rectifier. This increases the DC value at the output.

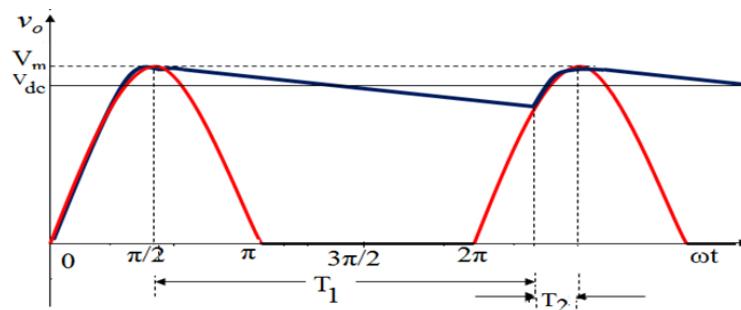


Figure 1.2.6: Illustration of filtering function

Note 4: The output of filter circuit is a DC signal with small AC component called ripples

**d) Regulator**

Regulation is defined as the ability of a system to provide near constant supply over a wide range of load and power line fluctuating conditions. A circuit that can provide constant DC voltage despite of variations in the mains AC power supply or load variations is called voltage regulator.

*Self test:*

1. List out the appliances or electronic products used in daily life which require DC power supply for their operation. Mention the DC values recommended for their operation.
2. List and classify the appliances or electronic products which use DC and AC power supply.
3. Explain the block diagram of a DC power supply.

### 1.2.3 Half Wave Rectifier

The half wave rectifier consists of a single semiconductor diode. The secondary voltage of transformer is fed as an input to the rectifier circuit. The output is measured across the load resistance  $R_L$ .

#### a) Working of an HWR circuit

For simplicity it is assumed that the diodes are ideal, and is represented as an open circuit when reversed biased and short circuit when forward biased. The circuit diagram of half wave rectifier and rectified output waveform with respect to secondary voltage waveform is as shown in Figure 1.2.7(a) and (b) respectively.

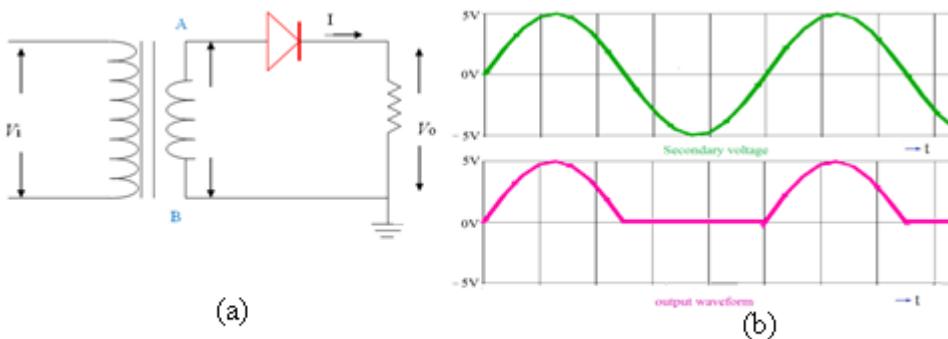


Figure 1.2.7: (a) Half wave rectifier circuit (b) Rectified output waveform

During the positive half cycle of the input waveform, voltage at node A is positive with respect to voltage at node B, which forces the diode to be forward biased and acts as a short. The equivalent circuit for the positive half cycle is as shown in Figure 1.2.8. This results in current flow through the load resistance  $R_L$ . Hence output voltage is approximately equal to the secondary voltage.

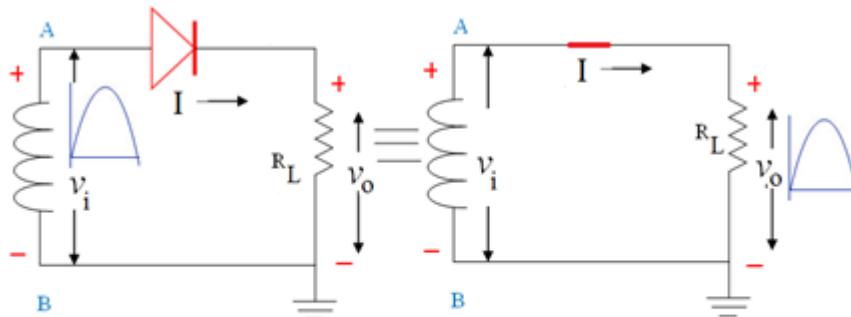


Figure 1.2.8: equivalent circuit of HWR when the input voltage between A and B is positive.

Similarly during the negative half cycle, the voltage at node A is negative with respect to node B that forces the diode to be reverse biased and acts as open circuit. The equivalent circuit is as shown in Figure 1.2.9. This results in no current flow through the load.

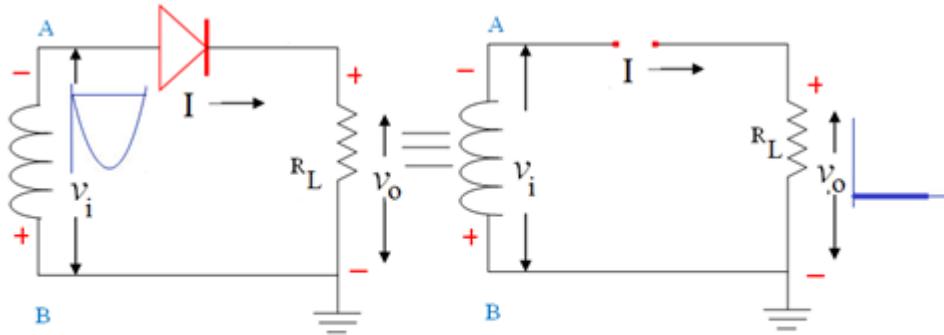


Figure 1.2.9: equivalent circuit of HWR when the input voltage between A and B is negative.

Thus the diode conducts only during positive half cycle and hence the circuit is referred as '*Half Wave Rectifier*'. The rectified voltage is *pulsating DC*, which can be smoothed using filter circuits.

### **Analysis of a HWR circuit**

Filter performance is measured using certain standard parameters. Few of them will be defined and analyzed for each of the filter circuits.

- A. Dc voltage  $V_{dc}$ :** The average value of the output voltage measured across the load resistor.
- B. Ripple Factor:** Ripple factor  $\gamma$  is defined as the ratio of rms value of AC component to DC component of the signal.

$$\gamma = \frac{V_{rms}}{V_{dc}} \quad (1.2.3)$$

where  $V_{rms}$  is the rms value of the ripple (AC component in the pulsating DC) and is given by

$$V_{rms} = \sqrt{V_{rms}^2 - V_{dc}^2} \quad (1.2.4)$$

Ripple factor actually measures the amount of AC content present as compared to the DC (or average) content in the pulsating dc.

Using 1.2.3 in 1.2.4, we can write

$$\gamma = \sqrt{\frac{V_{rms}^2}{V_{dc}^2} - 1} \quad (1.2.5)$$

- C. Efficiency:** Efficiency ( $\eta$ ) is the ratio of the DC output power to AC input power supplied by the secondary of the transformer.

$$\eta = \frac{dc\ output\ power}{ac\ input\ power} = \frac{P_{dc}}{P_{ac}}$$

Efficiency signifies the outcome of the rectifier circuit to output DC power in comparison to AC input power. Thus

$$\eta = \left( \frac{V_{dc}^2 / R_L}{V_{rms}^2 / R_L} \right) \quad (1.2.6)$$

**D. Peak Inverse Voltage (PIV):** It is defined as the peak value of input voltage across the reverse biased diode before the diode breaks down. It is essential that the diode used in rectifier circuit should be able to withstand the voltage available across it when it is reversed biased so that it acts as open circuit and does not enter into a breakdown state.

### Computation of DC Voltage and current for HWR

Consider input signal  $V_S = V_m \sin(\omega t)$  as the secondary voltage signal applied to the half wave rectifier. Then  $V_{av}$  the average or the DC component of the voltage across the load is can be computed using equation (1.2.1).

$$V_{av} = V_{dc} = \frac{1}{2\pi} \left[ \int_0^\pi V_m \sin(\omega t) d(\omega t) + \int_\pi^{2\pi} 0 d(\omega t) \right]$$

$$V_{dc} = \frac{V_m}{2\pi} [-\cos(\omega t)]_0^\pi = \frac{V_m}{\pi} \quad (1.2.7)$$

$$I_{dc} = \frac{V_{dc}}{R_L} = \frac{V_m}{\pi R_L} = \frac{I_m}{\pi} \quad (1.2.8)$$

### Computation of ripple factor for HWR

RMS voltage at the load resistance can be calculated using 1.2.2

$$V_{rms} = \sqrt{\frac{1}{2\pi} \int_0^\pi (V_m \sin \omega t)^2 d(\omega t)} \quad (1.2.9)$$

$$V_{rms} = \sqrt{\frac{V_m^2}{2\pi} \int_0^\pi \left( \frac{1-\cos 2\omega t}{2} \right) d(\omega t)} = \frac{V_m}{2} \quad (1.2.10)$$

Using 1.2.10, and 1.2.7 in 1.2.5, the ripple factor is

$$\gamma = \sqrt{\left( \frac{V_m/2}{V_m/\pi} \right)^2 - 1} = \sqrt{\frac{\pi^2}{4} - 1} = 1.21 \quad (1.2.11)$$

### Computation of efficiency for HWR

Consider  $V_S = V_m \sin(\omega t)$  as the secondary voltage signal applied to the half wave rectifier and using equation 1.2.7 and 1.2.10 in 1.2.6,

$$\eta = \left( \frac{V_{dc}^2 / R_L}{V_{rms}^2 / R_L} \right) = \frac{\left( \frac{V_m}{\pi} \right)^2}{\left( \frac{V_m}{2} \right)^2} = \frac{4}{\pi^2} = 0.406 \text{ or } 40.6\% \quad (1.2.12)$$

### Computation of PIV for HWR

For the HWR, during the negative half cycle the diode is reverse biased. The maximum voltage across the reverse biased diode will be equal to the secondary peak voltage. If the secondary voltage is  $V_s = V_m \sin(\omega t)$ , the PIV of the diode should be greater than  $V_m$ , peak of secondary voltage.

#### Self test 2:

1. Input AC signal of 25V peak value is to be rectified using HWR. For proper working it is essential to choose the diodes whose PIV rating is
   
 (a) 5V      (b) 15V      (c) 30V      (d) both a and b
2. In the circuit of figure 1.2.7 what happens when the diode connection is reversed? Draw the input and output waveforms. Will the values of PIV, ripple factor and efficiency for this circuit change?

### Solved Example

1. A sinusoidal secondary voltage of peak value 10V and frequency 50Hz is applied to half wave rectifier. If the load resistance is  $800\Omega$ . calculate average load current.

Given: peak voltage  $V_m = 10V$ ,  $f = 50\text{Hz}$ ,  $R_L = 800\Omega$

Solution:  $V_{dc} = \frac{V_m}{\pi} = 3.18309 \text{ V}$  (here V=volts)

$$I_{dc} = \frac{V_{dc}}{R_L} = 3.978 \text{ mA}$$
 (here mA=milli Amperes)

#### 1.2.4 Full Wave Rectifier

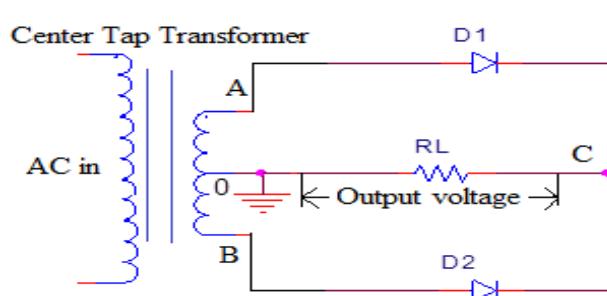


Figure 1.2.11: Circuit diagram of a center tapped FWR

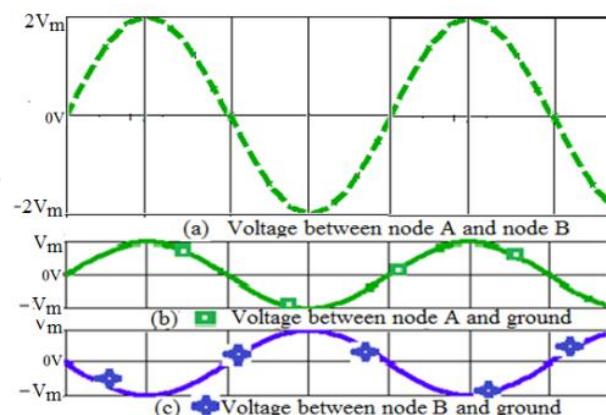


Figure 1.2.12: Waveform at different nodes of a center tapped Full wave rectifier

The circuit of full wave rectifier is shown in Figure 1.2.11 with the center tap of transformer grounded. The center tapped transformer consists of two input terminals (nodes) and three output nodes. The extreme nodes are labelled as node A and node B. The center node is usually used as a common reference voltage terminal relative to which all the voltages are measured. Hence the center output node of center tapped transformer is considered to be a common ground.

The secondary voltage observed between the extreme end nodes i.e. between node A and node B is a stepped down voltage as shown in Figure 1.2.12(a). The voltages measured between node A and center node (ground) or node B and center node is half in magnitude in comparison to the voltage measured between node A and node B. Also voltage at node B is  $180^\circ$  out of phase with the voltage at node A when measured relative to the center node (ground). All these secondary waveforms are shown in Figure 1.2.12(b) and (c).

### **Working of Center tapped Full wave rectifier circuit**

During the first half cycle, as shown in Figure 1.2.11, the voltage at node A is positive and voltage at node B is negative measured with respect to ground. Diode D1 gets forward biased and acts as a short whereas diode D2 is reverse biased and acts as open. Load resistor is connected at the junction of cathodes of the two diodes D1 and D2 with respect to ground. Considering ideal diodes, the equivalent circuit of a center tapped full wave rectifier for half cycle when voltage at node A with respect to node B is positive is as shown in Figure 1.2.13. This results in a current flow through upper half secondary windings of transformer, diode D1 and the load RL as shown in Figure 1.2.13. Direction of the current through the load is towards the ground from node C. Hence the output voltage measured at node C with respect

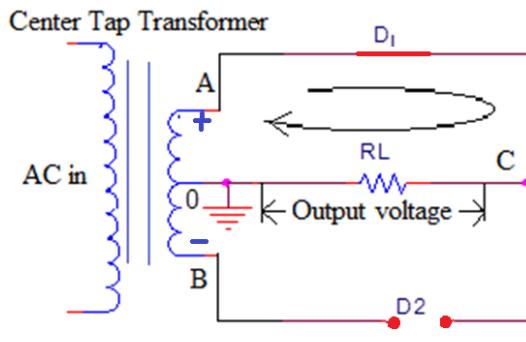


Figure 1.2.13: Equivalent circuit of a center tapped FWR when node A positive

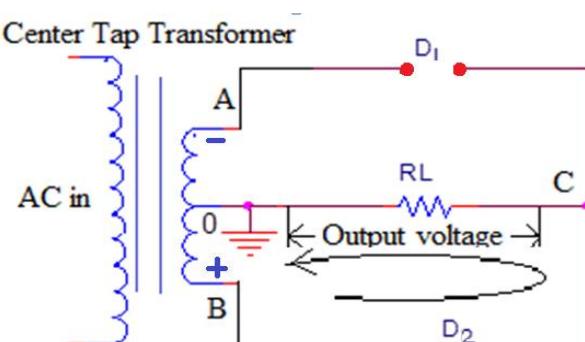


Figure 1.2.14: Equivalent circuit of a center tapped FWR when node B positive

to ground is equal to voltage at node A measured with respect to ground.

Similarly, during the second half cycle, as shown in Figure 1.2.12, the voltage at node A is negative and voltage at node B is positive measured with respect to ground. Diode D1 gets reverse biased and acts as an open whereas diode D2 is forward biased and acts as short. The equivalent circuit of a center tapped full wave rectifier for half cycle when voltage at node A

with respect to node B is negative is as shown in Figure 1.2.14. This results in a current flow through lower half secondary windings of transformer, diode D2 and the load RL as shown in Figure 1.2.14. Again note that the direction of the current through the load is towards the ground from node C. Hence the output voltage measured at node C with respect to ground is equal to voltage at node B measured with respect to ground.

The output waveform observed across load resistor along with voltage waveforms at node A and node B with respect to ground is shown in Figure 1.2.15.

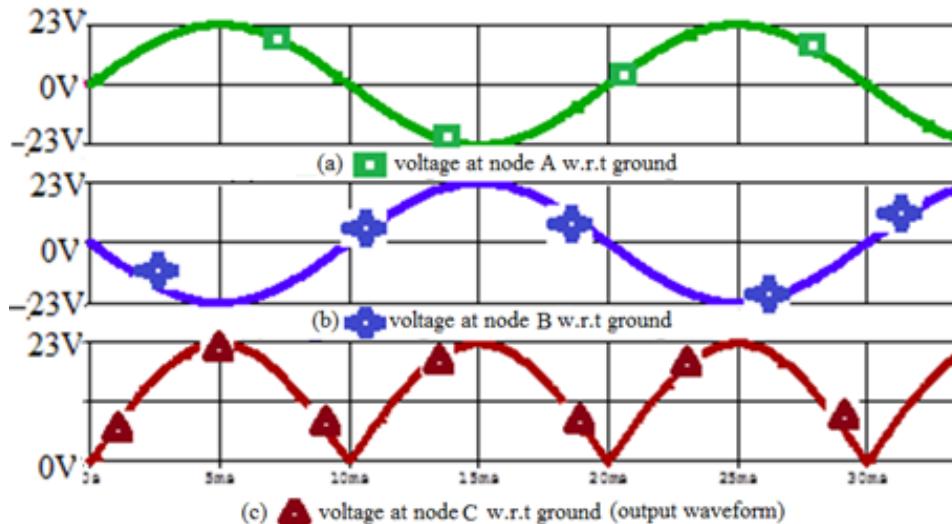


Figure 1.2.15: Secondary and output waveforms of center tapped

#### *Analysis of Center tap FWR circuit:*

##### **Computation of PIV for center-tapped FWR**

Consider Figures 1.2.15 and 1.2.16, the maximum voltage across the reverse biased diode will be twice the peak voltage measured between either of extreme ends and center node of secondary winding of a center tapped transformer. Hence *PIV* of the full wave rectifier is  $2V_m$ .

##### **Computation of Ripple Factor for center-tapped FWR**

The average or the DC component of the voltage across the load using 1.2.1 is given by

$$V_{av} = V_{dc} = \frac{1}{\pi} \left[ \int_0^\pi V_m \sin(\omega t) d(\omega t) \right] = \frac{V_m}{\pi} [-\cos(\omega t)]_0^\pi = \frac{2V_m}{\pi} \quad (1.2.13)$$

$$I_{dc} = \frac{V_{dc}}{R_L} = \frac{2V_m}{\pi R_L} = \frac{2I_m}{\pi} \quad (1.2.14)$$

RMS value of the voltage across the load using 1.2.2 is

$$V_{rms} = \sqrt{\frac{1}{\pi} \int_0^\pi (V_m \sin \omega t)^2 d(\omega t)} = \frac{V_m}{\sqrt{2}} \quad \text{and} \quad I_{rms} = \frac{I_m}{\sqrt{2}} \quad (1.2.15)$$

Using equation 1.2.13 and 1.2.15 in 1.2.5, the ripple factor is

$$\gamma = \sqrt{\left(\frac{V_m/\sqrt{2}}{2V_m/\pi}\right)^2 - 1} = \sqrt{\frac{\pi^2}{8} - 1} = 0.483 \quad (1.2.16)$$

### Computation of Efficiency for center-tapped full wave rectifier

From equation 1.2.6, the efficiency is calculated as

$$\eta = \left( \frac{V_{dc}/R_L}{V_{rms}/R_L} \right) = \frac{\left(2V_m/\pi\right)^2}{\left(V_m/\sqrt{2}\right)^2} = \frac{8}{\pi^2} = 0.812 \text{ or } 81.2\% \quad (1.2.17)$$

#### Self-test :

**Choose the correct answer: (T is the time period of the input signal)**

1. In center tapped FWR, each diode is forward biased for what duration of the time period?  
 (a) T/2            (b) T/4            (c) 3T/4            (d) T
2. In center tapped FWR, current through the load flows for what duration of the time period?  
 (a) T/2            (b) T/4            (c) 3T/4            (d) T
3. Input AC signal of 25V peak value is to be rectified using center tapped FWR. For proper working it is essential to choose the diodes whose PIV rating is  
 (a) 5V            (b) 15V            (c) 30V            (d) both a and b

### 1.2.5 Full Wave Bridge Rectifier

The bridge rectifier consists of four diodes in the form of a bridge. two parallel paths can be represented with each path having two diodes and all diodes are directed in the same direction as shown in Figure 1.2.16. The top path consists of diodes D3 and D1 whereas bottom path consists of diodes D2 and D4 respectively. Load resistor is connected between the ends of two parallel paths (i.e. node C and ground).

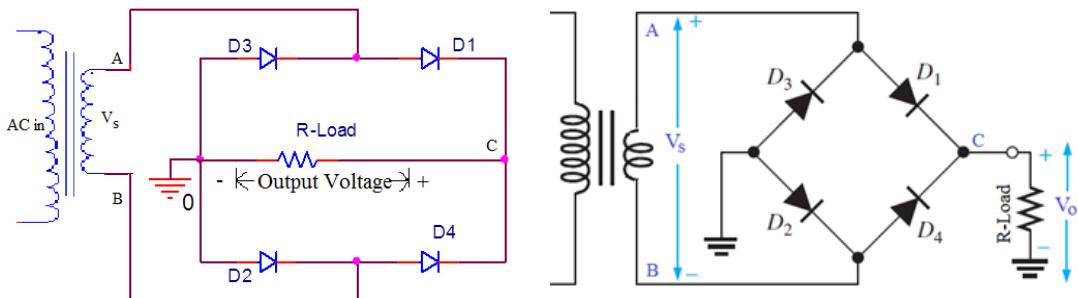


Figure 1.2.16: Bridge rectifier represented in two ways

If the secondary signal given to the bridge is as shown in Figure 1.2.17(a), the output voltage measured across load is expected to be as shown in Figure 1.2.17(b) assuming ideal diodes.

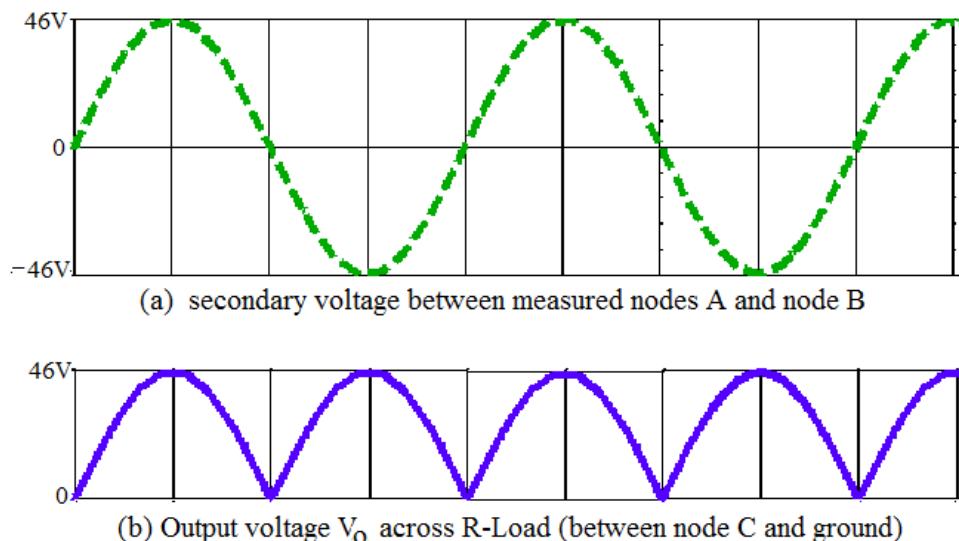


Figure 1.2.17: input and output waveforms of bridge

### **Working of Bridge FWR circuit**

During positive half cycle, node A is positive with respect to B, D1 and D2 are forward biased whereas D3 and D4 are reverse biased as shown in Figure 1.2.18. This results in current flow taking a closed path from node A, through D1, R-Load and D2 and from node B through the secondary coil as indicated in Figure 1.2.18. Note that current through the load resistor flows from node C to ground.

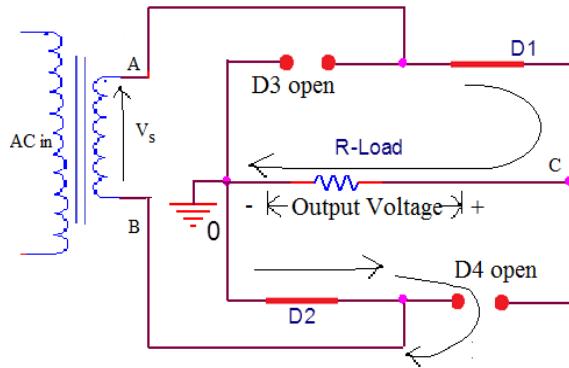


Figure 1.2.18: Equivalent circuit of a bridge rectifier when node A is positive

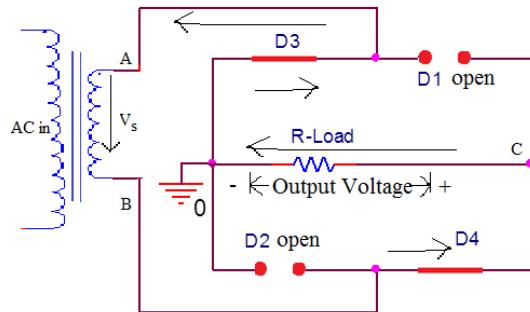


Figure 1.2.19: Equivalent circuit of a bridge rectifier when node B is positive

During negative half cycle, node B is positive with respect to A, D3 and D4 are forward biased whereas D1 and D2 are reverse biased as shown in Figure 1.2.19. This results in current flow taking a closed path from node B, through D4, R-Load and D3 and from node A through the secondary coil as indicated in Figure 1.2.19. Note that again current through the load resistor flows from node C to ground. Thus output voltage is unidirectional for both the half cycles.

#### **Analysis of Bridge FWR circuit**

Consider Figures 1.2.18 and 1.2.19, maximum voltage across the reverse biased diode will be secondary peak voltage. Hence PIV rating of the diode should be greater than  $V_m$ . The  $I_{dc}$ ,  $V_{dc}$ ,  $I_{rms}$ ,  $V_{rms}$ , Ripple factor, Efficiency are same as center tapped full wave rectifier.

**Self-test :**

**Choose the correct answer: (T is the time period of the input signal)**

1. In Bridge rectifier, each diode is forward biased for what duration of the time period?
  - (a)  $T/2$
  - (b)  $T/4$
  - (c)  $3T/4$
  - (d) T
  
2. In Bridge rectifier, current through the load flows for what duration of the time period?
  - (a)  $T/2$
  - (b)  $T/4$
  - (c)  $3T/4$
  - (d) T
  
3. Input AC signal of 25V peak value is to be rectified using bridge. For proper working it is essential to choose the diodes whose PIV rating is
  - (a) 5V
  - (b) 15V
  - (c) 30V
  - (d) both a and b

### Solved Example

2. Find the *PIV* rating of the diode used for proper working of the bridge rectifier when it is supplied with 230V, 50 Hz AC mains through a step down transformer with turns ratio equal to 10.

Given: Input AC mains peak voltage = 230V, turns ratio=10,

Solution: Secondary peak voltage  $V_m = 230/10 = 23V$

therefore *PIV* rating of diode  $\geq V_m = 23V$

### 1.2.6 Comparison of Rectifiers

Advantages of HWR over FWR

- Simple circuit
- Single diode
- PIV rating is  $V_m$

Disadvantages of HWR

- High ripple factor
- Low efficiency

Advantages of Center tapped FWR rectifier over HWR

- High Efficiency
- Low ripple factor

Advantages of bridge rectifier over to centre-tap full wave rectifier:

- PIV rating is  $V_m$
- Centre-tap transformer is not required

Disadvantage of bridge rectifier over to centre-tap full wave rectifier:

- Need for four diodes

Comparison of rectifier circuits based on the performance parameters is listed in Table 1.2.1, considering the input signal  $V_i(t) = V_m \sin(2\pi f_i t)$

Table 1.2.1: Comparison of Rectifiers

Parameters of rectified signal	HWR	Center-tapped FWR	Bridge FWR
$V_{dc}$	$\frac{V_m}{\pi}$	$\frac{2V_m}{\pi}$	$\frac{2V_m}{\pi}$
$V_{RMS}$	$\frac{V_m}{2}$	$\frac{V_m}{\sqrt{2}}$	$\frac{V_m}{\sqrt{2}}$
<i>PIV</i>	$\geq V_m$	$\geq 2V_m$	$\geq V_m$
Ripple factor	1.21	0.483	0.483
Efficiency	40.6%	81.2%	81.2%
Frequency $f_o$	$f_i$	$2f_i$	$2f_i$

### 1.2.7 Rectifier with Filter

A capacitor filter with HWR and FWR is as shown Figure 1.2.20. The capacitor allows AC component and blocks DC component. The capacitor filter minimizes the ripple and increases the average value of output voltage.

In each of the positive half cycle, the capacitor charges up to the peak value of the transformer secondary voltage,  $V_{in}$ . Capacitor tries to maintain this maximum value when the input drops to zero. The capacitor will discharge through the load resistance slowly until the input voltage again increases to a value greater than the capacitor voltage. The filtered output waveform for both HWR and FWR is as shown in figure 1.2.21.

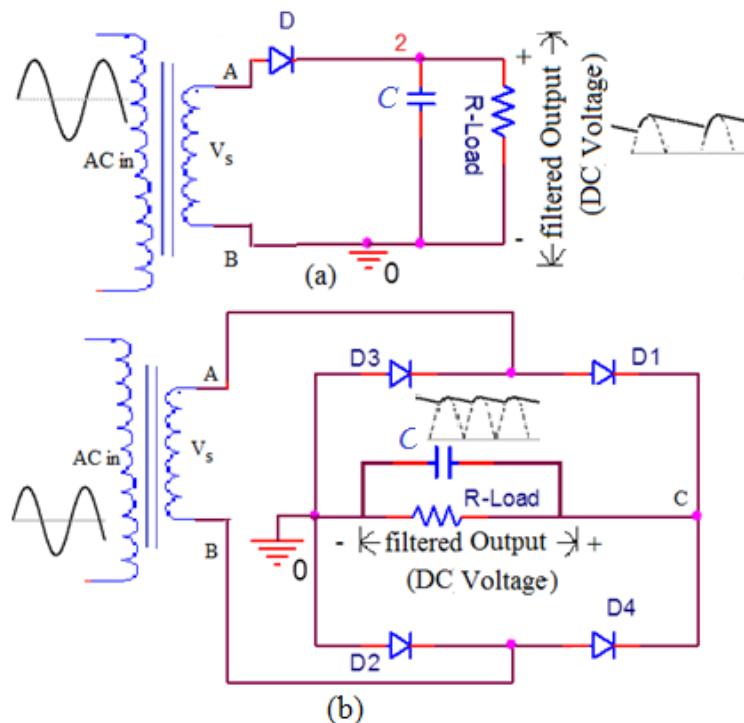


Figure 1.2.20: Filter circuit for (a) half wave rectifier and (b) full wave rectifier

Large value of the product “ $CR_L$ ” results in a small ripple factor. Thus increasing C or  $R_L$  (both) an approximate perfect DC voltage can be obtained.

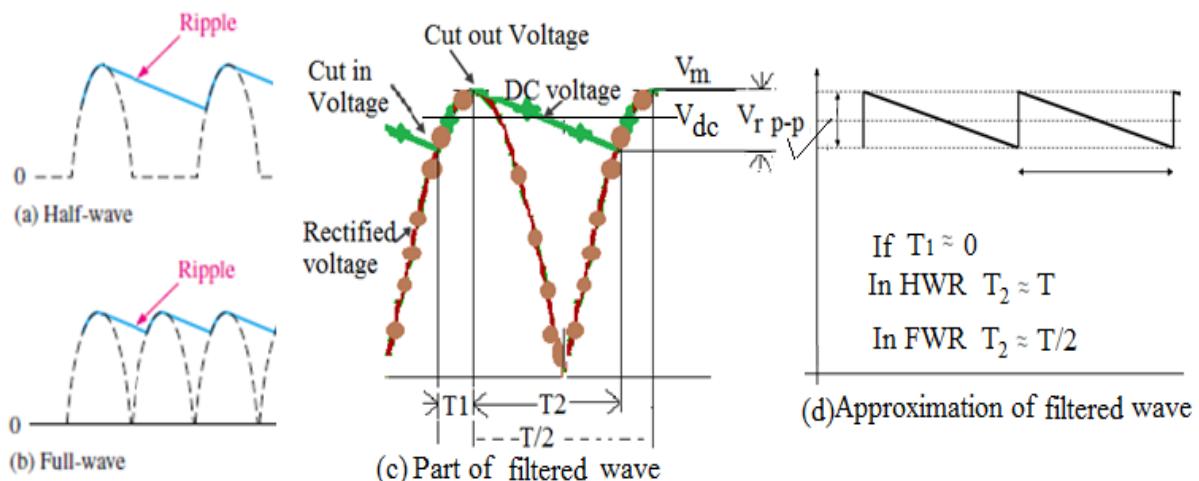


Figure 1.2.21: Filtered waveform (a) half wave rectifier and (b) full wave rectifier [floyd] (c) filtered and rectified wave (d) approximation of filtered wave

The performance of the filter circuits is measured by ripple factor. The approximate filtered waveform shown in Figure 1.2.21(d). The ripple factor for Capacitor filter for the HWR and FWR is given by equation (1.2.18) and (1.2.19) respectively.

$$r = \frac{1}{2\sqrt{3}fCR_L} \quad (1.2.18)$$

$$r = \frac{1}{4\sqrt{3}fCR_L} \quad (1.2.19)$$

The corresponding dc value for HWR and FWR is given by equation (1.2.20) and (1.2.21) respectively.

$$V_{dc} = \frac{2f CR_L}{1 + 2f CR_L} V_m \quad (1.2.20)$$

$$V_{dc} = \frac{4f CR_L}{1 + 4f CR_L} V_m \quad (1.2.21)$$

Note 5: Here ' $f$ ' in equation (1.2.18) to (1.2.21) is the frequency of the input signal

Comparison of ripple factor and the output dc voltage achieved from the input secondary  $V_i(t) = V_m \sin(2\pi ft)$  is listed in Table 1.2.2

Table 1.2.2: Comparison of C-filter with HWR and FWR

Parameters of rectified signal	HWR	FWR
$V_{dc}$	$\frac{2fCR_L}{1+2fCR_L}V_m$	$\frac{4fCR_L}{1+4fCR_L}V_m$
Ripple factor	$r = \frac{1}{2\sqrt{3}fCR_L}$	$r = \frac{1}{4\sqrt{3}fCR_L}$

**Solved Exercise:**

3. A sinusoidal voltage of peak value  $V_i = 20 \sin(2\pi 50t)$  V is applied to FWR. If the load resistance is  $1000\Omega$ . calculate the average and RMS value of load current, efficiency and ripple factor. Find the frequency of the output signal.

**Given**

$$V_m = 20V, f=50Hz, R_L = 1000 \Omega, R_f=10 \Omega$$

**Solution:**

$$I_m = \frac{V_m}{R_L} = 20mA , \quad I_{dc} = \frac{2I_m}{\pi} = 12.73 mA , \quad I_{rms} = \frac{I_m}{\sqrt{2}} = 14.142 mA$$

$$\text{Efficiency } \eta = \left( \frac{\frac{V_{dc}^2}{R_L}}{\frac{V_{rms}^2}{R_L}} \right) = 81.2\%$$

$$\text{Ripple factor} = \frac{I_{dc}}{I_{ac}} = 0.6365$$

FWR output signal frequency = 2 x frequency of the input signal = 100Hz.

4. A particular load has to be supplied with 10 mA average current at 5 V dc voltage, with ripple factor not more than 10%. Calculate the value of the filter capacitor that needs to be connected to the output of full wave bridge rectifier.

**Given:**

$$I_{dc} = 10 \text{ mA}, \quad V_{dc} = 5 \text{ V}, \quad r = 0.1,$$

Note: Input is not given but can be assumed appropriately. For example in INDIA, AC mains has  $V_{i(rms)} = 230 \times \sqrt{2}$  V,  $f = 50$  Hz

**Solution:**

$$I_m = \frac{V_m}{R_L} = 20mA , \quad I_{dc} = \frac{2I_m}{\pi} = 12.73\ mA , \quad I_{rms} = \frac{I_m}{\sqrt{2}} = 14.142\ mA$$

$$\text{Efficiency } \eta = \left( \frac{\frac{V_{dc}^2}{R_L}}{\frac{V_{rms}^2}{R_L}} \right) = 81.2\%$$

$$\text{Ripple factor} = \frac{I_{dc}}{I_{ac}} = 0.6365$$

FWR output signal frequency = 2 x frequency of the input signal = 100Hz.

5. A particular load has to be supplied with 10 mA average current at 5 V dc voltage, with ripple factor not more than 10%. Calculate the value of the filter capacitor that needs to be connected to the output of full wave bridge rectifier.

**Given:**

$$I_{dc} = 10\ mA, \ V_{dc} = 5\ V, \ r = 0.1,$$

Note: Input is not given but can be assumed appropriately. For example in INDIA, AC mains has  $V_{i(rms)} = 230 \times \sqrt{2}\ V, f = 50\ Hz$

$$R_L = V_{dc} / I_{dc} = 5\ V / 10\ mA = 500\ \Omega$$

$$\text{Solution: } r = \frac{1}{4\sqrt{3}fCR_L}. \text{ Hence, } C = \frac{1}{4\sqrt{3}frR_L} = 57.74\ \mu F$$

6. For the data provided in problem 4, find the turns ratio required for the transformer.

**Solution:**

Using equation (1.2.21), the peak value of the secondary is

$$V_m = \frac{V_{dc}(1+4f CR_L)}{4f CR_L} = 5.866\ V$$

By assumption AC mains  $V_{i(rms)} = 230 \times \sqrt{2}$  to get secondary peak voltage 5.866V a

$$\text{step down transformer. The turns ratio} = \frac{230 \times \sqrt{2}}{5.866} = 55$$

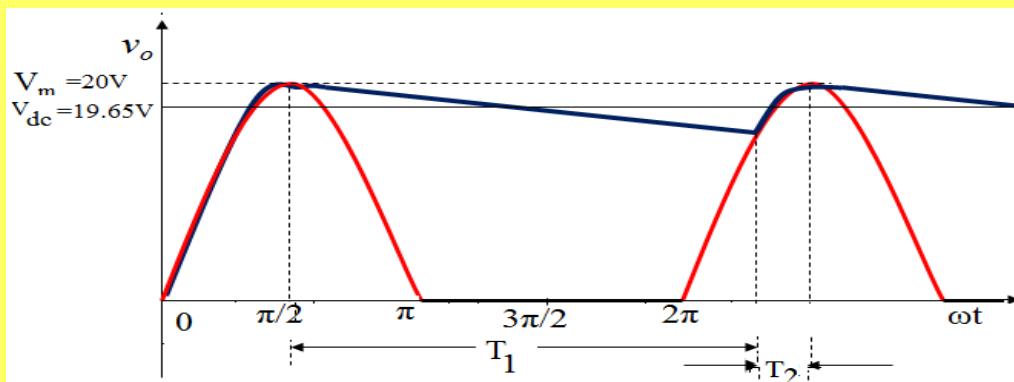
## Summary

1. A pure sinusoidal signal has an average value equal to zero. It means the dc value of this signal is zero.
2. A DC power supply unit consists of : Step down transformer, Rectifier circuits, Filter circuit and Regulator.
3. In a HWR the diode conducts only during positive half cycle and hence the circuit is referred as ‘Half Wave Rectifier’. The rectified voltage is pulsating DC, which can be smoothed using filter circuits.
4. In a FWR the diodes conduct in both half cycles and hence the name Full Wave Rectifier.
5. In a bridge rectifier there are four diodes.
6. A capacitor filter is used in rectifier circuits to remove DC components called as ripples.

**Exercise:**

1. Primary voltage is 120V, 60Hz. Turns ratio is 5:1. This transformer supplies to bridge rectifier employing 4 identical ideal diodes. The load resistance is  $1\text{k}\Omega$ . Calculate average and rms load voltage, efficiency, ripple factor, PIV rating and the frequency of output waveform.  
(Ans.: 108V, 120V, 81%, 0.484, 169.7V, 120HZ)
2. Repeat this problem for center tapped FWR. Comment on the results comparing it with results of problem 1.

3. A half wave rectifier with capacitor filter is supplied from a transformer having peak secondary voltage 20V and frequency 50Hz. The load resistance is  $560\Omega$  and capacitor used is  $1000\mu F$ . Calculate ripple factor and dc output voltage. Draw the filtered output and label peak and dc value. (Ans. 0.0103, 19.65V)



4. Repeat problem 3 for a full wave rectifier

5. A half wave rectifier with capacitor filter has to supply an average voltage of 30V to  $900\Omega$  load. Calculate the rms input voltage and value of capacitor needed to get ripple factor of 0.05, assuming  $f = 50\text{Hz}$ .  
 (Ans. : a) 23V, 128.3 $\mu F$ )

6. Repeat problem 5 for a full wave rectifier.

## Module – 3: Voltage Regulators

### Learning Outcomes:

At the end of this module, students will be able to:

1. Describe the working of Zener as voltage regulator
2. Understand the IC based voltage regulator.

#### 1.3.1 Zener Voltage Regulator

Zener diodes are widely used to regulate the voltage. It is connected in parallel with a variable voltage source in such a way that it is reverse biased, zener diode acts as an open circuit until the voltage reaches the diode's reverse breakdown voltage and the voltage at the output is equal to the voltage applied. Once the voltage goes beyond the breakdown voltage, the voltage across the diode remains constant at the value equal to the breakdown voltage. This behavior of the Zener diode make it a good reference voltage source and one of its major application is in voltage regulators. Figure 1.3.1 is a simple voltage control circuit using a Zener diode.

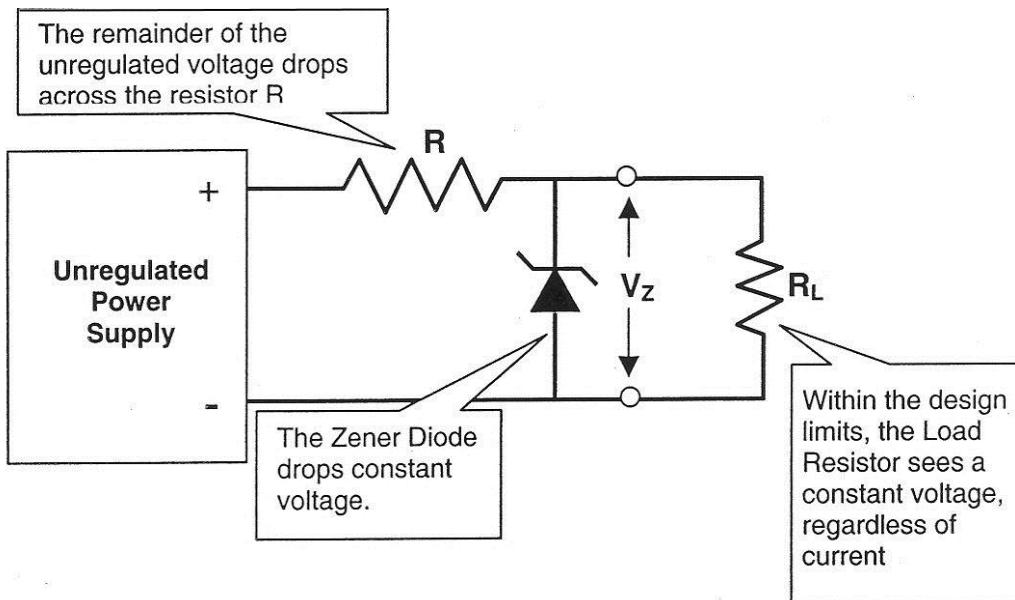


Figure. 1.3.1. Zener Voltage Regulator Circuit [http://rsandas.com/P2\_Session\_4-4.html].

In Figure 1.3.1, there is a power supply that outputs unregulated voltage of 12-volts. A load, represented by resistor  $R_L$ , requires a regulated 5-volt source. Using a 5-volt Zener diode as illustrated, this requirement can be met. In the circuit, load drops a portion of the source voltage. Since the Zener Diode is in series with resistor  $R$ , Zener diode will drop 5-volts ( $V_Z$ ) and series resistor  $R$  will drop the remaining 7 Volts. With the load,  $R_L$ , connected across the Zener Diode, it will provide a constant 5-volts; regardless of any variation of the power supply.

If the power supply voltage drops to 10-volts from initial 12 volts as mentioned above. The Zener diode still drops 5-volts (designed voltage) and the series resistor R will drop 5 volts. Again, because the load is connected across the Zener diode, it will produce 5-volts. Therefore, irrespective of change in line voltage, output voltage remains constant of 5V.

*Self test:*

1. If the series resistance increases in an unloaded Zener regulator, the Zener current
  - a. Decreases
  - b. Stays the same
  - c. Increases
  - d. Equals the voltage divided by the resistance
  
2. In a loaded Zener regulator, which is the largest current?
  - a. Series current
  - b. Zener current
  - c. Load current
  - d. None of these
  
3. If the load resistance increases in a Zener regulator, the Zener current
  - a. Decreases
  - b. Stays the same
  - c. Increases
  - d. Equals the source voltage divided by series resistance
  
4. A voltage regulator is a circuit which
  - a. Converts the ac voltage to dc voltage
  - b. Smoothens the ac variation in dc output voltage
  - c. Maintains a constant dc output voltage inspite of the fluctuations in ac input voltage or load current
  - d. None of the above

### Zener diode used for line and load regulation

A zener diode can be used as a *shunt voltage regulator* (*shunt* meaning connected in parallel), and voltage regulator being a class of circuit that produces a stable voltage across varying load and input voltages.

Consider the circuit for the voltage regulator as shown in Figure 1.3.2.

The circuit has to maintain constant voltage across a load resistor  $R_L$ . The circuit holds the voltage across the load  $R_L$  almost equal to the voltage across zener  $V_Z$  even after the input  $V_{in}$  and load resistor  $R_L$  undergo changes. If the unregulated DC voltage  $V_{in}$  rises, the current through R increases. This extra current is directed to the zener diode instead of flowing through the load. The zener diode voltage is virtually unaffected by the increase in this current and load voltage which is same as the diode voltage  $V_z$  remains constant. If the load

requires more current when  $R_L$  is decreased, the zener diode can supply the extra current without affecting the load voltage.

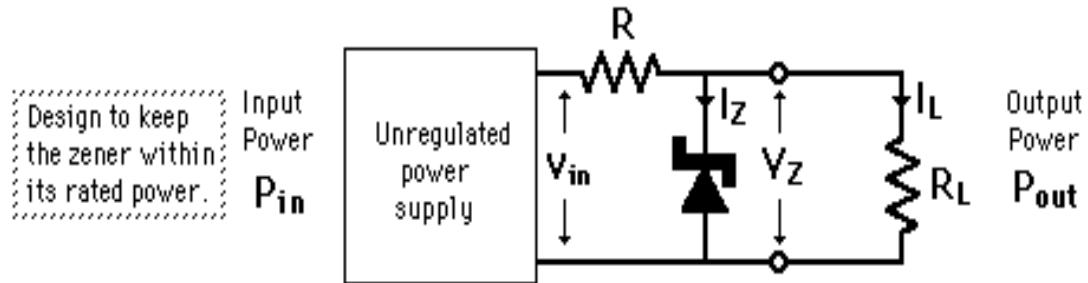


Figure 1.3.2. Zener voltage regulator

Let  $I$  be the current through the resistor  $R$ , and it can be written as

$$\begin{aligned} I &= I_z + I_L \\ I &= \frac{V_{in} - V_z}{R} \end{aligned} \quad (1.3.1)$$

$$\text{The power dissipated in the diode is } P_z = I_z V_z \quad (1.3.2)$$

The selection of  $R$  can be done by using the equation,

$$R = \frac{V_{in} - V_z}{I} \quad (1.3.3)$$

After substituting the value of  $I$ ,

$$R = \frac{V_{in} - V_z}{I_z + I_L} \quad (1.3.4)$$

(i) For Line regulation  $R_L$  is constant and  $I_L = \frac{V_z}{R_L}$  is also constant and  $V_{in}$  varies between  $V_{in(min)}$  to  $V_{in(max)}$

(ii) For Load Regulation,  $V_{in}$  is constant and  $R_L$  varies between  $R_{Lmin}$  and  $R_{Lmax}$  and load current is given by  $I_{Lmin} = \frac{V_z}{R_{Lmax}}$  and  $I_{Lmax} = \frac{V_z}{R_{Lmin}}$

When  $V_{in}=V_{in(min)}$ , and  $I_L$  is constant then

$$I_{\min} = \frac{V_{in(\min)} - V_z}{R}, \quad (1.3.5)$$

$$I_{\min} = I_{z(\min)} + I_L \quad (1.3.6)$$

Similarly when  $V_{in}=V_{in(\max)}$   $I_{\max} = \frac{V_{in(\max)} - V_z}{R}$ ,  $(1.3.7)$

$$I_{\max} = I_{z(\max)} + I_L \quad (1.3.8)$$

The selected R must be small enough to permit minimum zener current to ensure that the diode is in its breakdown region. That is R must be small enough to ensure that minimum current  $I_{Z(\min)}$  flows under worst condition. This is when  $V_{in}$  falls to its smallest possible value  $V_{in(\min)}$  and  $I_L$  is its largest possible value  $I_{L\max}$  (Load Regulation). At the same time R must be selected large enough to ensure that the current through the zener diode should not exceed the maximum zener current  $I_{Z(\max)}$  so that power desipation in the diode will not exceed  $P_z$ . That is the condition when  $V_{in}$  rises to the value of  $V_{in(\max)}$  and load current  $I_L$  to its minimum  $I_{L\min}$

Therefore,

$$R \leq \frac{V_{in(\min)} - V_z}{I_{z\min} + I_{L\max}} \quad \text{and} \quad R \geq \frac{V_{in(\max)} - V_z}{I_{z\max} + I_{L\min}} \quad (1.3.9)$$

### Applications:

- As Voltage regulators
- As Voltage Limiters
- Wave shaping
- Protection diode
- Fixed reference voltage

### Example of Zener Regulation with varying input voltage (Line regulation)

A Zener diode can be used as a voltage regulator. To illustrate this, let's use a Zener diode 1N4740A in the circuit as shown in Figure 1.3.3.

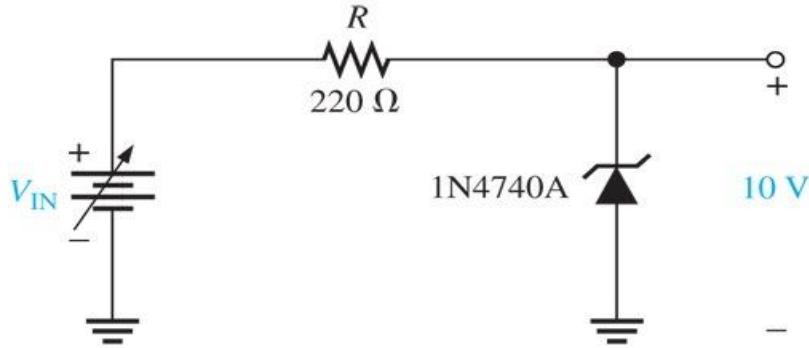


Figure 1.3.3 Zener diode as regulator for line regulation.

As  $V_{IN}$  changes  $I_Z$  changes, the limitations on the input voltage variation ( $V_{IN(\min)}$  and  $V_{IN(\max)}$ ) are set by the minimum and maximum current levels ( $I_{ZK}$  and  $I_{ZM}$ ) with which the Zener diode can operate.

The minimum current value  $I_{ZK} = 0.25\text{ mA}$  (from the 1N4740A Zener diode datasheet). Maximum current can be calculated from the power specification ratings,  $P_{D(\max)} = 1\text{ Watt}$  as follows:

$$I_{ZM} = P_{D(\max)} / V_Z = 1\text{W}/10\text{V} = 100\text{ mA} \quad (1.3.10)$$

For the minimum Zener current, the voltage across the  $220\Omega$  resistor is:

$$V_R = I_{ZK} \times R = 0.25\text{ mA} \times 220\Omega = 55\text{ mV} \quad (1.3.11)$$

Since  $V_{IN} = V_R + V_Z$ , then

$$V_{IN(\min)} = V_R + V_Z = 55\text{ mV} + 10\text{ V} = 10.055\text{ V} \quad (1.3.12)$$

For the maximum Zener current, the voltage across the  $220\Omega$  resistor is:

$$V_R = I_{ZM} \times R = 100\text{ mA} \times 220\Omega = 22\text{ V} \quad (1.3.13)$$

$$\text{Therefore } V_{IN(\max)} = V_R + V_Z = 22\text{ V} + 10\text{ V} = 32\text{ V} \quad (1.3.14)$$

This shows that the Zener diode 1N4740A can ideally regulate voltage from  $10.055\text{ V}$  to  $32\text{ V}$ , and maintain an approximate  $10\text{ V}$  output.

### Example of Zener Regulation with a variable load (load regulation)

Figure 1.3.4 shows a Zener voltage regulator with a variable load resistor across its terminal:

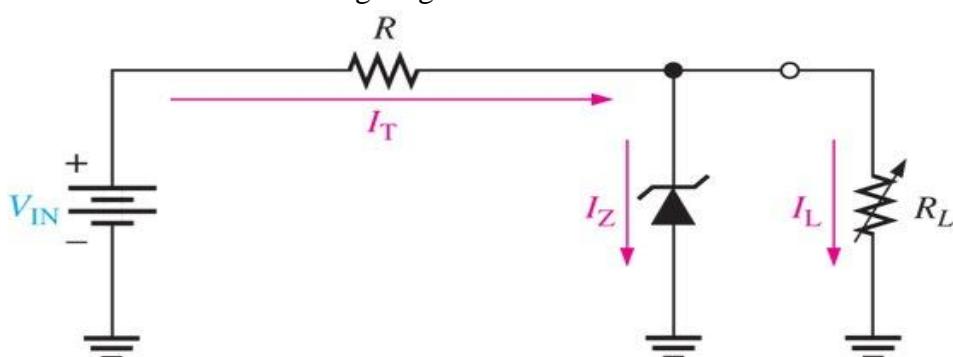


Figure 1.3.4 Zener diode as regulator for load regulation.

To understand the Zener Regulation with a variable load (load regulation), consider the following example.

Design a zener regulator circuit to meet the following specifications:

Load voltage=8V, input voltage=30 V, Load current=0-50 mA,  $I_{z\min}= 5 \text{ mA}$ ,  $P = 1\text{W}$

$$V_{in(\min)} = V_{in(\max)} = 30 \text{ V}, I_{L\min} = 0\text{A}, I_{L\max} = 50 \text{ mA}$$

$$I_{Z\max} = \frac{P_Z}{V_o} = \frac{1}{8} = 125 \text{ mA}, R_{L\min} = \frac{V_o}{I_{L\max}} = \frac{8}{50 \text{ mA}} = 160 \Omega$$

To find current limiting series resistance,

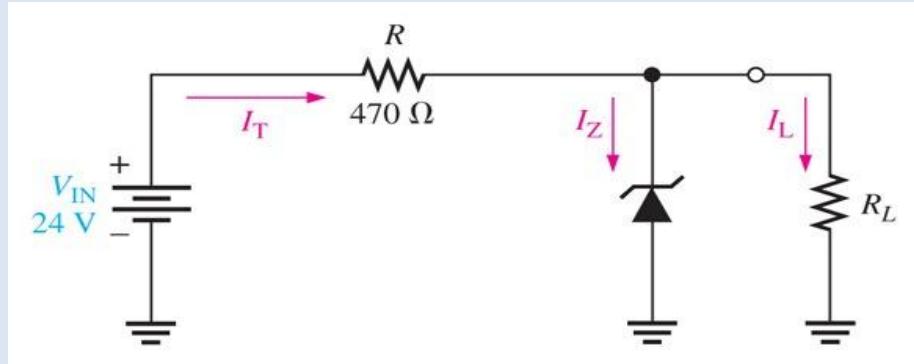
$$R_{\max} = \frac{V_{in\min} - V_o}{I_{L\max} + I_{Z\min}} = \frac{30 - 8}{0.05 - 0.005} = 400 \Omega$$

$$R_{\min} = \frac{V_{in\max} - V_o}{I_{L\min} + I_{Z\max}} = \frac{30 - 8}{0 + 0.125} = 176 \Omega$$

### Example Problem 1:

1. An 1N756 zener diode is used as a 12 V regulator in the circuit shown below:

What is the smallest load resistor that can be used before losing regulation? Assume the ideal model for the zener diode?



### Solutions:

The no load zener current

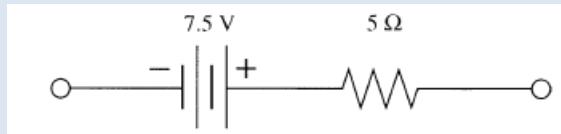
$$I_{NL} = (V_{IN} - V_Z) / R = (24 \text{ V} - 12 \text{ V}) / 470 \Omega = 25.5 \text{ mA}$$

This is the maximum load current in regulation, therefore the minimum value of load resistance  $R_{L(\min)} = V_Z / I_{NL} = 12 \text{ V} / 25.5 \text{ mA} = 470 \Omega$ .

Note that if  $R_L$  is less than  $470 \Omega$  it will draw more of the total current away from the zener diode and  $I_Z$  will be reduced below  $I_{ZK}$ . This will cause the zener diode to come out of break down and hence output will not be regulated.

2. A certain zener diode has a  $V_z = 7.5 \text{ V}$  and an  $R_z = 5 \Omega$  at a certain current. Draw the equivalent circuit.

**Solutions:** The equivalent circuit is shown below.

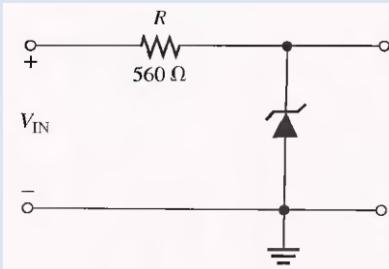


3. When the reverse current in a particular zener diode increases from 20 mA to 30 mA, the zener voltage changes from 5.6 V to 5.65 V. What is the resistance of this device?

**Solutions:**

$$R_z = \frac{\Delta V_z}{\Delta I_z} = \frac{5.65 - 5.6}{30mA - 20mA} = 5\Omega$$

4. Determine the minimum input voltage required for regulation to be established in the figure shown below. Assume an ideal zener diode with minimum zener current = 1.5 mA and  $V_z = 14 \text{ V}$ .



**Solutions:**

$$V_{IN(\min)} = V_z + I_{ZK}R = 14 + (1.5mA)(560\Omega) = 14.8V$$

### 1.3.2 IC Voltage Regulators

IC voltage regulators are versatile, relatively inexpensive and are available with features such as programmable output, current / voltage boosting IC voltage regulators are available as

1. Fixed voltage regulator
2. Adjustable voltage regulators

#### Fixed voltage regulators:

78XX series are three terminal positive voltage regulators. In 78XX, XX indicates the output voltage. They are available as 7805, 7806, 7808, 7815, 7818, and 7819. 79XX series are negative fixed voltage regulators which are complements to the 78XX series devices. MC7805 is a 3-terminal positive voltage regulator. It is designed for a wide range of applications. An example is shown in Figure 1.3.5.

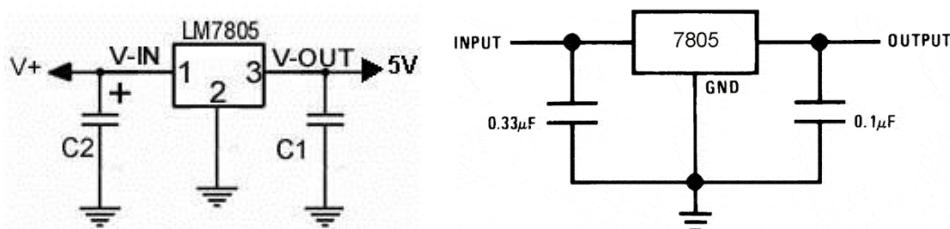


Figure 1.3.5: IC voltage regulators

The internal limiting and thermal shutdown features of this regulator makes it essentially immune to overload. When used as a replacement for a Zener diode-resistor combination, an effective improvement in output impedance can be obtained together with lower-bias current. For output current up to 1A, no external components are required. The input capacitor is used to cancel the inductive effects due to long distributive leads and the output capacitor to improve the transient response.

IC regulator like LM117, LM317, LM338 are adjustable voltage regulators. The LM117 series of adjustable 3-terminal positive voltage regulators is capable of supplying in excess of 1.5A over a 1.2V to 37V output range. They are exceptionally easy to use and require only two external resistors to set the output voltage. Further, both line and load regulations are better than standard fixed regulators. Normally, no capacitors are needed unless the device is situated more than 6 inches from the input filter capacitors in which case an input bypass is needed. An optional output capacitor can be added to improve transient response.

#### Summary

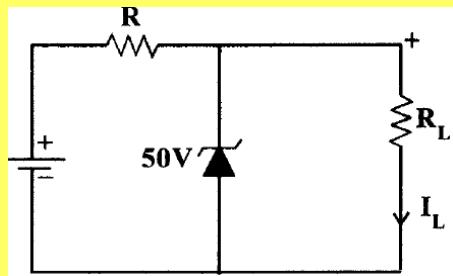
1. Zener diodes are used as voltage references, regulators, and limiters.
2. Zener diodes are available in many voltage ratings ranging from 1.8 V to 200 V.

### Exercise problems

1. Define Rectification, filtering and regulation.
2. List the applications of Varactor diode.
3. (a) Consider the circuit shown below.

The Zener Diode regulates at 50V over a range of diode current from 5 mA to 40 mA. Supply voltage  $V = 200V$ . Calculate the value of  $R$  to allow voltage regulation from a load current  $I_L = 0$  upto  $I_{max}$ ; the maximum possible value of  $I_L$ . What is  $I_{max}$ ?  
 (Ans:  $I_{max}=35$  mA)

- (b) If  $R$  is set as in part (a) and  $I_L = 25$  mA, what are the limits between which  $V$  may vary without loss of regulation in the circuit?



(Ans:  $V_{min}=162.5$  V,  $V_{max}=293.8$  V )

## Module – 4: Special purpose diodes

### Learning Outcomes:

At the end of this module, students will be able to:

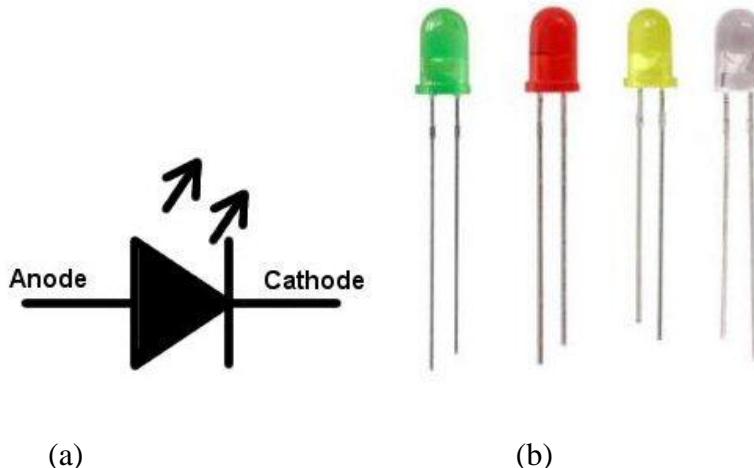
1. Explain the working of LEDs and photo-diode.
2. List the applications of LEDs and photo-diode.

#### 1.4.1 Light Emitting Diode

Light emitting diode (LED) is a diode that emits light in visible or invisible (infrared) range when forward biased. In any P-N junction there is a recombination of holes and electrons. During this process energy possessed by the free electron is transferred to another state. Some of this energy is transferred as heat and some in the form of photons. In silicon and germanium greater percentage is converted into heat and the emitted light is insignificant. However in materials like Gallium Arsenide(GaAs) the light emission dominates and hence used in LEDs.

By using elements like Gallium, Arsenic and Phosphorous, LEDs produce red, green, yellow, blue, orange or infrared (visible) light. LED's have replaced incandescent lamps in many applications because of their low voltage, long life, and fast on-off switching.

Diodes emitting light in the infrared region find applications in security systems, industrial processing, optical coupling etc. Figure 1.4.1 shows the circuit symbol and structure of a typical LED.



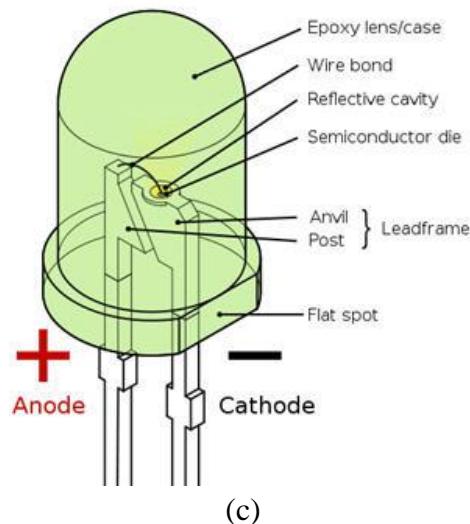


Figure 1.4.1 (a) Circuit symbol of LED (b) Different color LEDs (c) Structure of a typical LED

One of the common applications of the LED is in seven segment display. A common anode seven segment display arrangement is shown in Figure 1.4.2. It can be used to display any alphanumeric character.

LEDs are also used in burglar alarm system, digital meters, electronic display panels, optical communication system etc. LED's are much cheaper, last nearly indefinitely, and consume less energy. The biggest disadvantage is the cost of replacement to the consumer. For example, in the past, if a single instrument cluster bulb went out, it could be easily replaced. Today, if a single LED goes out on the instrument cluster, it is not replaceable. The entire instrument cluster must be replaced.

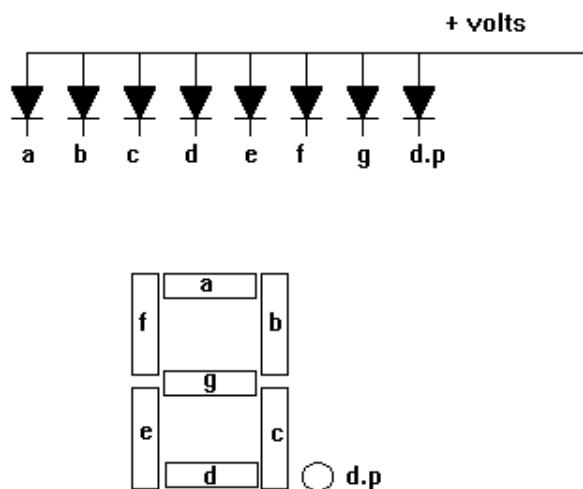


Figure 1.4.2. Seven segment display

#### 1.4.2 Photo Diodes and applications:

A photodiode consists of an active P-N junction which is operated in reverse bias as shown in Figure 1.4.3. When light falls on the junction, reverse current flows which is proportional to

the illuminance. The linear response to light makes it useful photodetectors for some applications. It is also used as the active element in light-activated switches.

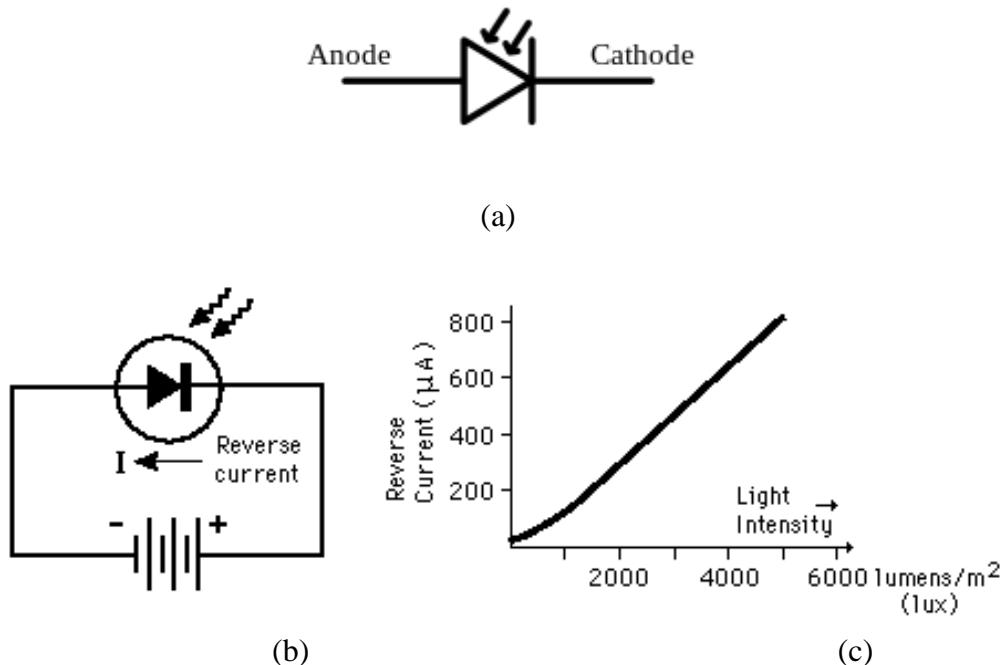


Figure 1.4.3. (a) Symbol of photo diode (b) Circuit using photodiode (c) Characteristic of photodiode. [<http://hyperphysics.phy-astr.gsu.edu/hbase/electronic/photdet.html>]

Photo diodes are used as/in light detectors, demodulators, encoders, high speed counting, switching circuits etc.

### 1.4.3 Optocoupler:

An optocoupler, also called opto-isolator, is an electronic component that transfers an electrical signal or voltage from one part of a circuit to another or from one circuit to another, while electrically isolating the two circuits from each other as shown in Figure 1.4.4. It consists of an infrared emitting LED chip that is optically in-line with a light-sensitive silicon semiconductor chip, all enclosed in the same package. The silicon chip could be in the form of a photo diode, photo transistor, photo Darlington, or photo SCR(silicon controlled rectifier).

The optocoupler application or function in the circuit is to:

- Monitor high voltage
- Output voltage sampling for regulation
- System control micro for power on/off
- Ground isolation

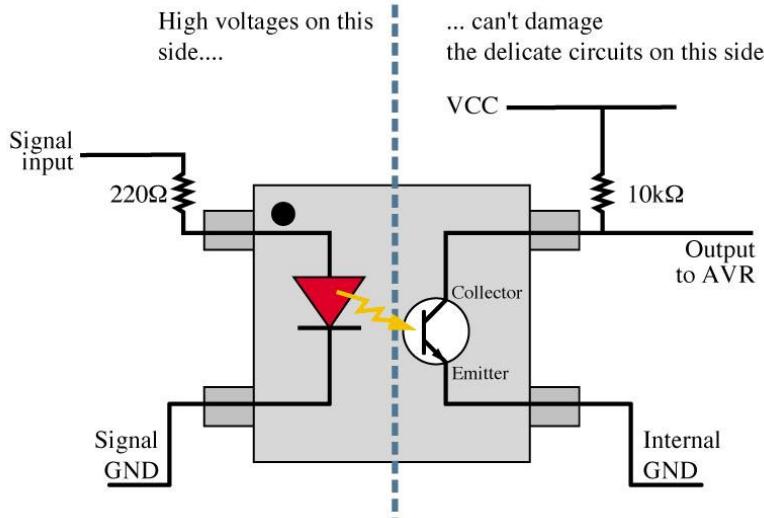


Figure 1.4.4. Schematic diagram of an opto-isolator  
[\[http://www.ustudy.in/node/7519\]](http://www.ustudy.in/node/7519)

#### *Self Test*

1. A LED is a diode that gives off .....when .....biased.
2. LED is manufactured using gallium arsenide gives ..... light.
3. How photo diode differs from rectifier diode?
4. What is dark resistance of photo-diode ?
5. To display the digit 0 in a common anode seven segment display
  - (i) A must be ON
  - (ii) F must be OFF
  - (iii) G must be ON
  - (iv) all segments except G should be ON

#### **1.4.4 Solar cell:**

A solar cell (also called a photovoltaic(PV) cell) is an electrical device that converts the energy of light directly into electricity by the photovoltaic effect. It is a form of photoelectric cell when exposed to light, can generate and support an electric current without being attached to any external voltage source, but do require an external load for power consumption.

#### **Structure of a Solar Cell**

A typical solar cell is a multi-layered unit as shown in Figure 1.4.5, consisting of a:

**Cover** - a clear glass or plastic layer that provides protection to external elements.

**Transparent Adhesive** - holds the glass to the rest of the solar cell.

**Anti-reflective Coating** - this substance is designed to prevent the light that strikes the cell from bouncing off so that the maximum energy is absorbed into the cell.

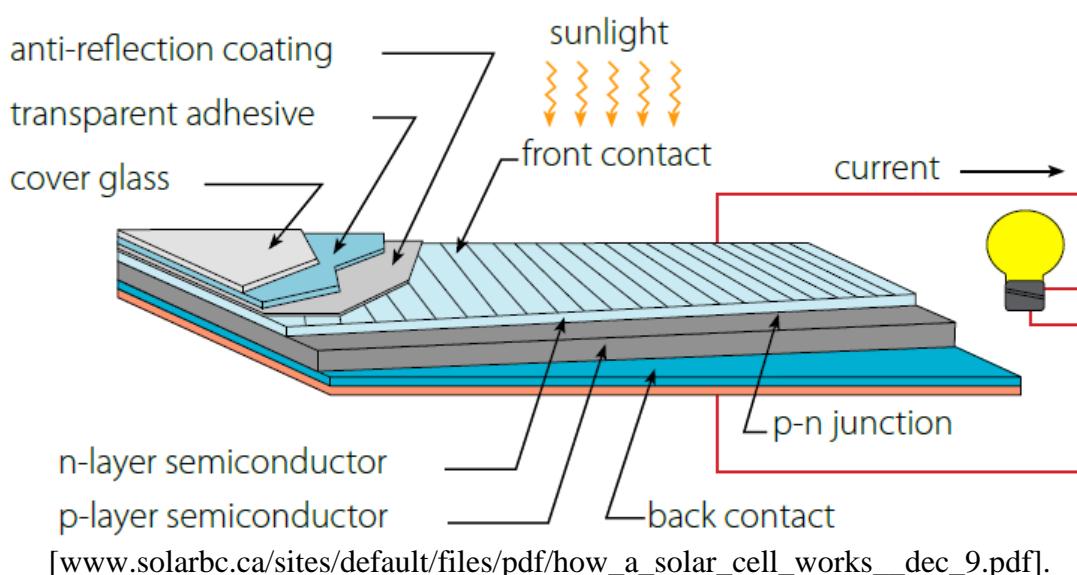
**Front Contact** - Transmits the electric current.

**N-Type Semiconductor Layer** - This is a thin layer of silicon which has been mixed with phosphorous to make it a better conductor.

**P-Type Semiconductor Layer** - This is a thin layer of silicon which has been mixed or doped with boron to make it a better conductor.

**Back Contact** - Transmits the electric current.

Figure 1.4.5. Structure of solar cell



[[www.solarbc.ca/sites/default/files/pdf/how\\_a\\_solar\\_cell\\_works\\_dec\\_9.pdf](http://www.solarbc.ca/sites/default/files/pdf/how_a_solar_cell_works_dec_9.pdf)].

**N-Layer**- is often formed from silicon and a small amount of Phosphorus. Phosphorus gives the layer of excess of electrons and therefore has a negative character. The N-layer is not a charged layer but it has an equal number of protons and electrons. Also some of the electrons are not held tightly to the atoms and are free to move.

**P-Layer**- is formed from Silicon and Boron and gives the layer a positive character because it has a tendency to attract electrons. The P-layer is not a charged layer and it has an equal number of protons and electrons.

**P-N Junction** - when the two layers are placed together, the free electrons from the N-layer are attracted to the P-layer. At the moment of contact between the two wafers, free electrons from the N-layer flow into the P-layer, then form a barrier to prevent more electrons from moving from one layer to the other. This contact point and barrier is called the P-N junction.

Once the layers have been joined, there is a negative charge in the P-layer and a positive charge in the n-layer section of the junction. This imbalance in the charge of the two layers at the P-N junction produces an electric field between the p-layer and the N-layer. If the PV cell is placed in the sun, radiant energy strikes the electrons in the P-N junction and energizes them, knocking them free of their atoms. These electrons are attracted to the positive charge in the N-layer and are repelled by the negative charge in the P-layer. A wire

can be attached from the P-layer to the N-layer to form a circuit. As the free electrons are pushed into the N-layer by the radiant energy, they repel each other. The wire provides a path for the electrons to flow away from each other. This flow of electrons constitutes electric current. The electron flow provides the current, and the cell's electric field causes a voltage. With both current and voltage, power is obtained, which is the product of the two. Solar array is shown in Figure 1.4.6.



Figure 1.4.6 Schematic of solar arrays.

#### **Equivalent circuit of a solar cell:**

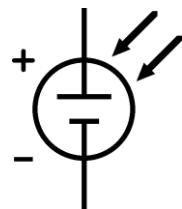


Figure 1.4.7 Circuit symbol of solar cell.

Figure 1.4.7 shows equivalent circuit of a solar cell. Solar cells are used to generate electricity. Incident sunlight can be converted into electricity by photovoltaic conversion using a solar panel. A solar panel consists of individual cells that are large-area semiconductor diodes, constructed so that light can penetrate into the region of the p-n junction. The junction formed between the n-type silicon wafer and the p-type surface layer governs the diode characteristics as well as the photovoltaic effect. Light is absorbed in the silicon, generating both excess holes and electrons. These excess charges can flow through an external circuit to produce power.

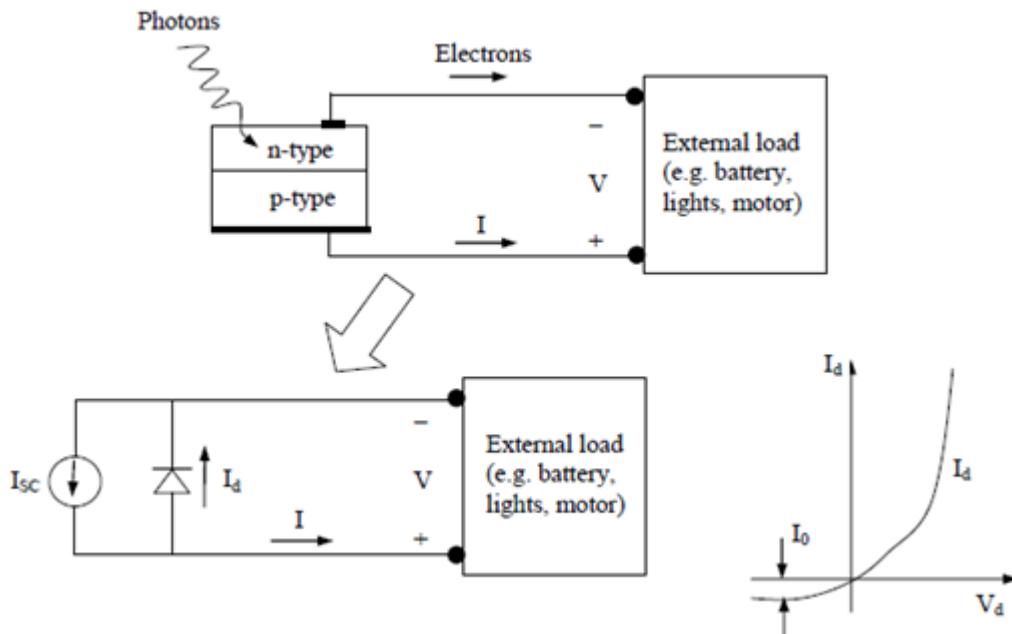


Figure 1.4.8. Equivalent circuit and I-V characteristic of a solar cell.

Figure 1.4.8. shows the equivalent circuit to describe a solar cell. The diode current  $I_D = I_o(e^{V_D/\eta V_T} - 1)$  comes from the standard I-V equation for a diode. It is clear that the current  $I$  that flows to the external circuit is  $I = I_{sc} - I_D$ , where  $I_{sc}$  is short circuit current. If the solar cell is open circuited, then all of the  $I_{sc}$  flows through the diode and produces an open circuit voltage  $V_{oc}$  of about 0.5-0.6V. If the solar cell is short circuited, then no current flows through the diode, and all of the short-circuit current  $I_{sc}$  flows through the short circuit.

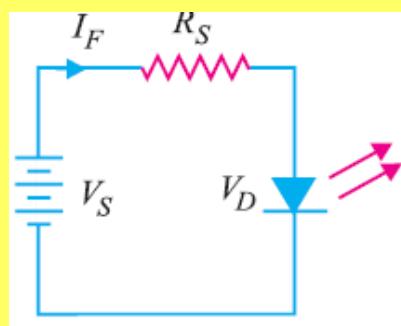
Since the  $V_{oc}$  for one solar cell is approximately 0.5-0.6V, then individual cells are connected in series as a “solar panel” to produce more usable voltage and power output levels. Most solar panels are made to charge 12 V batteries and consist of 36 individual cells (or units) in series to yield panel  $V_{oc} \approx 18\text{-}20$  V. The voltage for maximum panel power output is usually about 16-17 V.

### Summary

1. An LED emits light when forward-biased.
2. High-intensity LEDs are used in large-screen displays, traffic lights, automotive lighting, and home lighting.
3. LEDs are available for either infrared or visible light.
4. The photodiode exhibits an increase in reverse current with light intensity.

**Exercise Problems:**

1. Light Emitting Diodes (LED) is used in fancy electronic devices such as toys emit  
A. X-rays B. Ultraviolet light C. visible light D. radio waves
2. The maximum wave length of photons that can be detected by a photo diode made of a semiconductor of band gap 2 eV is about.....
3. What value of series resistor is required to limit the current through a LED to 20 mA with a forward voltage drop of 1.6 V when connected to a 10V supply? Given  $V_D=1.6$  V.

(Ans: 420  $\Omega$ )

5. Define sensitivity of the photo-diode.
6. How reverse biased voltage effects the capacitance of the varactor diode ? Explain with the help of the curve.

## Chapter 2

### BJT and Applications

The first Bipolar Junction Transistor (BJT) was demonstrated by a team of scientists at Bell laboratories in 1947. BJT has attractive features like, small in size, light weight, low power consumption and low operating voltages. These devices are used in applications such as signal conditioners, amplifiers, electronic switches, oscillators, etc.

#### **Module – 1: BJT Characteristics**

##### **Learning Outcomes:**

At the end of this module, students will be able to:

1. Explain the operation of BJT.
2. Draw CB and CE configurations of transistor.
3. Explain input and output characteristics of transistor in CB Configuration.
4. Explain input and output characteristics of transistor in CE Configuration.
5. Derive expressions for current gains in transistor.

#### **2.1.1 BJT Construction and Operation**

##### *a. BJT Structure*

It is a three terminal, three-layered, two junction semiconductor device.

Two types:

- Thin layer of n-type material is placed between two p-type materials (called PNP transistor)
- Thin layer of p-type material is placed between two n-type materials (called NPN transistor)

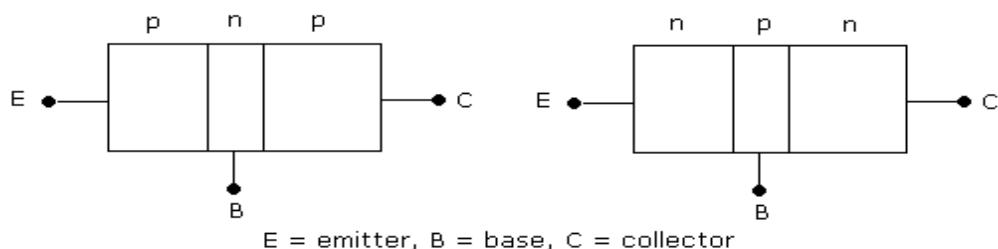


Figure 2.1.1: Transistor structure

Emitter is heavily doped, it supplies the charge carriers. Base is lightly doped, allows most of the charge carriers to pass through it. Collector is moderately doped, collects the charge carriers. The width of the collector region is less than the emitter region but more than that of the base region.

Two junctions of the BJT are:

- Emitter-base junction (or E-B diode)
- Collector-base junction (or C-B diode)

For normal operation, E-B junction is forward biased and C-B junction is reverse biased.

### b. BJT Operation

Working of NPN transistor is discussed here. Working of PNP transistor is similar (roles of free electrons and holes are interchanged and current directions are reversed).

As shown in Figure 2.1.2(a), EB junction is forward biased. The depletion region at EB junction is narrow. Similarly CB junction is reverse biased. The depletion region at CB junction is wide. The free electrons from emitter region cross the junction and reach base region. (Repelled by the negative potential at the emitter terminal). Some of these free electrons combine with the holes in the base region. They move towards the base terminal and form the base current. There are less number of holes available in base. Therefore, most electrons (about 99%) that comes out of emitter do not combine with holes, they enter the collector region (Attracted by the positive potential at the collector terminal). Hence the emitter emits electrons and the collector collects these electrons. Directions of three currents are indicated in Figure 2.1.2(a).

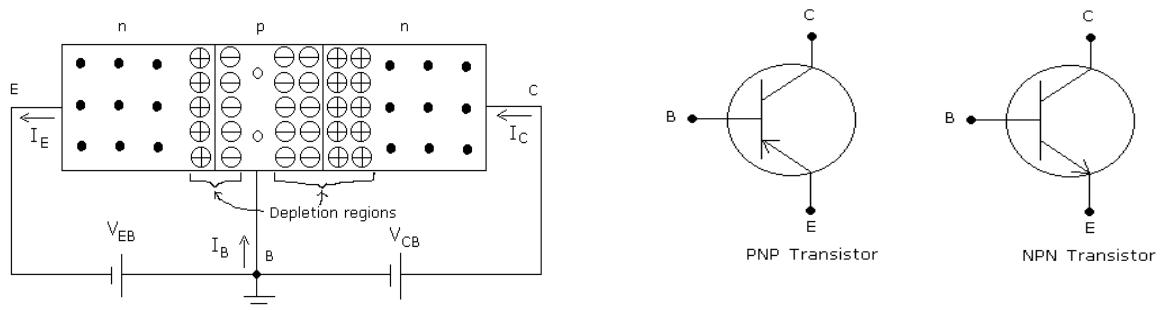


Figure 2.1.2(a): Transistor operation

Figure 2.1.2(b): Transistor symbols

Current directions are opposite to the flow of electrons.  $I_E$  is the emitter current,  $I_B$  is the base current and  $I_C$  is the collector current. Arrow head represents the direction of current flow through emitter in the transistor symbols shown in Fig 2.1.2(b).

### 2.1.2 BJT Configurations

Transistor is a three terminal device. For amplifier circuit, four terminals are required, two for input and two for output. Hence, one of the three terminals of transistor is made common to both input and output. Accordingly, there are 3 configurations:

- Common base (CB) configuration
- Common emitter (CE) configuration
- Common collector (CC) configuration

a. **Common base configuration**

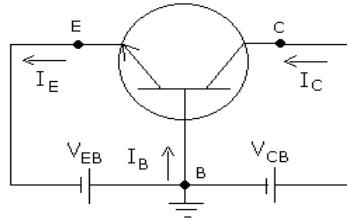


Figure 2.1.4: Common Base (CB) configuration

As shown in Figure 2.1.4, base is common to both emitter and collector. The emitter is the input terminal and the collector is the output terminal.

b. **Common emitter configuration**

In this emitter is common to both base and collector terminals. The base is input terminal and the collector is output terminal

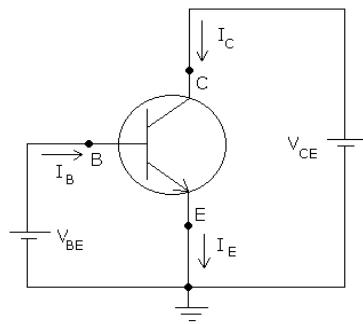


Figure 2.1.5: Common Emitter (CE) configuration

**Note:** Common collector configuration is not discussed here.

*Self test:*

1. Draw the circuit diagrams of transistor in CB and CE Configurations.

### 2.1.3 BJT Current Components

The emitter current relation with collector and base currents is given by

$$I_E = I_C + I_B \quad (2.1.1)$$

When the emitter circuit is opened, there is no supply of free electrons from emitter to collector. Even then, there will be small amount of collector current called reverse saturation collector current  $I_{CBO}$ . This is due to thermally generated electron-hole pairs.

Even during normal operation,  $I_{CBO}$  is present. So, the total collector current is:

$$I_C = \alpha_{dc} I_E + I_{CBO} \quad (2.1.2)$$

Where,  $\alpha_{dc}$  is fraction of emitter current that flows to the collector. From (2.1.2)

$$\alpha_{dc} = \frac{I_C - I_{CBO}}{I_E} \quad (2.1.3)$$

Since  $I_{CBO}$  is very small,

$$\alpha_{dc} \approx \frac{I_C}{I_E} \quad (2.1.4)$$

Also,

$$\beta_{dc} = \frac{I_C}{I_B} \quad (2.1.5)$$

### a. Relation between Current Gains $\alpha_{dc}$ and $\beta_{dc}$

From equation 2.1.1

$$I_E = I_C + I_B$$

Dividing throughout by  $I_C$ ,

$$\frac{I_E}{I_C} = 1 + \frac{I_B}{I_C} \quad (2.1.6)$$

From equations (2.1.4) and (2.1.5),

$$\frac{1}{\alpha_{dc}} = 1 + \frac{1}{\beta_{dc}} \quad (2.1.7)$$

Rearranging the terms,

$$\alpha_{dc} = \frac{\beta_{dc}}{\beta_{dc} + 1} \quad (2.1.8)$$

Also,

$$\beta_{dc} = \frac{\alpha_{dc}}{1 - \alpha_{dc}} \quad (2.1.9)$$

### 2.1.3.2 Relation between $I_{CBO}$ and $I_{CEO}$

From equation 2.1.2,

$$I_C = \alpha_{dc} I_E + I_{CBO}$$

Using equation 2.1.1,

$$I_C = \alpha_{dc}(I_C + I_B) + I_{CBO} \quad (2.1.10)$$

Rearranging the terms,

$$I_C(1 - \alpha_{dc}) = \alpha_{dc}I_B + I_{CBO} \quad (2.1.11)$$

Hence the collector current  $I_C$ ,

$$I_C = \left( \frac{\alpha_{dc}}{1 - \alpha_{dc}} \right) I_B + \frac{I_{CBO}}{(1 - \alpha_{dc})} \quad (2.1.12)$$

When  $I_B = 0$ ,  $I_C = I_{CEO}$  and is given by,

$$I_{CEO} = \frac{I_{CBO}}{(1-\alpha_{dc})} \quad (2.1.13)$$

*Self test:*

1. Obtain the relation between  $\alpha_{dc}$  and  $\beta_{dc}$ .
2. Obtain the relation between  $I_{CBO}$  and  $I_{CEO}$ .

#### 2.1.4 BJT Characteristics

##### a. Common Base Input and Output characteristics:

Input Characteristics of Common Base configuration is a plot of input current  $I_E$  versus input voltage  $V_{EB}$ , for various constant values of output voltage  $V_{CB}$ . As  $V_{EB}$  is increased,  $I_E$  increases. The characteristics curve is similar to diode characteristics. If  $V_{CB}$  is increased, then  $I_E$  shoots up early. This is due to the increase in electric field aiding the flow of electrons from emitter. The variation is shown in Figure 2.1.6 for various values of  $V_{CB}$ .

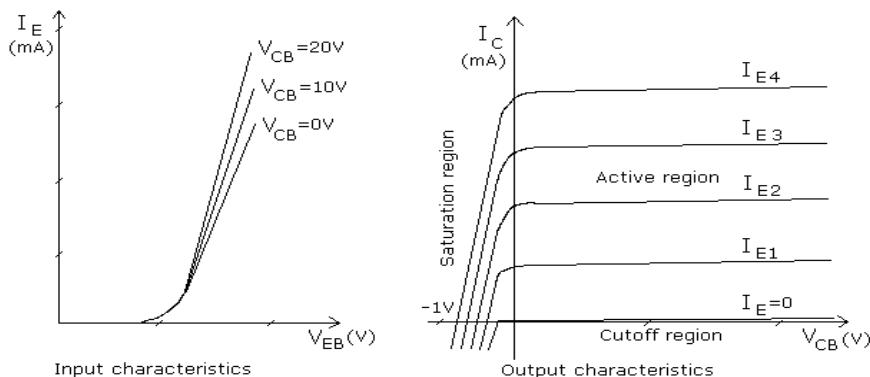


Figure 2.1.6: Input and Output Characteristics of Common Base Configuration

Output Characteristics is a plot of output current  $I_C$  versus output voltage  $V_{CB}$  for various constant values of input current  $I_E$ . The characteristics can be divided into three regions namely, Active, Saturation and Cut off regions.

- a. **Active region:** This is the region to the right of y-axis above  $I_E=0$  curve (Figure 2.1.6), where the curves are linear.  $I_E$  is positive non zero (i.e., E-B diode is forward biased) and  $V_{CB}$  is positive (i.e., C-B diode is reverse biased). When  $V_{CB}$  is increased,  $I_C$  increases slightly. This is because, when  $V_{CB}$  is increased, depletion region width at C-B junction increases, effectively base width decreases and hence  $I_B$  decreases. Due to this collector current  $I_C$  increases. This effect is known as Early effect (also called base width modulation). If  $I_E$  is increased to higher constant value,  $I_C$  also increases. When  $I_E=0$ ,  $I_C=I_{CBO}$  (reverse saturation collector current in common base with emitter open).  $I_{CBO}$  doubles for every ten degree rise in temperature.
- b. **Saturation region:** This is the region to the left of y-axis, above  $I_E=0$  curve (Figure 2.1.6). In this  $I_E$  is positive nonzero (E-B diode forward biased) and  $V_{CB}$  is negative (C-B diode is forward biased).  $I_C$  decreases exponentially in this region.

- c. **Cut-off region:** This is the region below  $I_E=0$  curve (Figure 2.1.6). In this emitter current  $I_E$  is less than zero (E-B diode is reverse biased) and collector to base voltage  $V_{CB}$  is positive (C-B diode is reverse biased). Transistor is said to be in OFF state since  $I_C$  is zero.

#### 2.1.4.2 Common Emitter Input and Output Characteristics

Input Characteristics of Common Emitter configuration is a plot of input current  $I_B$  versus input voltage  $V_{BE}$  for various values of output voltage  $V_{CE}$ . As  $V_{BE}$  is increased,  $I_B$  increases. The characteristics curve is similar to diode characteristics. If  $V_{CE}$  is increased to higher constant value, then  $I_B$  decreases slightly as shown in Figure 2.1.7. This is due to Early effect.

Output Characteristics of Common Emitter configuration is a plot of output current  $I_C$  versus output voltage  $V_{CE}$  for various values of input current  $I_B$ . The characteristics can be divided into three regions namely, Active, Saturation and Cut off regions.

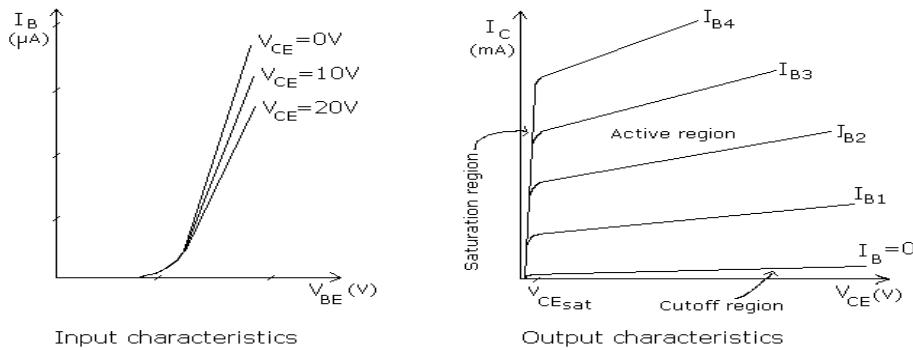


Figure 2.1.7: CE Configuration Input and Output characteristics.

- Active region:** The region to the right of  $V_{CE\text{ Sat}}$  above  $I_B=0$  curve (Figure 2.1.7), where the curves are linear. Note that,  $V_{CE} = V_{CB} + V_{BE}$ . If  $V_{CE} > V_{CE\text{ Sat}}$ , then  $V_{CB}$  becomes positive (i.e., C-B diode is forward biased).  $V_{CE\text{ Sat}}$  is around 0.3V for silicon transistor. If  $I_B > 0$ , then it means E-B diode is forward biased. When  $V_{CE}$  is increased,  $I_C$  increases slightly due to Early effect. Note that slope of curve is more than that of CB output characteristics. If  $I_B$  is increased,  $I_C$  also increases. When  $I_B=0$ ,  $I_C=I_{CEO}$ .
- Saturation region:** The region to the left of  $V_{CE\text{ Sat}}$  and right of y-axis (Figure 2.1.7). In this region E-B diode and C-B diode are both forward biased.
- Cut-off region:** The region below  $I_B=0$  curve (Figure 2.1.7). In this E-B diode and C-B diode are both reverse biased Transistor is said to be in OFF state since  $I_C$  is almost zero.

#### Self test:

1. Define input and output characteristics of CB and CE configuration transistor.
2. List the different regions in the output characteristics of a transistor?
3. Give the biasing conditions required for the different regions of transistor operation.
4. Define Early effect.

**Example Problem 1:**

1. A BJT has alpha (dc) 0.998 and collector-to-base reverse sat current  $1\mu\text{A}$ . If Emitter current is  $5\text{mA}$ . Calculate *i* Collector current *ii* Base current.

$$\begin{aligned}\text{Ans: } & \text{i} & I_C = \alpha_{dc} I_E + I_{CBO} \\ & = 0.998 * 5 * 10^{-3} + 10^{-6} \\ & = 4.99 \text{ mA.}\end{aligned}$$

$$\begin{aligned}& \text{ii} & I_B = I_E - I_C \\ & = 5\text{mA} - 4.99\text{mA} = 10 \mu\text{A.}\end{aligned}$$

**Summary**

In this module we have learnt

1. The operation of a transistor and the relevant current components.
2. To plot the input and output characteristics of CB and CE Configuration transistor.
3. To find current gains in different configurations of transistor.

**Exercises:**

1. An NPN transistor has collector current  $4\text{mA}$  and base current  $10 \mu\text{A}$ . Calculate the alpha and beta values of the transistor, neglecting the reverse saturation current  $I_{CBO}$ . (Ans: 0.9975, 400)
2. In a transistor, 99% of the carriers injected into the base cross over to the collector region. If collector current is  $4\text{mA}$  and collector leakage current is  $6 \mu\text{A}$ , Calculate emitter and base currents. (Ans:  $4.034 \text{ mA}$ ,  $34 \mu\text{A}$ )
3. In a transistor circuit, when the base current is increased from  $0.32 \text{ mA}$  to  $0.48 \text{ mA}$ , the emitter current increases from  $15 \text{ mA}$  to  $20 \text{ mA}$ . Find  $\alpha_{ac}$  and  $\beta_{ac}$  values. (Ans: 0.968, 30.25)

## Module 2: BJT Biasing

### Learning Outcomes:

At the end of this module, students will be able to:

1. Explain the need for biasing of BJT.
2. Analyse fixed bias circuit by plotting load line and fixing operating point.
3. Design a fixed bias circuit for given operating point.
4. Analyse and design Self bias circuit

Biasing means application of external voltage to the device so as to make it to operate in the required region. For transistor to work as an amplifier, it is biased in active region. Similarly if it has to work as a switch, it must be biased either in saturation or cut-off region.

### 2.2.1 DC Load line and need for Biasing

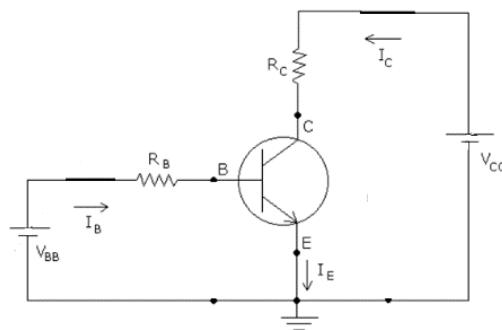


Figure 2.1.1(a): Transistor in CE configuration

Applying KVL to the collector loop, which consists of  $V_{CC}$ ,  $R_C$ , collector, emitter and ground, we get:

$$V_{CC} - I_C R_C - V_{CE} = 0 \quad \text{or} \quad V_{CE} = V_{CC} - I_C R_C$$

The output characteristics of the transistor also relate the same two variables  $I_C$  and  $V_{CE}$ . The characteristics of the BJT are superimposed on a plot of the network equation defined by the same axis parameters.

The most direct method of plotting the network equation on the output characteristics is to use the fact that a straight line is defined by two points. If we choose  $I_C$  to be 0 mA, we are specifying the horizontal axis as the line on which one point is located. By substituting  $I_C = 0$  mA we find that

$$V_{CE} = V_{CC} - (0)R_C$$

$$V_{CE} = V_{CC} \quad 2.2.9$$

If we now choose  $V_{CE}$  to be 0V, which establishes the vertical axis as the line on which the second point will be defined, we find that  $I_C$  is determined by the following equation:

$$0 = V_{CC} - I_C R_C$$

$$I_C = V_{CC}/R_C \quad 2.2.10$$

By joining these two points defined by 2.2.9 and 2.2.10, the straight line can be drawn on the output characteristics as shown in Figure 2.2.3. The resulting line on the graph is called the *load line* since it is defined by the load resistor  $R_C$ . The intersection of load line with the base current results in operating point or Q point.

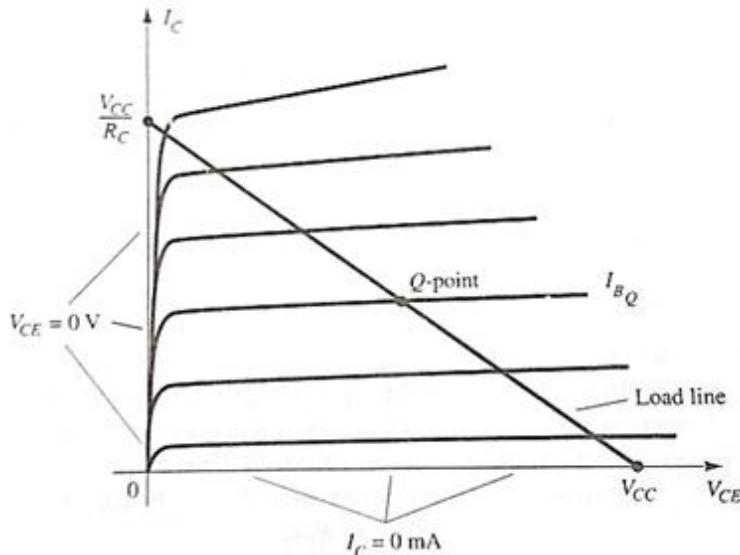


Figure 2.1.1(b): Load line and operating point

The variation of the Q-point up or down the load line with varying values of  $I_B$ ,  $R_C$  and  $V_{CC}$  are shown in Figure 2.2.4. With  $I_B$  variation Q point moves along the load line (Figure 2.2.1(a)). If  $R_C$  of the circuit is varied, the slope of the load line changes (Figure 2.2.1(b)).  $V_{CC}$  will shift the load line keeping the slope constant (Figure 2.2.1(c)).

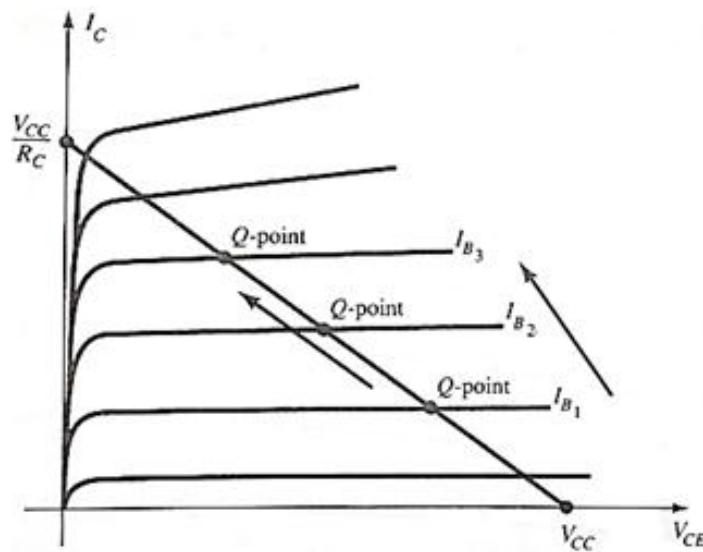
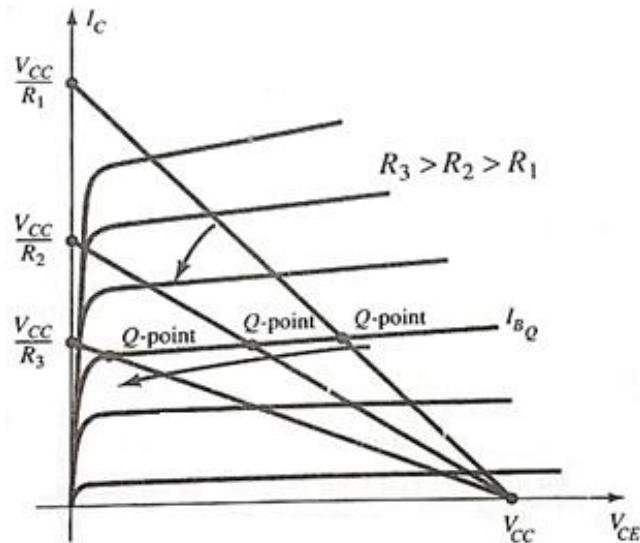
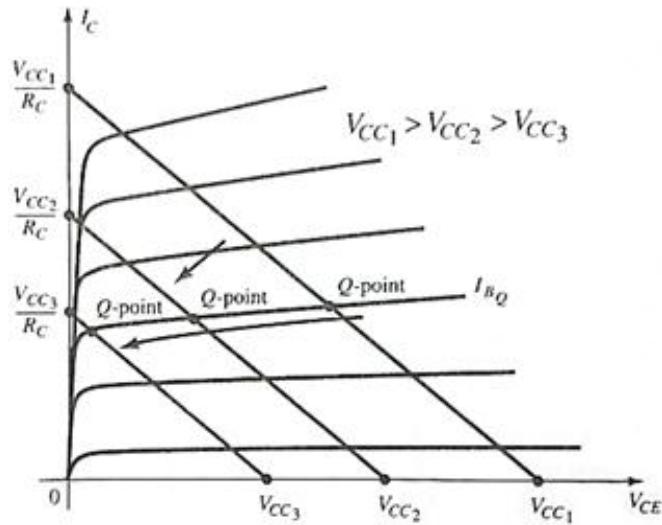


Figure 2.1.1(c): Movement of the Q-point with increasing value of  $I_B$ .

Figure 2.1.1(d): Effect of an increasing value of  $R_C$  on the load line and the Q-pointFigure 2.1.1(e): Effect of lower values of  $V_{CC}$  on the load line and the Q-point

*Self test:*

1. What do you mean by BJT biasing?
2. Explain the need for Operating point.
3. State the parameters which are affecting the load line.

## 2.2.2 Fixed Biasing

One of the simple way of biasing the transistor is by using fixed biasing technique.

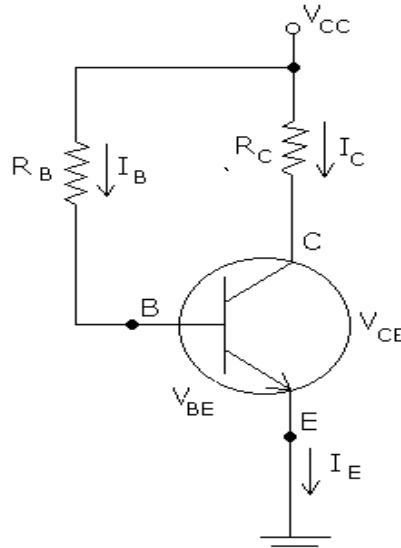


Figure 2.2.2: Fixed-bias circuit

The circuit diagram of a fixed bias transistor is as shown in Figure 2.2.2. Compared with CE configuration in Figure 2.1.1(a), here base resistor  $R_B$  is connected to  $V_{cc}$  (Instead of  $V_{BB}$ ). Negative terminal of  $V_{cc}$  is not shown. It is assumed to be at ground.

Applying KVL to the input side, (base emitter loop) which consists of  $V_{cc}$ ,  $R_B$ , base, emitter and ground, we get:

$$V_{cc} - I_B R_B - V_{BE} = 0 \quad 2.2.1$$

Rearranging, we get

$$I_B = \frac{V_{cc} - V_{BE}}{R_B} \quad 2.2.2$$

$V_{cc}$  is constant,  $V_{BE}$  is almost constant (0.7V for silicon). So by selecting proper  $R_B$ , we can fix  $I_B$  as required. Applying KVL to the collector loop, which consists of  $V_{cc}$ ,  $R_c$ , collector, emitter and ground, we get:

$$V_{cc} - I_C R_C - V_{CE} = 0 \quad \text{or} \quad V_{CE} = V_{cc} - I_C R_C \quad 2.2.3$$

$I_C$  is related to  $I_B$  by  $\beta$ .

$$I_C = \beta I_B \quad 2.2.4$$

So,  $V_{CE}$  can be fixed by selecting proper  $R_C$ .

$V_{CE}$  can also be written as

$$V_{CE} = V_C - V_E \quad 2.2.5$$

using single subscript notations, where  $V_C$  and  $V_E$  are voltages from collector and emitter to ground, respectively. In the case of fixed bias,  $V_E = 0V$  since emitter is grounded directly.

Therefore,

$$V_{CE} = V_C \quad 2.2.6$$

In a similar way,

$$V_{BE} = V_B - V_E \quad 2.2.7$$

and since  $V_E = 0V$ ,

$$V_{BE} = V_B$$

2.2.8

**Example Problem 1:**

1. For a fixed bias circuit using Si transistor,  $R_B = 500 \text{ k}\Omega$ ,  $R_C = 2 \text{ k}\Omega$ ,  $V_{CC} = 15 \text{ V}$  and  $\beta = 70$ . Find the collector current  $I_C$  and  $V_{CE}$ . Take  $V_{BE}$  as 0.7 V.

**Solution:**

From the input loop expression,

$$I_B = \frac{V_{CC} - V_{BE}}{R_B}$$

Substituting the given values,  $I_B = 28.6 \mu\text{A}$   
As  $I_C = \beta I_B$   
 $I_C = 2.002 \text{ mA}$ .

From the output loop expression,

$$V_{CE} = V_{CC} - I_C R_C, V_{CEQ} = 10.996 \text{ V}$$

**Exercises:**

1. A Si transistor is biased for a constant base current. If  $\beta = 80$ ,  $V_{CEQ} = 8 \text{ V}$ ,  $R_C = 3 \text{ k}\Omega$  and  $V_{CC} = 15 \text{ V}$ , find  $I_{CQ}$  and the value of  $R_B$  required. ( Ans: 2.33mA, 493K $\Omega$ )
2. Repeat problem 1 if the transistor is a Germanium device. ( Ans: 2.33mA, 507K $\Omega$ )

**2.2.2.1. Advantages of fixed bias**

- Simple to design
- Requires less circuit components

**2.2.2.2. Disadvantages of fixed bias**

- Q-point is not stable. i.e., if temperature varies,  $\beta$  will vary, hence  $I_C$  will vary.
- If transistor is replaced by another transistor having different  $\beta$  then Q-point will shift

**Self test:**

1. Mention the extreme end points of the load line for a fixed biased circuit.
2. Identify the operating region of the transistor when biasing circuit is used.

### 2.2.3 Voltage divider bias or Self bias

In voltage divider bias, as shown in Figure. 2.2.5, two resistors  $R_1$  and  $R_2$  are used instead of  $R_B$  as compared with the fixed bias circuit. These two resistors together will provide the required base supply for the circuit. There is another register  $R_E$  is connected between emitter and ground.

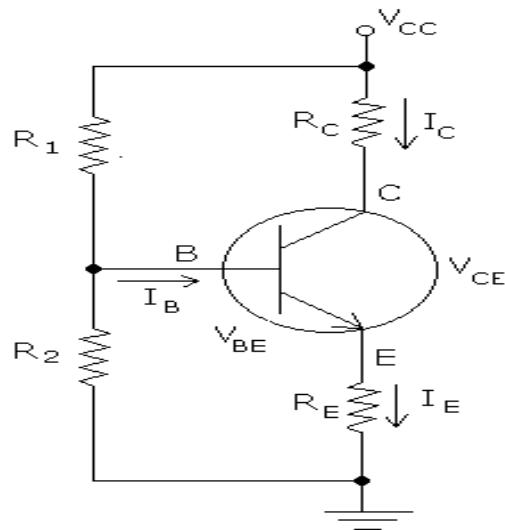


Figure. 2.2.5. Voltage Divider Bias circuit

Input side of the above circuit called as voltage divider circuit, is redrawn below in Figure 2.2.6.

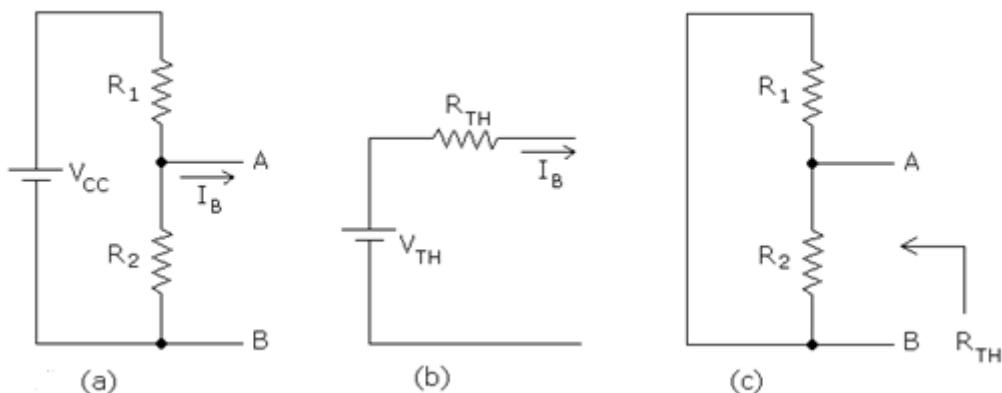


Figure 2.2.6. Equivalent circuit

Figure 2.2.6. a is replaced by equivalent circuit as shown in Figure 2.2.6.b.  $V_{TH}$  is the open circuit voltage between points A & B of Figure. 2.2.6. a which is given by:

$$V_{TH} = \frac{V_{CC}R_2}{R_1 + R_2} \quad 2.2.11$$

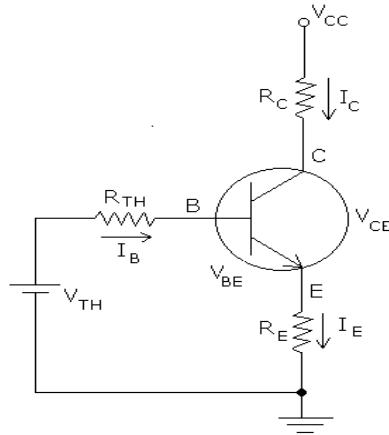


Figure. 2.2.7: Equivalent circuit

$R_{TH}$  is the resistance seen between the same points A & B with  $V_{CC}$  replaced by short circuit.

$$R_{TH} = R_1 \parallel R_2 = \frac{R_1 R_2}{R_1 + R_2} \quad 2.2.12$$

Self-bias circuit with its input loop replaced by equivalent circuit is shown in Figure 2.2.7.

Applying KVL to the input loop we get:

$$V_{TH} - I_B R_{TH} - V_{BE} - I_E R_E = 0 \quad 2.2.13$$

Substituting  $I_E = (\beta + 1)I_B$  and rearranging, we get

$$I_B = \frac{V_{TH} - V_{BE}}{R_{TH} + (\beta + 1)R_E} \quad 2.2.14$$

Applying KVL to the output loop, we get

$$V_{CC} - I_C R_C - V_{CE} - I_E R_E = 0 \quad 2.2.15$$

Rearranging, we get

$$V_{CE} = V_{CC} - I_C R_C - I_E R_E \quad 2.2.16$$

Also using single subscript notation,

$$V_C = V_{CC} - I_C R_C \quad 2.2.17$$

where,  $V_C$  is voltage from collector to ground and,

$$V_E = I_E R_E \quad 2.2.18$$

where,  $V_E$  is voltage from emitter to ground.

To show  $I_C$  is independent of  $\beta$ :

Since  $\beta \gg 1$ , we have  $(\beta + 1) \cong \beta$ . If  $\beta R_E \gg R_{TH}$ , then equation for  $I_B$  reduces to:

$$I_B \approx \frac{V_{TH} - V_{BE}}{\beta R_E} \quad 2.2.19$$

Now,

$$I_C = \beta I_B = \frac{V_{TH} - V_{BE}}{R_E} \quad 2.2.20$$

Since equation for  $I_C$  does not contain  $\beta$ , we say that  $I_C$  is independent of temperature variation and transistor replacement.

#### **Exercises:**

1. For a self-bias circuit using silicon transistor,  $R_E = 300 \Omega$ ,  $R_C = 500 \Omega$ ,  $V_{CC} = 15 \text{ V}$ ,  $\beta = 100$  and  $\beta R_E = 10R_2$ . Find the values of  $R_1$  and  $R_2$  to get  $V_{CEQ} = V_{CC} / 2$ .  
(Ans: 9.03KΩ)
2. For a self-bias circuit, the transistor is a Si device,  $R_E = 200 \Omega$ ,  $R_1 = 10R_2 = 10 \text{ k}\Omega$ ,  $R_C = 2 \text{ k}\Omega$ ,  $\beta = 100$  and  $V_{CC} = 15 \text{ V}$ . Determine the values of  $I_{CQ}$  and  $V_{CEQ}$ . ( Ans:  
(Ans: 3.13mA, 8.11V)

#### **2.2.3.1 Advantages of voltage divider bias**

- Collector current and Q-point is independent of  $\beta$ . Hence Q-point is stable against variation in temperature and replacement of transistor.

#### **2.2.3.2 Disadvantages of voltage divider bias**

- Design is relatively complex.
- More circuit components are required.

#### **Self test:**

1. What are the modifications done to fixed bias circuit to transform it into self bias circuit?
2. Explain how the stability of the Q point is achieved in self bias circuit.

#### **Summary**

In this module we have learnt:

1. Importance of biasing in transistors
2. Concept of load line, operating point and their significance
3. Operating point dependent on the circuit parameters such as supply voltage and resistor values as well as operating temperatures.
4. To analyse fixed bias circuit to obtain the operating point and plot the load line.
5. Finding the equivalent circuit for self bias circuit and analyse it to obtain the Q point.

## Module 3: Transistor Amplifier

### Learning outcomes:

At the end of this module students will be able to:

1. Explain the need of amplifier circuits.
2. Draw the circuit diagram of an RC coupled amplifier and explain its working.
3. Define gain of the amplifier.
4. Plot the frequency response of an amplifier and define the bandwidth.
5. Identify feedback signal and its effect on the circuit operation.
6. Describe the requirement of multistage amplifiers.

Amplifier is a circuit which increases the magnitude of input signal applied. Bipolar junction transistor basically amplifies current. In CE configuration, base current is the input current and the collector current will be the output current. As collector current is beta times more than the input current ( $i_b$ ), it is an amplified version of the input. This is the effect of transistor.

By suitably designing transistor circuit, we can get voltage amplification and power amplification also. To work as amplifier, transistor should be in active region throughout the input signal cycle. This is achieved by proper use of biasing circuit. Consider the working of the circuit shown in Figure 2.3.1: Batteries  $V_{BB}$  and  $V_{CC}$  ensure that transistor is operating in the active region. It causes direct currents  $I_B$ ,  $I_C$  and  $I_E$  to flow in the circuit.  $V_{in}$  is a weak input signal to be amplified. This causes an alternating current  $i_b$  to flow through input circuit. Total base current into the base terminal  $i_B$  is sum of  $I_B$  and  $i_b$ , which is a shifted sinusoidal signal. During positive quarter cycle of input waveform, as input voltage increases,  $i_b$  and hence  $i_B$  increases. Due to transistor action,  $i_C$  also increases. We have  $i_C = \beta i_B$ , where  $\beta$  is current amplification factor. Since  $\beta$  is very large, even for small increase in  $i_B$ , there is a large increase in  $i_C$ . Hence large alternating voltage  $i_C R_C$  develops across resistor  $R_C$ .  $V_{out} = V_{CC} - i_C R_C$  will decrease. During second quarter cycle of input waveform, as input voltage decreases,  $i_B$  decreases, and also  $i_C$  decreases. During negative half cycle of input waveform, E-B junction still remains forward biased because,  $V_{BB}$  is so chosen that it is greater than peak value of  $V_{in}$ . So, during negative half cycle when  $i_B$  decreases,  $i_C$  also decreases, and hence  $i_C R_C$  decreases. Thus output voltage  $V_{out}$  increases. Hence  $V_{out}$  is exact replica of input voltage  $V_{in}$ , but magnified many times with  $180^\circ$  out of phase with input.

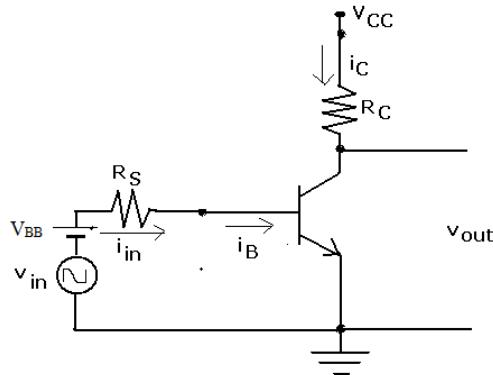


Figure 2.3.1: Common emitter amplifier

*Self test:*

1. What have you understood from amplification of a signal?
2. Mention the phase shift obtained between the input and output voltage signals of an amplifier.

### 2.3.1 RC coupled amplifier

A practical common-emitter transistor amplifier using voltage divider bias is shown Figure 2.3.2. The use of bias eliminates the need for two separate batteries  $V_{BB}$  and  $V_{CC}$ . Resistors  $R_1$ ,  $R_2$ ,  $R_C$ ,  $R_E$  and voltage source  $V_{CC}$  will fix the operating point in active region. This is a voltage divider bias circuit, which is already discussed.  $C_C$  is called as coupling capacitor. At input side, it blocks dc component of input voltage (or output of previous stage) from reaching the base of transistor. If dc is not blocked, then it will shift the operating point. At output side, it blocks dc component from entering into the load (or next stage).

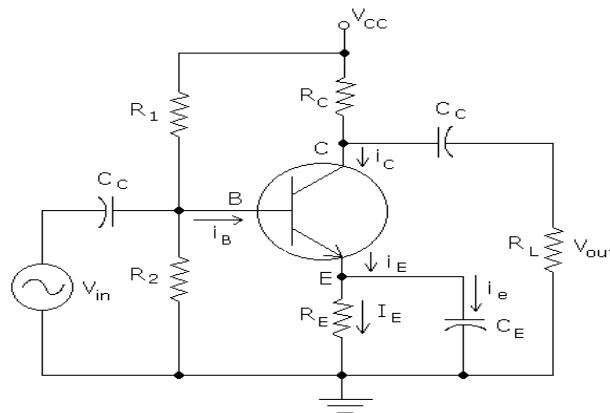


Figure 2.3.2: RC coupled amplifier

Feedback means application of a portion of the output signal fed back to the input. The signal fed back will be either added to or subtracted by the input signal applied to the circuit. Depending on whether the feedback signal increases or decreases the input signal applied to the transistor it is termed as positive or negative feedback respectively. Negative feedback will reduce the gain but increase the stability of the circuit and the available bandwidth. It is

generally used in amplifier circuits. Positive feedback will increase the gain but decrease the stability and bandwidth. It will be used in oscillator circuits (negative and positive feedback significance).

$C_E$  is called emitter by-pass capacitor. It is used to provide a negative feedback signal to the amplifier. The negative feedback in amplifiers will improve the performance such as stability, frequency response of the amplifier.  $C_E$  offers low reactance path for ac component, thus preventing ac component from passing through  $R_E$ . With this ac voltage drop across the resistor  $R_E$  is zero. The circuit is named as RC coupled amplifier without feedback as there is no feedback signal available to the input. On the other hand if the capacitor  $C_E$  is removed from the circuit, then ac signal passes through  $R_E$ , there will be ac voltage drop across it. As it is a negative feedback, this will decrease  $V_{BE}$ , bringing down output voltage. Hence circuit is named as RC coupled amplifier with feedback.  $R_L$  is the equivalent resistance of the load connected at output of amplifier. As explained earlier, when input voltage varies,  $i_B$  varies, this varies the  $i_C$  proportionally. Thus the output voltage is a amplified version of the input voltage, but with a phase shift of  $180^\circ$ .

### 2.3.2 Frequency response of an amplifier

Plot of amplifier gain versus frequency of input signal is called frequency response. Frequency of input signal is increased in steps. At each frequency, voltage gain is determined and then plotted. It is found that gain is very small at lower frequencies and at higher frequencies. Gain remains constant at mid frequencies. For audio amplifier, it is required that gain should be constant over the audio frequency range from 20 Hz to 20 kHz. Bandwidth of amplifier is defined as range of frequencies over which gain is either equal or greater than 0.707 (or  $1/\sqrt{2}$ ) times the maximum gain. Since  $20 \log_{10}(0.707) = -3$ , bandwidth is also defined as range of frequencies over which gain is within 3 dB of maximum gain (in dB). Figure 2.3.3 shows the frequency responses of RC coupled amplifier with and without feedback. It can be observed that without feedback circuit has a larger gain but smaller bandwidth as compared with the circuit with feedback.

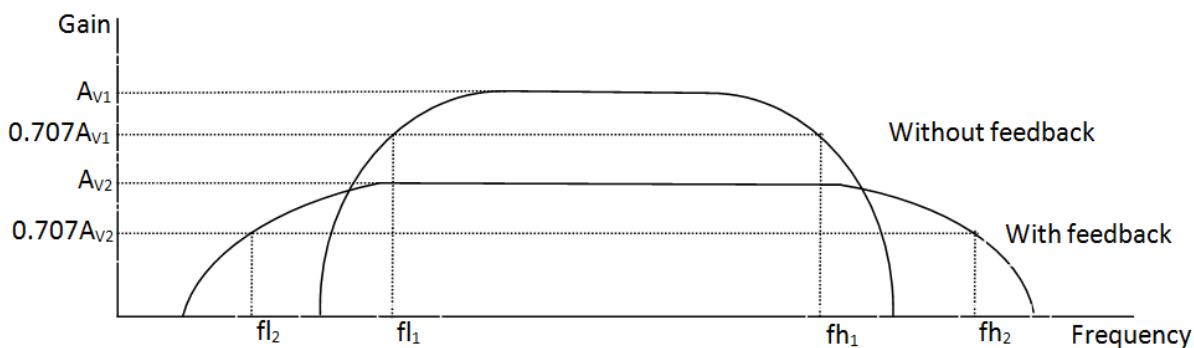


Figure 2.3.3: Frequency response of amplifier with and without feedback.

Here,  $f_l$  is called lower cut-off frequency;  $f_h$  is called upper cut-off frequency. These are also called 3 dB frequencies

Bandwidth is defined as

$$BW = f_h - f_l \quad 2.3.1$$

### 2.3.2.1 Analysis of frequency response curve of RC coupled amplifier

At low frequencies, reactance of coupling capacitors is high. Hence a part of input does not reach the transistor, gain reduces. Also at low frequencies reactance of emitter bypass capacitors is high. Hence ac component of emitter current is not fully bypassed that results in ac voltage drop across  $R_E$ , reducing the gain.

At high frequencies, reactance of shunt capacitances due to wiring and reactance of junction capacitances will become low. This offers low reactance path for signal to ground, thus reducing voltage gain.

At mid frequencies, reactance of coupling and emitter bypass capacitors is low, reactance of shunt capacitances is high. Hence there is no loss of signal, gain remains constant.

### 2.3.3 Multistage amplifier

If high gain is required, then amplifier stages are cascaded as shown in Figure. 2.3.4. Overall gain  $A_V$  is product of individual gains:

$$A_V = A_{V1} \cdot A_{V2} \cdots \cdots A_{VN} \quad (2.3.2)$$

In decibels, overall gain is sum of individual gains:

$$(A_V)_{dB} = (A_{V1})_{dB} + (A_{V2})_{dB} + \cdots \cdots + (A_{VN})_{dB} \quad (2.3.3)$$

However, practically the gain will be less than calculated  $A_V$  due to loading effects..

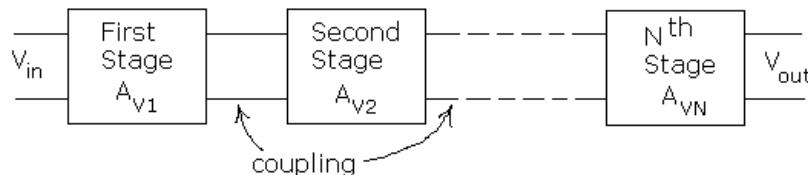


Figure 2.3.4: Multistage amplifier

*Self test:*

1. How to obtain the frequency response of an amplifier?
2. Define three dB frequency and bandwidth of an amplifier.
3. What have you understood by midband gain of an amplifier?
4. What is the purpose of using multistage amplifiers?
5. What is the purpose of using the coupling capacitors in the amplifier circuits?

### Summary

In this module we have learnt:

1. Amplifier circuits are used for increasing the strength of weak signal.
2. The working of an RC coupled amplifier with and without feedback.
3. The Gain of the amplifier defined as ratio of output signal to input signal.
4. The frequency response is a plot of frequency v/s gain of the amplifier and defines the bandwidth.
5. Gain of multistage amplifiers will be obtained by multiplying gain of each individual stage.

**Exercises:**

1. An amplifier is known to have a power gain of 40 dB. If the output power is 4 watts, determine the input power.(Ans: 0.4mW)
2. What output power is obtained from an amplifier whose power gain is 55 dB, when the input power is 1 mW? (Ans: 316.23W)
3. In a three-stage amplifier, the voltage gain of first stage is 40 dB, gain of second stage is 200 (not in dB) and that of third stage is 0 dB. Find the overall gain of the amplifier.(Ans: 86.02dB)

## Module 4: Transistor as switch

### Learning outcomes:

At the end of this module, students will be able to:

1. Explain how the transistor can be used as switch.
2. Application of transistor switch as LED driver and inverter.

In the previous section we have explored the use of transistor as an amplifier, where it was configured in active region. In this module we will study the use of transistor configured in other regions of operation.

#### 2.4.1 Introduction

Transistor can be made to operate as “ON/OFF” solid state switches. Transistor switches can be used for controlling high power devices such as motors, solenoids or lamps, as well as they can be used in low power digital electronics and logic gate circuits. To be specific the transistor must operate in the extreme ends of the load line curve: i.e. in cut-off and saturation regions. To review, in cut-off region both junctions of the transistor are reverse biased, ( $V_{BE} < 0.7V$  and  $I_C = 0$ ) whereas, in saturation region both junctions were forward biased, ( $V_{BE} > 0.7V$  and  $I_C = \text{Maximum}$ ). The operating conditions of the transistor in the cut off region and saturation regions are listed in the table given below. Therefore transistor in cut off region

Parameters	Cut off Region	Saturation region
$I_B$	Zero	Maximum
$I_C$	Zero	Maximum
$V_{CE}$	Maximum	Zero

acts as a “Fully-OFF” switch. In saturation region the transistor acts as a “Fully-ON” switch.

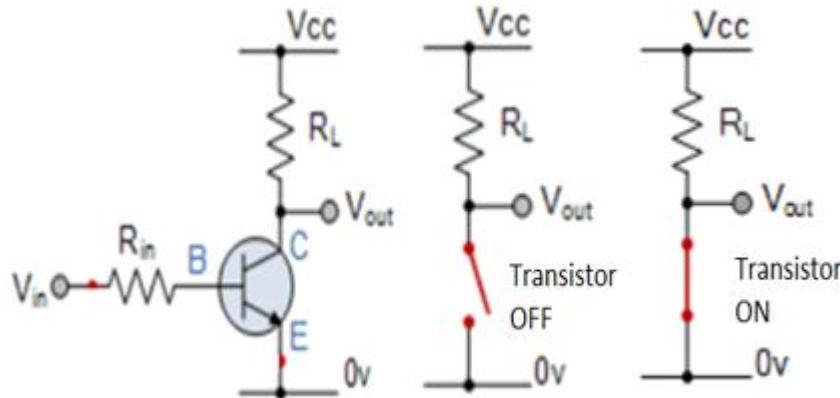


Figure 2.4.1 Transistor as switch. (a) circuit diagram (b) and (c) are equivalent circuits when the switch is OFF and ON respectively.

The circuit diagram of the transistor switch is as shown in Figure 2.4.1 with the equivalent circuits. With a zero  $V_{in}$  signal applied to the Base of the transistor it turns “OFF” acting like

an open switch and zero collector current flows. With a positive  $V_{in}$  signal applied to the Base of the transistor, it turns “ON” acting like a closed switch and maximum circuit current flows through the device, provided the base current is large enough to drive the transistor in to saturation.

#### 2.4.2 Transistor as LED driver:

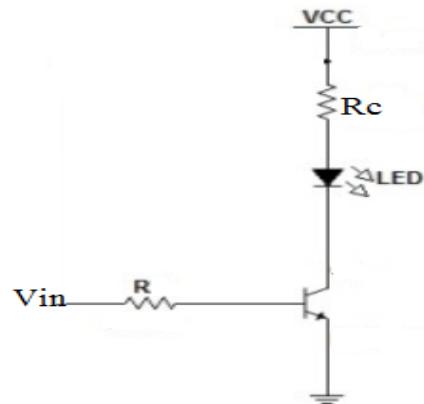


Figure 2.4.2: LED driver circuit

The circuit of an LED driver is as shown in Figure. 2.4.2 The series resistor  $R$  is used to provide the required base current of the transistor. With the input voltage  $V_{in} = 0V$ , no current flows through the base of the transistor, and hence the transistor is in cut-off region. Therefore, with the collector current zero, the LED does not turn on. When  $V_{in} \approx V_{CC}$ , the flow of base current through the transistor drives it to saturation and behaves like a closed switch. The amplified collector current turns the LED on. If the  $V_{in}$  is directly connected to LED through the resistor  $R$ , without any transistor, the current through the LED would be  $I = V_{in}/R$ , which is not sufficient to drive the LED.

#### 2.4.3 Transistor as Inverter.

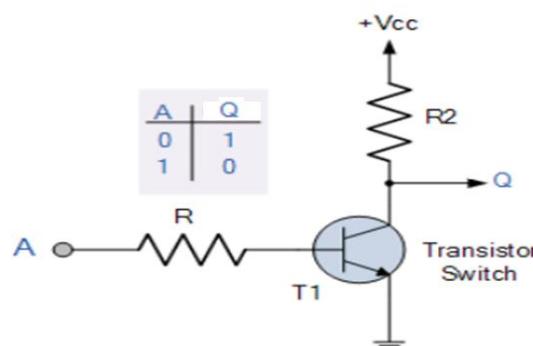


Figure 2.4.3: Transistor inverter circuit

Inverter operates with the condition that input voltage  $A$  can take only two possible values i.e. high (logic 1/ 5V DC) or low (logic 0/ 0V). Correspondingly output voltage  $Q$  can take two possible values either low or high. With input voltage high, transistor is in saturation condition and output voltage is low (logic 0/ 0V). With input voltage low, transistor is in cut-off condition and output voltage is high (logic 1 / 5V DC).

**Self test:**

1. State the operating regions required for the transistor to be used as a switch.
2. Which circuits are named as driver circuits?
3. With what value of  $V_{in}$  the switch is said to be closed?

**Summary**

In this module we have learnt:

1. The transistor can be used as switch by operating it in either saturation region (when the switch is said to be ON) or in cut off region (when the switch is said to be OFF).
2. The applications of transistor switch as LED driver and inverter.

**Chapter: 3****Operational Amplifier and Applications****Module-1: Operational Amplifier**

Operational Amplifiers, or Op-amps as they are usually called, are one of the basic building blocks of Electronic Circuits. Op-Amps are one of the widely used ICs (Integrated Circuits) in electronics. The very name Operational Amplifier comes from the fact that they are used to build circuits to perform variety of mathematical operations such as addition, subtraction, integration, differentiation etc., Operational amplifiers exhibit properties of nearly ideal DC amplifier and are therefore employed in a wide range of applications.

The integrated operational amplifier has gained wide acceptance as a versatile, predictable, and economic system building block because of its small size, high reliability, and reduced cost.

**Learning Outcomes:**

At the end of this module, students will be able to:

1. Draw the internal block diagram of an OP-AMP and briefly describe the functions
2. List and define key parameters of an OP-AMP.
3. Discuss OP-AMP based amplifier topologies.
4. Design OP-AMP based circuits for simple mathematical operations.

### 3.1.1 Introduction

An operational amplifier is a high gain direct coupled amplifier which can amplify signals over a wide range of frequencies. The circuit symbol of op-amp is shown in Fig. 3.1.1, which has two inputs and a single output. The input terminal that is marked as positive is called non-inverting terminal and that marked as negative is known as inverting terminal. The output signal of an Operational Amplifier is the amplified version of the difference between the two signals being applied to the two inputs. One of the common IC versions of op-amp is  $\mu$ A741.

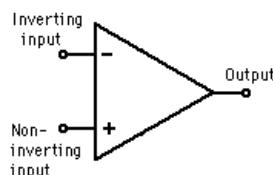


Fig 3.1.1 op-amp symbol

Figure 3.1.2 shows the pin diagrams of different op-amp ICS.

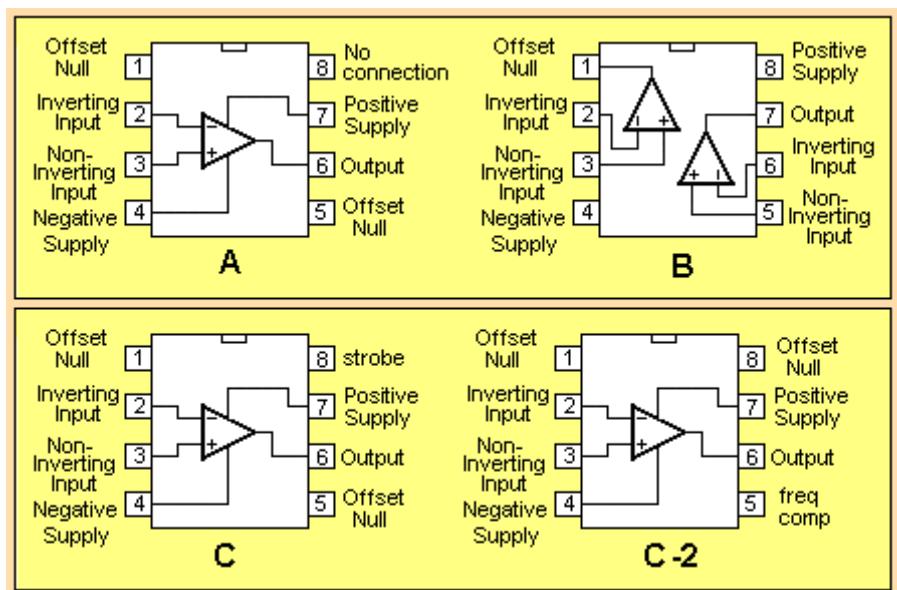


Figure 3.1.2 A: LF353N, B: LF351N, C: CA5130E, C-2: NE531N, D:  
[source: <http://www.talkingelectronics.com/ChipDataEbook-1d/html/OpAmpList.html#C>]

### 3.1.2 Internal Block Diagram of op-amp:

The internal block schematic of op-amp is shown in Figure 3.1.3

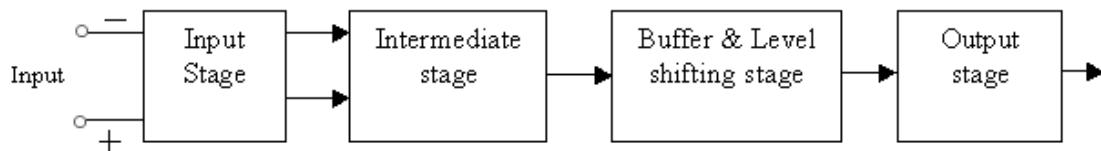


Figure.3.1.3 Internal block diagram of op-amp

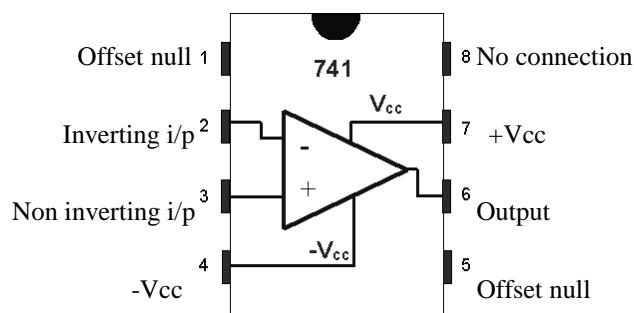
**Input stage:** It is a dual input, dual output differential amplifier. Its function is to amplify the difference between the two input signals. It provides high differential gain, high input impedance and low output impedance. The differential amplifier mainly helps to minimize the effect of noise.

**Intermediate stage:** The overall gain requirement of an op-amp is very high. Since the input stage alone cannot provide such a high gain, an intermediate stage is used to provide the required additional voltage gain.

**Buffer and Level shifting stage:** The dc quiescent voltage level of previous stages may get amplified and applied to the next stage causing distortion at the output. Hence the level shifting stage is used to eliminate the dc level. Buffer is a unity voltage gain amplifier usually used for impedance matching.

**Output stage:** This stage contributes to the overall gain of the op-amp and also provides low output impedance.

The pin diagram for a typical  $\mu$ A741 op-amp with 8 pin DIP (Dual In-line Package) is shown in Figure. 3.1.4.

Figure. 3.1.4 Pin diagram of a typical  $\mu$ A741 op-amp

### Differential Amplifier

The circuit shown in Figure 3.15 shows a generalized form of a differential amplifier with two inputs marked  $V_1$  and  $V_2$ . The two identical transistors  $TR_1$  and  $TR_2$  are both biased at the same operating point with their emitters connected together and returned to the common rail,  $-V_{EE}$  by way of resistor  $R_E$ .

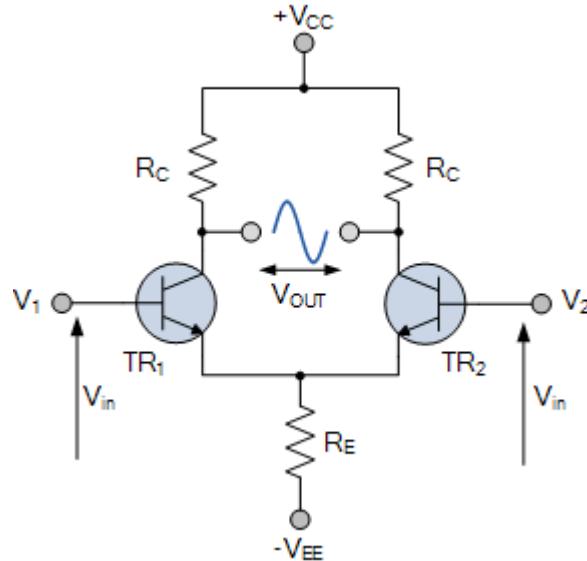


Figure 3.1.5 Differential amplifier circuit diagram

The circuit operates from a dual supply +V<sub>CC</sub> and -V<sub>EE</sub> which ensures a constant supply. The voltage that appears at the output, V<sub>OUT</sub> of the amplifier is the difference between the two input signals as the two base inputs are anti-phase with each other.

When the forward bias voltage of transistor, TR<sub>1</sub> is increased, the forward bias voltage of transistor TR<sub>2</sub> is reduced and vice versa. Then if the two transistors are perfectly matched, the current flowing through the common emitter resistor, R<sub>E</sub> will remain constant.

Like the input signal, the output signal is also balanced and since the collector voltages either swing in opposite directions (anti-phase) or in the same direction (in-phase) the output voltage signal, taken between the two collectors is, (assuming a perfectly balanced circuit) the zero difference between the two collector voltages.

This is known as the *Common Mode of Operation* with the **common mode gain** of the amplifier being the output gain when the input is zero.

Ideal Operational Amplifiers also have one output (although there are ones with an additional differential output) of low impedance that is referenced to a common ground terminal. The op-amp rejects any common mode signals that are appearing at the inputs. That means, if an identical signal is applied to both the inverting and non-inverting inputs then the voltage at the output terminals due to such inputs should be zero. This is measured by a parameter called Common Mode Rejection Ratio (CMRR).

**Common Mode Rejection Ratio (CMRR)** is the ratio of the differential gain to the common mode gain of Op-Amp.

An operational amplifier only responds to the difference between the voltages applied at its two input terminals, known commonly as the “*Differential Input Voltage*”. If the same voltage is applied to both the input terminals the resultant output will be zero. An Operational Amplifiers gain is commonly known as the **Open Loop Differential Gain**, and is represented as (A<sub>o</sub>).

### **Op-amp specifications:**

*Output offset voltage ( $V_{oo}$ ):* The output voltage, when both the inputs are zero is called the output offset voltage. It is due to input offset voltage and input bias current.

*Input bias current ( $I_b$ ):* It is the average of the current that flows into the inverting and non-inverting input terminals when both of the two inputs are grounded.

*Input offset current ( $I_{io}$ ):* It is the algebraic difference between the currents flowing into non-inverting and inverting terminals of balanced op-amp.

*Input resistance ( $R_i$ ):* It is the equivalent resistance that can be measured at either the inverting or non-inverting terminal with the other terminal connected to ground.

*Slew Rate(SR):* It is defined as the maximum rate of change of output voltage per unit time.  
i.e:  $SR = \left( \frac{\Delta V_o}{\Delta t} \right)_{max}$ .

*Supply Voltage Rejection Ratio (SVRR):* The change in op-amp input offset voltage caused by variations in one of the power supply voltage is called SVRR.

*Output resistance ( $R_o$ ):* The equivalent resistance observed between the output terminal and the ground.

*Common Mode Rejection Ratio (CMRR):* This is a figure of merit for an op-amp. It is defined as the ratio of the magnitude of differential gain to the common mode gain.

The CMRR in deciBels is given by

$$CMRR = 20 \log_{10} \left( \frac{A_d}{A_c} \right) \text{ dB} \quad (3.1.1.1)$$

Where,  $A_d$  is the differential gain,  $A_d = \frac{1}{2}(A_1 - A_2)$

$A_c$  is the common-mode gain,  $A_c = (A_1 + A_2)$ . Where  $A_1(A_2)$  is the voltage amplification from input 1(2) to the output under the condition that input 2(1) is grounded. This is an important specification, as it indicates how much of the common-mode signal  $V_c$  gets rejected from the input. A high CMRR is desirable.

The output of a differential amplifier is given by

$$V_o = A_d V_d + A_c V_c. \quad (3.1.1.2)$$

Where,  $V_d = (V_1 - V_2)$  and  $V_c = (V_1 + V_2) / 2$  are differential and common mode inputs respectively. Note:  $V_1$  and  $V_2$  are the non-inverting and inverting input voltages respectively.

### **3.1.3 Op-Amp characteristics**

The characteristics of an ideal op-amp are given in the Table 3.1.1. However, it cannot be realised in practice. As an example, an op-amp IC,  $\mu$ A 741 has typical values of the parameters as given in the table.

Table 3.1.1 Values of an ideal and practical op-amp.

Sl. No.	Characteristics	Ideal	Practical
1	Open loop voltage gain, $A_{vo}$	$\infty$	$2*10^5$
2	Unity gain BW	$\infty$	1MHz
3	Input resistance, $R_i$	$\infty$	$2M\Omega$
4	Output resistance. $R_o$	zero	$75\Omega$
5	CMRR	$\infty$	90dB
6	Slew rate	high	0.5 V/ $\mu$ s
7	SVRR	zero	$150\mu$ V/V
8	Input offset voltage	0	6mv (max)
9	Input offset current	0	200 nA (max)
10	Differential mode output voltage gain $A_d$	$\infty$	50000
11	Common mode output voltage gain $A_c$	0	10

For ideal op-amp, the characteristic do not change with temperature. Ideally, the op-amp is perfectly balanced, if  $V_o = 0$ , when  $V_1 = V_2$

#### ***Concept of Virtual ground:***

The input impedance of an ideal op-amp is infinite ( $R_i = \infty$ ), that means there is no current flowing into the op-amp. As the differential voltage gain of an ideal op-amp is infinite,  $V_1 - V_2$  tends to zero. This is equivalent to virtual short between two input terminals and hence if one of the terminals is grounded the other terminal also experiences the same potential even though they are not electrically connected. Therefore, it is called virtual ground.

#### ***Transfer Characteristics of a typical op-amp:***

The transfer characteristics of op-amp is as shown in Figure 3.1.6. In the linear region, any change in the input difference voltage,  $\pm V_{id}$  produces a proportional output voltage. The range of input difference voltage to operate the op-amp in linear region is approximately equal to 100 mV. Beyond 100mV of  $\pm V_{id}$ , the output becomes  $\pm V_{sat}$  because of very high gain offered by the op-amp. The output will be at  $+V_{sat}$  if it is used in non-inverting mode or  $-V_{sat}$  if it is configured in inverting mode.

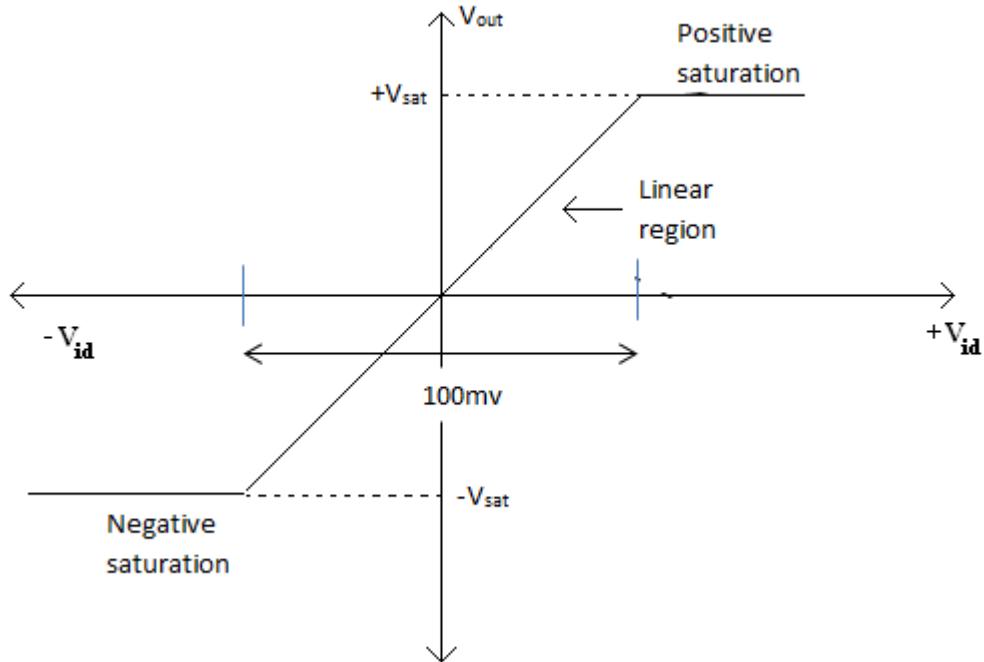


Fig.3.1.6 Transfer Characteristics of a typical op-amp

In linear applications, the op-amp is operated as a closed loop amplifier in the active region. The difference between the input voltages is maintained around 100mV so as to produce a linear output voltage. In nonlinear applications, the op-amp is driven to saturation either in open loop or closed loop configuration by applying a difference input voltage exceeding 100mV.

**Self test:**

1. A linear integrated circuit responds to
  - a) Analog signal
  - b) Digital signal
  - c) Neither a) nor b)
  - d) Both a) and b)
2. An op-amp can amplify
  - a) only ac signals
  - b) only dc signal
  - c) Neither (a) nor (b)
  - d) Both (a) and (b)
3. The ability of an op-amp to reject the common mode signal is termed as its
  - a) common mode gain
  - b) differential mode gain
  - c) offset voltage
  - d) CMRR
4. The CMRR of an op-amp is usually expressed in
  - a) volts
  - b) decibels (dB)
  - c) volts/sec
  - d) volts/mS
5. The potential at the virtual node w.r.t ground in an ideal op-amp is
  - a) 0V
  - b) 5V
  - c) 10V
  - d) 100mV

**Summary:**

1. An operational amplifier is a high gain direct coupled amplifier which can amplify signals over a wide range of frequencies.
2. The amplified output signal of an Operational Amplifier is the difference between the two signals being applied to the two inputs. Hence it is termed as differential amplifier.
3. The ratio of the change to the output voltage with respect to the change in the common mode input voltage is called the Common Mode Rejection Ratio or CMRR.
4. The input impedance of an ideal op-amp is infinite ( $R_i = \infty$ ), that means there is no current flowing into the op-amp.
5. In linear applications, the op-amp is operated as a closed loop amplifier in the active region.
6. In nonlinear applications, the op-amp is driven to saturation either in open loop or closed loop configuration by applying a difference input voltage exceeding 100mV.

**Module-2: Linear Applications of op-amp**

Linear applications of op-amps include mathematical operations such as inversion, addition, subtraction, integration, differentiation and multiplication etc., some of them will be discussed here.

**Learning Outcomes:**

At the end of this module, students will be able to:

1. Discuss OP-AMP based amplifier topologies.
2. Analyze basic OP-AMP circuits.
3. Design OP-AMP based circuits for implementing simple mathematical operations.

**3.2.1 Inverting amplifier:** The circuit diagram for an inverting amplifier is as shown in Figure 3.2.1.

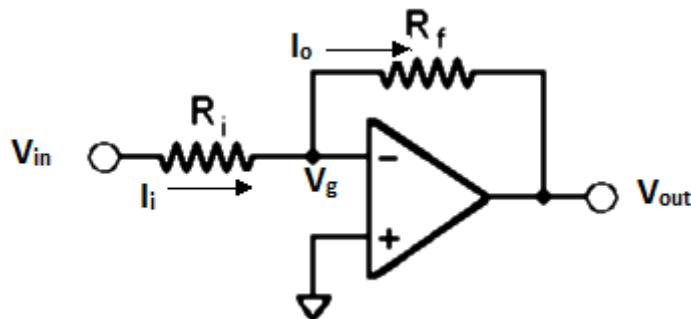


Figure 3.2.1 Inverting amplifier

By applying KCL to the inverting input node,  $I_i = I_o$

$$\frac{V_{in} - V_g}{R_i} = \frac{V_g - V_{out}}{R_f} \quad (3.2.1)$$

The inverting input terminal is at virtual ground, i.e.  $V_g = 0$ . Substituting in the above expression,

$$\frac{V_{in}}{R_i} = -\frac{V_{out}}{R_f} \quad (3.2.2)$$

$$V_{out} = -\left(\frac{R_f}{R_i}\right)V_{in} \quad (3.2.3)$$

The closed loop voltage gain

$$A_V = \frac{V_{out}}{V_{in}} = -\left(\frac{R_f}{R_i}\right) \quad (3.2.4)$$

The negative sign indicates that there is  $180^\circ$  phase difference between input and output signals. The voltage gain depends only on the resistor values as long as the op-amp is in linear region.

### 3.2.2 Non-inverting amplifier:

The circuit diagram for a non-inverting amplifier is shown in Figure 3.2.2

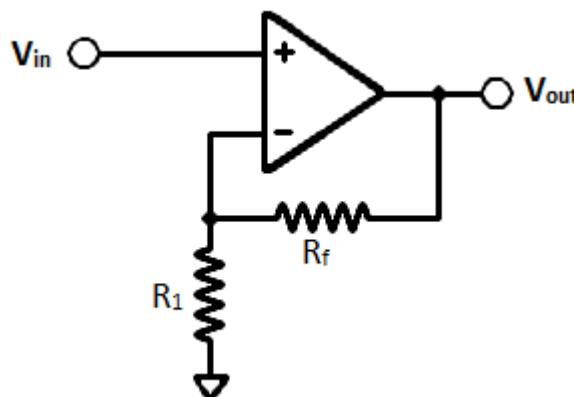


Figure. 3.2.2: Non-inverting amplifier

Since there is virtual short between inverting and non-inverting input terminals,  $V_{in}$  appears across  $R_1$ .

$$\text{Applying KCL, } V_{in} / R_1 + (V_{in} - V_{out}) / R_f = 0$$

Solving for  $V_{out}$ ,

$$V_{out} = V_{in} \left( 1 + \frac{R_f}{R_1} \right) \quad (3.2.5)$$

The closed loop voltage gain for non-inverting amplifier is given by,

$$A_v = (1 + R_f / R_1) \quad (3.2.6)$$

### 3.2.3 Voltage follower

Voltage follower is a special case of non-inverting amplifier with unity gain as shown in Figure 3.2.3. It has high input impedance and very low output impedance and hence it is used as a buffer amplifier to achieve impedance matching.

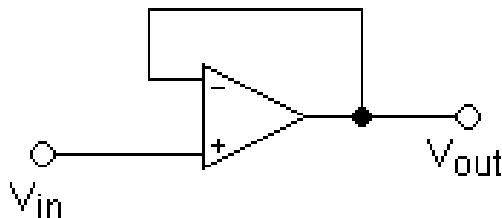


Figure 3.2.3: Voltage follower

Substituting  $R_f = 0$  and  $R_1 = \infty$  in Eqn. (3.2.6),

$$V_{out} = V_{in} \quad (3.2.7)$$

### 3.2.4 Inverting summing amplifier

The circuit diagram for an inverting summing amplifier is shown in Figure 3.2.4

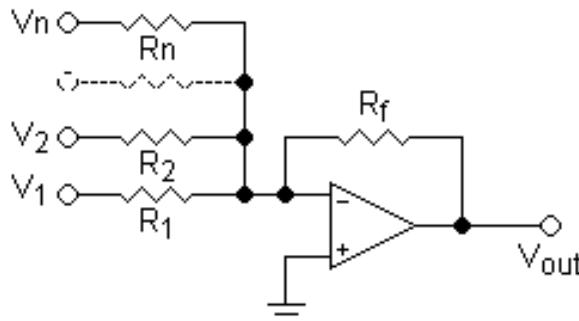


Figure.3.2.4: Inverting summing amplifier

By applying KCL to the inverting input node and making use of the virtual ground concept, the output voltage is,

$$V_{\text{out}} = -R_f \left( \frac{V_1}{R_1} + \frac{V_2}{R_2} + \dots + \frac{V_n}{R_n} \right) \quad (3.2.8)$$

Thus, summing amplifier produces an output voltage which is an inverted (in sign), weighted sum of all inputs.

If  $R_1 = R_2 = \dots = R_n = R_f$ , then,

$$V_{\text{out}} = -\left(\frac{R_f}{R_1}\right) (V_1 + V_2 + \dots + V_n) \quad (3.2.9)$$

If  $R_1 = R_2 = \dots = R_n = R_f$  then

$$V_{\text{out}} = -(V_1 + V_2 + \dots + V_n) \quad (3.2.10)$$

The circuit is therefore acts as an adder or summer. Strictly speaking, this circuit is acting as an inverting adder.

### 3.2.5 Difference amplifier

The circuit configuration for a difference amplifier is shown in Figure.3.2.5. It produces an output voltage which is proportional to the difference between the two inputs.

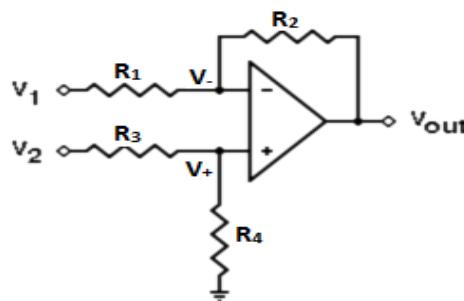


Figure 3.2.5: Difference Amplifier

Using Superposition theorem

$$\text{When } V_2=0, \quad V_{\text{out1}} = -V_1 \frac{R_2}{R_1} \quad (3.2.11)$$

$$\text{When } V_1=0, \quad V_{\text{out2}} = V_+ \left( 1 + \frac{R_2}{R_1} \right) = V_2 \frac{R_4}{R_3 + R_4} \left( 1 + \frac{R_2}{R_1} \right) \quad (3.2.12)$$

$$\text{Then, } V_{out} = V_{out2} + V_{out1} = V_2 \frac{R_4}{R_3 + R_4} \left( 1 + \frac{R_2}{R_1} \right) - V_1 \frac{R_2}{R_1} \quad (3.2.13)$$

$$\text{If } R_2/R_1 = R_4/R_3 \text{ then, } V_{out} = \frac{R_2}{R_1} (V_2 - V_1) \quad (3.2.14)$$

The circuit is called a difference amplifier and if,  $R_1 = R_2 = R_3 = R_4$ , the above equation simplifies to  $V_{out} = V_2 - V_1$  and the circuit acts as a subtractor.

### 3.2.6 Integrator

The circuit for an integrator is shown in Figure. 3.2.6. It produces an output voltage which is proportional to the integral of input voltage.

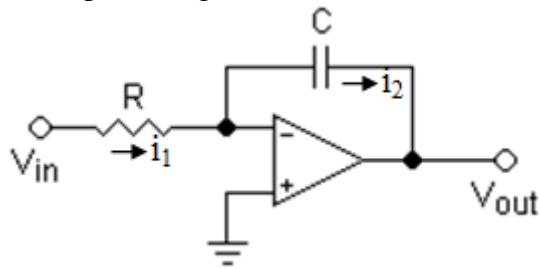


Figure. 3.2.6: Integrator

Applying KCL at inverting terminal,  $i_1 = i_2$

$$\frac{V_{in} - 0}{R} = C \frac{d(0 - V_{out})}{dt} \quad (3.2.15)$$

Integrating on both sides

$$V_{out} = -\frac{1}{RC} \int_0^t V_{in} dt \quad (3.2.16)$$

### 3.2.7 Differentiator

The differentiator circuit is shown in Figure 3.2.7. It produces an output voltage which is proportional to the differential of input voltage.

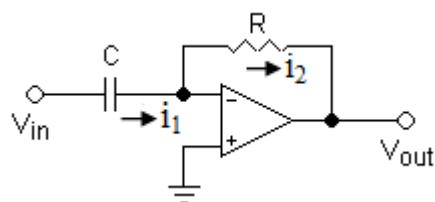


Fig.3.2.7: Differentiator

Applying KCL at inverting terminal  $i_1 = i_2$

$$C \frac{d(V_{in} - 0)}{dt} = \frac{0 - V_{out}}{R} \Rightarrow V_{out} = -RC \frac{dV_{in}}{dt} \quad (3.2.17)$$

**Exercises:**

1. Realize each of the following equations using single OPAMP.
  - (i)  $V_o = -5V_1$
  - (ii)  $V_o = +5V_1$  (iii)  $V_o = -(5V_1 + 7V_2)$  (iv)  $V_o = V_1 - 0.5V_2$
2. Realize the equation using OPAMP  $V_o = 3V_1 - 0.8V_2 + 0.5V_3$
3. Sketch the output waveform for an inverting integrator if the input signal is square wave with Amplitude is 5V and frequency 1KHz.
4. A 200mV peak to peak sine wave form voltage is applied to an OPAMP inverting amplifier with  $R_f/R_1 = 10$ . Sketch the output.

**Summary:**

1. The closed loop voltage gain  $A_v = -(R_f/R_i)$ . The negative sign indicates that there is  $180^\circ$  phase difference between input and output signals.
2. Voltage follower is a special case of non-inverting amplifier with unity gain.
3. Summing amplifier produces an output voltage which is an inverted (in sign), weighted sum of all inputs.
4. The circuit of an integrator produces an output voltage which is proportional to the integral of input voltage.
5. The differentiator circuit produces an output voltage which is proportional to the differential of input voltage.

### Module-3: Non Linear Application

#### Learning Outcomes:

At the end of this module, students will be able to :

1. Discuss different types of OP-AMP based Comparators.
2. Draw the circuit of square wave generator using op-amp.

#### 3.3.1 Voltage Comparator:

The op-amp in an open loop configuration is shown in Figure 3.3.1

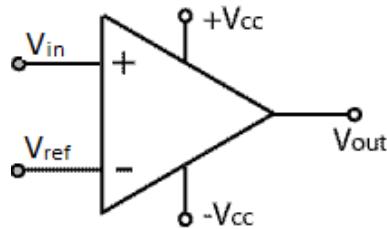


Figure 3.3.1: Voltage comparator

The output voltage of the circuit / topology is given by,

$$V_{out} = +V_{sat} \quad \text{if } V_{in} > V_{ref} \quad (3.3.18)$$

$$V_{out} = -V_{sat} \quad \text{if } V_{in} < V_{ref} \quad (3.3.19)$$

#### 3.3.2 Square wave generator

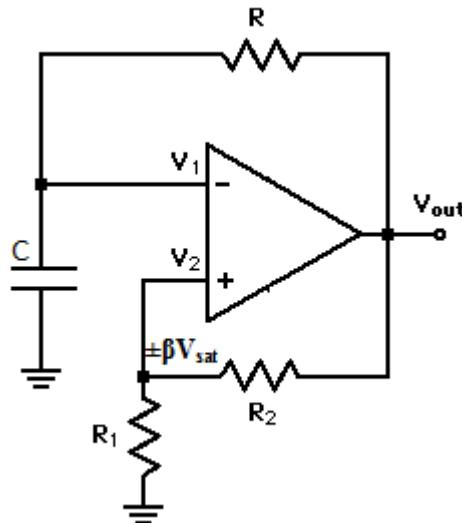


Fig.3.3.2a: Circuit of Square wave generator

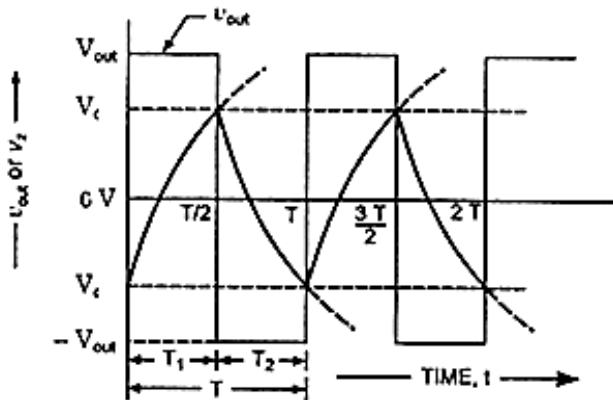


Fig.3.3.2b: Waveform across capacitor and output

The output of the op-amp will be at either positive or negative saturation voltages ( $\pm V_{sat}$ ) depending on  $V_1$  and  $V_2$  i.e.

$$\begin{aligned} V_{out} &= +V_{sat} & V_1 < V_2 \\ &= -V_{sat} & V_1 > V_2 \end{aligned}$$

Here  $V_1$  is the voltage across the capacitor

$$V_2 = V_{out} \frac{R_1}{R_1 + R_2} = \beta V_{out} \quad (3.3.20)$$

$$\text{where } \beta \text{ is feedback factor given by } \beta = \frac{R_1}{R_1 + R_2} \quad (3.3.21)$$

When  $V_{out} = +V_{sat}$ , capacitor will charge towards  $+V_{sat}$ . When the voltage across capacitor  $V_1$  exceeds  $\beta V_{sat}$ , output makes a transition to  $-V_{sat}$  and capacitor will start discharging towards  $-V_{sat}$ . When the voltage across capacitor  $V_1$  becomes slightly less than  $(-\beta V_{sat})$ , output makes a transition to  $+V_{sat}$  and this action repeats.

The time period of the square wave generated is

$$T = 2RC \ln \left[ \frac{1+\beta}{1-\beta} \right] \quad (3.3.22)$$

$$\text{Where } \beta = \frac{R_1}{R_1 + R_2} \quad (3.3.23)$$

Peak to peak amplitude of the square wave generated is

$$V_{o(p-p)} = +2V_{sat} \quad (3.3.24)$$

**Exercises:**

1. Design a square wave generator using OP-AMP for the following specifications:  
Frequency of oscillation = 1KHz,  $V_0(p-p) = 12.4V$ .
2. What should be the  $T_{ON}$  and  $T_{OFF}$  of a square wave signal of frequency 2KHz and duty cycle of 50% ?
3. Explain the working of OP-AMP as a Comparator.

**Summary:**

1. An operational amplifier (op-amp) has a well-balanced difference input and a very high gain. This parallels the characteristics of comparators.
2. The output of the op-amp in a square wave circuit will be at either positive or negative saturation voltages ( $\pm V_{sat}$ ) depending on  $V_1$  and  $V_2$

## Part –II

# DIGITAL ELECTRONICS

### Chapter - 4 : Number systems and codes

Electronic circuits and systems can be broadly classified into analog and digital. Analog circuits are those in which voltages and currents show continuous variations with respect to time and can take any arbitrary value of magnitude within a specified range. A digital circuit is one in which the voltage levels assume a finite number of distinct values. In all modern digital circuits, normally there are two discrete voltage levels, called logic ‘0’ and logic ‘1’.

#### **Learning Outcomes:**

At the end of this module, students will be able to:

1. Describe different types of number system.
2. Represent data in the respective number system.
3. Convert the number from one system to another system.
4. Perform binary arithmetic using ones and two's complement

#### **Module 1- Number systems:**

We discuss binary, decimal, octal and hexadecimal number systems in this module. A radix or base is an important part of any number system. The total number of symbols in every number system is equal to its base or radix. In fact the name of the number system is derived from its base.

##### *Decimal Number system:*

This system has 10 symbols, namely 0,1,2,3,4,5,6,7,8 and 9. The decimal number system is also called the base ‘10’ system as it has ‘10’ digits. Example:  $(781)_{10}$ ,  $(82.901)_{10}$

##### *Binary Number system:*

A number system that uses only two symbols ‘0’and‘1’ is called binary number system or base 2 system or radix-2 system. The symbols are called bits. Example:  $(100010)_2$ ,  $(0.1011)_2$ .

*Octal Number System:* A number system that uses 8 symbols (0-7) is called an octal number system. The radix or base of octal number system is 8. Example:  $(723)_8$ ,  $(6.76)_8$ .

##### *Hexadecimal Number System:*

The hexadecimal number system has base 16. It has 16 distinct symbols. It uses the digits 0-9 in addition to alphabets A,B,C,D, E and F as 16 symbols to represent the numbers.

*Self test:*

1. The hex numbering system has a base of \_\_\_\_\_, and the binary numbering system has a base of \_\_\_\_\_.
2. The value of a particular digit in a number is determined by its relative position in a sequence of digits. (T/F)
3. A single hexadecimal digit can represent how many binary bits: (a) two, (b) three, (c) four?
4. The bases of the binary and decimal numbering systems are multiples of 2. (T/F)

**4.1 Conversion of numbers:** The decimal system is a more familiar system than the other systems. So it is essential to understand the conversion of a number from any base to decimal and vice versa. The computer systems accept the data in decimal form, whereas they store and process the data in binary form. Therefore, it becomes necessary to convert the numbers represented in one system into the numbers represented in another system.

#### 4.1.1: Decimal Number System

**Decimal to Binary:** The given decimal number is repeatedly divided by 2, which is the base number of binary system till quotient becomes ‘0’ and the remainder is collected from bottom to top. To convert the fractional part into binary, fraction part is multiplied by 2 repeatedly and any carry in integer place is recorded. The string of integer obtained from top to bottom gives the equivalent fraction in binary number system.

$$\text{Ex1: } (37)_{10} = (100101)_2$$

37	divided by	2	Q=18	R=1
18	divided by	2	Q=9	R=0
9	divided by	2	Q=4	R=1
4	divided by	2	Q=2	R=0
2	divided by	2	Q=1	R=0

Note the way the bits are read to form the binary number.

**Decimal to Octal:** The given decimal number is repeatedly divided by 8, which is the base number of octal system till quotient becomes ‘0’ and the remainder is collected from bottom to top. To convert the fractional part into octal, fraction part is multiplied by 8 repeatedly and any carry in integer place is recorded. The string of integer obtained from top to bottom gives the equivalent fraction in the octal number system.

$$\text{Example - 1: } (97)_{10} = (141)_8$$

97	divided by	8	Q=12	R=1
12	divided by	8	Q=1	R=4

**Decimal to Hexadecimal:** The given decimal number is repeatedly divided by 16, which is the base number of hexadecimal system till quotient becomes ‘0’ and the remainder is collected from bottom to top. To convert the fractional part into hexadecimal, fraction part is multiplied by 16 repeatedly and any carry in integer place is recorded. The string of integer obtained from top to bottom gives the equivalent fraction in the hexadecimal number system.

Ex1:  $(546)_{10} = (222)_{16}$

546	divided by	16	Q=34	R=2
34	divided by	16	Q=2	R=2

#### 4.1.2: Binary Number System

**Binary to decimal:** Multiply the number by its equivalent binary weights. Add the products to get the decimal number.

Ex1:  $110_2 = 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = 4 + 2 + 0 = 6$

Ex2:  $0.101_2 = 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} = 0.5 + 0 + 0.125 = 0.625$

Ex3:  $110.101_2 = 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 + 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} = 4 + 2 + 0 + 0.5 + 0 + 0.125 = 6.625$

**Binary to Octal:** Each octal digit is represented by three bits of binary.

Ex1:  $(110101010)_2 = (652)_8$     110 101 010    6    5    2

Ex2:  $(0.1011010)_2 = (550)_8$     101 101 0    5    5    0

Ex3:  $(110101010. 1011010)_2 = (652.550)_8$

**Binary to Hexadecimal:** Each group of 4 binary bits is equal to 1 hexadecimal digit.

Ex1:  $(11110110010)_2 = (FB2)_{16}$     1111 1011 0010    15    11    2 ;    F    B    2

Ex2:  $(0.01010111011)_2 = (0.57B)_{16}$     0101 0111 1011;    5    7    11 ;    5    7    B

#### 4.1.3: Octal Number system:

**Octal to Binary:** Each octal digit is represented by three binary bits.

Ex1:  $(347)_8 = (011100111)_2$

Ex2:  $(0.245)_8 = (0.010100101)_2$

Ex3:  $347.245_8 = (011100111.010100101)_2$

**Octal to decimal:** Multiply the number by its equivalent octal weights. Add the products to get the decimal number.

$$\text{Ex1: } (457)_8 = 4 \times 8^2 + 5 \times 8^1 + 7 \times 8^0 = 4 \times 64 + 5 \times 8 + 7 \times 1 = 256 + 40 + 7 = (303)_{10}$$

$$\begin{aligned} \text{Ex2: } (0.246)_8 &= 2 \times 8^{-1} + 4 \times 8^{-2} + 6 \times 8^{-3} = 2 \times 0.125 + 4 \times 0.015625 + 6 \times 0.001953125 \\ &= (0.267969)_{10} \end{aligned}$$

$$\text{Ex3: } (457.246)_8 = (303.267969)_{10}$$

**Octal to Hexadecimal:** To convert an octal number to hexadecimal number, the octal number is first converted into binary. The binary bits are grouped such that each group consists of 4 bits and a group starts from the LSB. The 4 bits group is represented by its equivalent hexadecimal number.

Ex1:  $(235)_8 = (09D)_{16}$

$$(235)_8 = (\underline{0} \underline{1001} \underline{1101})_2 = (09D)_{16}$$

#### 4.1.4 Hexadecimal Number Systems:

**Hexadecimal to Binary:** Each group of 4 binary bits is equal to one hexadecimal digit.

$$\text{Ex1: } (A7D)_{16} = (\underline{1010} \underline{0111} \underline{1110})_2$$

**Hexadecimal to Octal:** To convert an hexadecimal number to octal number, the hexadecimal number is first converted into binary. They are grouped into 3 bits each starting from the LSB. The 3 bits group is represented by its equivalent octal number.

Ex1:  $(C4)_{16} = (11\ 000\ 100)_2 = (304)_8$

Ex2:  $(0.26A)_{16} = (0.001\ 001\ 101\ 010)_2 = (0.1152)_8$

Ex3:  $(C4.26A)_{16} = (11\ 000\ 100.001\ 001\ 101\ 010)_2 = (304.1152)_8$

**Hexadecimal to Decimal:** Multiply the number by its equivalent hexadecimal weights. Add the products to get the decimal number.

Ex1:  $(B40)_{16} = 11 \times 16^2 + 4 \times 16^1 + 0 \times 16^0 = 11 \times 256 + 4 \times 16 + 0 \times 1 = (2880)_{10}$

Ex2:  $(0.237)_{16} = 2 \times 16^{-1} + 3 \times 16^{-2} + 7 \times 16^{-3} = (0.138427)_{10}$

Ex3:  $(B40.237)_{16} = (2880.138427)_{10}$

*Self test:*

1. The binary equivalent of decimal number 255 is \_\_\_\_\_.
2. The binary equivalent of hexadecimal number 1C is \_\_\_\_\_.
3. The decimal equivalent of hexadecimal number 1B6 is \_\_\_\_\_.
4. The hexadecimal equivalent of decimal number 129 is \_\_\_\_\_.
5. The decimal equivalent of binary number 110101 is \_\_\_\_\_.
6. The hexadecimal equivalent of binary number 1001 is \_\_\_\_\_.
7. The binary equivalent of decimal number 28 is \_\_\_\_\_.
8. The binary equivalent of hexadecimal number 35 is \_\_\_\_\_.
9. The decimal equivalent of hexadecimal number 7 is \_\_\_\_\_.
10. The hexadecimal equivalent of decimal number 49 is \_\_\_\_\_.
11. The decimal equivalent of binary number 110110110 is \_\_\_\_\_.
12. The hexadecimal equivalent of binary number 1110 is \_\_\_\_\_.

#### 4.1.5 Number System - Arithmetic:

**Addition in Binary system:** It is a simple task to add two binary numbers and it is very similar to addition of decimal numbers.

Ex1:  $1010 + 0111 = 10001$

$$1010 = 10$$

$$+ 0111 = 7$$

$$10001 = 17$$

Augend	Addend	Sum	Carry
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

**Addition in Octal system:** Add the digit in each column in decimal and convert this sum into octal. Write the sum in that column and carry the carry term to the next higher significant column.

Ex1: Add  $(334.65)_8$  to  $(671.14)_8 = (1226.01)_8$

Augend	Addend	Sum	Carry
5	4	1 ( $5+4=9-8$ )	1(8 is subtracted once)
6	1	0 ( $6+1+1=8-8$ )	1(8 is subtracted once)
4	1	6 ( $4+1+1$ )	0(since sum <8)
3	7	2( $3+7=10-8$ )	1(8 is subtracted once)
3	6	2( $3+6+1=10-8$ )	1(8 is subtracted once)

**Addition in Hexadecimal system:** Add the digit in each column in decimal and convert this sum into hexadecimal number. Write the sum in that column and carry term to the next higher significant column.

Ex1: Add  $(7AB.67)_{16}$  to  $(15C.71)_{16} = (907.D8)_{16}$

Augend	Addend	Sum	Carry
7	1	8	0(since sum<16)
6	7	6+7 = 13 (D)	0(since sum <16)
B(11)	C(12)	11+12=23 – 16= 7	1(16 subtracted once)
A(10)	5	10+5+1 = 16 – 16=0	1(16 subtracted once)
7	1	7+1+1 = 9	0(since sum <16)

*Self test:*

1. The result of  $101_2 + 11_2$  is \_\_\_\_\_ (in binary).
2. The result of  $A1_{16} + BC_{16} + 10_{16}$  is \_\_\_\_\_ (in hexadecimal).
3. The result of  $60_{10} + F1_{16} - 1001001_2$  is \_\_\_\_\_ (in decimal).
4. The result of  $11_2 + 27_8 + 93_{10} - B_{16}$  is \_\_\_\_\_ (in decimal).

#### 4.1.6 Complementation of numbers:

The subtraction operation and logical manipulations become easy in digital computers by using the concept of complements. For a given number ‘N’ in base-‘r’, two types of complements are defined, namely, r’s complement and (r-1)’s complement.

The main advantage of performing subtraction by complement is that it requires less hardware to implement.

*Subtraction using complements in Binary system:*

The following rules are used in performing complement subtraction in binary system.

1. Identify the minuend and the subtrahend.
2. Find the 2’s complement or the 1’s complement of the subtrahend.
3. Add the complemented subtrahend to the minuend.
4. Check if there is a carry. If there is no carry in either of the form (1’s or 2’s complement) then find the corresponding complement of the result.
5. In case of 2’s complement if there is a carry then neglect it. In case of 1’s complement if there is a carry then add it to the LSB of the result.

**Note:** To find 1’s complement of a binary number we need to invert each bit.

To find 2’s complement of a binary number find 1’s complement and add 1 to the result.

**Ex1:** Subtract  $(0011)_2$  from  $(0101)_2$  using 1’s complement.

Step1: 0101 is minuend and 0011 is subtrahend.

Step2: 1’s complement of 0011 is 1100.

Step3:  $0101 + 1100 = 1\ 0001$  (there is a carry)

Step4:  $0001 + 1 = 0010$  ( $5 - 3 = 2$ )

**Exercises:**

1. Subtract  $(AEF3.6D)_{16}$  from  $(445.63)_8$  using 1's complement method.

2. Perform the following

(i)  $(257.75)_{10} - (128.825)_{10}$  using binary 2's complement arithmetic

(ii)  $(ABCD)_{16} = (?)_{10} = (?)_2 = (?)_8$

3. Perform the following addition in the binary system.

$$1101 + 1001$$

$$1101.01 + 101.11$$

$$11.011 + 10.111$$

$$1001 + 111$$

4. Determine the base  $b$  of the number system such that  $225_{(b)} = 89_{(10)}$

5. Convert each of the following hexadecimal number into its equivalent in the binary number system

i) 1C.3

ii) F2.C

iii) 450.B

iv) 8EA.59

6. Perform the following conversions

i)  $(7305)_{10}$  to hexadecimal

ii)  $(7305)_8$  to decimal

7. Perform the following:

i)  $(F69.D3)_{16} - (A25.68)_{16} = (?)_{10}$

ii)  $(13.25)_{10} - (26.37)_{10}$  using Binary 1's complement arithmetic

### Summary

1. A numeral system (or system of numeration) is a writing system for expressing numbers, that is, a mathematical notation for representing numbers of a given set, using digits or other symbols in a consistent manner.
2. The computer systems accept the data in decimal form, whereas they store and process the data in binary form. Therefore, it becomes necessary to convert the numbers represented in one system into the numbers represented in another system.
3. The advantage of performing subtraction by complement is reduction in the hardware. Instead of having separate digital circuits for addition and subtraction, only adding circuits are needed.

## Module 2 - Codes

### Learning Outcomes:

At the end of this module, students will be able to:

1. Discuss different types of binary codes.
2. Explain error detection using parity bit.
3. Describe error correction using hamming code.

When numbers, letters or words are represented as a specific group of symbols based on certain rules, it is said to be encoded. The group of symbols is called a code. Codes are represented, stored and transmitted in the form of binary bits. The codes may also use alphanumeric characters.

Advantages of Binary Code:

- Suitable for computer applications.
- Used in digital communications.
- Ease of circuit implementation.

Classification of binary codes

Generally codes are classified into following categories.

- Weighted Codes
- Non-Weighted Codes
- Alphanumeric Codes
- Error Detecting Codes
- Error Correcting Codes

#### **4.2.1 Weighted binary codes:**

Weighted binary codes are those binary codes which obey the positional weight principle. Each position of the bit or digit in the code has a specific weight. Decimal digits 0-9 can be represented using 4 bit binary code. One such example is 8421 BCD (Binary coded Decimal), where each group of 4 bits is a weighted code as shown in Figure 4.2.1.

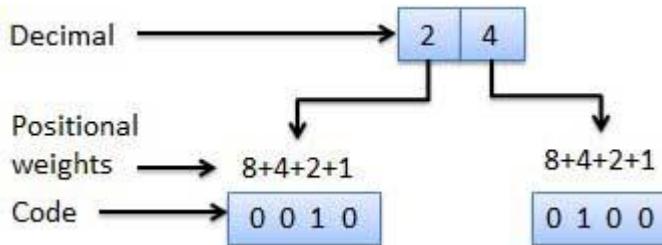


Figure 4.2.1: 8421 BCD code Illustration

Value of the number given by the code can be obtained using the following equation.

$$N = d_3w_3 + d_2w_2 + d_1w_1 + d_0w_0 \quad (4.2.1)$$

$w_3, w_2, w_1$  and  $w_0$  are the weights selected for a given BCD code:

<b>w<sub>3</sub></b>	<b>w<sub>2</sub></b>	<b>w<sub>1</sub></b>	<b>w<sub>0</sub></b>
<b>8</b>	4	2	1

Figure 4.2.2 weights of a BCD code.

For example the decimal number 9 would be represented by 1001 where  $9 = 1 \times 8 + 0 \times 4 + 0 \times 2 + 1 \times 1$ . The most commonly used binary coded decimal (BCD) is shown in Table 4.2.1

Table 4.2.1 Decimal to Other BCD forms

Decimal	8421	7421	4221	8421
0	0000	0000	0000	0000
1	0001	0001	0001	0111
2	0010	0010	0010	0110
3	0011	0011	0011	0101
4	0100	0100	1000	0100
5	0101	0101	0111	1011
6	0110	0110	1100	1010
7	0111	1000	1101	1001
8	1000	1001	1110	1000
9	1001	1010	1111	1111

#### **4.2.2 Non-Weighted Codes:**

Non-weighted code is one in which the positions in the code do not have a specific weight.

#### **EXCESS-3 CODE**

The Excess-3 code is also called XS-3 code. It is non-weighted code used to express decimal numbers. The Excess-3 code words are derived from the 8421 BCD code words by adding  $(0011)_2$  or  $(3)_{10}$  to each code word in 8421.

#### **GRAY CODE**

The Gray code is neither a decimal code, nor is it an arithmetic code. The essential feature of a Gray code is that there is only a single bit difference between successive code words.

Table 4.2.2 BCD, Excess-3 and Gray code equivalent for decimal numbers

Decimal	BCD = 8421	Excess-3	Gray
0	0000	0011	0000
1	0001	0100	0001
2	0010	0101	0011
3	0011	0110	0010
4	0100	0111	0110
5	0101	1000	0111
6	0110	1001	0101
7	0111	1010	0100
8	1000	1011	1100
9	1001	1100	1101

#### **4.2.3 Self complementing code:**

When arithmetic operations are performed, it is often required to take the “complement” of a given number. If the logical complement of a coded number is also its arithmetic complement, it will be convenient from hardware point of view. Self-complementing codes are applicable in such scenario. A code is said to be self-complementing, if the code word of the 9's complement of N can be obtained from the code word of N by converting all the 0's into 1's and all 1's into 0's.

Example: Consider the 2421 code.

The 2421 code of  $(4)_{10}$  is 0100.

Its complement is 1011 which is 2421 code for  $(5)_{10} = (9 - 4)_{10}$ .

#### **4.2.4 Error detection and correction codes:**

When data is transmitted in digital form from one place to another place through a transmission channel some data bits may be lost or modified. As there are needs to transmit millions of bits per second, the data integrity should be very high. The error rate cannot be reduced to zero. A simple process of adding a special code bit to a data word can improve its integrity. This extra bit will allow detection of a single error in a given code word in which it is used, and is called the ‘Parity Bit’.

##### **a) Parity (Error detection code)**

The simplest technique for error detection is that of adding an extra bit known as parity bit, to each word being transmitted. There are two types of parity – odd parity and even parity.

- i. Odd parity: the parity bit is set to 0 or 1 at the transmitter such that the total number of 1 bit in the word including the parity bit is an odd number.
- ii. Even parity: the parity bit is set to 0 or 1 at the transmitter such that the total number of 1 bit in the word including the parity bit is an even number.

Table 4.2.3 BCD with Odd and Even parity

Decimal	8 4 2 1 BCD	Odd Parity	Even Parity
0	0 0 0 0	00001	00000
1	0 0 0 1	0	1
2	0 0 1 0	0	1
3	0 0 1 1	1	0
4	0 1 0 0	0	1
5	0 1 0 1	1	0
6	0 1 1 0	1	0
7	0 1 1 1	0	1
8	1 0 0 0	0	1
9	1 0 0 1	1	0

***Self test:***

1. *What is the simplest technique for detecting error in the codes?*
2. *What is a parity bit?*
3. *What is a self-complementing code? Give an example.*

1. Example: In an even parity scheme, which of the following words contain an error?
  - a. 10101010 Ans: No
  - b. 1110110 Ans: Error
  - c. 10111001 Ans: Error
2. Example: In an odd parity scheme, which of the following words contain an error?
  - a) 10110111 Ans: Error
  - b) 10011010 Ans: Error
  - c) 11101010 Ans: No

***b) Hamming Codes (Error correction code)***

A code is said to be an error-correcting code, if there is a built in capability to correct errors occurred in the code word by some means, and thereby the correct code word can always be recovered from an erroneous word. For a code to be a single – bit error correcting code, the minimum distance of that code must be three. The minimum distance of a code is the smallest number of bits by which any two code words must differ. A code with minimum distance of three cannot only correct single bit errors, but also detect two bit errors. One type of error correcting code is hamming code.

In this code, to each group of  $m$  information,  $k$  parity checking bits denoted by  $P_1$  to  $P_k$  located at position  $2^{(k-1)}$  from left are added to form an  $(m+k)$  bit code word. To correct the error,  $k$  parity checks are performed on selected bits of each code word, and the position of the error bit is located by forming an error word and the error bit is then complemented. The  $k$  bit error word is generated by putting a 0 or a 1 in the  $2^{(k-1)}$ th position depending upon whether the check for parity involving the parity bit  $P_k$  is satisfied or not.

***7 Bit Hamming code***

To transmit 4 data bits, 3 parity bits located at positions  $2^0$ ,  $2^1$  and  $2^2$  from left are added to make a 7 bit code word which is then transmitted. The word format

P1 P2 D3 P4 D5 D6 D7

Where the D bits are data bits and P bits are parity bits.

P1 is to be set to 0 or 1 so that it establishes even parity over bits 1,3,5 and 7.

P2 is to be set to 0 or 1 so that it establishes even parity over bits 2,3,6 and 7.

P4 is to be set to 0 or 1 so that it establishes even parity over bits 4,5,6 and 7.

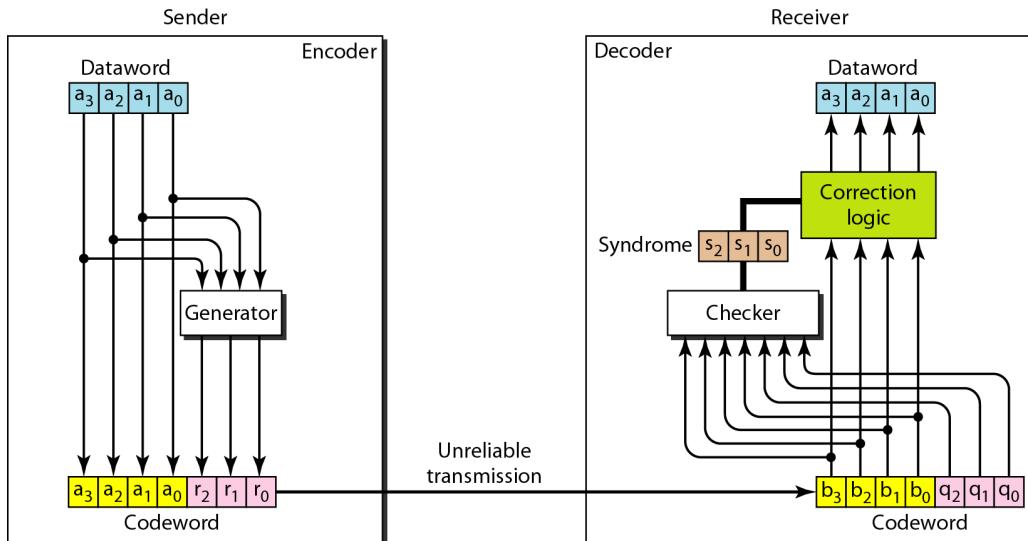


Figure 4.2.3 Coding and Decoding of Hamming

7 bit hamming code for the decimal digit coded in BCD and EXCESS 3 codes is shown in Table 4.2.4. The minimum distance of the 7 bit hamming code for the BCD code is 3 as observed Table 4.2.4.

Table 4.2.4 Hamming code for BCD and Excess-3

Decimal Digit	Hamming Code bits													
	For BCD				For EXCESS 3									
	P1	P2	D3	P4	D5	D6	D7	P1	P2	D3	P4	D5	D6	D7
0	0	0	0	0	0	0	0	1	0	0	0	0	1	1
1	1	1	0	1	0	0	1	1	0	0	1	1	0	0
2	0	1	0	1	0	1	1	0	1	0	0	1	0	1
3	1	0	0	0	0	1	1	1	1	0	0	1	1	0
4	1	0	0	1	1	0	0	0	0	0	1	1	1	1
5	0	1	0	0	1	0	1	1	1	0	0	0	0	0
6	1	1	0	0	1	1	0	0	0	1	1	0	0	1
7	0	0	0	1	1	1	1	1	0	1	1	0	1	0
8	1	1	1	0	0	0	0	0	1	1	0	0	1	1
9	0	0	1	1	0	0	1	0	1	1	1	1	0	0

At the receiving end, the message received in the hamming code is decoded to see if any errors have occurred. Bits 1,3,5, 7 and bits 2,3,6,7 and 4,5,6,7 are all checked for even parity. If they all check out, there is no error. If there is an error, the error bit can be located by forming a 3 bit binary number C3 C2 C1

$$C3 = D4 \oplus D5 \oplus D6 \oplus D7$$

$$C2 = P2 \oplus D3 \oplus D6 \oplus D7$$

$$C1 = P1 \oplus D3 \oplus D5 \oplus D7$$

After 3 parity checks , the error bit is corrected by complementing it.

Example: Encode data bits 1101 into 7 bit even parity hamming code.

P1	P2	D-3	P4	D5	D6	D7
1			1	0	1	

Bits 1,3,5,7 (i.e. P1 1 1 1) must be even parity. So P1 must be 1.

Bits 2,3,6,7 (i.e. P2 1 0 1) must be even parity. So P2 must be 0.

Bits 4,5,6,7 (i.e. P4 1 0 1) must be even parity. So P4 must be 0.

Therefore , the final code is 1 0 1 0 1 0 1

**Exercises:**

1. Consider the sequence of digits 10001001010110000011. Determine the number being represented in each of the following BCD coding schemes:
  - a. 8421 code
  - b. Excess-3 code
  - c. 2 4 2 1 code
2. The message coded in the 7 bit hamming code 1001001 is transmitted through a noisy channel. Decode the message assuming that at most a single error occurred.

## Chapter -5: Boolean Algebra and Logic Gates

### Module – 1 : Introduction to Boolean Algebra

#### Learning Outcomes:

At the end of this module, students will be able to:

1. State the laws of Boolean algebra.
2. Simplify Boolean expressions using Boolean algebraic theorems.

#### 5.1.1 Boolean Algebraic Theorems

George Boole in 1854 invented a new kind of algebra known as Boolean algebra. It is sometimes called switching algebra. Boolean algebra is the mathematical frame work on which logic design is based. It is used in synthesis and analysis of binary logical function.

#### History

The algebraic system known as Boolean algebra named after the mathematician George Boole. George Boole Invented multi-valued discrete algebra (1854) and E. V. Huntington developed its postulates and theorems (1904). Historically, the theory of switching networks (or systems) is credited to Claude Shannon, who applied mathematical logic to describe relay circuits (1938). Relays are controlled electromechanical switches and they have been replaced by electronic controlled switches called logic gates. A special case of Boolean Algebra known as Switching Algebra is a useful mathematical model for describing the combinational circuits.

#### a. Boolean Algebraic theorems:

These have been derived by using Boolean postulates. These laws are used to design and analyze logic circuit mathematically. Table 5.1.1 summarizes the Boolean theorems.

Table 5.1.1: Boolean theorems

Laws of union
Law1: $A + 0 = A$ Law2: $A + 1 = 1$
Laws of intersection
Law3: $A \cdot 1 = A$ Law4: $A \cdot 0 = 0$
Laws of tautology
Law5: $A + A = A$ Law6: $A \cdot A = A$
Laws of complements
Law7: $A + A' = 1$ Law8: $A \cdot A' = 0$
Law of double complement
Law9: $A'' = A$
Laws of commutation
Law10: $A + B = B + A$ Law11: $A \cdot B = B \cdot A$
Laws of association
Law12: $A + (B + C) = (A + B) + C$ Law 13: $A \cdot (B \cdot C) = (A \cdot B) \cdot C$
Laws of distribution
Law14: $A(B + C) = AB + AC$ Law15: $(A+B)(C+D) = AC + AD + BC + BD$
Laws of absorption
Law16: $A(A+B) = A$ Law17: $A + AB = A$ Law18: $A(A'+B) = AB$ Law19: $AB + B' = A+B'$ Law20: $AB' + B = A+B$

**Self test:**

1. Prove laws of absorption using appropriate Boolean laws

**b) Principle of Duality:** One more important property of Boolean algebra is called Duality principle. The Dual of any expression can be obtained easily by the following rules.

1. Change all 0's to 1's
2. Change all 1's to 0's
3. AND's (dot's) are replaced by OR's (plus)
4. OR's (plus) are replaced by AND's (dot's)

**Example 1:**

No.	Boolean Expression	Dual of the expression
1	$X + 0 = X$	$X \cdot 1 = X$
2	$X + Y = Y + X$	$X \cdot Y = Y \cdot X$
3	$X + 1 = 0$	$X \cdot 0 = 1$

**c) De Morgan's Theorems:** It is one of the important properties of Boolean algebra. It is extensively useful in simplifying complex Boolean expression.

**De Morgan's First Theorem:** It states that “the complement of product of variables is equal to sum of the complements of individual variable”.

i.e.  $\overline{AB} = \overline{A} + \overline{B}$ . The proof is given below.

A	B	$\overline{A}$	$\overline{B}$	$A \cdot B$	$\overline{AB}$	$\overline{A} + \overline{B}$
0	0	1	1	0	1	1
0	1	1	0	0	1	1
1	0	0	1	0	1	1
1	1	0	0	1	0	0

**De Morgan's Second Theorem:** It states that “the complement of sum of variables is equal to product of complement of two individual variables”.

i.e.  $\overline{A+B} = \overline{A} \cdot \overline{B}$ . The proof is shown using a truth table given below.

A	B	$\overline{A}$	$\overline{B}$	$A+B$	$\overline{A+B}$	$\overline{A} \cdot \overline{B}$
0	0	1	1	0	1	1
0	1	1	0	1	0	0
1	0	0	1	1	0	0
1	1	0	0	1	0	0

### 5.1.2 Simplification of Boolean algebraic expressions

It will be always desirable to simplify any Boolean Expression to its simplest form before converting to a logic circuit so that the circuit is constructed with minimum number of logic gates. It makes cost effective and more reliable due to lesser number of interconnections.

Example 2: Simplify

$$\text{i. } F = X'Y'Z + X'YZ \quad \text{ii) } F = X(X' + Y) \quad \text{iii) } F = B(A+C) + C$$

Solution:

$$\begin{aligned} \text{i. } F &= X'Y'Z + X'YZ \\ &= X'Z(Y' + Y) \quad (Y' + Y = 1) \\ &= X'Z \end{aligned}$$

$$\begin{aligned} \text{ii. } F &= X(X' + Y) \\ &= X \cdot X' + XY \quad (X \cdot X' = 0) \\ &= XY \end{aligned}$$

$$\begin{aligned} \text{iii. } F &= B(A+C) + C \\ &= BA + BC + C \\ &= BA + C(1+B) \quad (1+B = 1) \\ &= BA + C \end{aligned}$$

Example 3:

- i. Prove that  $Y(WZ' + WZ) + XY = Y(W+X)$
- ii. Prove that  $ABC + A'BC + AB'C + ABC' = AB + BC + CA$

i) L.H.S. =  $Y(WZ' + WZ) + XY$

$$= Y(WZ' + WZ) + XY$$

$$= YW(Z+Z') + XY \quad (Z+Z'=1)$$

$$= YW + XY$$

$$= Y(X+W)$$

ii) L.H.S =  $ABC + A'BC + AB'C + ABC'$

$$= BC(A+A') + AB'C + ABC'$$

$$= BC + AB'C + ABC'$$

$$= C(B+AB') + ABC' \quad (B+B'A = B+A)$$

$$= C(B+A) + ABC'$$

$$= BC + AC + ABC'$$

$$= BC + A(C+C'B)$$

$$= BC + A(C+B)$$

$$= BC + AC + AB$$

**Exercise:**

1. Simplify the following

i)  $F = XYZ + XYZ' + X'YZ \quad (\text{Ans : } XY + YZ)$

ii)  $F = XYZ + X'Y + XYZ' \quad (\text{Ans: } Y)$

2. Prove that  $(A+B)(A+C) = A+BC$

## Summary

### In this module we have learnt to:

1. List the Boolean algebraic theorems
2. Simplify the given Boolean expressions by applying Boolean Algebra theorems
3. The need for simplification of Boolean Expression.
4. The Dual of any expression can be obtained easily by the following rules - 1. Change all 0's to 1's, 2. Change all 1's to 0's, 3. AND's (dot's) are replaced by OR's (plus) and 4. OR's (plus) are replaced by AND's (dot's).
5. De Morgan's First Theorem: It states that "the complement of product of variables is equal to sum of the complements of individual variable".
6. De Morgan's Second Theorem: It states that "the complement of sum of variables is equal to product of complement of two individual variables".
7. Simplification of Boolean algebraic expressions makes cost effective and more reliable logic circuits due to lesser number of interconnections.

## Module 2: Logic Gates

Logic gate is a physical device or hardware used to implement a Boolean function. It performs the Boolean operation as per the design. Logic gates are primarily implemented using diodes or transistors operating as electronic switches. In this module, we will discuss basic logic functions, corresponding Gates and implementation of combinational circuits using logic gates.

### Learning Outcomes:

At the end of this module, students will be able to:

1. Describe basic logic gates and the concept of universal logic.
2. Build a logic circuit for the given Boolean expressions.
3. Write Boolean expressions for the given logic circuit.
4. Differentiate combinational and sequential circuits.

### 5.2.1 Logic gates and operation

Logic gate is an electronic circuit, which makes logic decisions. It is a digital circuit with one or more input signal and only one output signal. It produces one output level when some combinations of input levels are present and a different output level when other combinations of input levels are present. All input or output signals are either low voltage or high voltage. A digital circuit is referred to as logic gate for simple reason that, it can be analyzed based on Boolean algebra. To make logical decisions, three basic gates are used. They are OR, AND and NOT gate. These logic gates are building blocks, which are available in the form of an Integrated circuit (IC). The input and output of the binary variables for each gate can be represented in a tabular column called as truth table.

#### a) Basic gates

**i) OR Gate:** The OR gate performs logical additions commonly known as OR function. The OR gate has two or more inputs and only one output. The operation of OR gate is such that a HIGH (logic 1) on the output is produced when any of the input or both the inputs are HIGH. The output is LOW(logic 0) only when all the inputs are LOW. If A and B are the input variables of an OR gate and Y is its output, then  $Y=A+B$ . Similarly for more than two variables, the OR function can be expressed as  $Y=A+B+C$ .



Figure 5.2.1: Logical Symbol: Two Input OR gate

Table 5.2.1: Truth table for two input OR gate:

Input		Output
A	B	$Y = A + B$
0	0	0
0	1	1
1	0	1
1	1	1

Realization of OR gate using diodes: Two input OR gate using "diode-resistor" logic is shown in Figure 5.2.2, where X, Y are the inputs and F is the output.

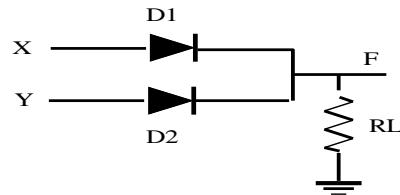


Figure 5.2.2: OR gate using Diode Resistor Logic

- If  $X = 0$  and  $Y = 0$ , then both the diodes D1 and D2 are reverse biased and thus both the diodes are in cut-off and thus F is LOW.
- If  $X = 0$  and  $Y = 1$ , D1 is reverse biased, thus does not conduct. But D2 is forward biased, thus conducts and thus pulling F to HIGH
- If  $X = 1$  and  $Y = 0$ , D2 is reverse biased, thus does not conduct. But D1 is forward biased, thus conducts and thus pulling F to HIGH.
- If  $X = 1$  and  $Y = 1$ , then both the diodes D1 and D2 are forward biased and thus both the diodes conduct and thus F is HIGH

**ii) AND Gate:** The AND gate performs logical multiplication, and commonly known as AND function. The AND gate has two or more inputs and a single output. The output of an AND gate is HIGH only when all the inputs are HIGH. Even if any one of the input is LOW, the output will be LOW. If A and B are input variables of an AND gate and Y is its output, then  $Y = A \cdot B$ .



Figure 5.2.3. Logical Symbol: Two input AND gate

Table 5.2.2: Truth table for two input AND gate

Input		Output
A	B	$Y = A \cdot B$
0	0	0
0	1	0
1	0	0
1	1	1

Realization of AND gate using diodes: Two input AND gate using "diode-resistor" logic is shown in Figure 5.2.4, where X, Y are inputs and F is the output.

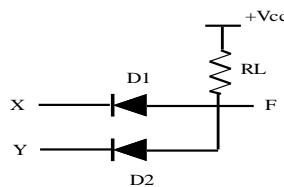


Figure 5.2.4 AND gate using Diode Resistor Logic

- If  $X = 0$  and  $Y = 0$ , then both the diodes D1 and D2 are forward biased and thus both the diodes conduct and pulls F to LOW.
- If  $X = 0$  and  $Y = 1$ , D1 is reverse biased, thus does not conduct. But D2 is forward biased, thus conducts and pulls F to LOW.
- If  $X = 1$  and  $Y = 0$ , D2 is reverse biased, thus does not conduct. But D1 is forward biased, thus conducts and pulls F to LOW.
- If  $X = 1$  and  $Y = 1$ , then both the diodes D1 and D2 are reverse biased and thus both the diodes are in cut-off and there is no drop in voltage at F. Thus F is HIGH.

**iii) NOT Gate (Inverter):** The NOT gate performs the basic logical function called inversion or complementation. The purpose of his gate is to convert one logic level into the opposite logic level. It has one input and one output. When a HIGH level is applied to an inverter, a LOW level appears at the output and vice-versa.



Figure 5.2.5: Logic symbol of NOT gate

Table 5.2.3: Truth table for NOT gate

Input	Output
A	$Y = \bar{A}$
0	1
1	0

Realization of NOT gate using Transistors: A NOT gate using a transistor is shown in Figure 5.2.6. ‘A’ represents the input and ‘F’ represents the output.

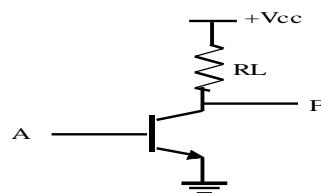


Figure 5.2.6: Realization of NOT gate using transistor.

- If  $A = 0$ , then the transistor is OFF thus pulling F to HIGH.
- If  $A = 1$ , then the transistor is ON thus driving F to HIGH.

### b) Other type of gates

i) **NAND Gate:** The output of a NAND gate is LOW only when all inputs are HIGH and output of the NAND is HIGH if one or both inputs are LOW.



Figure 5.2.7: Logical Symbol: Two input NAND gate

Table 5.2.4 : Truth table of NAND gate

Input		Output
A	B	$Y = \bar{AB}$
0	0	1
0	1	1
1	0	1
1	1	0

**ii) NOR Gate:** The output of the NOR gate is HIGH only when all the inputs are LOW.

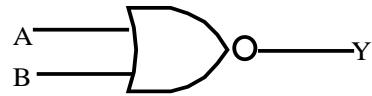


Figure 5.2.8: Logical Symbol: Two input NOR Gate

Table 5.2.5: Truth table for NOR gate

Input		Output
A	B	$Y = \overline{A + B}$
0	0	1
0	1	0
1	0	0
1	1	0

**iii) XOR Gate or Exclusive OR gate:** In this gate output is HIGH only when any one of the input is HIGH. The circuit is also called as inequality comparator, because it produces output when two inputs are different.

$$Y = A \oplus B = A \bar{B} + \bar{A} B$$



Figure 5.2.9: Logical Symbol: Two input XOR Gate

Table 5.2.6: Truth table for an XOR gate

Input		Output
A	B	$Y = A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

**iv) XNOR Gate or Exclusive NOR Gate:** XNOR gate is a gate with two or more inputs and one output. XNOR operation is complementary of XOR operation. i.e. The output of XNOR gate is High, when all the inputs are identical; otherwise it is low.

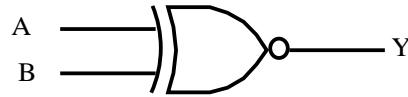


Figure 5.2.10: Logical Symbol of Two input XNOR Gate

Table 5.2.7 Truth table for XNOR gate

Input		Output
A	B	$Y = \overline{A} \overline{B} + AB$
0	0	1
0	1	0
1	0	0
1	1	1

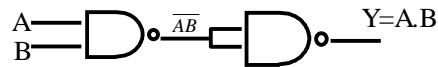
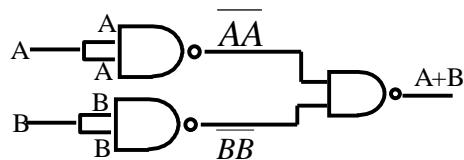
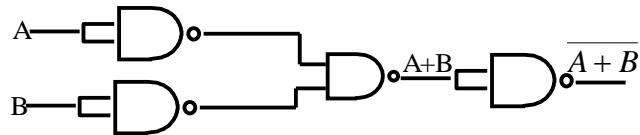
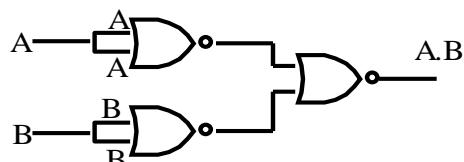
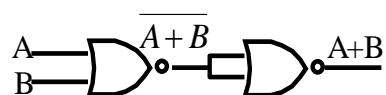
**Self test:**

How many two input AND and OR gates are required to realize the following expressions in addition to inverters?

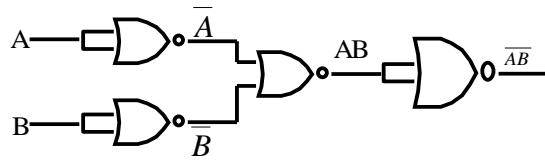
- a)  $F_1 = ABC + AB'CD + EF' + AD$
- b)  $F_2 = A(B+C+D')(B'+C+E')(A+B'+C+E)$
- c) What is the uniqueness of XOR logic? Explain briefly.
- d) List two typical applications of XOR logic.

### 5.2.2. Concept of Universal Logic

NAND and NOR gates are called Universal gates or Universal building blocks, because both can be used to implement any gate like AND, OR and NOT gates or any combination of these basic gates.

*a) NAND gate as Universal gate:***1.** NOT operation:**2.** AND operation:**3.** OR operation:**4.** NOR operation:*b )NOR gate as Universal gate:***1.** NOT operation:**2.** AND operation:**3.** OR operation:

**4. NAND operation:**



**Exercise:**

1. Draw the logic circuit for the Boolean expression.  $\underline{Y = BC + A' C + AB'C}$ .
2. Show that  $AB + (A+B)$  is equivalent to  $A \text{ O } B$ . Also construct the corresponding logic diagrams.
3. The most suitable gate to check whether the number of 1's in a digital word is even or odd is
  - a) X-OR
  - b) NAND
  - c) NOR
  - d) AND, OR and NOT
4. Realize NOR and NAND gate using discrete components.
5. Realize NOR and NAND gate using Basic gates.

### 5.2.3 Classification of digital circuits

Digital circuits are circuits constructed using digital gates and operate as per digital logic. Basically digital circuits can be classified into two types.

- Combinational Circuits
- Sequential Circuits

*a) Combinational Digital Circuits:*

Logic circuits whose output at any instant of time is entirely dependent upon the input signals present at that time are known as combinational digital circuits as shown in Figure 5.2.11. In particular, the output of the combinational circuit doesn't depend upon any past input or output. The circuit doesn't possess any memory. The output signals of combinational circuits are not fed back to any other part of the circuit.

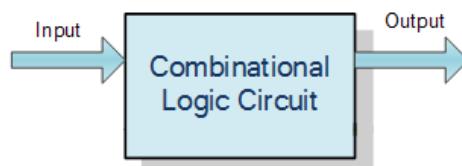


Figure 5.2.11 Combinational Circuit

Combinational circuits are generally faster, since the operation need not have to be performed in sequences. Combinational circuits can be constructed using Sum of Products (SOP) or

4. Sum of products (SOP): Sum of product is the logical expression in which OR of multiple product terms are present. Each product term is the logical AND of literals.  $Y+XY'+XYZ$  is an example of SOP.
5. Product of Sums (POS): Product of sums is the logical expression in which AND of multiple sum terms are present. Each sum term is the logical OR of literals.  $(X+Y')(X'Y+Z)(X+Y+Z)$  is an example of POS.
6. Minterm: It is a special type of product (AND) term. It is a product term which contains all the input variables that make up a Boolean expression.
7. Maxterm: It is a special type of sum (OR) term. A maxterm is a sum term that contains all the input variables that make up a Boolean expression.
8. Canonical form: Canonical is defined as “conforming to a general rule”. The rule for Boolean logic is that each term used in a boolean expression must contain all the variables.
9. Canonical Sum of Products: A canonical SOP is a complete set of minterms that defines when an output variable is a logical ‘1’. Each minterm corresponds to the row in the truth table when the output function is 1.
10. Canonical Product of Sums: A canonical POS is a complete set of maxterms that defines when an output variable is a logical ‘0’. Each maxterm corresponds to the row in the truth table when the output function is 0.

Boolean algebra is used to simplify the Boolean expressions, thus reducing redundancy and designing low cost logic circuits. Any logic expression can be implemented by logic gates.

**b) Examples of combinational circuits:**

**i) Half Adder:** An electronic combinational circuit which performs the arithmetic addition of two binary bits is called Half Adder. In the half adder circuit, there are two inputs, addend and augend and two outputs are Sum and Carry. The logic symbol of half adder is shown in Figure 5.2.13 and the truth table of half adder is given in table 5.2.8.

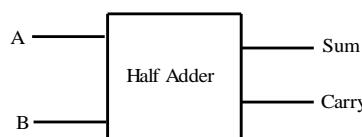


Figure 5.2.13: Half Adder Logic Symbol.

Table 5.2.8: Truth table of half adder.

Input		Output	
Augend	Addend	Sum	Carry
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

The expression for sum and carry is given in expressions 5.2.1 and 5.2.2 respectively.

$$\text{Sum} = \bar{A}B + A\bar{B} = A \oplus B \quad (5.2.1)$$

$$\text{Carry} = A \cdot B \quad (5.2.2)$$

The circuit for Half Adder using Basic Gates is shown in Figure 5.2.14

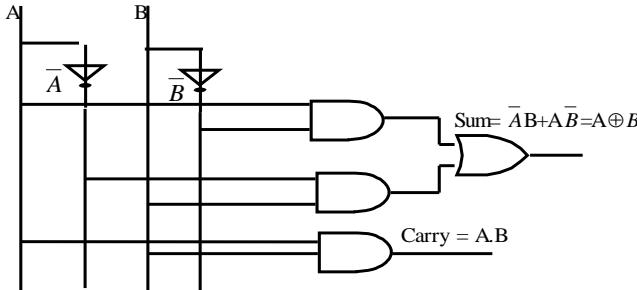


Figure 5.2.14 Half adder using basic gates.

The circuit for Half Adder using XOR gate is shown in Figure 5.2.15.

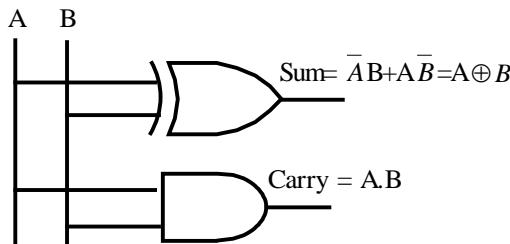


Figure 5.2.15: Half adder using XOR gate.

**ii) Full Adder:** Full adder is a combinational circuit that performs the arithmetic sum of three input bits. It consists of three inputs and two outputs. Two of the inputs are variables, denoted by A and B, which represent the two significant bit to be added. The third input ‘Cin’ represents carry from the previous lower significant position. The logical symbol of full adder is shown in Figure 5.2.16 and the truth table is given in table 5.2.9.

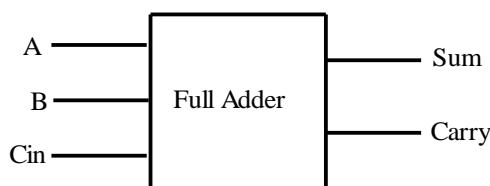


Figure 5.2.16: Full Adder logic symbol.

Table 5.2.9: Truth table for Full Adder

Input			Output	
A	B	Cin	Sum	Carry
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

The expression for sum and carry of full adder is given in expressions 5.2.3 and 5.2.4 respectively.

$$\begin{aligned}
 \text{Sum} &= \overline{A} \overline{B} \text{Cin} + \overline{A} B \overline{\text{Cin}} + A \overline{B} \overline{\text{Cin}} + ABC\text{Cin} \\
 &= \overline{A} [\overline{B} \text{Cin} + B \overline{\text{Cin}}] + A[\overline{B} \overline{\text{Cin}} + BC\text{Cin}] \\
 &= \overline{A} [B \oplus \text{Cin}] + A[\overline{B} \oplus \overline{\text{Cin}}] \\
 &= A \oplus B \oplus \text{Cin}
 \end{aligned} \tag{5.2.3}$$

$$\begin{aligned}
 \text{Carry} &= \overline{A} BC\text{Cin} + A \overline{B} \text{Cin} + AB \overline{\text{Cin}} + ABC\text{Cin} = \overline{A} BC\text{Cin} + A \overline{B} \text{Cin} + AB(\overline{\text{Cin}} + \text{Cin}) \\
 &= \overline{A} BC\text{Cin} + A \overline{B} \text{Cin} + AB = \overline{A} BC\text{Cin} + A(\overline{B} \text{Cin} + B) = \overline{A} BC\text{Cin} + AB + AC\text{Cin} \\
 &= B(\overline{A} \text{Cin} + A) + AC\text{Cin} = B(A + \text{Cin}) + AC\text{Cin} \\
 &= AB + BC\text{Cin} + AC\text{Cin}
 \end{aligned} \tag{5.2.4}$$

The logic circuit of full adder is shown in Figure 5.2.17.

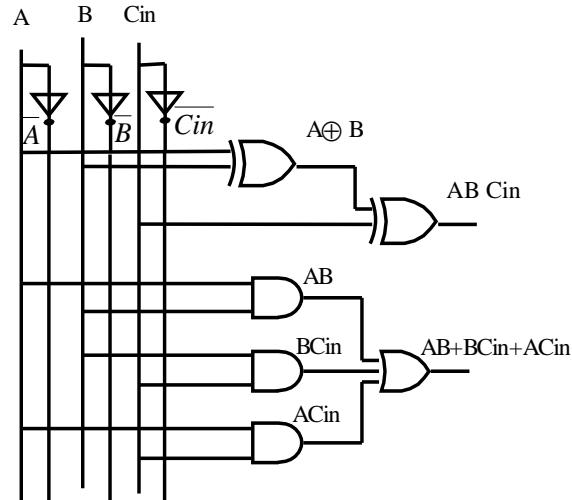
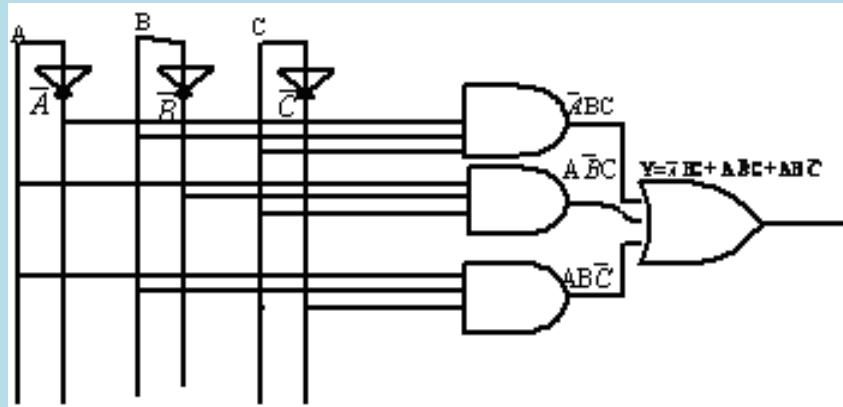


Figure 5.2.17: Full Adder circuit

Ex1: Draw the logic circuit for the Boolean expression.  $Y = \bar{A}BC + A\bar{B}C + ABC'$



### Summary

1. Logic gates are fundamental building blocks of digital systems
2. The basic set of logic gates are AND, OR and NOT and this set is called Universal set.
3. NAND and NOR are called Universal gates.
4. Inputs and outputs of logic gates can occur in two levels. These two levels are termed as HIGH and LOW, or TRUE and FALSE, or ON and OFF, or simply 1 or 0.
5. Logic circuits whose output at any instant of time is entirely dependent upon the input signals present at that time are known as combinational digital circuits.
6. Logic circuits whose output at any instant of time depend, not only on the present input but also on the past outputs are called Sequential Circuits.

**Exercise:**

1. Implement Full Subtractor using Basic gates.
2. Implement full adder using two half adders and an OR gate.
3. Simplify and realize using only NAND gates
4.  $XYZ + XYZ + YZ + \bar{Z}$ .
5.  $(A + \bar{B} + C)(\bar{A} + B + \bar{C})(A + \bar{B})$
6. Simplify and realize using only NOR gates
  - a.  $Y = A \bar{B} \bar{C} + \bar{A} \bar{B} \bar{C} + \bar{B} \bar{C} + A \bar{C}$

### Module 3: Simplification of Boolean functions using K-map

Maurice Karnaugh, a telecommunications engineer, developed the Karnaugh map at Bell Labs in 1953 while designing digital logic based telephone switching circuits. The Karnaugh map, also known as the K-map, is a graphical method to simplify Boolean algebraic expressions. Karnaugh maps reduce a logic function more quickly and easily compared to Boolean algebra and thus reduces cost of the circuit in terms of inputs and gates.

The required Boolean results are transferred from a truth table onto a two-dimensional grid. The cells are addressed by Gray code, and each cell represents one combination of input conditions ( Minterm or Maxterm), while each cell value represents the corresponding output of the function. Optimal groups of 1s or 0s are identified, which represent the terms of a canonical form of the logic in the original truth table. These terms can be used to write a minimal Boolean expression representing the required logic.

#### Learning Outcomes:

At the end of this module, students will be able to:

1. Describe standard form of Boolean expressions.
2. Apply K-map method for simplification of Boolean expressions.

#### 5.3.1 Standard forms of Boolean expressions

Sum of product (SOP) is the logical expression in which OR of multiple product terms are present. Each product term is the logical AND of literals.  $Y+XY'+XYZ$  is an example of SOP. In SOP all the individual terms do not involve all literals. If each term in SOP expression contains all the literals then that SOP is known as standard or canonical SOP. Every individual term in the standard SOP is a Minterm.

Product of sums (POS) is the logical expression in which AND of multiple sum terms are present. Each sum term is the logical OR of literals.  $(X+Y')((X'Y+Z)(X+Y+Z))$  is an example of POS. In POS all the individual terms do not involve all literals. If each term in POS expression contains all the literals then that POS is known as standard or canonical POS. Every individual term in the standard POS is a Maxterm.

##### a) Steps to convert SOP to canonical SOP:

1. Find the missing literal in each product term.
2. AND each product term having missing literals with terms by OR ing the literal and its complement.
3. Expand the terms and reduce the expression by removing repeated terms.

**Ex1:**  $F(A,B,C) = AC + AB + BC$ 

In first term B is missing literal, in second term C is missing literal and in third term A is missing literal.

$$F = AC(B+B') + AB(C+C') + BC(A+A')$$

$$F = ACB + ACB' + ABC + ABC' + ABC + A'BC$$

$$F = ABC + A'BC + AB'C + ABC'$$

**b) Steps to convert POS to canonical POS:**

1. Find the missing literal in each sum term.
2. OR each sum term having missing literals with terms by AND ing the literal and its complement.
3. Expand the terms and reduce the expression by removing repeated terms.

**Ex2:**  $F(A, B, C) = (A+B)(B+C)(C+A)$ 

In first term C is missing literal, second term A is missing literal and third term B is missing literal.

$$F = (A+B+C.C')(B+C+A.A')(C+A+B.B')$$

$$F = (A+B+C)(A+B'+C)(A+B+C)(A'+B+C)(A+B+C)(A+B'+C)$$

$$[A+BC=(A+B)(A+C) \dots \text{distributive property} \& X.X=X]$$

$$F = (A+B+C)(A'+B+C)(A+B'+C)(A+B+C')$$

The canonical SOP and Canonical POS representation of a Boolean function are complementary. The variables which exists in SOP representation do not appear in POS representation. Similarly all the variables which exist in POS representation do not appear in SOP representation. To convert from one canonical form to other canonical form, the symbols  $\Sigma$  and  $\Pi$  will be interchanged and the list of variables will be present in new form which is actually missing from original form.

**Ex3:** Determine the Boolean function from the truth table in terms of minterms. Also give canonical POS form of the expression.

Inputs			Output
A	B	C	Y
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	0

### 5.3.2 Introduction to Karnaugh Map (K-map)

K-map is a matrix of squares and each square represents a minterm or maxterm of a Boolean expression. It is used to simplify the Boolean expression. Each map lists  $2^n$  product terms that can be formed from 'n' variables, each in a different square. A product term in 'n' variables is called a minterm.

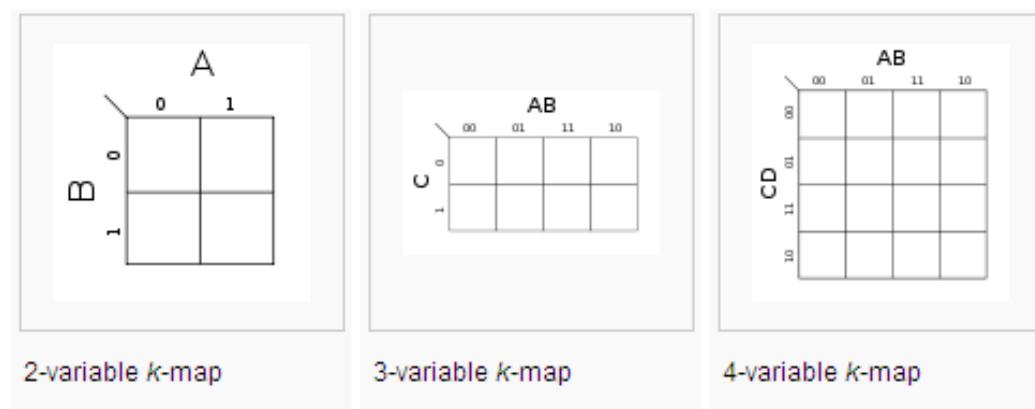


Figure 5.3.1: Variable sized K Maps.

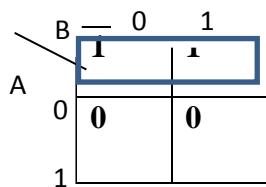
**Size of K-map:** The size of K-map with  $n$  Boolean variables is determined by  $2^n$ . The size of the group within a K-map with  $n$  Boolean variables and  $k$  number of terms in the resulting Boolean expression is determined by  $2^{nk}$ . Common sized maps are of 2 variables which is a

$2 \times 2$  map, 3 variables which is a  $2 \times 4$  map and 4 variables which is a  $4 \times 4$  map. These maps are shown in Figure 5.3.1.

### 5.3.3 Simplification of Boolean expressions using K-map

The K-Map method may theoretically be applied for the simplification of any Boolean expression regardless of its number of variables, but is most often used when there are fewer than six variables because K-Maps of expressions with more than six variables are complex and tedious to simplify. Each variable contributes two possibilities: the complemented and un-complemented forms. It therefore organizes all possibilities of the system. The variables are arranged in Gray code in which only one variable changes between two adjacent grid boxes.

#### a) Two variable K-map

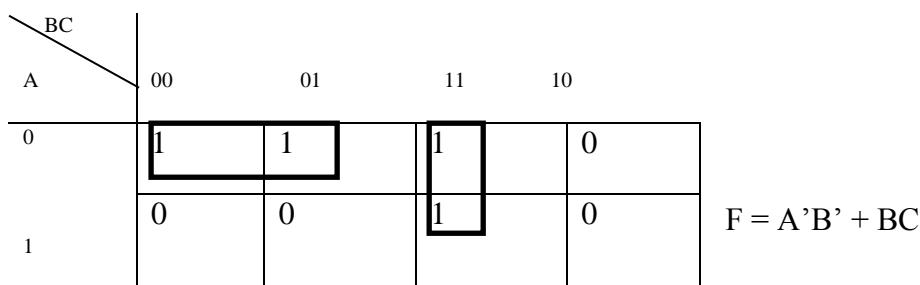


$$f(A,B) = A'$$

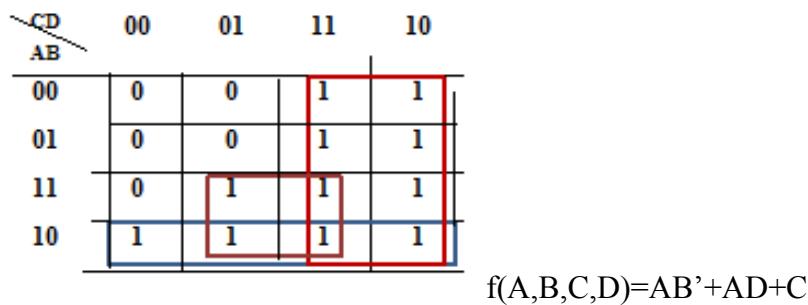
#### b) Three variable K-map

Simplify the following Boolean expression using K Map.

$$F(A,B,C) = \Sigma m(0,1,3,7)$$



#### c) Four variable K-map



Ex: For the following truth table construct the K Map

Sl.No.	A	B	C	D	f
0	0	0	0	0	0
1	0	0	0	1	0
2	0	0	1	0	0
3	0	0	1	1	0
4	0	1	0	0	0
5	0	1	0	1	0
6	0	1	1	0	1
7	0	1	1	1	0
8	1	0	0	0	1
9	1	0	0	1	1
10	1	0	1	0	1
11	1	0	1	1	1
12	1	1	0	0	1
13	1	1	0	1	1
14	1	1	1	0	1
15	1	1	1	1	0

K-map generally becomes more cluttered and hard to interpret when the number of variables increase. For expressions with larger numbers of variables, we have other algorithms. One such algorithm is called Quine McCluskey algorithm and is suitable for automation.

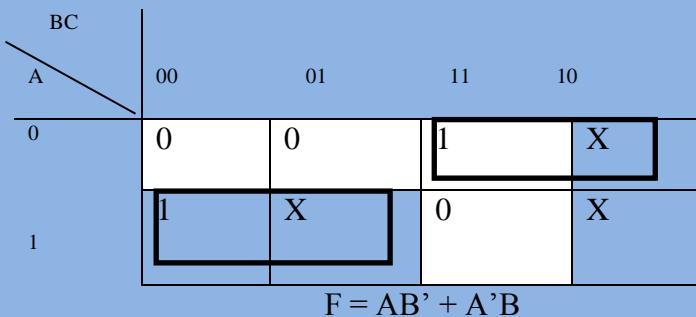
**Note:**

- When moving horizontally or vertically, only one variable changes between adjacent squares. This property of K-map is unique and accounts for its unusual numbering system.
- In all K-maps, left and right edges are a common edge, while top and bottom edges are also common. Thus top and bottom rows are adjacent, as are the left and right columns.

So far, the expressions considered have been completely specified for every combination of the input variables, that is, each minterm (maxterm) has been specified as a 1 or a 0. For certain input combinations, the value of the output is unspecified either because the input combinations are invalid or because the precise value of the output is of no consequence. The combinations for which the values of the expression are not specified are called don't care combinations. The don't care terms are denoted by 'd' or 'X'. During the process of design using SOP map, each don't care is treated as a 1 if it is helpful in map reduction; otherwise it is treated as a 0 and left out. During the process of design using a POS map, each don't care is treated as 0 if it is useful in reduction, otherwise it is treated as a 1 and left out.

**Ex1:** Simplify the following Boolean expression using K Map.

$$F(A,B,C) = \Sigma m(3, 4) + d(2,5,6)$$



### Summary

1. Any Boolean expression can be expressed in a standard or canonical or expanded sum (OR) of products (AND) form –SOP form or in a standard or canonical or expanded product (AND) of sums (OR) form – POS form.
2. The K-map is a chart or a graph, composed of an arrangement of adjacent cells, each representing a particular combination of variables in sum of product form.
3. The K-map is useful in Boolean expression simplification and hence further reduces the cost for logic realization.
4. It is a means of showing the relationship between the logic inputs and desired output.
5. K-map is limited to 6 variables.

**Exercise:**

1. Consider the truth table of a function. Transfer the outputs to the K map and write the Boolean expression.

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

2. Simplify the following Boolean expressions using K maps.
  - (a)  $F = \Sigma m(0,2,4,6)$
  - (b)  $F = \Sigma m(0,2,4,6) + d(5,7)$ .

## Chapter -6: SEQUENTIAL CIRCUITS

Sequential circuits are digital circuits whose output at a given time is dependent on the input at that time and also on the earlier inputs. Thus sequential circuits always have a memory associated with them. The operation of sequential circuits is expressed in terms of their states. The basic memory element in sequential circuits is the flip-flop. A flip-flop can store a 1 or a 0 and hence is called one bit memory.

### Module – 1 : Flip-Flops and its Applications

#### Learning Outcomes:

At the end of this module, students will be able to:

1. Differentiate latches and Flip flops.
2. Draw the circuit of SR, D, JK and D flip-flop using NAND gates and explain its working principle with its truth table.
3. Design ripple up/down counter using flip flop
4. Explain the working principle of SISO, SIPO shift registers.

#### 6.1.1 Introduction to Flip-Flops

It is an electronic circuit or device which is used to store data in binary form. A flip-flop is a sequential circuit used as a one bit memory element. A flip-flop circuit can be constructed from NAND gates or NOR gates. It has two stable outputs Q and Q' and hence it is also known as Bi-stable circuit.

#### History of Flip-Flop

The first electronic flip-flop was invented in 1918 by William Eccles and F. W. Jordan. It was initially called the *Eccles–Jordan trigger circuit* and consisted of two active elements (vacuum tubes) now it is realized using transistors.

A “clock” is derived using a special circuit that sends electrical pulses, called clock in the context of digital applications as shown in Figure 6.1.1. Each pulse has a precise width and there is a precise interval between pulses – known as clock cycle time. Figure 6.1.2 shows various parts of the clock signal for which the circuit will respond.



Figure 6.1.1 Clock Signal

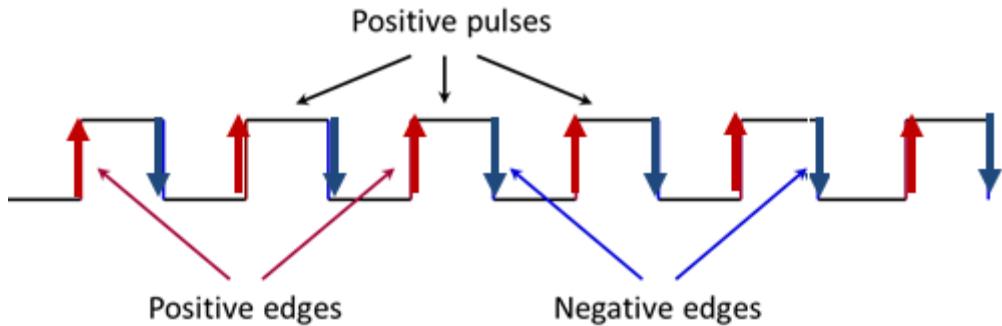


Figure 6.1.2 Edges and levels of a Clock Signal

### b) Flip-Flop and Latch

The word latch is mainly used for storage elements, while clocked devices are described as flip-flops. A latch is level-sensitive, whereas a flip-flop is edge-sensitive. Based on the triggering condition, Flip flops are divided into positive edge-triggered (active when transition of clock is 0 to 1) and negative edge-triggered (active when transition of clock is 1 to 0).

### c) SR Flip-Flop:

The SR flip-flop can be considered as one of the most basic sequential logic circuit possible. This simple flip-flop is basically a one-bit memory bi-stable device that has two inputs, one which will “SET” the device (output  $Q = "1"$ ), and is labeled S and another which will “RESET” the device (output  $Q = "0"$ ), labeled R.

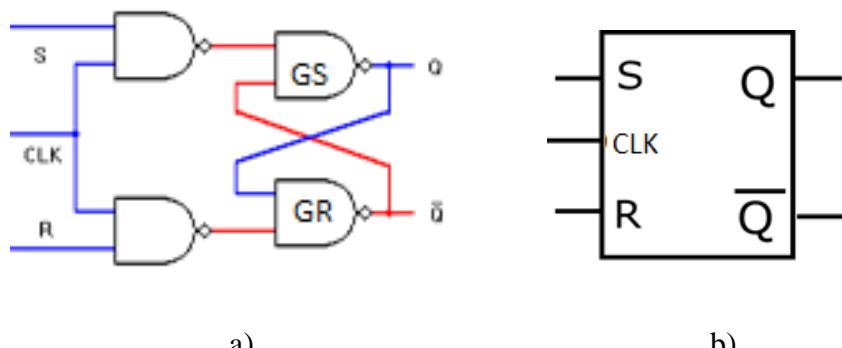


Figure 6.1.3 (a): Logic diagram of Clocked SR Flip flop (b) Logical Symbol

A basic NAND gate SR flip-flop circuit provides feedback from both of its outputs back to its opposing inputs and is commonly used in memory circuits to store a single data bit as shown in Figure 6.1.3 (a). The SR flip-flop actually has three inputs, Set, Reset and its current output Q relating to it's current state or history. The term “Flip-flop” relates to the actual operation of the device, as it can be “flipped” into one logic state (SET) or “flopped” back into the opposing logic state (RESET). Figure 6.1.3 (b) shows symbolic representation of SR Flip-Flop.

The operation of S-R Flip Flop can be described as follows::

**When Clock is High:**

- When S input is made '1' and R input is made '0', the Q output takes the state '1'. This is known as SET condition.
- When S input is made '0' and R input is made '1', the Q output takes the state '0'. This is known as RESET condition.
- When  $S = R = '0'$ , the SR flip flop exhibits the memory. i.e.  $Q_{n+1} = Q_n$ ; Holds the state.
- $S = R = 1$  is the Not allowed in S-R Flip flop.

**When Clock is Low:**

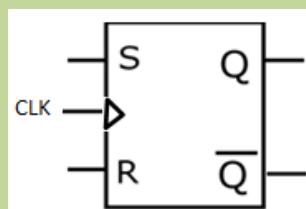
- Irrespective of S and R inputs, the SR flip flop exhibits the memory. i.e.  $Q_{n+1} = Q_n$  : No change in the state.

The function of the circuit is described by the table 6.1.1. This table is called Truth Table.  $Q_{(n+1)}$  represents the next state of the output and  $Q_n$  corresponds to the previous state.

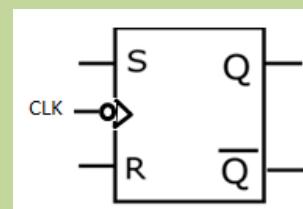
Table 6.1.1: Truth Table of SR Flip-Flop

<b>CLK</b>	<b>S</b>	<b>R</b>	<b>Q(n+1)</b>	<b>Mode</b>
1	0	0	$Q_n$	Previous Output
1	0	1	0	Reset
1	1	0	1	Set
1	1	1	Invalid	Invalid
0	X	X	$Q_n$	Previous Output

Note: Practical available Flip flops are edge triggered, that means flip flop responds for clock edges either positive edge or negative edge. Following are the symbolic representation of positive edge and negative edge triggered flip flops. Observe the 'not' (bubble) symbol at the clock of negative edge triggered flip flop.



a) Positive edge triggered Flip flop



b) Negative edge triggered Flip flop

**d) D Flip-Flop:**

The D flip-flop shown in Figure 6.1.4 (a) is a modification of the clocked SR flip-flop. The D input goes directly into the S input and the complement of the D input goes to the R input. The D input is sampled during the occurrence of a clock pulse. If it is 1, the flip-flop is switched to the set state (unless it was already set). If it is 0, the flip-flop switches to the clear state. The R and S inputs will be always in complimentary form, thus preventing  $R = S = 1$  condition from occurring. When the clock (CLK) is low, previous data is stored, i.e., it exhibits memory as shown in table 6.1.2.

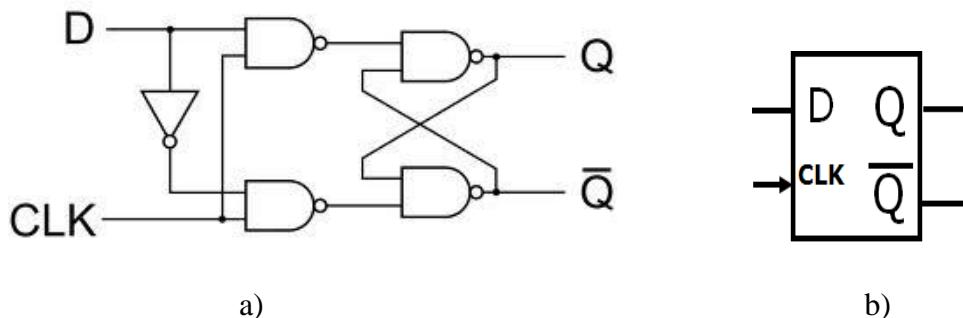


Figure 6.1.4: (a) Logic diagram of D Flip flop (b) Logic Symbol

Table 6.1.2: Truth Table

En	D	$Q(n+1)$	Mode
1	0	0	Reset
1	1	1	Set
0	X	$Q_n$	Previous Output

**e) JK Flip-Flop:**

A JK flip-flop is a refinement of the SR flip-flop, in which the indeterminate state of the SR type is defined in the JK type as shown in figure 6.1.5 (a). Inputs J and K behave like inputs S and R to set and clear the flip-flop (note that in a JK flip-flop, the letter J is for set and the letter K is for clear). When logic 1 inputs are applied to both J and K simultaneously, the flip-flop switches to its complement state, i.e., if  $Q=1$ , it switches to  $Q=0$  and vice versa.

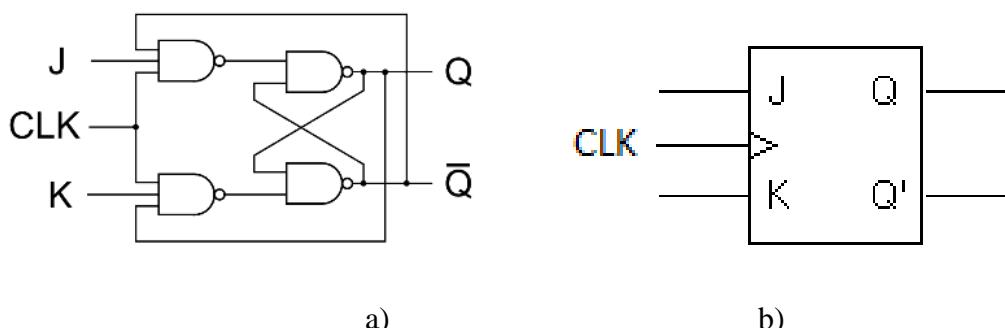


Figure 6.1.5: (a) Logic Diagram of JK Flip-flop (b) Logic symbol

JK flip flop is one of the most versatile flip flop. The operation of JK flip flop can be explained as follows:

#### When Clock is High:

- $J = K = 0$  causes memory condition.
- $J = 0 \& K = 1$  causes reset condition.
- $J = 1 \& K = 0$  causes set condition.
- $J = K = 1$  causes output to toggle. i.e.  $Q_{n+1} = Q_n'$

#### When Clock is Low:

- Irrespective of J & K inputs, the J-K flip flop exhibits the memory condition.

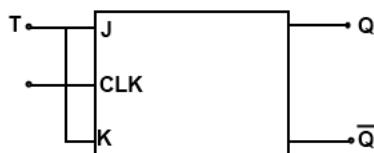
All these characteristics are summarized in Table 6.1.3

Table 6.1.3 Truth Table of JK flip-flop

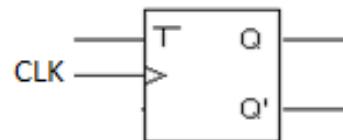
<b>CLK</b>	<b>J</b>	<b>K</b>	<b>Q(n+1)</b>	<b>Mode</b>
1	0	0	$Q_n$	Previous Output
1	0	1	1	Reset
1	1	0	0	Set
1	1	1	$Q_n'$	Toggle
0	X	X	$Q_n$	Previous Output

#### f) T Flip-flop:

Figure 6.1.6 shows the logic circuit of T flip-flop in which J and K inputs of a JK flip flop are combined and taken as a single input T. When  $T = 0$ , output of T flip-flop will remain as it was previously. When  $T = 1$ , output of T flip-flop will be complement of its previous output and hence this circuit is known as toggle circuit. The excitation table is as shown in Table 6.1.4.



(a)



(b)

Figure 6.1.6: (a) Logic Diagram of T Flip-flop (b) Logic symbol of T Flip-flop

Table 6.1.4: Truth table of T flip-flop

<b>CLK</b>	<b>T</b>	<b>Q(n+1)</b>	<b>Mode</b>
1	0	Q <sub>n</sub>	No change
1	1	Q <sub>n'</sub>	Toggle
0	X	Q <sub>n</sub>	Previous Output

**Self test:**

1. Implement clocked SR flip flop and JK flip flop using only NOR gates and write the truth table.
2. Convert SR flip flop to D flip flop and T flip flop. Show using appropriate logic diagram.
3. Compare a) SR and JK flip flop b) T and D flip-flop.
4. Draw the timing diagram for a) SR Flip flop and b) JK Flip flop

**6.1.2 Counters:**

A digital counter is a set of flip flops whose states change in response to pulses applied at the input to the counter. In binary counter, the flip flops are interconnected such that their combined state at any time is the binary equivalent of the total number of pulses that have occurred up to that time. Thus a counter is used to count pulses. Each of the counts of the counter is called state of the counter. The number of states through which the counter passes before returning to the starting state is called modulus of the counter. In general an n-bit counter will have 'n' flip flops and  $2^n$  states and divides the input frequency by  $2^n$ . Hence it is a divide by  $2^n$  counter.

**a) Classification of counters****i) Synchronous and Asynchronous counters :**

Counters may be synchronous or asynchronous counters as shown in Figure 6.1.7. In Asynchronous counters, flip flops are not triggered simultaneously. Synchronous counters are clocked such that each flip flop in the counter is triggered at the same time. This is accomplished by connecting the same clock line to each stage of the counter. Synchronous counters are faster than asynchronous counters because the propagation delay involved is less.

**ii) Up counter or Down Counter:**

A counter may be an up counter or a down counter. An up counter is a counter which counts in the upward direction i.e 0, 1, 2, 3,..., N. A down counter is a counter which counts in the downward direction i.e N, N-1, N-2,...,1, 0.

Note:

Practical available Flip flops are edge triggered, so counters can be realized using positive edge triggered or negative edge triggered flip flops.

We can realize other than  $2^n$  counter by appropriately setting the asynchronous inputs of flip flop using logic gates. For example Decade counter counts 10 states of a counter which is not a  $2^n$

**b) Applications of counter:**

Counters are used

1. As Frequency divider,
2. To perform the timing function as in digital watches
3. To create time delays
4. To produce non-sequential binary counts
5. To generate pulse trains
6. To act as frequency counters.

**c) Ripple Counters:**

Asynchronous counters are also called as ripple counters. In ripple counters the flip flops within the counter are not made to change the states at exactly the same time. In this counter, JK flip flops or T flip flops are used with J and K inputs or T input of the flip flops connected to 1. This makes the flip flop output to toggle when a clock pulse is applied. It is also called as serial or series counters.

**Example 3:**

Design Two bit ripple up counter using negative edge triggered JK flip flops:

**Steps:**

1. Select 2 JK flip flops (number of flip flops depends upon number of bits to count).
2. Connect JK inputs to high
3. Apply the –ve edge clock pulse to first JK flip flop
4. Write the truth table of 2 bit up counter as shown where Q1 and Q2 are the outputs of 2 JK flip flops.

Clk pulse	Q2	Q1(LSB)
1	0	0
2	0	1
3	1	0
4	1	1

5. Observe when higher bit (Q2) is changing. In this case, Q2 is changing when Q1 is changing from logic 1 to logic 0 (At clock pulse 3). Since we require –ve edge counter which responds when clock is changing from 1 to 0 is required, connect Q1 as the clock for the next flip flop as shown in Figure 6.1.7 (a)

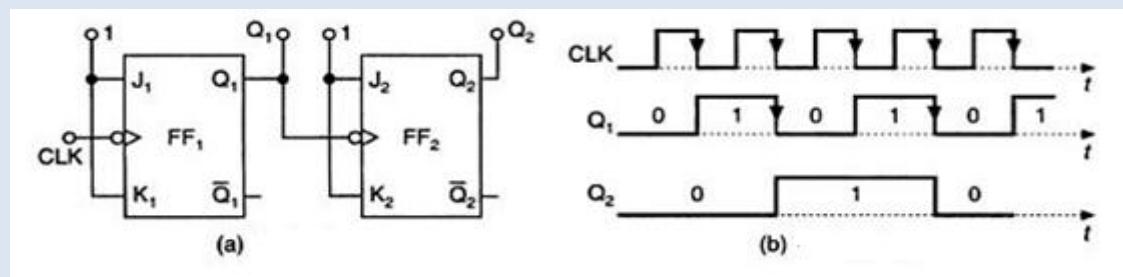


Figure 6.1.7 (a) Asynchronous 2 bit up counter using -ve edge triggered flip flops (b) Timing diagram

**Working principle:** For the clock pulse applied to FF1, the output of FF1 toggles. For the Q1 pulse applied to FF2, the output of FF2 toggles. Thus the 2 bit up counter counts in the order 0, 1, 2, 3, 0, 1, ....Figure 6.1.7 shows a 2 bit ripple up counter using negative edge triggered JK flip flops and its timing diagram.

#### Example 4:

Design Two bit ripple up counter using positive edge triggered JK flip flops:

Steps:

Steps 1and 2 are same as example 3

3 Apply the +ve edge clock pulse to first JK flip flop (Observe the Clk signal in Figure 6.1.7

4 Write the truth table of 2 bit up counter as in example 6.3

5 Observe when higher bit (Q2) is changing. In this case, Q2 is changing when Q1 is changing from logic 1 to logic 0. Since we require +ve edge counter which responds when clock is changing from 0 to 1 is required, connect Q1' (Q1 complement) as the clock for the

next flip flop (because when  $Q_1$  is changing 1 to 0  $Q_1'$  will change 0 to 1 which is the positive edge of the clock) as shown in Figure 6.1.8

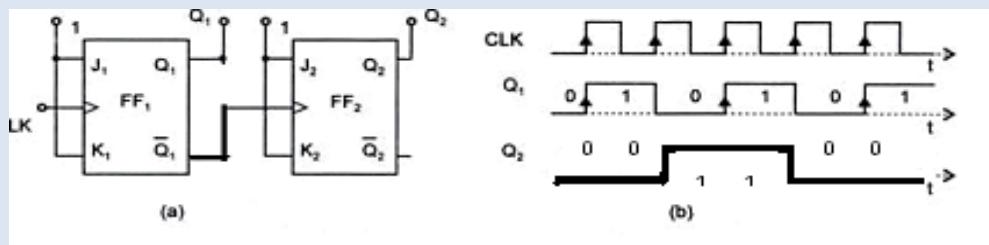


Figure 6.1.8 (a) Asynchronous 2 bit up counter using +ve edge triggered flip flops. (b) Timing diagram.

**Working principle:** The output  $Q_1'$  of FF1 is connected to the clock input of FF2. For the clock pulse applied to FF1, the output of FF1 toggles. For the  $Q_1'$  pulse applied to FF2, the output of FF2 toggles. Thus the 2 bit up counter counts in the order 0, 1, 2, 3, 0, 1,

#### Example 5:

Design Two bit ripple down counter using negative edge triggered JK flip flops:

**Steps:** Steps 1, 2 and 3 are same as example 3

4 Write the truth table of 2 bit down counter as shown

Clk pulse	$Q_2$	$Q_1$ (LSB)
1	1	1
2	1	0
3	0	1
4	0	0

5 Observe when higher bit ( $Q_2$ ) is changing. In this case,  $Q_2$  is changing when  $Q_1$  is changing from logic 0 to logic 1. Since we require -ve edge counter which responds when clock is changing from 1 to 0, connect  $Q_1'$  (which changes 1 to 0 at clock pulse 3) as the clock for the next flip flop as shown in Figure 6.1.9

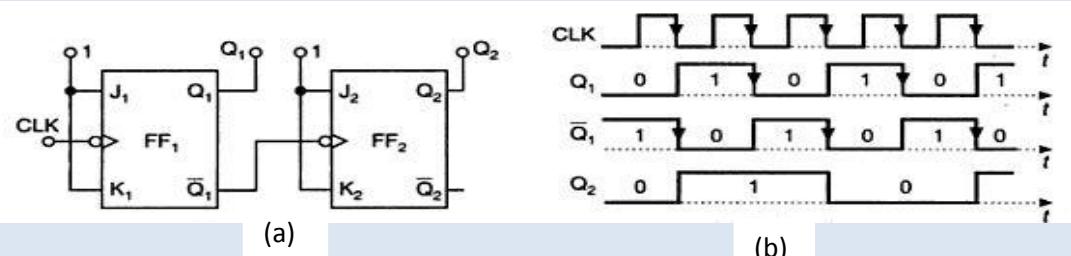


Figure 6.1.9 (a) Asynchronous 2 bit down counter using -ve edge triggered flip flops. (b) Timing diagram.

**Working principle:** The output  $Q_1'$  of FF1 is connected to the clock input of FF2. For the clock pulse applied to FF1, the output of FF1 toggles. For the  $Q_1'$  pulse applied to FF2, the output of FF2 toggles. Thus the 2 bit down counter counts in the order 0, 3, 2, 1, 0, 3 ... Figure 6.1.8 shows a 2 bit ripple down counter using negative edge triggered JK flip flops and its timing diagram.

#### Self test:

- 1 What is the modulus of 2 bit counter?
- 2 Why ripple counters are known as divide by n counter (In case of 3 bit counter, it is divide by 3 counter) ? (\* Hint: analyze the frequency at the output of  $n^{\text{th}}$  flip flop with respect to clock signal)
- 3 Realize 2 bit +ve edge triggered up counter using T flip flop.

### 6.1.3 Shift Register

Registers are digital circuits which are used to store ‘n’ bits information in the same time. Generally registers are built with D flip flops. Each flip flop can store a one bit, so a register composed of ‘n’ flip flops can store ‘n’ bit number. All these flip flops are driven by a common clock and they are set or reset simultaneously. Therefore, the data processing happens sequentially. The output of one flip flop is fed as input to the next flip flop.

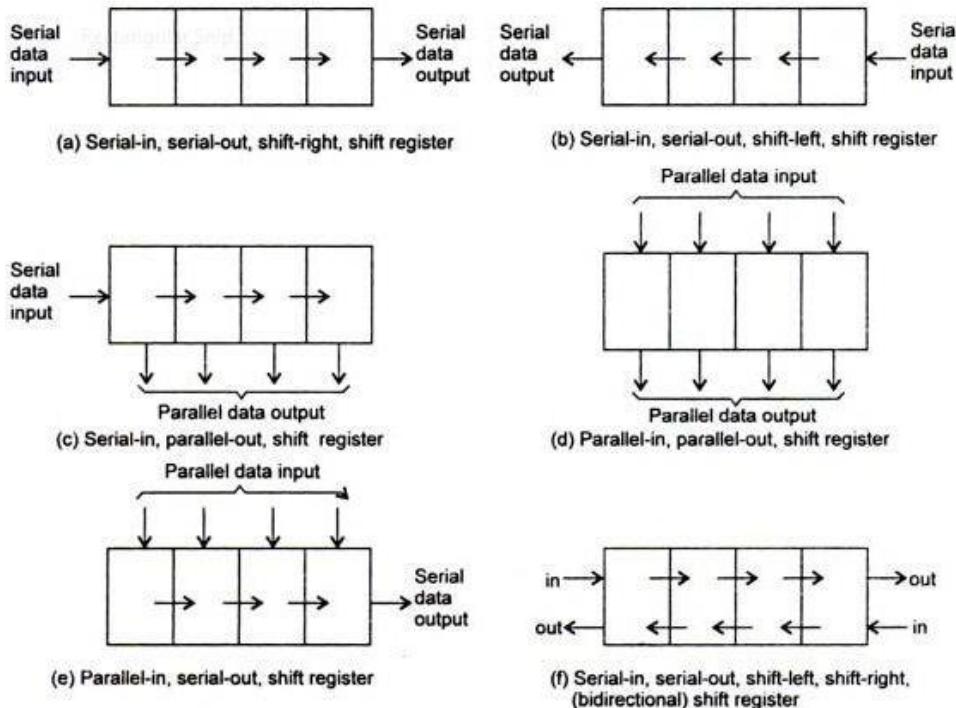


Figure 6.1.10 Data transmission in shift register

In a register data can be entered in serial form or data can also be output in serial form. Then this register is called as shift register since data bits are shifted in the flip flops with each clock pulse. Data can be shifted either towards right or left or even in both the directions. When the data is shifted from left to right, it is known as right shift register. If data is shifted right to left, it is called as left shift register. In bidirectional shift register, data can be shifted either left to right or right to left, depending upon the mode control signal. Data transmission in shift register is shown in Figure 6.1.10.

There are four basic types of shift registers namely, Serial in Serial out (SISO), Serial in Parallel out (SIPO), Parallel in Parallel out (PIPO) and Parallel in Serial out (PISO) shift registers.

#### a) Serial In Serial Out Shift Register (SISO) :

In SISO shift register, data input is in serial form and clock pulses are applied to each flip flop. After each clock pulse, data moves by one position. The output can be obtained in serial form. In this type of shift register data moves either in left or right direction.

The logic diagram of a 4 bit SISO shift right shift register is shown in Figure 6.1.11 with four flip flops, the register can store up to four bits of data. Serial data is applied at the D input of the FF1. The Q output of FF1 is connected to the D input of FF2, the Q output of FF2 is connected to the D input of FF3 and the Q output of FF3 is connected to D input of FF4. When serial data is transferred into a register, each new bit is clocked into the first flip flop FF1 at the positive going of each clock pulse. The bit that was previously stored by FF1 is transferred to FF2. The bit that was stored by FF2 is transferred to FF3 and so on. The bit that was stored by the last FF4 is shifted out.

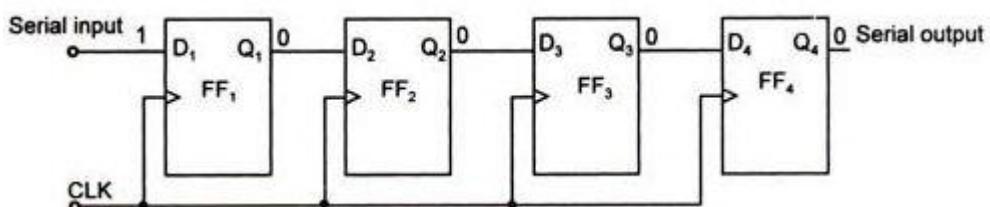


Figure 6.1.11 SISO Shift Register

Table 6.1.5 shows the operation of 4 bit SISO shift register to store  $1010_{(2)}$ . The LSB bit is entered first in the register D1 and 4 clock pulse required to store 4 bit data and 5<sup>th</sup> clock pulse Data present in D4 will be taken out. So SISO requires  $(2n-1)$  clock pulses to take out n bit data i.e in above example MSB bit 1 is taken out from Q<sub>4</sub> at 7<sup>th</sup> clock pulse.

Table 6.1.5 Data shifting in 4 bit SISO shift register for data 1010

Clk Pulse	Q1	Q2	Q3	Q4
Before the CLK	#	#	#	#
Clk 1	0	#	#	#
Clk 2	1	0	#	#
Clk 3	0	1	0	#
Clk 4	1	0	1	0
Clk 5	#	1	0	1
Clk 6	#	#	1	0
Clk 7	#	#	#	1
Clk 8	#	#	#	#

Data is loaded

MSB is available

Note: # indicates any random data

**b) Serial In Parallel Out Shift Register (SIPO):**

SIPO shift register is shown in Figure 6.1.12. In this type of register, the data bits are entered into the register serially, but the data stored in the register is shifted out in parallel form. Once the data bits are stored, each bit appears on its respective output line and all bits are available simultaneously, rather than on a bit-by-bit basis as with the serial output. The serial-in parallel-out shift register can be used as a serial-in serial-out shift register if the output is taken from the Q terminal of the last FF.

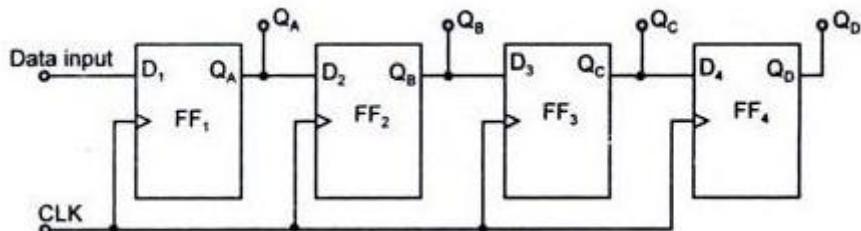


Figure 6.1.12: SIPO Shift Register

Table 6.1.6 Data shifting in 4 bit SIPO shift register for data 1010

Clk Pulse	Q1	Q2	Q3	Q4
Before the CLK	#	#	#	#
Clk 1	0	#	#	#
Clk 2	1	0	#	#
Clk 3	0	1	0	#
Clk 4	1	0	1	0

Data is loaded and can be taken out

### c) Applications of shift Registers:

Shift registers are commonly used

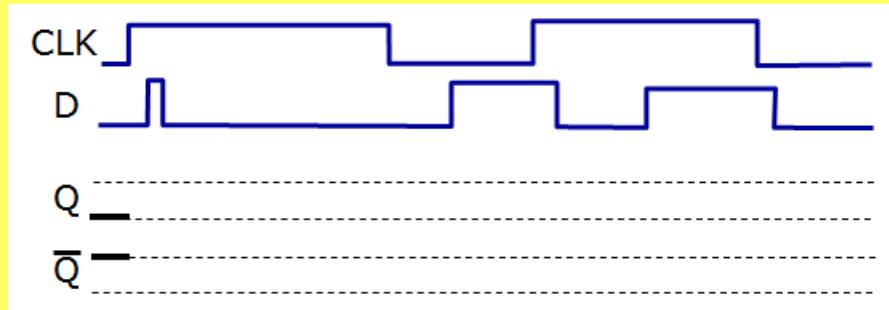
1. To store digital data during arithmetic and logical operations.
2. To build Shift register counters.
3. To generate timing and control waveforms.

#### Summary

- 1 latch is a level triggered bi-stable multi-vibrator circuit while a flip flop is an edge triggered.
- 2 Flipflop is an electronic circuit or device which is used to store data in binary form. . Latch is a class of flip-flops whose output responds immediately to appropriate changes in the input.
- 3 We have studied the working principle of SR, D, JK and T flip flops using NAND gates.
- 4 In binary counter, the flip flops are interconnected such that their combined state at any time is the binary equivalent of the total number of pulses that have occurred up to that time.
- 5 Asynchronous counters are also called as ripple counters
- 6 Registers are digital circuits which are used to store 'n' bits information in the same time
- 7 Shift registers are commonly used to store digital data during arithmetic and logical operations.

**Exercises:**

1. Complete the timing diagram of D flip-flop



- 2 Design mod 4 down counter using +ve edge triggered T flip flop.
- 3 Design 4 bit up counter using -ve edge triggered JK flip flop with the neat timing diagram
- 4 Compare SISO and SIPO
- 5 Consider data 101110 is given to SISO, SIPO. Data is entered from LSB. After how many clock pulse, MSB is retrieved?

**Part –III****PRINCIPLES OF ELECTRONIC COMMUNICATION****Chapter -7: Fundamentals of Analog Communication**

Communication is the process of transferring information from one point to another. Based on the type of the signal transmitted, communication can be analog or digital. The analog communication involves transmitting of an analog signal such as music, video etc. from source to destination. In electronic communication, some kind of modulation of a known signal called carrier with the information or message signal before transmission is performed. In modulation one of the parameters of carrier such as amplitude, frequency or phase is varied in accordance with the instantaneous amplitude of the message signal. Depending on the parameter that is varied, there are three basic types of modulation called amplitude, frequency and phase modulation.

**Module – 1: Amplitude Modulation****Learning Outcomes:**

At the end of this module, students will be able to:

1. Explain the principle of electronic communication using a block diagram.
2. Define modulation and discuss the need for modulation.
3. Explain amplitude modulation using suitable waveforms and define modulation index.
4. Draw the spectrum of AM DSB signal, identify sidebands and estimate bandwidth.
5. Derive expression for power content of AM signal.
6. List and describe different types of AM signal and compare them.
7. Explain the principle of AM demodulation process.
8. Explain with suitable block diagram the principle of Super- heterodyne receiver.

**7.1.1 Principles of Electronic Communication**

In a broad sense, the term electronic communication refers to the sending, receiving and processing of information by electronic means. It can also be defined as the process of transmitting the information or signal from one point known as the source to another point known as the destination. Information can be continuous such as speech, music, image, picture etc. or discrete signals like data from computer etc.

The block diagram of basic communication system is shown in Figure 7.1.1.

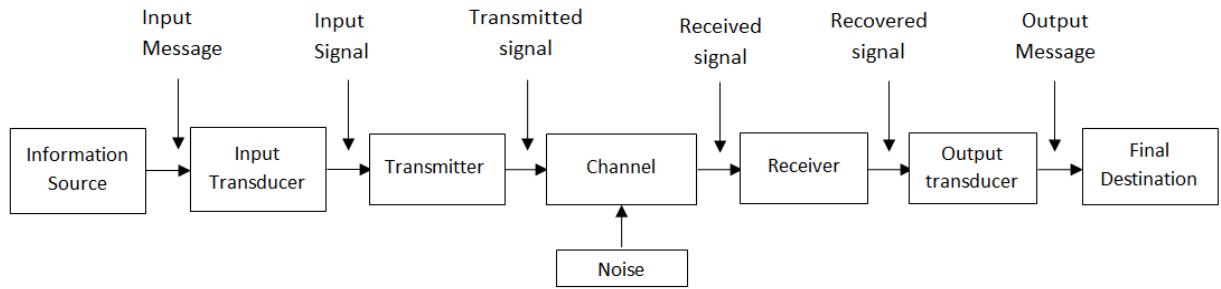


Figure 7.1.1: Block Diagram of Electronic Communication System

**a. Information Source:** The first stage of communication system is the information source because a communication system transmits information from an information source to the destination. The physical form of information is represented by a message that is originated by an information source. The examples of message are voice, live scenes, music, image, written text and e-mail etc.

**b. Input/output transducer:** The input transducer converts physical quantity (non-electrical) to an electrical signal. This electrical signal is called as baseband signal/ message signal. For example, voice is converted to electrical signal using microphone. Similarly at the destination, output transducer is used to convert electrical signal back to physical quantity. For example a loudspeaker is used to convert electrical signal back to voice. Likewise for video, the input transducer can be camera and output transducer can be any picture display unit such as Cathode Ray Tube (CRT) or Liquid Crystal Display (LCD).

**c. Transmitter:** The baseband signal generated by the transducer may not be in the form suitable for the transmission. Hence some kind of processing and signal conditioning is required to make it suitable for transmission. The transmitter section processes the signal before transmission, which mainly consists of filters, amplifiers, modulator and transmitting antenna (for wireless transmission). The filter is used to eliminate the unwanted component of signal generated by the transducer. The desired signal is further amplified to the required level using amplifier. The baseband signal is applied to the modulator, which translates the baseband signal from its low frequency to high frequency, makes it suitable for the transmission in the chosen environment. The modulated signal is then transmitted through a transmission channel.

**d. Transmission channel:** It is a medium over which the electronic signal is transmitted from one point to another. The type and characteristics of the channel along with the noise power decides the transmitter and receiver selection and design and hence the cost of the communication system. The communication medium can be either wired or wireless. An example for wired communication is telephony, where a pair of physical wires is running parallel between transmitter and receiver. Now-a-days optical fibers are used in between transmitter and receiver in which light carries the information. Similarly an example for wireless communication is radio communication, where free space is used as a transmission channel. Antennas are used to couple the signal to and from the channel to the communication system.

**e. Noise:** It is a random, undesirable electrical energy that interferes with the transmitted signal. Noise is a highly undesirable part of the communication system which must be minimized. The noise introduced in the transmission medium is called external noise and the noise introduced in the transmission and reception equipment is called as internal noise. The main cause of internal noise is the thermal agitation of atoms and electrons of the electronic components used in the equipment.

**f. Receiver:** The receiver block mainly consists of receiving antenna, filter, demodulator and amplifier. The signal received from receiving antenna is filtered and desired signal is amplified. It is further demodulated to get back the information signal. Finally an output transducer is employed to convert back the information in electrical form to physical form.

### 7.1.2 Need for Modulation

In wireless communication, the free space is the transmission medium through which electromagnetic waves carrying information propagates. A transmitting antenna at the transmitter radiates energy into the free space and it is received at the receiver using receiving antenna. The original signal which is called base band signal is not suitable for direct transmission/radiation over a long distance and hence modulation is usually performed.

Modulation is a process of varying some of the characteristics of high frequency carrier wave in accordance with the instantaneous amplitude of base band signal. After modulation the baseband signal of low frequency is transferred to the high frequency carrier, which carries information in the form of some variations. The three parameters of a sinusoidal carrier that can be varied are: amplitude, phase and frequency. A given modulation scheme can result in the variation of one or more of these parameters.

- **Modulation for long distance communication:** According to electromagnetic theory, the amount of radiation emitted from an antenna depends on the frequency of the signal current supplied to it. For message signals of low frequency, radiation is poor and they get highly attenuated when transmitted over a longer distance. Therefore modulation is necessary to effectively increase the frequency of the signal to be radiated and thus increases the distance over which signal can be transmitted faithfully.
- **Modulation to reduce the height of the antenna:** The height of the antenna required to transmit and receive radio waves is a function of wavelength of the signal. Since wavelength ( $\lambda$ ) is inversely proportional to frequency ( $\lambda = c/f$ ), for message signal having low frequency,  $\lambda$  is high and hence the height of the antenna required for transmission is very large and impractical. Therefore high frequency carrier is used to transmit the information which requires antenna of lesser height.
- **Modulation for multiplexing:** Multiplexing is a method of transmitting more than one information signal simultaneously over a single channel. This allows several users to use the same channel if they are assigned with different carrier frequencies. Therefore it allows the maximum possible utilization of the available bandwidth of the system.

***Self-test:***

1. List the key components required for electronic communication.
2. List the basic functions of radio transmitter and the corresponding functions of the receiver.
3. Signals travel in air as -----
4. ----- is another name for simplex communication and ---- is another name for duplex communication.
5. Signals travel over copper cable as-----
6. Signals travel over fiber optic cable as-----

**7.1.3 Amplitude Modulation (AM)**

AM is defined as the process of varying the amplitude of the carrier wave proportional to the instantaneous amplitude of modulating signal. In practice, the carrier may be higher frequency while modulating signal may be lower frequency. Figure 7.1.2 shows the modulating, carrier and amplitude modulated wave. Here the modulating signal considered is multi-tone in nature.

**Note:**

- A multi-tone signal is the superposition of several sine waves or tones or frequency components, whereas single-tone signal has only one frequency component in it.

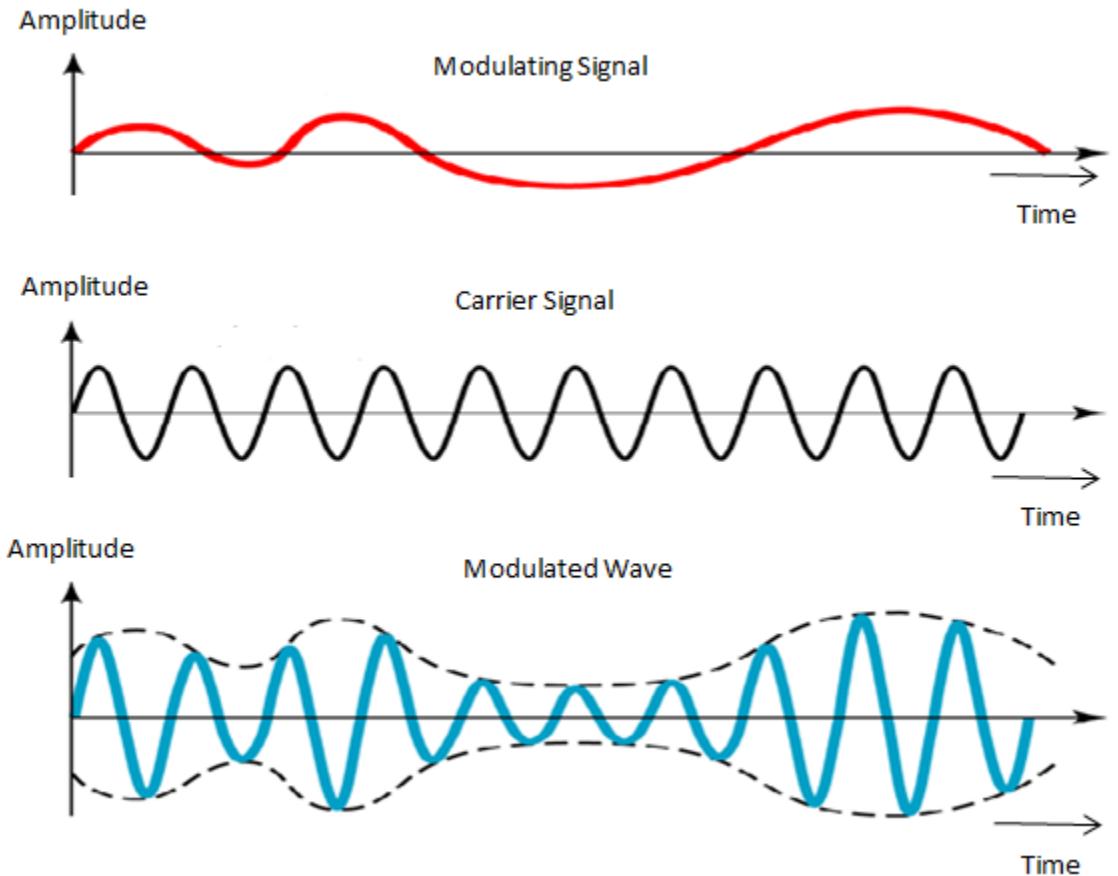


Figure 7.1.2: Waveforms of amplitude modulation

#### 7.1.4 Time domain analysis

The Figure 7.1.3 shows the time domain representation of the AM wave.

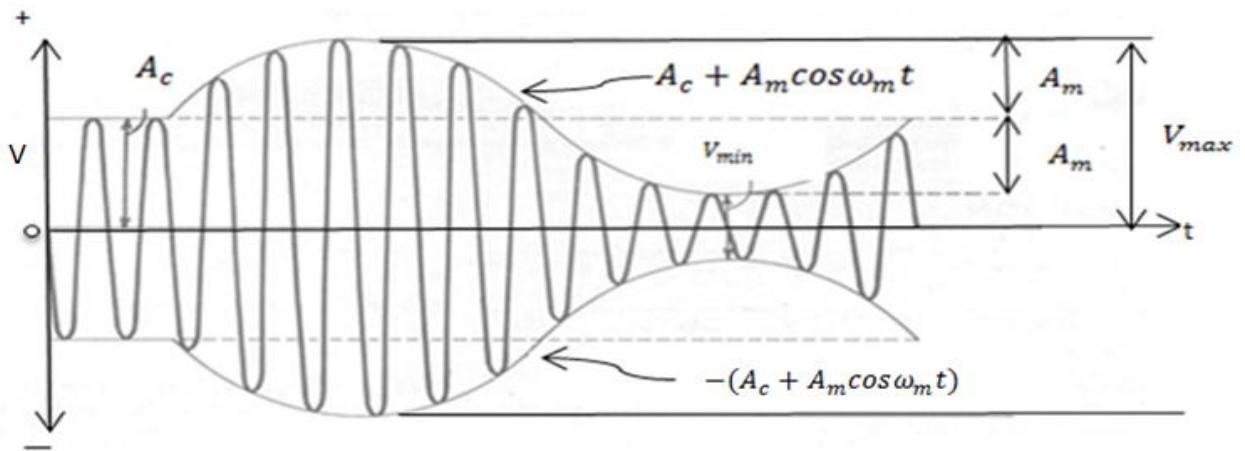


Figure 7.1.3: Time domain representation of AM wave

Let the equation of carrier signal be  $c(t) = A_c \cos(2\pi f_c t)$ , where  $A_c$  is the peak amplitude of carrier signal and  $f_c$  is the frequency of the carrier signal.

Let the equation of modulating signal be  $m(t) = A_m \cos(2\pi f_m t)$ , where  $A_m$  is the peak amplitude of modulating signal and  $f_m$  is the frequency of the modulating signal.

Then by the definition of AM:

$$V_{AM}(t) = [A_c + A_m \cos(2\pi f_m t)] \cos(2\pi f_c t) \quad (7.1.1)$$

$$\begin{aligned} &= A_c \cos(2\pi f_c t) + A_m \cos(2\pi f_m t) \cos(2\pi f_c t) \\ &= A_c \cos(2\pi f_c t) + \frac{A_m}{2} [\cos[2\pi(f_c + f_m)t] + \cos[2\pi(f_c - f_m)t]] \end{aligned}$$

$$\begin{aligned} V_{AM}(t) &= \\ &A_c \cos(2\pi f_c t) + \frac{m A_c}{2} [\cos[2\pi(f_c + f_m)t] + \cos[2\pi(f_c - f_m)t]] \quad (7.1.2) \end{aligned}$$

Where, ' $m$ ' is the modulation index of AM signal which is defined as ratio of the amplitude of modulating signal to that of carrier signal i.e.  $\frac{A_m}{A_c}$ . The significance of modulation index is, it decides the depth of modulation. If it is less than one, then AM signal is known as under modulated signal. If it is more than one, then AM signal is known as over modulated signal. If it is equal to one, then AM signal is known as perfect modulated signal. To obtain the original information, modulation index should always be less than or equal to one. The effect of modulation index on AM wave is illustrated in Figure 7.1.4

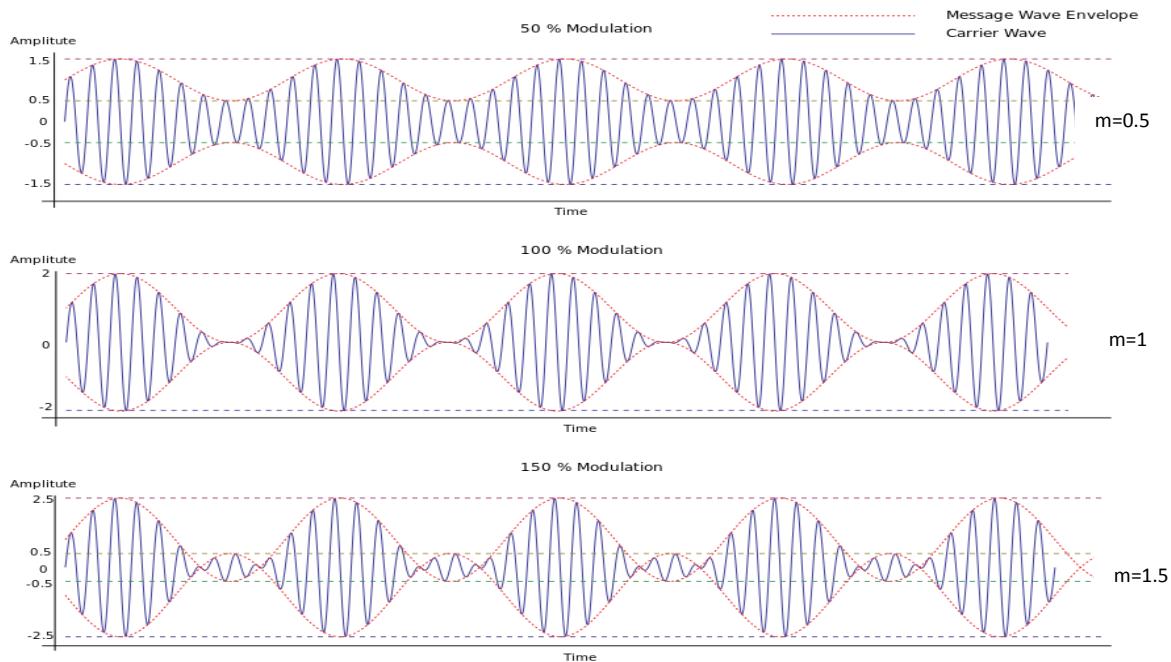


Figure 7.1.4: Effect of modulation index on AM waveform

### 7.1.5 Spectrum of AM signal

Another way of representing the AM signal is in the frequency domain known as frequency spectrum, which shows all the frequency components present in a signal. It is the plot of frequency versus their respective amplitudes. Spectrum of AM signal is shown in Figure 7.1.5.

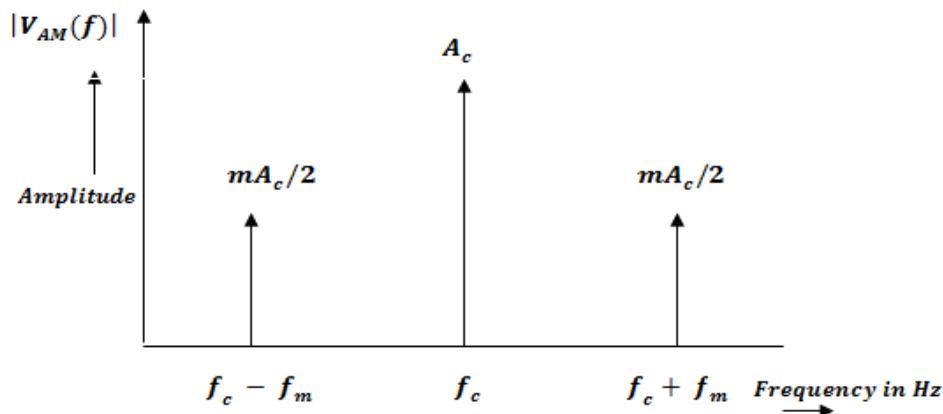


Figure 7.1.5: Frequency Spectrum of AM signal

As shown in Figure 7.1.5, the spectrum of AM consists of three frequency components, one at  $f_c$  and other two at  $f_c + f_m$  and  $f_c - f_m$  respectively. The frequencies  $f_c + f_m$  and  $f_c - f_m$  are known as sideband frequencies i.e.  $f_c + f_m$  is Upper Side Band (USB) and  $f_c - f_m$  is the Lower Side Band (LSB). For this reason this is called AM DSB (Amplitude Modulation with Double Side Band) system. The difference between the two side band frequencies is defined as bandwidth of AM signal. Therefore the bandwidth of AM signal is  $2f_m$ .

#### Self-test:

1. What is a spectrum? What are all the information obtained from the spectrum of AM signal?
2. Imagine that there is a signal with two frequency components,  $f_1$  and  $f_2$  with  $f_2 > f_1$ . If they modulate a carrier with frequency  $f_c$ ,
  - Plot the frequency spectrum of the modulated signal.
  - What is the bandwidth required for such a signal?

### 7.1.6 Power Content of AM

The total power ( $P_T$ ) of AM signal is given by

$$P_T = P_C + P_{USB} + P_{LSB} \quad (7.1.3)$$

Where,  $P_C$  is the carrier power,  $P_{LSB}$  and  $P_{USB}$  are the side band signal powers.

$$P_T = \frac{A_c^2}{2R} + \frac{m^2 A_c^2}{8R} + \frac{m^2 A_c^2}{8R} \quad (7.1.4)$$

$$= \frac{A_c^2}{2R} \left\{ 1 + \frac{m^2}{2} \right\} \quad (7.1.5)$$

Where R is the load resistance.

Thus,

$$P_T = P_C \left\{ 1 + \frac{m^2}{2} \right\} \quad (7.1.6)$$

For 100% modulation,  $m = 1$ ,

$$\text{Therefore, } P_T = P_C \left\{ 1 + \frac{1}{2} \right\} = \frac{3}{2} P_C$$

$$P_C = \frac{2}{3} P_T \quad \text{or}$$

$$P_C = 66.67\% P_T \quad (7.1.7)$$

i.e. 66.67% of total power is carried by the carrier and only 33.33% of total power is available in the sidebands. As the information is available only in the sidebands, and carrier does not in any way contribute to the information, 66.67 % of power is wasted if AM DSB with full carrier is used.

- Modulation by several sine waves:**

When the modulating signal consists of several sine waves,

$$m(t) = A_{m1} \cos(2\pi f_{m1} t) + A_{m2} \cos(2\pi f_{m2} t) + \dots$$

The overall modulation index will be ,

$$m_t = \sqrt{\frac{A_{m1}^2 + A_{m2}^2 + \dots}{A_c^2}} \quad (7.1.8)$$

Or

$$m_t = \sqrt{m_1^2 + m_2^2 + \dots} \quad (7.1.9)$$

Therefore, the total power is,

$$P_T = P_C \left\{ 1 + \frac{m_t^2}{2} \right\} \quad (7.1.10)$$

**Example Problem 1:**

1. An audio signal of  $10\sin(2\pi 1000t)$  volts amplitude modulates a carrier of  $40\sin(2\pi 2000t)$  volts. Find
- Modulation index
  - Sideband frequencies
  - Bandwidth
  - Total power delivered if  $R_L = 1\text{k}\Omega$
  - Amplitude of each side band components

**Solution:**

- Modulation index:  $m = \frac{A_m}{A_c} = \frac{10}{40} = 0.25$
- Sideband frequencies :  
Upper side band =  $f_C + f_m = 3000\text{Hz}$   
Lower side band =  $f_C - f_m = 1000\text{Hz}$
- Bandwidth =  $2f_m = 2\text{kHz}$
- Total power delivered:  
$$P_T = \frac{A_c^2}{2R} \left( 1 + \frac{m^2}{2} \right) = \frac{1600}{2000} \left( 1 + \frac{(0.25)^2}{2} \right) = 0.825 \text{ Watts}$$
- Amplitude of each sideband =  $m \frac{A_c}{2} = 0.25 * \frac{40}{2} = 5\text{V}$ .

**Example Problem 2:**

Certain AM transmitter radiates 9 kW of power with carrier unmodulated and 10.125kW of power when carrier is sinusoidally modulated. Calculate the modulation index. If another sine wave corresponding to 40% modulation is transmitted simultaneously, determine the total power radiated.

**Solution:**

- Given:  $P_C = 9\text{kW}$   
 $P_T = 10.125\text{kW}$   
$$P_T = P_C \left\{ 1 + \frac{m^2}{2} \right\}$$
  
$$m = \sqrt{2\left(\frac{P_T}{P_C} - 1\right)} = 0.5.$$
- $m_1 = 0.5, m_2 = 0.4, P_C = 9\text{kW}$ .  
$$m_t = \sqrt{m_1^2 + m_2^2} = 0.64.$$
  
$$P_T = P_C \left\{ 1 + \frac{m_t^2}{2} \right\} = 10.84\text{kW}$$

**Exercises:**

1. Show that modulation index =  $\frac{V_{MAX} - V_{MIN}}{V_{MAX} + V_{MIN}}$ , where  $V_{MAX}$  and  $V_{MIN}$  are maximum and minimum voltage values of the envelope of AM signal.
2. A 360W carrier is simultaneously modulated by two audio waves with percentage modulation of 55 and 65 respectively. Find the modulation index, total power radiated and power in each sideband. Assume  $R_L=1\Omega$ . [Ans:  $m_t = 0.85$ ,  $P_T = 490W$ ,  $P_{USB} = P_{LSB} = 65W$ ].
3. A broadcast AM transmitter radiates 10kW when the modulation percentage is 60. How much of this is the carrier power? [Ans:  $P_c=8.47\text{ kW}$ ]

**7.1.7 Different Types of AM Signals**

The basic form of AM signal is called Double Side Band with Carrier (AM-DSB), where it contains two side bands namely, LSB, USB and the carrier signal in its unmodulated form. The transmission efficiency of the AM-DSB transmitter is very poor because the maximum efficiency that can be achieved is only 33.33%. The remaining 66.67% power is lost in the unmodulated carrier signal, which does not carry any useful information. The two side bands carry similar information and together they carry 33.33% of the transmitted power. Significant amount of transmitting power can be saved if the AM signal is modified before its transmission, accordingly several forms of AM signals have been developed which are explained below.

- ***Double Side Band- Suppressed Carrier (DSB-SC) Signal:***

It is the first modified version of AM-DSB signal. In a AM-DSB signal carrier does not contain any useful information. Therefore, the two side bands are transmitted by suppressing the carrier signal leading to a savings of 66.67% of the total transmitting power. This modified AM signal is called as Double Side Band Suppressed Carrier (DSB-SC) signal.

- ***Single Side Band (SSB) Signal:***

The two side bands transmitted in DSB-SC signal are identical and carry similar information. If only one side band is transmitted, even then the information will be satisfactorily communicated. This gives rise to another form of AM known as Single Side Band (SSB) modulation. Thus in SSB, only one side band is transmitted by suppressing the other side band and the carrier signal. This also results in reduced transmitter power compared to DSB-SC. The main disadvantage of AM-SSB is, it requires complex receiver.

- ***Vestigial Sideband (VSB) Signal:***

It is another form of AM modulation, where one sideband is completely present, and a part (vestige) of other sideband is retained. If the carrier signal is transmitted along with the side bands, then the recovery of the baseband signal becomes easier. This also reduces the complexity of the receiver circuit and which in turn reduces its cost.

**Note:**

- AM-DSB signal is also referred as DSB with full carrier (DSB-FC) and the AM without carrier is called DSB with suppressed carrier (DSB-SC) signal.
- Watch this Video for animation of amplitude modulated output with different modulation indices (DSB): <http://www.youtube.com/watch?v=1wUjLWNgqMs>

Table 7.1.1 Comparision of different types of AM for  $m=1$

AM Scheme	Bandwidth	Carrier power	Side band power	% of power saving as compared to AM-DSB	Typical Applications
AM-DSB	$2f_m$	66.67%	33.33%	NIL	AM radio broadcast
DSB-SC	$2f_m$	NIL	33.33%	66.67%	Non-commercial systems
SSB	$f_m$	NIL	16.67%	83.33%	Carrier telephony systems, military applications

*Self-test:*

1. Compare different types of Modulation techniques based on a) bandwidth required and b) power content of the signal?
2. Calculate the percentage of power saving when the carrier and one of the side bands are suppressed in an AM wave modulated to a depth of (a) 100% (b) 50%
3. Explain why AM-DSB is preferred for commercial radio broadcasting?

### 7.1.8 Detection of AM Signal

Amplitude modulation or AM is one of the most straight forward ways of modulating a radio signal or carrier. In the process of demodulation (detection), the audio signal is removed from

the radio carrier in the receiver. Demodulation is a process of recovering the original base band signal (information) from the modulated signal. The simple and highly effective method for demodulation is by using envelope detector.

Envelope detector produces an output signal that follows the input signal waveform exactly. Figure 7.1.6 shows the circuit diagram of an envelope detector that consists of a diode and a resistor-capacitor filter.

This is essentially a half wave rectifier with filter circuit, which allows only half of the alternating waveform through. The capacitor bypasses the high frequency carrier component and allows low frequency message signal to go to the output. This demodulator is applicable only for AM-DSB and the main advantage of this form of AM demodulator is that it is very simple and cost effective.

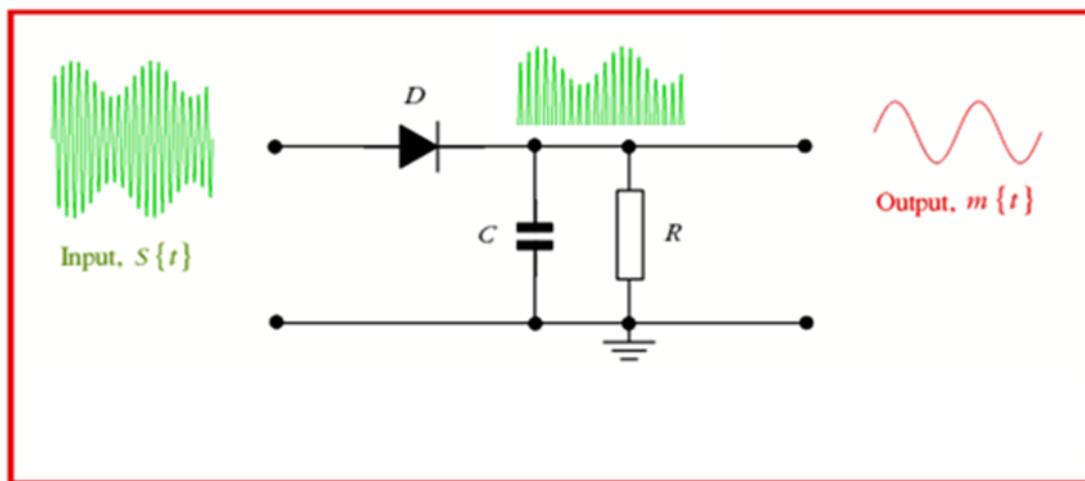


Figure 7.1.6: Envelope Detector

### 7.1.9 Super-heterodyne principle of AM Detection

#### Antenna

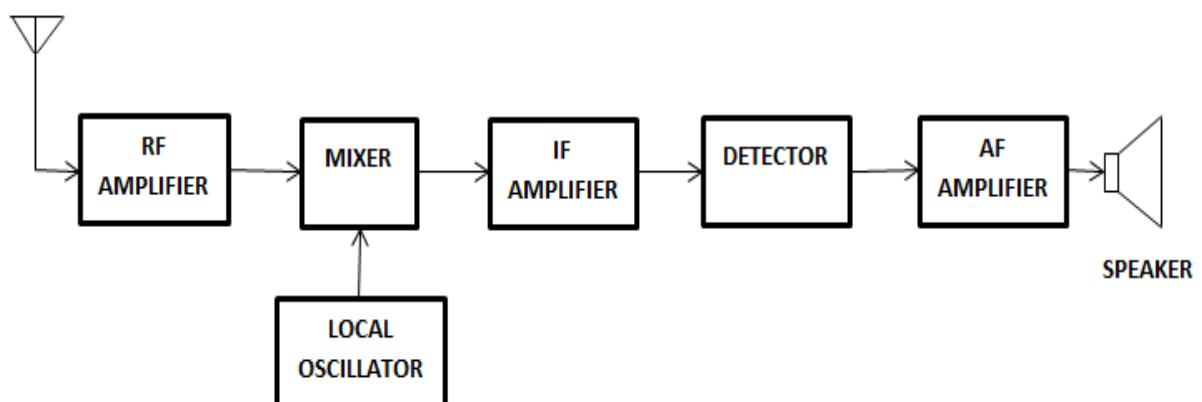


Figure 7.1.7: Block diagram of AM Super-heterodyne Receiver

There are a great variety of receivers in communication systems based on the requirements such as the modulation scheme, the operating frequency and its range. One of them is *super-heterodyne* type, which uses frequency mixing or heterodyning to convert a received signal to get a fixed intermediate frequency (IF). This allows the processing of signal easier as the circuits after IF needs to be designed for narrow band of frequency. The functional block diagram is shown in Figure 7.1.7.

The received signal from the antenna is amplified by the RF amplifier and is fed to the mixer stage, which performs the heterodyning of the incoming signal with the local oscillator signal to produce the sum and the difference of those two frequencies. The IF amplifier will be tuned to the difference frequency as it is smaller among the two, and is known as the intermediate frequency (IF). A typical value of IF for an AM communication receivers is 455 KHz. The difference frequency is at a lower frequency than either the RF input or oscillator frequencies.

Once the IF stage/stages have amplified the intermediate frequency to a sufficient level, it is fed to the detector. The detector is used to demodulate the signal and to get back the message. The detector stage consists of a rectifying device and filter, which respond only to the amplitude variations of the IF signal. This develops an output voltage varying at an audio-frequency rate. The output from the detector is further amplified in the audio amplifier and is used to drive a speaker or earphones.

## Summary

1. Basic principle of electronic communication, which has essentially three components:  
a) Transmitter, b) Receiver and c) Channel
2. Definition of modulation, which is nothing but varying some parameter of a known signal called carrier in accordance with the amplitude of the message signal called modulating signal.
3. Modulation is necessary for the following reasons:  
a) Ease of radiation b) efficient transmission and c) supporting multiplexing
4. To draw the waveforms for amplitude modulated signal with respect to the chosen modulating and carrier signals.
5. Modulation index gives the depth of modulation or the extent to which the carrier is modulated by the signal and is given by  $\frac{A_m}{A_c}$
6. Draw the spectrum of AM-DSB signal for a single tone modulation and identify two sidebands and the carrier. The bandwidth of AM DSB is given by  $2f_m$ , where  $f_m$  is the maximum frequency component of the modulating signal.
7. The power content of AM DSB is given by  

$$P_T = P_C \left\{ 1 + \frac{m^2}{2} \right\}$$
8. The different types of AM signal are AM DSB with carrier, DSB SC, SSB SC, SSB with carrier and VSB.
9. AM DSB can be demodulated by relatively simple process of envelope detection.
10. One of the popular AM reception method is called super heterodyne principle, where in the input RF signal is translated to an IF signal by mixing or beating it with the output of local oscillator. Since the local oscillator frequency is maintained above the incoming signal frequency it is called Super- heterodyning.

## Module – 2: Frequency Modulation

In the previous module, we have learnt amplitude modulation for transmission of analog signals. Here, we will discuss about another type of analog modulation called Frequency Modulation (FM) where the frequency of the carrier is varied in accordance with the instantaneous amplitude of the modulating signal.

### Learning Outcomes:

At the end of this module, students will be able to:

1. Define Frequency Modulation(FM)
2. Explain frequency modulation using suitable waveforms and define modulation index.
3. Illustrate the concept of FM graphically
4. Write the time domain equation for FM and determine the bandwidth of FM using Carson rule
5. Distinguish between AM &FM.

### 7.2.1 Frequency Modulation

Frequency Modulation is defined as a process of altering the frequency of the carrier signal with respect to the instantaneous amplitude of the modulating signal. It is illustrated in Figure 7.2.1.

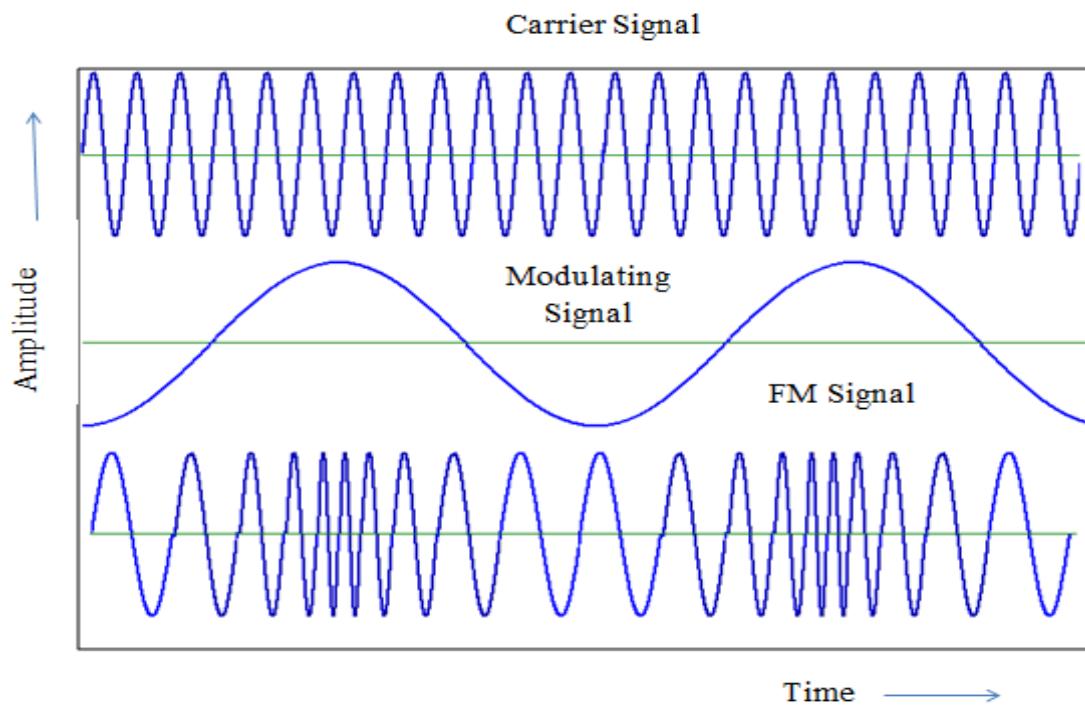


Fig.7.2.1: Frequency Modulation waveforms

### 7.2.2 Time domain analysis of FM signal

- **Time domain analysis of FM signal**

From definition, the frequency of the FM modulated signal is,

$$f_{FM} = f_c + K_f m(t) \quad (7.2.1)$$

Where  $K_f$  is the frequency sensitivity,  $f_c$  is the carrier frequency and  $m(t)$  is the message signal or modulating signal.

Let the message signal  $m(t) = A_m \cos(2\pi f_m t)$ , where  $A_m$  is the peak amplitude of the modulating signal and  $f_m$  is the modulating signal frequency. Substituting for  $m(t)$  in equation (7.2.1), the equation for the FM signal is,

$$\begin{aligned} f_{FM} &= f_c + K_f A_m \cos(2\pi f_m t) \\ f_{FM} &= f_c + \Delta f \cos(2\pi f_m t) \end{aligned} \quad (7.2.2)$$

where  $\Delta f$  is the frequency deviation. It signifies the amount by which the carrier frequency gets deviated.

Multiplying by  $2\pi$  on both sides of equation (7.2.2)

$$\begin{aligned} 2\pi f_{FM} &= 2\pi f_c + 2\pi \Delta f \cos(2\pi f_m t) \\ \Delta \omega_{FM} &= \Delta \omega_c + \Delta \omega \cos(2\pi f_m t) \end{aligned} \quad (7.2.3)$$

Since,  $\omega_{FM} = \frac{d\theta(t)}{dt}$ , integrating both sides of equation (7.2.3)

$$\omega_{FM}(t) = \omega_c(t) + \Delta \omega \int \cos(2\pi f_m t) dt \quad (7.2.4)$$

$$\begin{aligned} \theta(t) &= \omega_c(t) + \Delta \omega / \omega_m \sin(2\pi f_m t) \\ \omega_{FM}(t) &= \theta(t) \end{aligned} \quad (7.2.5) \quad (\text{since } \omega_{FM} = \frac{d\theta}{dt})$$

Therefore, the equation of FM Signal is given by,

$$\begin{aligned} V_{FM}(t) &= A_c \cos[\theta(t)] \\ &= A_c \cos[\omega_c t + \Delta \omega / \omega_m \sin(2\pi f_m t)] \\ &= A_c \cos[\omega_c t + 2\pi \Delta f / 2\pi f_m \sin(2\pi f_m t)] \\ &= A_c \cos[\omega_c t + \Delta f / f_m \sin(2\pi f_m t)] \\ &= A_c \cos[\omega_c t + \beta \sin(2\pi f_m t)] \end{aligned} \quad (7.2.6)$$

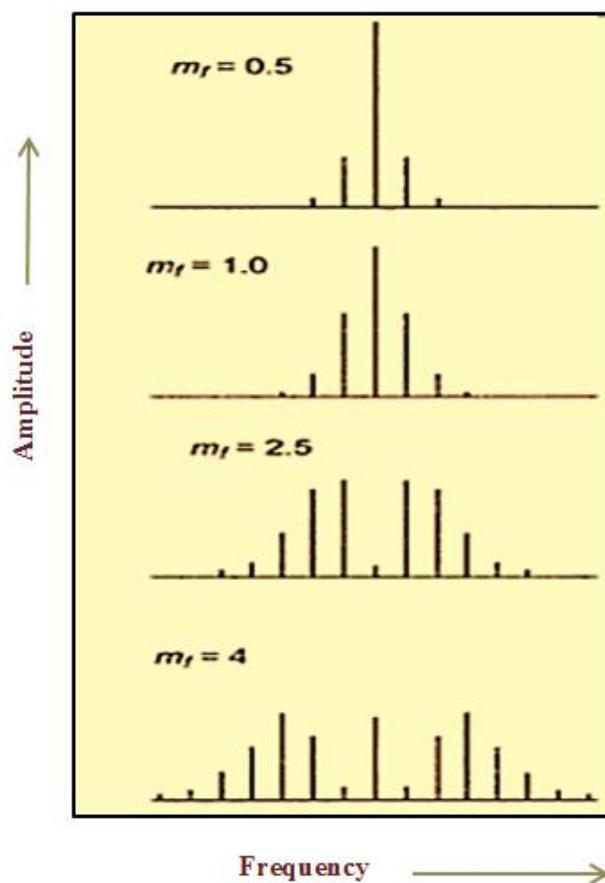
Where  $m_f = \beta = \frac{\Delta f}{f_m}$  is defined as modulation index of FM. Unlike AM modulation index,  $m_f$  is not restricted to one. It can be more than unity.

- **Spectrum of FM**

Figure 7.2.2 shows the spectrum of FM for different values of  $m_f$ . It is seen that as the modulation index,  $m_f$ , increases, more number of sidebands appear. Therefore any FM signal with large  $m_f$  will have a large number of sidebands and hence larger bandwidth. Ideally, FM signal has infinite bandwidth. However, for practical purpose, Carson's rule is followed, which says that for good reception of FM, it is enough if those many side bands which constitutes 98% of power is taken. This acts as the basis for estimation of bandwidth for FM.

The bandwidth for FM as per Carson's rule is given by,

$$\text{Bandwidth} \approx 2(\Delta f + f_m) \quad (7.2.7)$$



*Self-test:*

- 1 Define frequency modulation.
- 2 Write the time domain expression for frequency modulation and explain.
- 3 Define modulation index and write the expression for the same.
- 4 List the factors that affect the bandwidth of FM signal?

**Example Problems:**

- 1 Given a FM equation  $V_{FM}(t) = 10 \cos [2\pi 10^8 t + 5 \sin(2\pi 15000t)]$ , calculate carrier frequency, modulating frequency, frequency deviation and bandwidth.

Solution:

$$\text{Carrier frequency: } f_c = 10^8 \text{ Hz}$$

$$\text{Modulating frequency: } f_m = 15 \text{ kHz}$$

$$\text{Frequency deviation: } \Delta f = \beta \times f_m = 5 * 15 = 75 \text{ kHz}$$

$$\text{Bandwidth} = 2(\Delta f + f_m) = 2(75 + 15) = 180 \text{ kHz}$$

- 2 In an FM system when the audio frequency is 500Hz , modulating voltage is 2.5V , the deviation produced is 5kHz. If the modulating voltage is now increased to 7.5V, calculate the new value of frequency deviation. If the AF voltage is raised to 10V while the modulating frequency is dropped to 250Hz, what is the frequency deviation produced. Also calculate modulation index in each case.

Solution:

$$\text{Given: } f_m = 500 \text{ Hz}, A_m = 2.5 \text{ V}, \Delta f = 5 \times 10^3 \text{ Hz.}$$

$$\text{i) Modulation index: } \beta = \frac{\Delta f}{f_m} = \frac{5 \times 10^3}{500} = 50$$

$$\text{If } A_m = 7.5 \text{ V, } \Delta f = ?$$

$$K_f = \frac{\Delta f}{A_m} = \frac{5 \times 10^3}{2.5} = 2 \text{ kHz/V}$$

$$\Delta f = K_f \times A_m = 2 \times 7.5 = 15 \text{ kHz}$$

$$\text{Modulation index: } \beta = \frac{\Delta f}{f_m} = \frac{15 \times 10^3}{500} = 30$$

$$\text{ii) } \Delta f = K_f \times A_m = 2 \times 10 = 20 \text{ kHz}$$

$$\text{iii) Modulation index: } \beta = 80$$

**Exercises**

- 1 A carrier of amplitude 5V and frequency 90MHz is frequency modulated by a sinusoidal voltage of amplitude 5V and frequency 15 KHz. The frequency sensitivity is 1Hz/V. Calculate the frequency deviation and modulation index. (Ans:  $\Delta f = 5 \text{ Hz}$ ,  $\beta = 0.0003$ )

- 2 The carrier frequency in an FM modulator is 1000 KHz. If the modulating frequency is 15 KHz, what are the first three upper sideband and lower sideband frequencies?

(Ans: 955kHz, 970kHz, 985kHz, 1015kHz, 1030kHz, 1045kHz)

### 7.2.3 Comparison of AM and FM

Sl. no	Parameter	AM	FM
1	Amplitude of the modulated wave	Varies instantaneously with the modulating signal amplitude	constant
2	Frequency of the modulated wave	Contains Carrier and sideband frequency components	Contains carrier and infinite sideband frequency components
3	Modulation Index	$0 < m_a \leq 1$	$m_f > 0$
4	Noise immunity	Less	More
5	Adjacent channel interference	More	Less due to guard bands
6	Bandwidth	Less	More
7	Circuit complexity	Less	More
8	Coverage area	More	Less

*Self-test:*

1. Noise interference in AM is greater than in FM (True/False)
2. AM systems require more bandwidth than FM (True/False)
3. FM signals reach longer distances than AM (True/False)
4. AM systems are more complex to build than FM (True/False)

#### Exercises

- 1 Explain why FM waves cover shorter distances as compared to AM?
- 2 What is guard band? Explain.

**Watch this Video for animation of amplitude modulation:**

1. <http://www.youtube.com/watch?v=5JyiFWLn-w>

**Watch this Video for animation of frequency modulation:**

1. <http://www.youtube.com/watch?v=SmW4z76KgNQ>

### Summary

In this module we have learnt:

1. The definition of Frequency Modulation is a process of altering the frequency of the carrier signal with respect to the instantaneous amplitude of the modulating signal.
2. To draw the waveforms for frequency modulated signal with respect to the chosen modulating and carrier signals.
3. Modulation index gives the depth of modulation and is given by  $m_f = \beta = \frac{\Delta f}{f_m}$
4. The bandwidth is estimated approximately by Carlson Rule given by  
Bandwidth =  $2(\Delta f + f_m)$

## Chapter-8

### Introduction to digital communication

Type your te

In analog communication, the message signals are continuous in nature and can take infinite amplitude levels. When these signals are transmitted over long distances, even a small disturbance (noise) can cause distortion in the signal. Once the signal is distorted, the noise cannot be removed and as a result the original signal cannot be recovered perfectly at the receiver. With digital techniques, these disturbances can be removed by detecting and correcting the errors. Hence digital communication has more benefits as compared to analog communication. In digital communication, the message signal is in discrete form with finite amplitude levels. If the message signal is analog, it must be converted to digital form by the process of Analog-to-Digital Conversion.

#### **Module-1: Digitization of Analog signals**

For any analog information to be transmitted using digital communication system, the signal must be converted to digital form. The Analog-to-Digital conversion process comprises of sampling, quantizing and encoding the analog signal. The digitized signal is then modulated using digital modulation techniques.

##### **Learning Outcomes:**

At the end of this module, students will be able to:

1. State Nyquist Sampling Theorem and explain.
2. Explain qualitatively pulse amplitude modulation technique with the help of suitable waveforms.
3. Draw and explain the general block diagram of Digital Communication system.
4. Explain qualitatively different types of digital modulation techniques with the help of suitable waveforms.

#### **8.1.1 Basic principle of Sampling**

The basic principles of Analog to Digital Conversion are as shown in Fig 8.1.1. Sampling of analog signals is the first step used to digitize analog information. Sampling can be observed in numerous real life applications. For example, music CDs (Compact Discs) are produced by sampling music signal at frequent intervals followed by quantizing and encoding each sample. Even in digital photography, periodic snapshots (samples) are taken to capture continuous phenomena. If the sampling rate is fast enough, the human sensory organs cannot discern the gaps between each snapshot when they are played back. This is the principle behind motion pictures. If the sampling rate is not fast enough, there will be distortion in the reconstructed picture obtained from the digitized samples. Therefore, while sampling an analog signal, there is a minimum sampling rate requirement, called the Nyquist

Sampling rate that avoids distortion in the reconstructed signal. Harry Nyquist proved the sampling theorem which states that “It is possible to reconstruct a band-limited analog signal from periodic samples, as long as the sampling rate ( $f_s$ ) is at least twice the highest frequency component ( $f_m$ ) of the signal”.

Mathematically it can be expressed as

$$f_s \geq 2f_m \quad (8.1)$$

Where  $f_s$  is the sampling frequency and  $f_m$  is highest frequency component in the signal. This theorem is also commonly called the Nyquist criteria for sampling or Sampling theorem.

This means that , for example if a voice signal has frequencies ranging from 0 to 4kHz (Low pass signal), then according to the Nyquist Sampling Theorem, in order to sample this signal without distortion, the minimum required sampling rate is equal to 8kHz. If an analog signal has frequency components ranging from 2 kHz to 5 kHz (Band pass signal), then according to the sampling theorem, the Nyquist sampling rate is equal to twice that of the bandwidth of the signal.i.e.  $2*(5-2)$  kHz = 6 kHz and not 10 kHz.

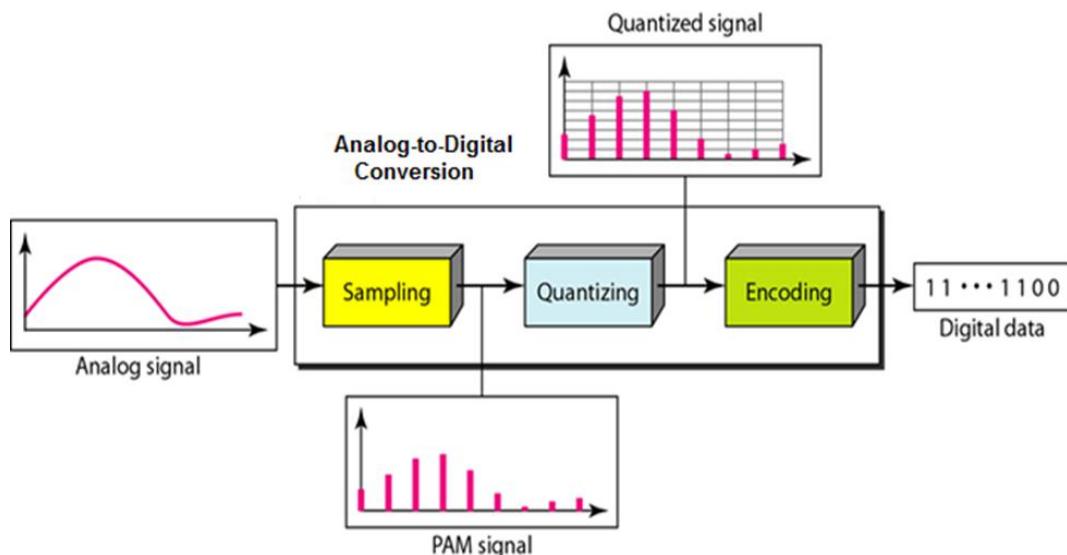


Fig 8.1.1 Steps involved in Analog-to Digital Conversion

#### **Self -test:**

- 1 If a signal is composed of frequency components up to 10 kHz., what is the minimum frequency at which the signal should be sampled as per Nyquist criteria?

**Example Problem:**

- 1 Consider the analog signal  $x(t) = 3\cos 100\pi t$ . Determine the minimum sampling rate required to avoid aliasing.

Solution:

The frequency of the analog signal can be calculated as  $2\pi f_c = 100\pi$ .

Therefore  $f_c = 50\text{Hz}$ . According to the Nyquist sampling rate, the minimum sampling rate required to avoid aliasing is  $f_s = 100 \text{ Hz}$ .

**Exercise:**

- 1 Consider the analog signal  $x(t) = 3\cos 50\pi t + 10\sin 300\pi t - \cos 100\pi t$ . What is the Nyquist rate of sampling for this signal? (Ans: 300Hz)

**Note:** The effect of incorrect sampling rate can be seen when the rotation of a helicopter blade is observed. As the speed of the blade increases, our eyes are under sampling the true speed of the blade with a rate which is limited by the human brain. Similarly, in movies, when the motion of car wheels with increasing speed is observed, the movie camera is under sampling the motion of car wheels by sampling at a rate equal to the fixed frames per second of the camera. In both the examples it is observed that as the speed increases, it creates an illusion of backward rotation. This is because in both cases the actual speed is under sampled.

For Analogy of sampling to Wagon wheel effect, visit the following link:

<http://www.youtube.com/watch?v=6XwgbHjRo30>

**Introduction to pulse modulation techniques**

Pulse modulation involves communication of information using a train of pulses as carrier. It may be used to transmit either analog information such as continuous speech or digital data. If the modulating signal is in analog form, it is called analog pulse modulation technique. Some of analog pulse modulation techniques are: Pulse Amplitude Modulation (PAM), Pulse Width Modulation (PWM) and Pulse Position Modulation (PPM). If digital data is used to modulate a train of pulses, then it is called digital pulse modulation technique. Some of the digital pulse modulation techniques are: Pulse Code Modulation (PCM) and Delta Modulation (DM).

Analog Pulse modulation is a process in which continuous waveforms are sampled at regular intervals using a train of recurrent pulses. Information regarding the signal is transmitted only at the time of sampling along with any synchronizing pulses that may be

required. At the receiving end, the original waveforms can be reconstructed from such samples, if the samples are taken as per the sampling theorem or Nyquist criteria . In analog pulse modulation, the sample amplitude may be infinitely variable while in digital pulse modulation such as PCM and DM, a code which indicates the sample amplitude that is assigned the nearest predetermined discrete amplitude level is sent.

A pulse train has three parameters, namely, Pulse Amplitude, Pulse Width and the instant of occurrence of the pulse – Pulse Position. The information to be transmitted can be used to vary any of these parameters according to the instantaneous amplitude of the modulating signal. This leads to three different types of pulse modulation and they are Pulse Amplitude Modulation, Pulse Width Modulation and Pulse Position Modulation as shown in Fig 8.1.2.

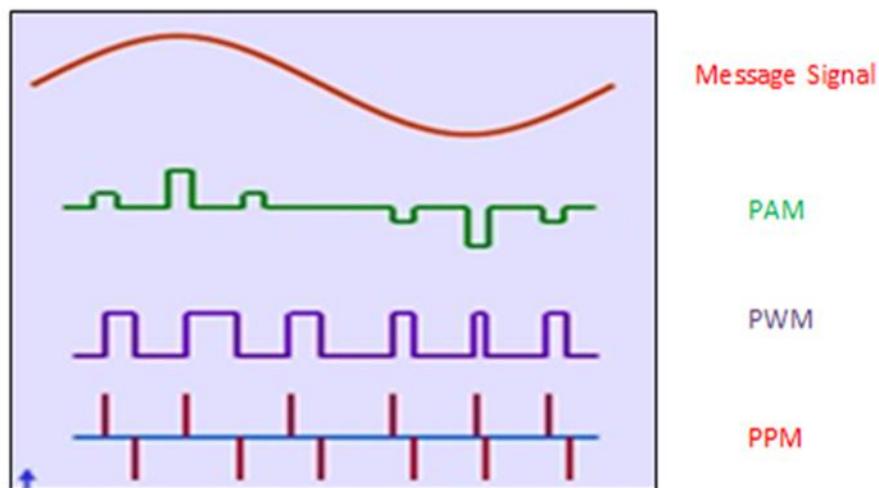


Fig 8.1.2: Different types of pulse modulation signals

### 8.1.2 Pulse Amplitude Modulation (PAM)

Pulse Amplitude Modulation (PAM) is the simplest form of pulse modulation as shown in Fig 8.1.3. Here the signal is sampled at regular intervals and each sample is made proportional to the amplitude of the message signal at the instant of sampling.

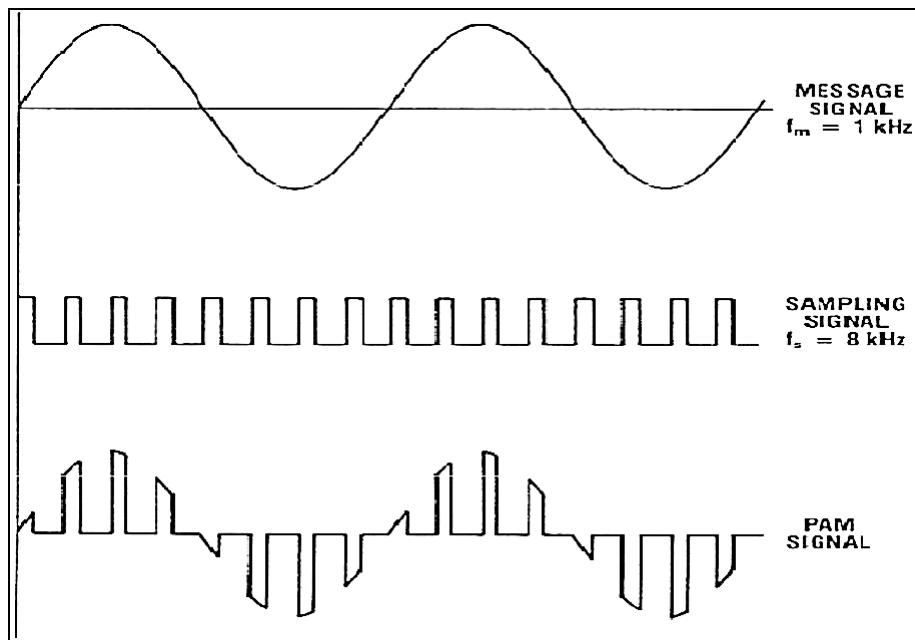


Fig 8.1.3: PAM signal

### 8.1.3 Pulse Width Modulation

The Pulse Width Modulation is also called Pulse Duration Modulation (PDM). Here the amplitude and starting time of each pulse is fixed but the width of each pulse is made proportional to the amplitude of the modulating analog signal at that instant (Refer Fig 8.1.2). PWM has lot of real world applications. For example, the transmission of voice or music can be performed by using PWM. Pulse Width Modulation is also used in dimmer circuits to control the intensity of light. PWM is also used in many control applications such as control of speed and torque of DC motors etc.,

### 8.1.4 Pulse Position Modulation (PPM)

In this method, both the amplitude and the duration are kept constant while the position of each pulse in relation to the position of a recurrent reference pulse is varied by each instantaneous sampled value of the modulating signal (Refer Fig 8.1.2). PPM is used in both analog and digital data transmission. It is commonly used in optical fibre communication, remote controls for TV, toys etc.

### 8.1.5 Digital Communication System

The pulse modulation techniques discussed in the previous section are used to transmit message signals over short distances. They are also called baseband modulation techniques. If message signals have to be transmitted over longer distances, then the pulse modulation technique is not suitable because the modulated signal is in digital form (information is contained in either the amplitude or width or position of the train of pulses). For this reason, digital modulation techniques are used, also called band pass modulation techniques. Here a continuous signal such as a high frequency sinusoidal

signal acts as carrier. Digital modulation is achieved by varying either the amplitude or frequency or phase of the carrier in accordance with the digital data to be transmitted.

### **ELEMENTS OF DIGITAL COMMUNICATION SYSTEMS:**

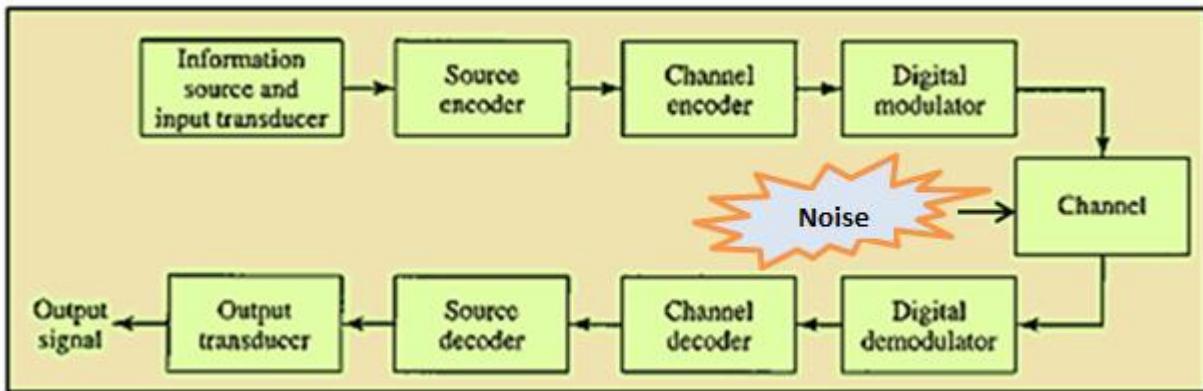


Fig 8.1.4 Block Diagram of a Digital Communication System

Fig 8.1.4 shows the functional elements of a digital communication system. The function of each block is explained as follows:

#### **a. Information Source and Input transducer.**

Many of the real world signals are physical in nature. The device used to convert these physical parameters to corresponding electrical signals is called input transducer. Examples of physical parameters are voice, speech, image etc. The input transducer used to convert voice, speech or music signal or image to an electrical signal. Examples of input transducers are microphone, camera etc. Usually, the output signal from the transducer will be analog in nature. This analog signal is converted into digital form by using analog –to-digital converter. The analog- to-digital conversion consists of sampling, quantizing and encoding. In the case of the output data of a computer, the signal is available in digital form directly.

#### **b. Source Encoder / Decoder**

The aim of the source coding is to represent the digital signal efficiently with as much less number of bits as possible. This will reduce the bandwidth required for transmission. Ex: Huffman coding. The source decoder performs the inverse operation of source encoder. ie. It is used to get back the data in its original representation.

#### **c. Channel Encoder / Decoder**

Channel coding consists of systematically adding extra bits in a known manner to the digital data to be transmitted. These extra bits do not convey any information but help the receiver to detect and / or correct some of the errors in the received data. Channel encoding is done by using either Block Coding or Convolution Coding methods. The channel decoder performs the inverse operation of channel encoder. ie. It is used to extract the digital data from its encoded form with minimum possible error. The decoder helps in detecting and/or correcting the errors in the received data that gets introduced during transmission.

#### ***d. Modulator/ Demodulator***

The Modulator converts the input digital information into an electrical waveform suitable for transmission over the communication channel. Mainly there are three types of Digital modulation techniques viz., Amplitude Shift Keying (ASK), Frequency Shift Keying (FSK) and Phase Shift Keying (PSK).The extraction of the digital data from the received signal is accomplished by the demodulator.

#### ***e. Channel***

The channels are either wired such as pair of wires, coaxial cable and optical fiber or wireless (free space) such as radio channel, satellite channel or combination of any of these. The communication channels have only finite bandwidth and the signal often suffers amplitude and phase distortion as it travels over the channel in addition to attenuation of signal. It may also get corrupted by unwanted, unpredictable electrical signals referred to as noise. The two important parameters used to measure the channel characteristics are Signal to Noise power Ratio (SNR) and usable bandwidth.

- ***Advantages of Digital Communication***

1. The effect of distortion, noise and interference is less in a digital communication system. This is because the disturbance must be large enough to change the pulse from logic 0 to logic 1 or vice versa.
2. Regenerative repeaters can be used at fixed distance along the link, to identify and regenerate a pulse before it is degraded to an ambiguous state.
3. Digital circuits are more reliable and cheaper compared to analog circuits.
4. The Hardware implementation is more flexible than analog hardware because of the use of microprocessors, VLSI chips etc.
5. Signal processing functions like encryption, compression can be employed to maintain the secrecy of the information.
6. Error detecting and Error correcting codes improve the system performance by reducing the probability of error.

- ***Disadvantages of Digital Communication***

1. Large System Bandwidth: - Digital transmission requires a large system bandwidth to communicate the same information in a digital format as compared to analog format.
2. System Synchronization: - Digital detection requires system synchronization whereas the analog signals generally have no such requirement.

### 8.1.6 Digital Modulation Techniques

There are three basic types of digital modulation techniques. They are:

- i. Amplitude Shift Keying (ASK)
- ii. Frequency Shift Keying (FSK)
- iii. Phase Shift Keying (PSK)

In all these techniques, amplitude, frequency or phase of a sinusoidal carrier is varied to represent the information which is to be sent. Here the digitized information is mapped into any one of the above three aspects of sine wave and then transmitted. The sine wave at the receiver is remapped back to the information. The digital modulation techniques are widely used in MODEMs ( MODulator DEModulator) , mobile communication etc. Usually, FSK and PSK modulations are more frequently used than ASK.

#### i. *Amplitude Shift Keying (ASK)*

Here the amplitude of the carrier is changed in accordance with the binary information to be transmitted while the frequency and phase of the carrier are kept fixed. Bit 1 is transmitted by a carrier of one particular amplitude. Bit 0 is transmitted by changing the amplitude to 0 Volt or no signal in this case. The ASK also called On-Off keying (OOK) is as shown in Fig 8.1.5.

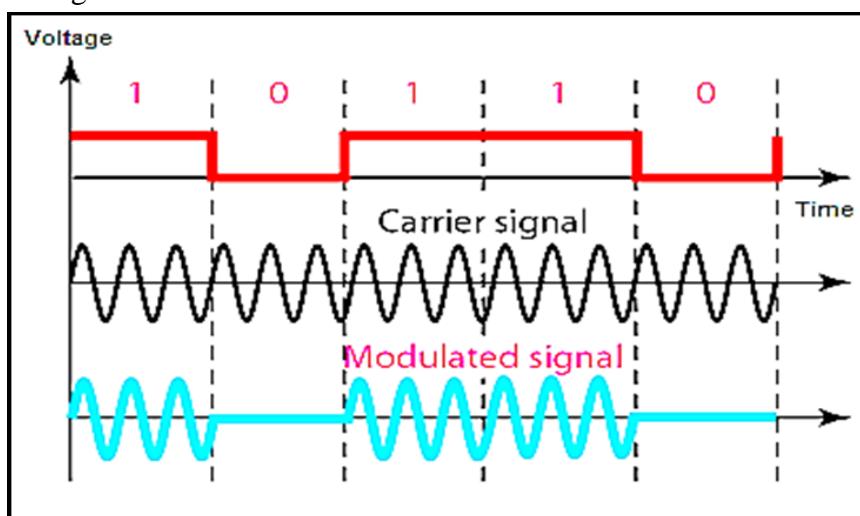


Fig 8.1.5 Binary ASK signal

#### ii. *Frequency Shift Keying (FSK)*

Here the frequency of the carrier is changed in accordance with the binary information. ie. one frequency for bit 1 and another frequency for bit 0. FSK signal is as shown in Fig8.1.6.

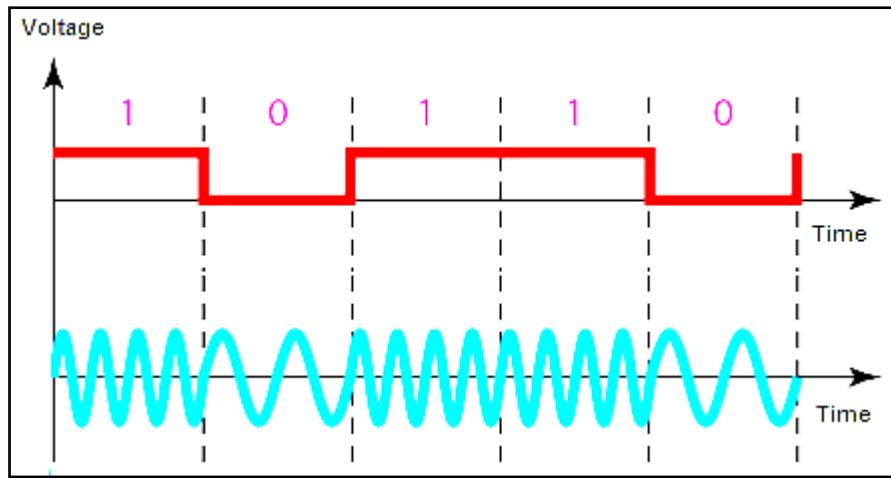


Fig 8.1.6 Binary FSK signal

### *iii. Phase Shift Keying (PSK)*

Here the phase of the sinusoidal carrier is changed in accordance with the binary information. Phase in this context is the starting angle at which the sinusoid starts. To transmit bit 0, the phase of the sinusoid is shifted by  $180^\circ$  whereas to transmit bit 1, the phase of the sinusoidal carrier is shifted by another  $180^\circ$ . Thus the carrier phase shift represents the change in the state of the information. Fig 8.1.7 shows the binary PSK representation.

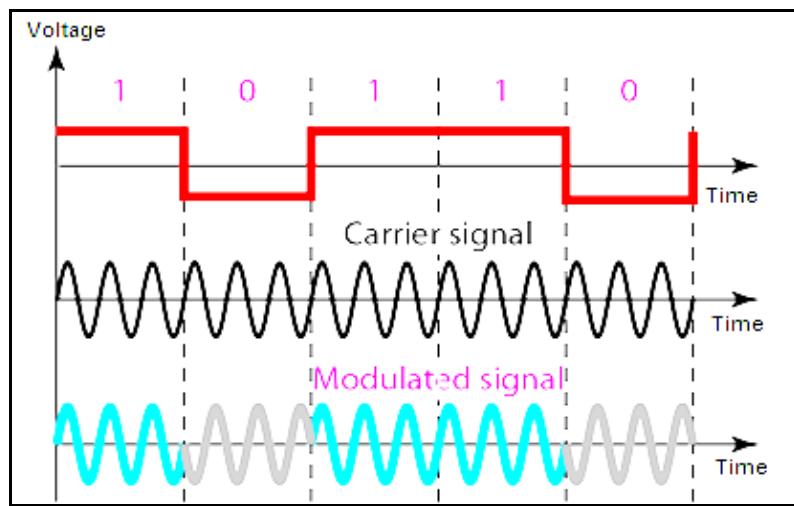


Fig 8.1.7 Binary PSK Signal

**Summary**

1. Sampling theorem states that “It is possible to reconstruct a band-limited analog signal from periodic samples, as long as the sampling rate is at least twice the highest frequency component of the signal”.
2. Pulse modulation involves communication of information using a train of pulses as carrier.
3. Some of analog pulse modulation techniques are: Pulse Amplitude Modulation (PAM), Pulse Width Modulation (PWM) and Pulse Position Modulation (PPM).
4. Digital modulation is achieved by varying either the amplitude or frequency or phase of the carrier in accordance with the digital data to be transmitted.
5. There are three basic types of digital modulation techniques. They are:
  - Amplitude Shift Keying (ASK)
  - Frequency Shift Keying (FSK)
  - Phase Shift Keying (PSK)

## Chapter- 9: Communication Networks

Data communication is the exchange of data between two or more devices via some form of transmission medium. For data communication to occur, the communicating devices must be part of a communication system, which is made up of a combination of hardware and software, and works based on a set of rules called protocol.

A network is a set of devices interconnected by a communication medium. Each device is referred as a node. A node can be a computer, a printer or any other computing device. Networking is a process by which people or group of people share information or services.

### Module – 1: Introduction to Communication Networks:

#### Learning Outcomes:

At the end of this module, students will be able to:

1. Describe the functions performed by communication networks.
2. Explain different types of data communication and network switching techniques.
3. Explain various networks used in data communication.
4. Draw and explain different topology used in communication networks.
5. Describe the need for layered approach and protocol standards.
6. Draw and explain the seven layer architecture of ISO-OSI reference model.

**9.1.1 Introduction to Data communication:** The objective of any data communication system is to facilitate communication between devices (e.g. computers) which may be located in adjacent rooms or located in different corners of the world. The information they may exchange can be in different forms like text, graphics, voice, or video.

The various functions to be carried out in a data communication system are multiplexing, multiple access, error detection and correction, source coding, switching and signaling.

- **Multiplexer:** Combines the signals from different sources to transmit on the channel. At the receiving end, a demultiplexer is used to separate the signals.
- **Multiple access:** When two or more users share the same channel, each user has to transmit the signal only at a specified time or using a specific frequency band.
- **Source coding:** If the channel has a lower bandwidth than the input signal bandwidth, the input signal has to be processed to reduce its bandwidth so that it can be accommodated on the channel.
- **Error detection and correction:** If the channel is noisy, the received data will have errors. Detection, and if possible correction, of the errors has to be done at the receiving end. This is done through a mechanism called channel coding.

- **Switching:** If a large number of users have to be provided with communication facilities, as in a telephone network, the users are to be connected based on the numbers dialed. This is done through a mechanism called switching.
- **Signaling:** In a telephone network, dialing a particular telephone number implies that the network is informed about the destined user. This is called signaling information. The telephone switch (or exchange) will process the signaling information to carry out the necessary operations for connecting to the called party.

**9.1.2 Introduction to Communication Networks:** A communication network is the infrastructure that allows two or more computers to communicate with each other.

Based on the requirements, communications can be of different types:

- a. **Point-to-point communication:** In this type, communication takes place between two end points. For instance, in the case of voice communication using telephones, there is one calling party and one called party. Hence the communication is point-to-point.
- b. **Point-to-multipoint communication:** In this type of communication, there is one sender and multiple recipients. For example, in voice conferencing, one person will be talking but many others can listen. The message from the sender has to be multicasted to many others.
- c. **Broadcasting:** In a broadcasting system, there is a central location from which information is sent to many recipients, as in the case of audio or video broadcasting. In a broadcasting system, the listeners are passive, and there is no reverse communication path.

Three types of switching are generally used in communication networks.

- a. Circuit switching
- b. Message switching
- c. Packet Switching

**a. Circuit switching:** In circuit switching, to transfer the data, circuit must be established so that the data transfer can take place. Applications which use circuit switching may have to go through three phases:

- Establish a circuit
- Transfer the data
- Disconnect the circuit

Example : Telephone networks

**b. Message switching:** In message switching data is routed in its entirety from source to the destination, one hop at a time.

Example: Telegraph, Military applications

**c. Packet switching:** In packet switching the entire message is broken down into smaller chunks called packets. The switching information is added in the header of each packet and transmitted independently. The main advantage of packet switching is that each packet can take a different route to reach the destination, and it is not necessary that all packets follow the same route. If one communication link fails, packets can take a different route. If one communication link has too much traffic, the packets can take a route with less traffic.

Different packets may take different amounts of time to reach the destination and the order of arrival of packets may be different from the one that has been sent. The packets are collected at the destination and reordered before it is delivered.

Example: Internet

### **9.1.3 Types of Communication Networks:**

Different networks can be used for communication purposes.

a. **Local Area Network(LAN)**

A local area network is a network of computers in a localized area, such as in an office or a school. All the computers are connected to each other through the LAN via a hub or a switch.

b. **Wide Area Network(WAN)**

A wide-area network covers a large geographical area and usually consists of multiple computer networks. The Internet is a WAN which relies on a large global network of service providers who use routers, switches, modems and servers to provide connectivity to people and organizations around the world. It is a network of interconnected computers that carry data, media and web pages.

c. **Metropolitan Area Network (MAN)**

A metropolitan area network is a larger network that usually spans several buildings in the same city or town. It is larger than a LAN but smaller than a WAN. A MAN is typically owned and operated by a single entity such as a government body or large corporation.

d. **Public Switched Network**

The public switched network is essentially the telephones' version of the Internet. It is a network of public circuit-switched telephones. The network today is largely digital and includes services for both cellular and landline phones.

e. **Wireless Networks**

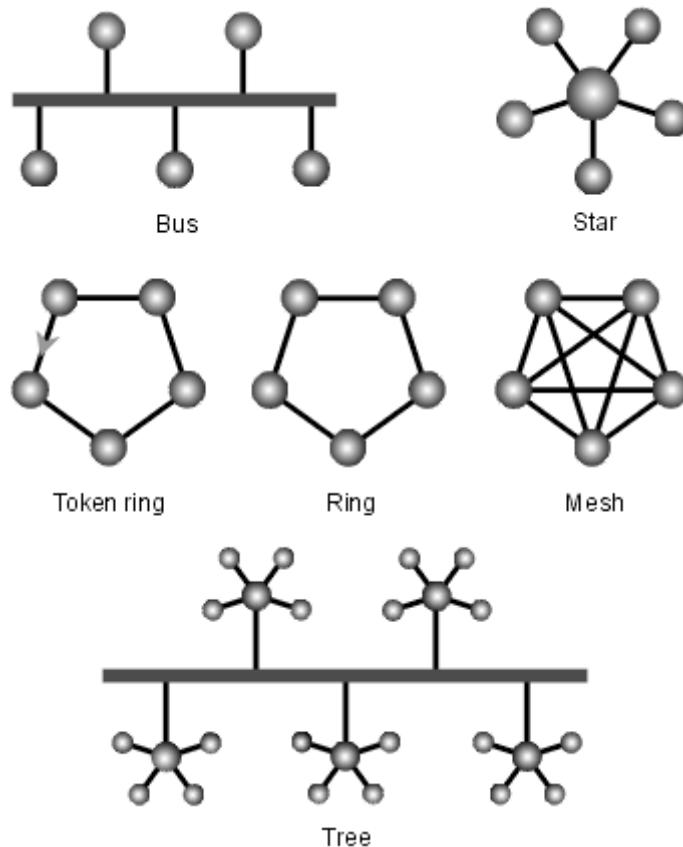
Wireless networks provide information transmission and network connectivity to devices without cables or wires. Some examples of a wireless network include broadcast radio, which sends data over long geographical distances and is available to anyone with a radio who knows the frequency. Wi-Fi is a wireless network for computers, which can access the network remotely. Bluetooth, which connects with a nearby mobile phone, is a shorter-range version of a wireless network, which supports transmission of voice and data but only at a distance of a few feet from the communication device with which it works.

f. **Satellite Networks**

Satellite networks come in a number of different varieties. Phone companies use satellites for data and voice transmission to mobile phones on the ground. Some satellite networks provide navigation information (e.g.: GPS-Global Positioning System), military surveillance or weather data. Still others provide television programming, radio broadcasts and even broadband Internet service.

**9.1.4 Communication Network topology:** In communication networks, a topology is a usually a schematic description of the arrangement of a network, including its nodes and connecting lines. There are two ways of defining network geometry: the physical topology and the logical (or signal) topology.

The physical topology of a network is the actual geometric layout of workstations. There are several common physical topologies, as shown in figure 9.1.1.



**Figure 9.1.1: Network Topology**

In the bus network topology, every workstation is connected to a main cable called the bus. Therefore, in effect, each workstation is directly connected to every other workstation in the network.

In the star network topology, there is a central computer or server to which all the workstations are directly connected. Every workstation is indirectly connected to every other through the central computer.

In the ring network topology, the workstations are connected in a closed loop configuration. Adjacent pairs of workstations are directly connected. Other pairs of workstations are indirectly connected, with the data passing through one or more intermediate nodes.

If a Token Ring protocol is used in a star or ring topology, the signal travels in only one direction, carried by a so-called token from node to node.

The mesh network topology employs either of two schemes, called full mesh and partial mesh. In the full mesh topology, each workstation is connected directly to each of the others. In the partial mesh topology, some workstations are connected to all the others, and some are connected only to those other nodes with which they exchange the most data.

The tree network topology uses two or more star networks connected together. The central computers of the star networks are connected to a main bus. Thus, a tree network is a bus network of star networks.

Logical (or signal) topology refers to the nature of the paths the signals follow from node to node. In many instances, the logical topology is the same as the physical topology. But this is not always the case.

**9.1.5 Network protocols and Reference models:** A protocol is a set of rules that governs how two communicating parties are to interact. In the Web browsing example, the HTTP(Hypertext transfer protocol) protocol specifies how the Web client and server are to interact. Many protocols are required in computer communication to tackle different issues.

Few examples for protocols are

- DNS –Domain Name System
- FTP –File Transfer Protocol
- HTTP- Hypertext Transfer Protocol
- TCP and UDP- Transmission Control Protocol and User Datagram Protocol
- SMTP- Simple Mail Transfer Protocol
- IP and Ipv6- Internet Protocol and Internet Protocol version 6

Network protocols like HTTP, TCP/IP, and SMTP provide a foundation on which the Internet is built on.

*Self -test:*

1. *List the different networks used for communication.*
2. *Draw the various topology used in communication networks.*
3. *Define protocol and explain the need for network protocols.*
4. *For N devices in a network, what is the number of cable links necessary for mesh, ring, bus and star networks.*

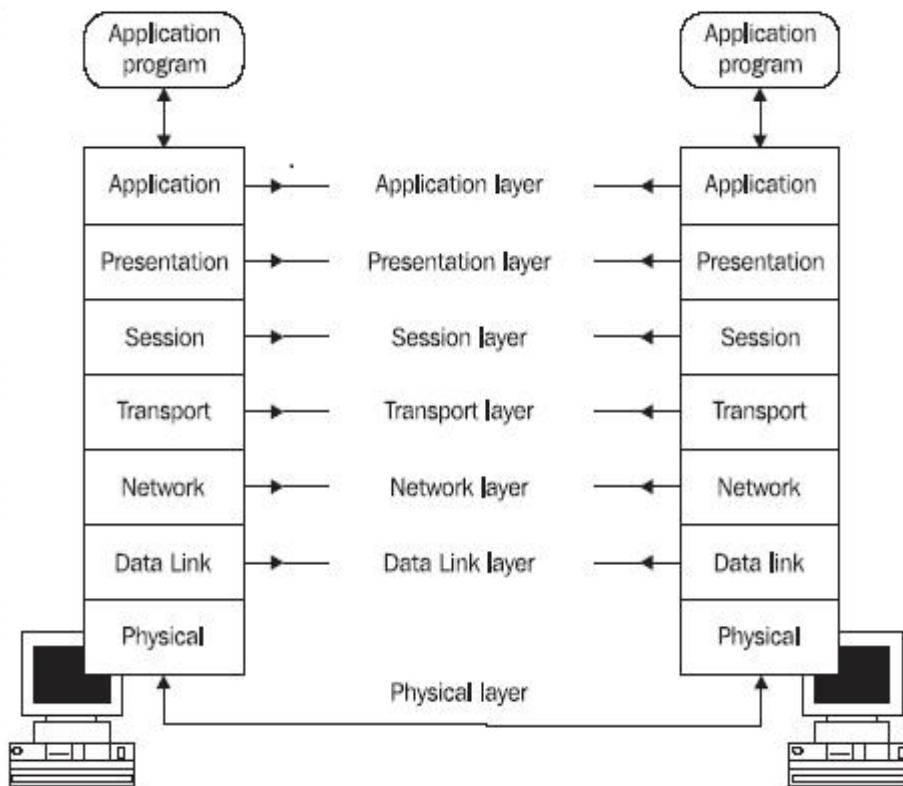
**a. Need for layered approach and protocol standards:**

One way of achieving computer communication is to write monolithic software for all the protocols to be implemented. This approach, being not modular, leads to lots of problems in debugging while developing the software and also in maintenance. On the other hand, a "layered approach" leads to modularity of the software. In a layered approach, each layer is used only for some specific protocols. Layered approach has many advantages:

- Every layer will perform well-defined, specific functions.
- Due to changes in the standards or technology, if there are modifications in one layer's functionality or implementation, the other layers are not affected and hence changes are easier to handle.
- If necessary, a layer can be divided into sub-layers for handling different functions
- If necessary, a layer can be eliminated or bypassed.
- If the protocols for different layers are based on international standards, software or hardware can be procured from different vendors. This multi-vendor approach has a major advantage because of the competition among the vendors, the prices will be competitive and in the bargain, the end user will be benefited.

For computer communication, the ISO/OSI protocol suite and the TCP/IP protocol suite follow the layered approach. The ISO/OSI protocol suite is a seven-layer architecture, whereas TCP/IP is a five-layer architecture.

The International Organization for Standardization (ISO) has developed the Open Systems Interconnection (OSI) protocol architecture, which is a seven-layer architecture for computer communications. This model is considered as the reference model for computer communications. The ISO-OSI model is shown in figure 9.1.2.



**Figure 9.1.2 : Seven-layer architecture of ISO OSI**

The seven layers are as follows:

1. Physical layer
2. Data link layer
3. Network layer
4. Transport layer
5. Session layer
6. Presentation layer
7. Application layer

Each layer performs a specific set of functions. Each protocol layer adds a header and passes the packet to the layer below. Because the header is interpreted only by the corresponding layer in the receiving system, the communication is called peer-to-peer communication. Peer means a layer at the same level.

1. **Physical layer:** The physical layer specifies the physical interface between devices. This layer describes the mechanical, electrical, functional, and procedural characteristics of the interface. An example of the physical layer is Electronic Industries Association (EIA) RS232, which specifies the serial communication interface. A modem is connected to the PC through the RS232 interface.
2. **Data link layer:** The data link layer's job is to activate the link, maintain the link for data transfer and deactivate the link after the data transfer is complete. Error detection and control, and flow control are also done by the data link layer.
3. **Network Layer:** The important function of the network layer is to relieve the higher layers of the need to know anything about the underlying transmission and switching technologies. The functions of the network layer are:
  - Switching and routing of packets
  - Management of multiple data links
  - Negotiating with the network for priority and destination address
4. **Transport Layer:** The transport layer can provide two types of services namely, connection-oriented and connectionless. In connection-oriented service, a connection is established between the two end systems before the transfer of data. The transport layer functionality is to ensure that data is received error-free, packets are received in sequence, and that there is no duplication of packets. The transport layer also has to ensure that the required quality of service is maintained. Quality of service can be specified in terms of bit error rate or delay. In connectionless service, the packets are transported without any guarantee of their receipt at the other end
5. **Session Layer:** The session layer specifies the mechanism for controlling the dialogue in the end systems. Session layer functionality is as follows:
  - Dialogue discipline: whether the communication should be full duplex or half duplex.

- Grouping: to group data into logical units.
- Recovery: a mechanism for recovery in case of intermittent failures of the links.

6. **Presentation Layer:** The functions of the presentation layer are:

- Provide selection of data formats to be exchanged between applications
- Encryption
- Data compression

7. **Application Layer:** The application layer provides management functions to support distributed applications such as e-mail, file transfer, remote login, and the World Wide Web.

The ISO/OSI reference model is fundamental to computer communication because it serves as a reference model for all protocol suites.

**Exercises:**

1. Compare circuit switching and message switching with packet switching.
2. Give an advantage for each type of network topology.
3. What is LAN, WAN and MAN? .Give one example for each.

**Summary**

1. A communication network is the infrastructure that allows two or more computers to communicate with each other.
2. Three types of switching are generally used in communication networks namely: Circuit switching, Message switching and Packet Switching.
3. Some of the types of communication networks are LAN, WAN, MAN, satellite networks, wireless networks etc.
4. The physical topology of a network is the actual geometric layout of workstations.
5. A protocol is a set of rules that governs how two communicating parties are to interact.
6. Network protocols like HTTP, TCP/IP, and SMTP provide a foundation on which the Internet is built on.
7. The ISO/OSI reference model is fundamental to computer communication because it serves as a reference model for all protocol suites.

## Chapter-10: Introduction to Mobile Communication

### Module-1: Introduction to Mobile Communication:

Mobile radio communication generally refers to any radio communication link between two terminals, of which one or both are in motion or stationary in unspecified location. That means, the system is able to establish communication even when the terminal is on the transit or situated in a new location. The very objective of electronic communication is to enable communication with a person, at any time, at any place and in any form. In general, Mobile communication meets this objective and enhances personal communication. Mobile communication allows the user to be able to establish communication even when he is on the transit by providing mobility to the terminal or device. Terminal or device mobility is enabled by wireless access. Personal mobility can be supported by providing unique number to the user and creating dynamic connection with the terminal. Additional facility such as service portability can also be provided by making use of intelligent network (IN) capabilities. In short, Mobile communication provides unlimited reachability, accessibility and rich services to the user.

With the increase in the number of users, accommodating them within the limited available frequency spectrum became a major problem. To resolve this problem, the concept of cellular communication was developed. In cellular communication a basic geographical area of a particular dimension called cell is defined. Each cell is in the form of a hexagon and consists of base station acting as a transceiver located at the center. To accommodate multiple users various multiple access technology is employed.

In the recent years, mobile communication has revolutionized the communication by opening up innovative applications that are limited only to one's imagination. Today, mobile communication has become the backbone of the society and various features available in mobile communication technology have changed the way of living.

### Learning Outcomes:

At the end of this module, students will be able to:

1. Explain the concept of basic cellular mobile communication.
2. Explain the multiple access techniques for cellular system.
3. Draw and explain the architecture of GSM system.
4. Explain the role of base station subsystem (BSS) and mobile switching center (MSC).

### 10.1.1 Principle of Cellular communication:

A cellular radio system provides a wireless connection to the public telephone network for any user location within the radio range of the system. Cellular systems accommodate a large number of mobile units over a large area within a limited frequency spectrum. Figure 10.1.1 shows a basic cellular structure used in mobile communication system. These are conceptualised by hexagonal seven cell structures. The cell number 1 in the Figure 10.1.1 (shaded) has six additional neighbouring cells surrounding it. This concept is extremely important to understand the frequency reuse is concerned. The basic seven cells use 7 frequencies for communication and the next set of seven cells can reuse the same set of frequencies to communicate as there will be sufficient distance between the cells.

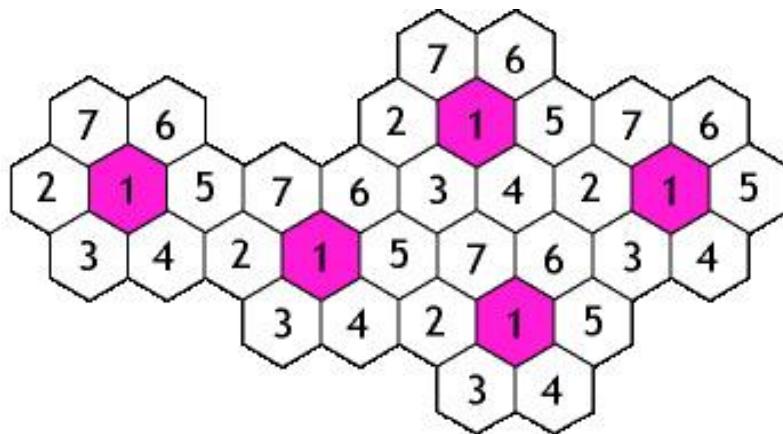


Figure 10.1.1: An illustration of a cellular system

Mobile phone networks are divided into thousands of overlapping, individual geographic areas or cells each with a base station. Each mobile communicates via radio with one or more base stations. An illustration of mobile to mobile communication is as shown in the Figure 10.1.2.



Figure 10.1.2: Illustration of Mobile to Mobile Communication

Each mobile contains a transceiver (transmitter and receiver), an antenna, and control circuitry. The base stations consist of several transmitters and receivers, which simultaneously handle full duplex communications and generally have towers that support several transmitting and receiving antennas. The base station connects the simultaneous mobile calls via telephone lines, microwave links, or fiber-optic cables to the switching center. The switching center coordinates the activity of all of the base stations and connects the entire cellular system to the public telephone network.

The channels used for transmission from the base station to the mobiles are called forward or downlink channels, and the channels used for transmission from the mobiles to the base station are called reverse or uplink channels. The two channels responsible for call initiation and service request are the forward control channel and reverse control channel.

Once a call is in progress, the switching center adjusts the transmitted power of the mobile (this process is called power control) and changes the channel of the mobile and base station (handoff) to maintain call quality as the mobile moves in and out of range of a given base station. A call from a user can be transferred from one base station to another during the call. This process of transferring a call from one base station frequency to another is called handoff.

When roaming users enter an area outside their home region, special procedures are required to provide the cellular phone service. To automatically provide roaming service, a series of interaction is required between the home network and the visited network, using the telephone signaling system. When the roamer enters a new area, the roamer registers in the area by using the setup channels.

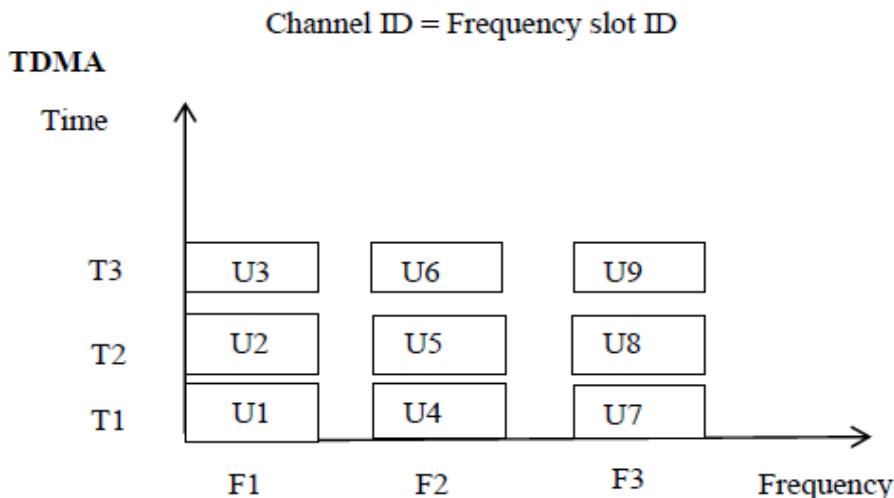
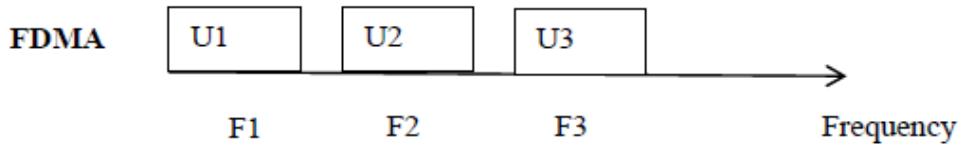
#### **10.1.2. Multiple Access Technology:**

Generally a fixed amount of frequency spectrum is allocated to a cellular system by national regulator (In India, the Department of Telecommunications (DoT) conducts auctions of licenses for electromagnetic spectrum). Multiple access techniques are then deployed so that many users can share the available spectrum in an efficient manner.

In 2010, 3G and 4G telecom spectrum were auctioned in a highly competitive bidding. The winners were awarded spectrum in September, and Tata Docomo was the first private operator to launch 3G services in India. The Government earned 677 billion from the 3G spectrum auction. While the broad band wireless spectrum auction generated a revenue of 385 billion. The Government earned total revenue of over 1062 billion from both auctions. The auction took place over 34 days and consisted of 183 rounds of bidding. The five most expensive circles were Delhi, Mumbai, Karnataka, Tamil Nadu and Andhra Pradesh. They accounted for 65.56% of the total bids. So the spectrum is very scarce and valuable resource.

The three basic multiple access methods are :

- Frequency division multiple access (FDMA)
- Time division multiple access (TDMA)
- Code division multiple access (CDMA)



$$\text{Channel ID} = \text{Frequency slot ID} + \text{time slot ID}$$

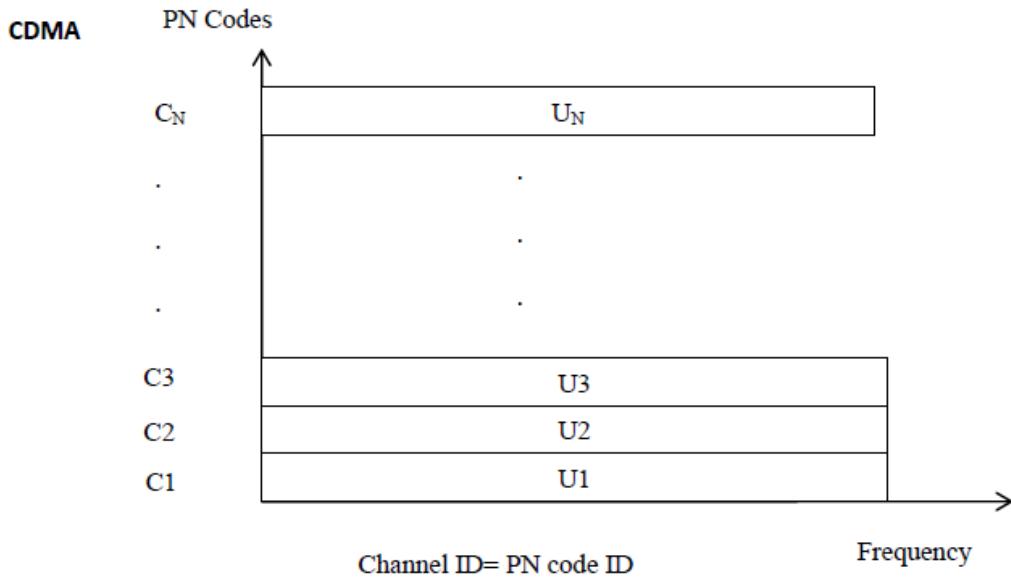


Figure 10.1.3: Basic multiple access techniques

Figure 10.1.3 illustrates the three basic multiple access methods used in mobile communication.

In case of FDMA, users share the available spectrum in the frequency domain and each user is allocated a part of the frequency band, on a demand basis.

In TDMA available spectrum is partitioned into narrow frequency bands (as in FDMA), which in turn are divided into a number of time slots. An individual user is assigned a time slot that permits access to the frequency channel for the duration of the time slot.

The CDMA system utilizes the spread spectrum technique, whereby a spreading code (called a Pseudo-random Noise or PN code) is used to allow multiple users to share a block of frequency spectrum.

*Self-test:*

1. How call set up takes place between two mobile subscribers?
2. Explain the need for cellular concept in mobile communication.
3. Explain frequency reuse concept in cellular systems.

### 10.1.3. Global System for Mobile Communications (GSM):

Global System for Mobile Communications (GSM), is a standard developed by the European Telecommunications Standards Institute (ETSI) to describe protocols for second generation digital cellular networks used by mobile phones.

**Features:**

- Higher digital voice quality.
- Low cost alternatives to making calls such as SMS
- Ability to deploy equipment from different vendors
- GSM allows network operators to offer roaming services

**GSM Architecture:** Figure 10.1.4 illustrates the architecture of GSM system.

**Mobile Station (MS):** The Mobile Station is made up of two entities:

1. Mobile Equipment (ME):
  - The Mobile Equipment is the hardware used by the subscriber to access the network .
  - Uniquely identified by an IMEI (International Mobile Equipment Identity)
2. Subscriber Identity Module (SIM):
  - Smart card containing the International Mobile Subscriber Identity (IMSI)
  - Allows user to send and receive calls and receive other subscribed services
  - Protected by a password or personal identification number (PIN)

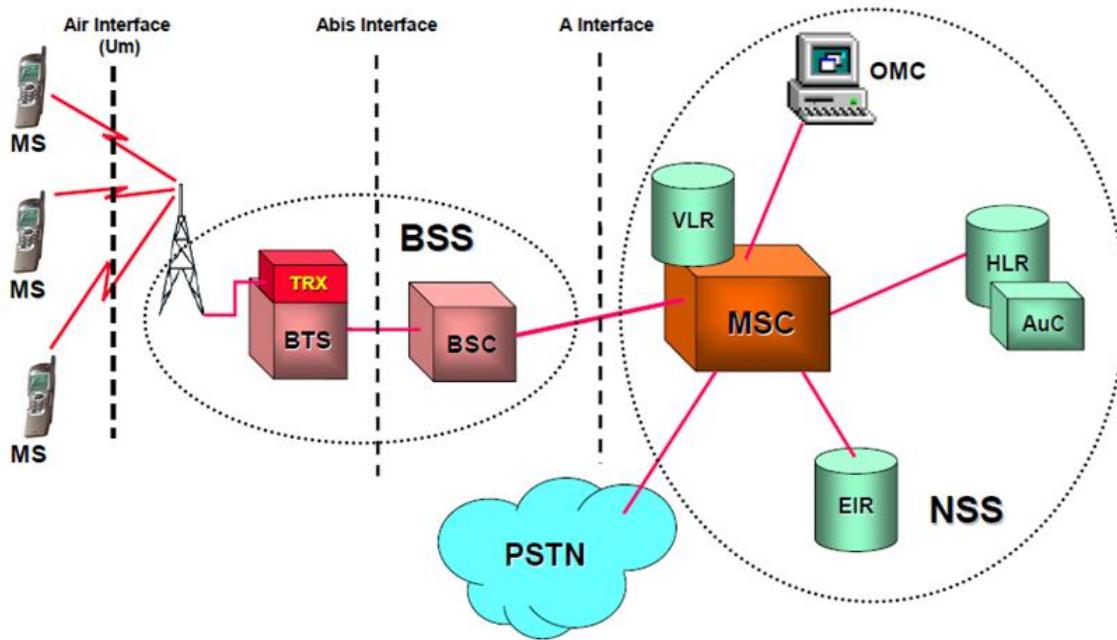


Figure 10.1.4:GSM architecture

**Base Station Subsystem (BSS):** Base Station Subsystem is composed of two parts:

- Base Transceiver Station (BTS)
- Base Station Controller (BSC)

**Base Transceiver Station (BTS):** Houses the radio transceivers that define a cell and handles radio-link protocols with the Mobile Station

**Base Station Controller (BSC):** The tasks performed by BSC are

- Manages Resources for BTS
- Handles call set up
- Location update
- Handover for each MS

**Network subsystem (NSS):** The NSS provides the link between the cellular network and the public switched telephone network (PSTN). The NSS controls handoffs between cells in different BSSs, authenticates users and validates their accounts, and includes functions for enabling worldwide roaming of mobile users.

**Mobile Switching Center (MSC):**

- The central component of the Network Subsystem
- Handles billing activities
- Handover management
- Communication with HLR, VLR, MSC'S
- Controlling of connected BSC'S

**Home Location Registers (HLR):** The HLR stores information, both permanent and temporary, about each of the subscribers that “belongs” to it. It is the most important database.

**Visitor Location Registers (VLR):** The VLR maintains information about subscribers that are currently physically in the region covered by the switching center.

**Authentication Center (AuC):** This database is used for authentication activities of the system; for example, it holds the authentication and encryption keys for all the subscribers in both the home and visitor location registers.

**Equipment Identity Register (EIR):** The EIR keeps track of the type of equipment that exists at the mobile station. It also plays a role in security e.g., blocking calls from stolen mobile stations and preventing use of the network by stations that have not been approved.

### Summary

1. A cellular radio system provides a wireless connection to the public telephone network for any user location within the radio range of the system.
2. Mobile phone networks are divided into thousands of overlapping, individual geographic areas or cells each with a base station.
3. The channels used for transmission from the base station to the mobiles are called forward or downlink channels, and the channels used for transmission from the mobiles to the base station are called reverse or uplink channels.
4. A call from a user can be transferred from one base station to another during the call. The process of transferring is called handoff.
5. The three basic multiple access methods are: Frequency division multiple access (FDMA), Time division multiple access (TDMA), Code division multiple access (CDMA).
6. Global System for Mobile Communications (GSM), is a standard developed by the European Telecommunications Standards Institute (ETSI) to describe protocols for second generation digital cellular networks used by mobile phones.

### Exercise:

1. Compare different generations of cellular mobile communication in terms of underlying modulation schemes, data rate, and applications offered etc.
2. Compare CDMA with that of GSM system.
3. Explain the process of how a call is established and routed in GSM