Syllabus

F.Y.B.Sc. (IT)

Electronics and Communication Technology

Unit I : Concept of Conductor, Senmiconductor, Insulator, Semiconductor Diode, Forward bias, Reverse Bias, Application of Diode as Rectifier, Zener diode and its applications, Introduction to Transistor (BJP, FET), PNP, NPN Transistors their Characteristic. Application to Transistor as amplifier and as a Switch.

Unit II : Concept of amplification, amplifier notations, Av, Ai, Ap Zi, Zo), Application of BJT as single stage Amplifier, Frequency response of single stage Amplifier. Multistage Amplifiers :- (Basics concepts) RC coupled, cascade, Darlington pair, DC amplifier.

Unit III : Concept of Feeback :- Negative Feedback and its advantage in Amplification, Positive Feedback :- Oscillators, RC Phase Shift Oscillator, LC Oscillator, Switching Circuits Multivibrators :- Monostable using IC 555 and Astable using IC 555 (including problems)

Unit IV : Introduction :- Need for modulation system, Concept of Modulation. AM, Definition of AM, Modulation index, Power relation in AM, Generation, Demodulation of AM. SSB :- Power requirement in comparison with AM, Advantages of SSB over AM, Concept of Balanced Modulator, Generation SSB, Pilot Carrier System, Independent Side System, Vestigial Sideback Transmission.

Unit V : FM :- Definition of FM, Bandwidth, Noise triangle, Per-emphasis and De- emphasis. PM :- Definition of PM. Difference between AM and FM. Radio receivers. Pulse Modulation :- Sampling Theorem, PAM, PTM, PWM, PP. pulse code modulation, Quantization noise, companding, PCM system differential PCM, Delta modulation. Multiplexing :- FDM/ TDM. Television Scanning, Composite Video Signal, Television Transmitter, Television receiver.

Unit VI : Introduction to Digital Communication : PSK, ASK, FSK. Introduction to fibre optics system : Propagation of light in optical fibre; Types of fibre : Single mode, steps index. Graded index, Signal distortion: attenuation, dispersion, Optical sources : LED, LASETS. Optics Detectors and optics links. Link Budget.

References :

Allen Mottershead, “Electronic Devices and Circuits”, PHI

Boylstead and Neshelesky, “Electronics Devices and Circuits”, 4th, PHI, 1999.

Simon Haykin, “An Introduction to Analog and Digital communications”, John Siley and Sons.

R.B. Carlson, “Communication System”, MacGraw Hill.

George Kennedy, “Electrical Communication Systems”, Tata McGraw Hill 1993.

Roody Collin, “Electronics Communication”, PHI

J. Millman and A Grabel, “Microelectronics” MacGraw Hillm, 1998.

Proakis J.J., “Digital Communications” McGraw Hill.

Digital Communications by TAUB Schilling

Electronic Communication Systems, Roy Blake Delmar, Thompson Learning

Introduction to telecommunications, Anu A Gokhale, Delmar Thompson Learning

Term Work and tutorial

Should contain 5 assignments and two class tests

Practical : Should contain minimum 8 experiments.

List of Practicals :

1. Study of Zener diode characteristics
2. Study of Half wave and full wave rectifiers
3. Study of bridge rectifier
4. Study of Transistor as a switch
5. Monostable multivibrator using IC 555 timer
6. Astable multivibrator using CI 555 timer
7. Study of Wien bridge oscillator
8. Frequency Response of single transistor amplifier
9. Study of Amplitude Modulation
10. Study of Frequency Modulation
11. Study of Fibre Optic transmission
12. Study of Pulse Amplitude Modulation
13. Study of transistor DC Amplifier

❖❖❖❖

UNIT I SEMICONDUCTOR DEVICES

1

SEMICONDUCTOR ELECTRONICS PART (I)

Unit structure :

1.0 OBJECTIVE :

In this lesson we are focussing mainly on the semiconductors. The conduction phenomenon in semiconductors. Their uses and different applications in electronics such as applications of diodes as rectifiers , Zener diode and its use as regulator.

1.1 INTRODUCTION :

- A solid is a large collection of atoms. The energy levels of an atom get modified due to the presence of other surrounding atoms and the energy levels in the outermost shells of all the atoms form valence band and the conduction band separated by a **forbidden energy gap.**
- The energy band formed by a series of energy levels containing valence electrons is called **valence band.** At 0 K, the electrons start filling the energy levels in valence band starting from the lowest one. The highest energy level, which an electron can occupy in the valence band at 0 K, is called **Fermi level.** The lowest unfilled energy band formed just above the valence band is called conduction band.
- At 0 K, the Fermi level as well as all the lower energy levels are completely occupied by the electrons. As the temperature rises, the electrons absorb energy and get excited. The excited electrons jump to the higher energy levels. These electrons in the higher energy levels are comparatively at larger distances from

the nucleus and are more free as compared to the electrons in the lower energy levels.

- Depending upon the energy gap between valence band and the conduction band, the solids behave as **conductors, insulators and semiconductors.**

1.2 CLASSIFICATIONS IN SOLIDS

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

1.2.1 Metals (Conductors) :

- Conductors *(e.g.* copper, aluminium) are those substances which easily allow the passage of electric current through them. It is because there are a large number of free electrons available in a conductor.
- In terms of energy band, the valance and conduction bands overlap each other as shown in Fig. **1.1**. Due to this overlapping, a slight potential difference across a conductor causes the free electrons to constitute electric current. Thus the electrical behaviour of conductors can be satisfactorily explained by the band energy theory of solids.
- For example, (i) in sodium, the conduction band is partially filled, while the valence band is completely filled. (ii) The valence band is completely filled and the conduction band is empty but the two overlap each other. Zinc is an example of band overlap metals.

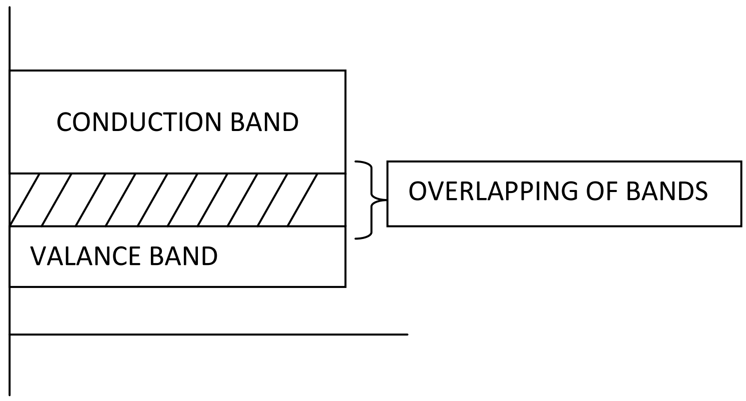


Fig. 1.1 Energy Bands in Metals (Conductors)

- In both the situations, it can be assumed that there is a single energy band, which is partially filled. Therefore, on applying even a small electric field, the metals conduct electricity.

1.2.2 Insulators :

- Insulators are those substances which do not allow the passage of electric current through them.
- In terms of energy band, the valence band is full while conduction band is empty.
- Further, the energy gap between valence and conduction band is very high (say 6 eV or above). Therefore, a very high electric field is required to push the electrons to the conduction band. For this reason, the electrical conductivity of such materials is extremely small and may be regarded as nil.

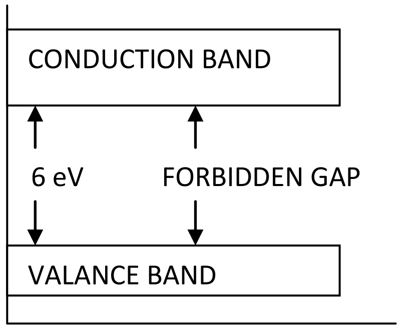


FIG. 1.2 Energy band in insulators

- In insulators, the forbidden energy gap is quite large For example, the forbidden energy gap for diamond is 6 eV, which means that a minimum of 6 eV energy is required to make the electron jump from the completely filled valence band to the conduction band. When electric field is applied across such a solid, the electrons find it difficult to acquire such a large amount of energy and so the conduction band continues to be almost empty. No electron flow occurs i.e. no current flows through such solids. So they behave as insulators.

1.2.3 Semiconductors.

- Semiconductors *(e.g.* germanium,. silicon etc.) are those substances whose electrical conductivity lies in between conductors and insulators.
- In terms of energy band, the valence band is almost filled and conduction band is almost empty. The energy gap between valence and conduction bands is very small as shown in Fig. therefore comparatively smaller electric field (smaller than insulators but much greater than conductors) is required to push the electrons fro valence band to the conduction band. In short, a semiconductor has :
    I) almost full valence band
    II) almost empty conduction band
    III) small energy gap(1 eV) between valence and conduction bands.
- At low temperature, the valence band is completely full and conduction band is completely empty. Therefore, a semiconductor virtually behaves as an insulator at low temperatures.
- The energy band structure of the semiconductors is similar to the insulators but in their case the size of the forbidden gap is much smaller than insulators. For examples forbidden gap os silicon is 1.1 eV and that of germanium is 0.69eV.

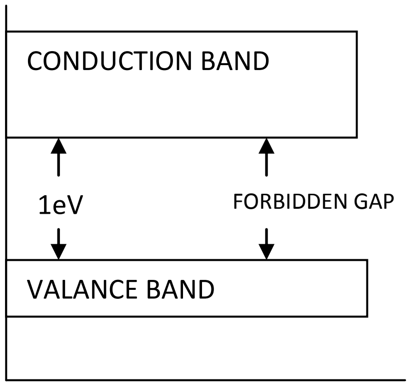


Fig. 1.3 Energy bands in semiconductors

- In semiconductors due to smaller width of forbidden energy gap the electrons in find easier to shift to the conduction band. So the conductivity in semiconductors lies between that of metals and insulators.

1.3 TYPES OF SEMICONDUCTORS :

The semiconductors are classified as intrinsic and extrinsic semiconductors on the basis of their purity.

1.3.1. Intrinsic Semiconductors :

- In a pure semiconductor, each atom behaves as if there are 8 electrons in its valence shell (due to formation of covalent bonds) and therefore the entire material behaves as an insulator low temperature.

- A semiconductor atom needs energy of the order of 1·1 e V to shake off the valence electron. This energy becomes available to the semiconductor even at room temperature. Due to thermal agitation of the crystal structure, electrons from a few covalent bonds come out. The bond from which electron is freed, a vacancy is created there. The vacancy in the covalent bond (where there should have been an electron) is called a hole. This hole can be filled by some other electron in a covalent bond.

- As an electron from a covalent bond moves to fill the hole, the hole is created in the covalent bond from which the electron has moved. In other words, the hole shifts from one covalent bond to another in a similar way as an electron does in an attempt to fill the hole. Since the direction of movement of the hole is opposite to that of the negative electron, a hole behaves as a positive charge carrier.

- Thus, at room temperature, a pure semiconductor will have electrons and holes wandering in random directions. ***These electrons and holes are called intrinsic carriers and such a semiconductor is called intrinsic semiconductor.***

- As the crystal is electrically neutral, the number of free electrons will be equal to the number of holes. If we apply potential difference across the semiconductor, the electrons will move towards positive terminal and the holes towards the negative terminal of the battery.

- The electrons and holes are not current in themselves but act only as the negative and positive charge carriers of the current respectively.

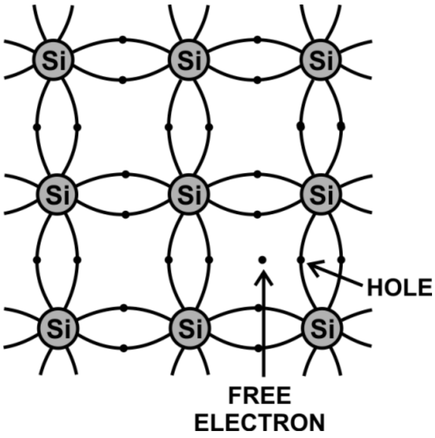


Fig. 1.4 Intrinsic Silicon structure

- Also, when an electron is raised from the valence band to the conduction band, a vacancy created in the valence band. This vacancy created in the valence band (where electron was present before moving to conduction band) acts as the hole.

<u>1.3.2 Extrinsic Semiconductors :</u>

- A pure semiconductor at room temperature possesses free electrons and holes but their number is so small that conductivity offered by the pure semiconductor cannot be made of practical use.

- By the addition of impurities to the pure semiconductor in a very small ratio, its conductivity can be remarkably improved. **<u>The process of adding impurity to a pure semi conductor crystal (Si or Ge-crystal) so as to improve its conductivity, is called doping.</u>**

- The impurity atoms are of two types:

I. **<u>Pentavalent</u>** impurity atoms i.e. atoms having 5 valence electrons such as antimony (Sb) or arsenic (As). Such atoms, when added to a pure semiconductor, produce excess of free electrons i.e. donate electrons to the semiconductor. For this reason, pentavalent impurity atoms are called **donor impurity** atoms. The semiconductor so produced is called **n-type extrinsic semiconductor**.

II. **<u>Trivalent</u>** impurity atoms i.e. atoms having 3 valence electrons such as indium (In) or gallium such atoms on being added to a pure semiconductor, instead of producing free electrons, accept electrons from the semiconductor. For this reason, trivalent impurity atoms are called **acceptor impurity** atoms. The semiconductor so produced is called **p-type extrinsic semiconductor** .

1.3.2.1 n-Type Semiconductor :

- Fig 1.5 shows the effect of adding pentavalent impurity arsenic to silicon crystal.
- When the arsenic impurity atoms are added to the silicon crystal in a small, its atoms replace the silicon atoms here and there. The four electrons out of five valance electrons of As-atom take part in covalent bonding with four silicon atoms surrounding it. The fifth electron is set free.
- Obviously the extra free electrons created in the crystal will be as many as the number of the pentavalent impurity atoms added. As the pentavalant impurity increases the number of free electrons also increases, hence, it is called **donor impurity.**
- The silicon crystal so obtained is termed as **n-type** Si-crystal. The electrons so set free in the silicon crystal are called **extrinsic carriers** and the n-type Si-crystal is **n-type extrinsic semiconductor.**

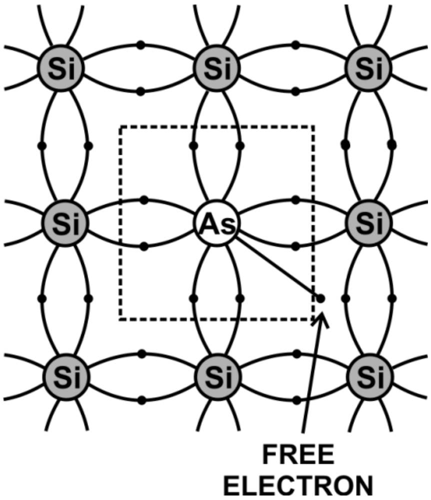


Fig. 1.5 Si-Crystal as n-type Semiconductor

- Due to thermal agitation, even the pure silicon crystal possesses a few electrons and holes. Therefore, n- type Si-crystal will have a large number of free electrons (majority carriers) and small number of holes (minority carriers).

1.3.2.2 p-Type Semiconductor :

- Fig 1.6 shows the effect of adding trivalent impurity indium (In) to silicon crystal.
- The four silicon atoms surrounding the In-atom can share one electron each with the In-atom, which has got three valance electrons. All three of the In-atom's valence electrons are used in the covalent bonds; and, since four electrons are required, a hole results when each trivalent atom is added.
- Thus, for every trivalent impurity atoms added, an extra hole will be created. As the impurity atoms accept electrons from the silicon it is called **acceptor impurity**. The Si-crystal so obtained is called **p-type** as it contains free holes. Each hole is equivalent to positive charge. The holes so created are **extrinsic**

carriers and p-type Si-crystal so obtained is called **p-type extrinsic semiconductor.**

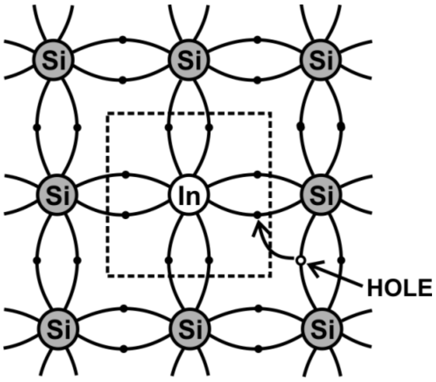


Fig. 1.6 p-Type Si semiconductor

1.4 FORMATION OF p-n JUNCTION :

- A p-n junction is a basic semiconductor device.

- When a p-type crystal is placed in contact with n-type crystal so as to form one piece, the assembly so obtained is called **p-n junction or junction diode or crystal diode**. The surface of contact of p and n- type crystals is called **junction.**
- In the p-section, holes are the majority carriers; while in n-section the majority carriers are electrons.
- Due to the high concentration of different types of charge carriers in the two sections, holes from p-region diffuse into n-region and electrons from n-region diffuse into p-region. In both cases, when an electron meets a hole, the two cancel the effect of each other and as a result, a thin layer at the junction becomes devoid of charge carriers. This is called **depletion layer**.
- The diffusion continues back and forth until the number of electrons which have crossed the junction have a large enough electrical charge to repel or prevent any more charge carriers from crossing over the junction. Soon enough, a state is reached where the pn junction is electrically neutral due to the creation of a potential barrier. The potential difference developed across the junction due to migration of majority charge carriers is called **potential barrier**.
- It opposes the further diffusion of charge carriers. The magnitude of the potential barrier is different for ex. For germanium junction diode for silicon junction diode. However, the value of potential barrier depends on the magnitude of doping of the semiconductor crystal.

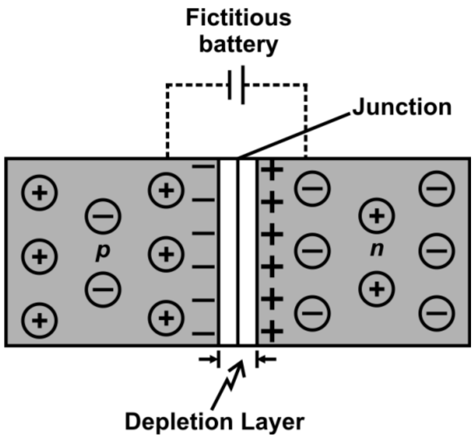


Fig. 1.7 Formation of p-n junction Diode.

- It may be pointed out that across the junction; a very large electric field is set up due to potential difference developed across it.

- Thus, the formation of p-n junction results in a very strong electric field (or potential gradient) across the junction.
- The junction diode is represented by the symbol as shown in Fig. 1.10

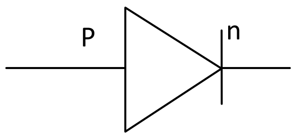


Fig 1.10 Junction Diode

- The arrow-head represents the p-section of the junction diode and points in the direction in which the hole current or conventional current will flow, when junction diode is forward biased. The electron current or the electronic current will flow in opposite direction.

1.5 FORWARD AND REVERSE BIASING ON A JUNCTION DIODE

A junction diode can be biased in the following two ways:

1.5.1. Forward Bias.

- *When an external D.C. source is connected to the junction diode with p-section connected to positive pole and n-section to the negative pole, the junction diode is said to be forward biased*.

- **Action of junction diode:**
  I. The p-n junction is forward biased as shown in Fig. 1.8 When forward biased, the positive holes in the p-section are repelled by the positive pole of the battery towards the p-n junction. Simultaneously, the negative electrons in the n-section are repelled by negative pole of the battery towards the junction.

II. However, the movement of electrons and holes across the junction is opposed by the barrier voltage or depletion voltage (0·3 V to 0·7 V) developed across the junction. Just near the p-n junction, electrons and holes combine and cease to exist as mobile charge carriers after the potential barrier is overcome by the applied potential.

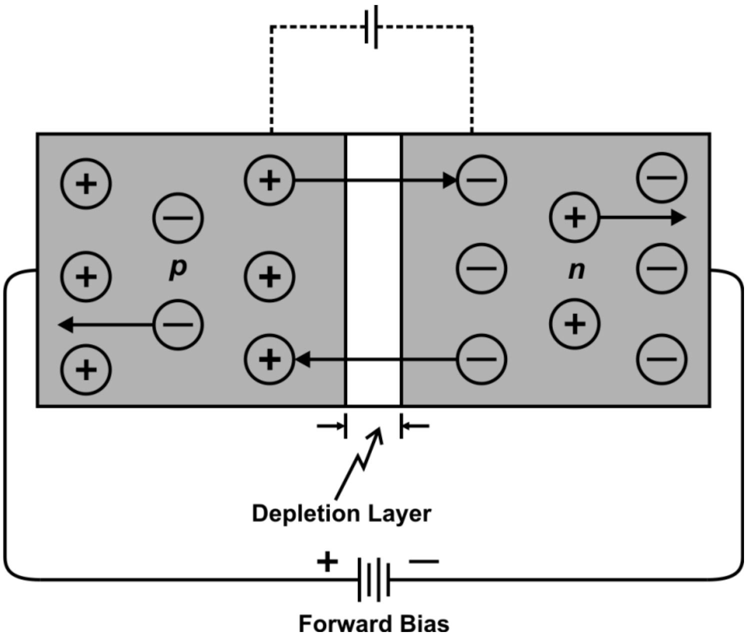


Fig. 1.8 Forward Biased p-n junction diode

III. For each electron-hole combination that takes place near the junction, a covalent bond breaks in the p-section near the positive pole of the battery. Of the electron and the hole produced, the electron is captured by the positive terminal, while the hole moves towards the junction.

IV. On the other hand, as soon as the hole is created in the p-section due to the breaking of a covalent bond, an electron is released from the negative terminal of the battery into the n-section to replace the electron lost by the combination with a hole at the junction. These electrons move towards the junction, where they again get neutralised on meeting the new holes coming from left. As a consequence, a relatively large current, called **forward current** flows through the junction.

V. The current in the external circuit is due to the electrons and is from negative terminal of battery to positive terminal through the junction diode.

VI. During the forward bias, the applied D.C. voltage opposes the barrier voltage developed across the p-n junction. Due to this, the potential drop across the junction decreases and as a result, the diffusion of holes and electrons across the junction increases. It makes the depletion layer thin and the junction diode offers low resistance during forward bias.

1.5.2 Reverse Bias:

- *When a battery is connected to junction diode with p-section section connected to negative pole and n-section connected to the positive pole, the junction diode is said to be reverse biased*.

- **Action of junction diode**.

I. When the p-n junction is reverse biased as shown in Fig. 1.9 the holes (majority carriers) in the p-section get attracted towards the negative terminal of battery and therefore, the holes move away from the junction. At the same time, the electrons (majority carriers) in the n-section get attracted towards the positive terminal and move away from the junction.

II. As a very small number of holes and electrons (minority carriers) are left in the vicinity of the junction, practically no flow of current cakes place. However, due to thermally generated electron-hole pairs within p-region as well as n-region, a small current (a few microamperes) still flows. Some covalent bonds always break because of the normal heat energy of the crystal molecules. Electrons liberated by this process in the p-region move to the left across the junction, while holes generated in the n-region move to the right under the electric field produced by the battery.

III. Thus, a small electron-hole combination current, called **reverse current** is maintained by the minority carriers. If the reverse bias is made very high, all the covalent bonds near the junction break and a large number of electron-hole pairs are liberated and the reverse current increases abruptly to a relatively high value.

IV. The maximum reverse potential difference, which a diode can tolerate without breakdown is called **reverse break down voltage or zener voltage**.

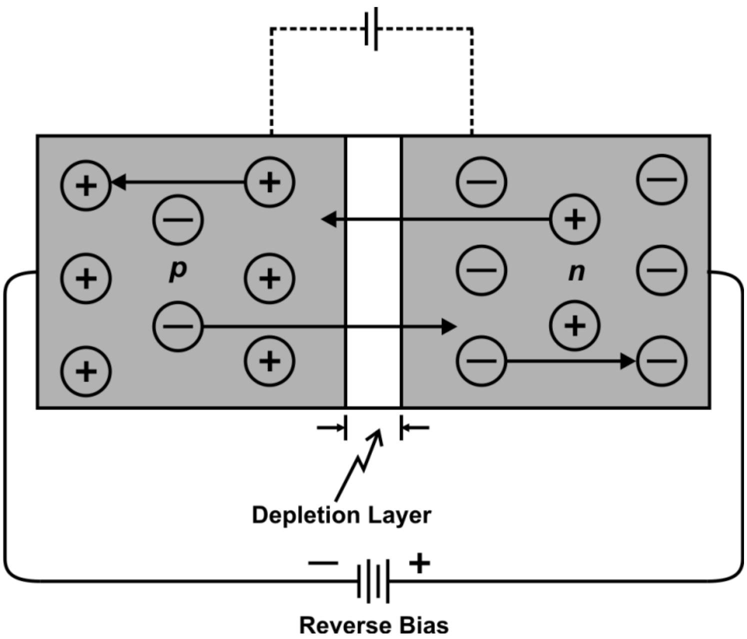


Fig 1.9 reverse Biased p-n junction diode

V. During the reverse bias, the applied D.C. voltage adds to the barrier voltage developed across the junction. Due to this, the potential drop across the junction increases and as a result, the diffusion of holes and electrons across the junction decreases. It makes the depletion layer thick and the junction diode offers high resistance during reverse bias.

VI. It may be noted that the potential barrier opposes the forward current, while it adds the reverse current.

1.6 ADVANTAGES OF SEMICONDUCTORS :

Because of their merits over vacuum tubes, semiconductor devices (junction diodes, transistors, integrated circuits) have practically completely replaced them in all the fields of electronics. Some of the advantages of the semiconductor devices are as given below:

Advantages:

1. As semiconductor devices have no filaments, hence no power is needed to heat them to cause the emission of electrons.
2. Since no heating is required, semiconductor devices are set into operation as soon as the circuit is switched on.
3. During operation, semiconductor devices do not produce any humming noise.
4. Semiconductor devices require low voltages for their operation as compared to vacuum tubes.
5. Owing to their small sizes, the circuits involving semiconductor devices are very compact.
6. Semiconductor devices are shock proof.
7. Semiconductor devices are cheaper as compared to vacuum tubes.
8. Semiconductor devices have almost unlimited life.
9. As no vacuum has to be created in semiconductor devices, they have no vacuum deterioration trouble.

Disadvantages:

1. Noise level is higher in semiconductor devices as compared to that in the vacuum tubes.
2. Ordinary semiconductor devices cannot handle as much power as ordinary vacuum tubes can do.
3. In high frequency range, they have poor response.
4. The semiconductor devices are temperature-sensitive. The maximum temperature, the semiconductor devices can withstand, is very low (about 50° C). Even a small over-heating damages the semiconductor device. This is because,

at a higher temperature, the covalent bonds break up and the semiconductor material forming the semiconductor device becomes conducting.

1.7 JUNCTION DIODE AS RECTIFIER :

- **An electronic device which converts A.C. power into DC. power is called a rectifier.**
- The junction diode offers a low resistance path, when forward biased and a high resistance path when reverse biased. This feature of the junction diode enables it to be used as a rectifier.
- The two half cycles of alternating input e.m.f. provide opposite kinds of bias to the junction diode.
- If the junction diode gets forward biased during first half cycle, it will get reverse biased during the second half cycle and vice-versa.
- In other words, when an alternating e.m.f. signal is applied across a junction diode, it will conduct only during those alternate half cycles, which bias it in forward direction.

1.7.1. HALF WAVE RECTIFIER :

I. ***A rectifier, which rectifies only one half of each A.C. input supply cycle, is called a half wave rectifier***.

II. **Principle**: It is based on the principle that junction diode offers low resistance path, when forward biased and high resistance when reverse biased. When A.C. input is applied to a junction diode it gets forward biased during one half cycle and reverse biased during the next opposite half cycle. Thus output is obtained during alternate half cycles of the A.C. input.

III. **Arrangement**: The A.C. supply is fed across the primary coil P of a step-down transformer. The secondary coil S of the transformer is connected to the junction diode and a load resistance R_L shown in Fig. 1.11. The output D.C. voltage is obtained across the load resistance R_L.

IV. **Theory**:
   - Suppose that during the first half of the input cycle, the junction diode gets forward biased. The conventional current will flow in the direction of the arrow heads.
   - The upper end of R_L will be at positive potential w.r.t. the lower end. The magnitude of output across R_L during first half cycle at any time will be proportional to the magnitude of current through
   - Hence, during the first half of the input cycle, when junction diode conducts, output across R_L vary in accordance with A.C. input.

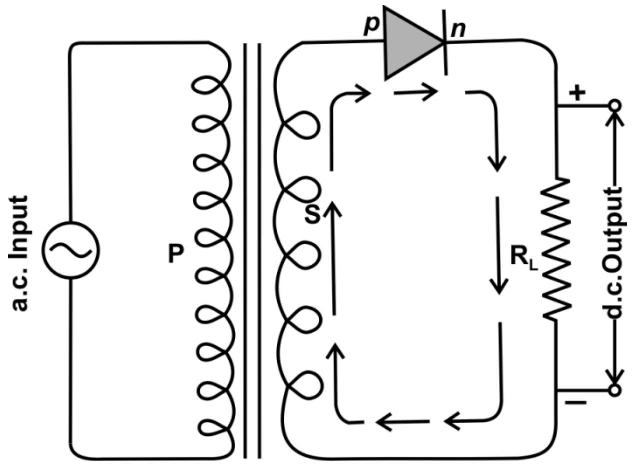


Fig 1.11A Input and Output wave forms of half wave rectifier

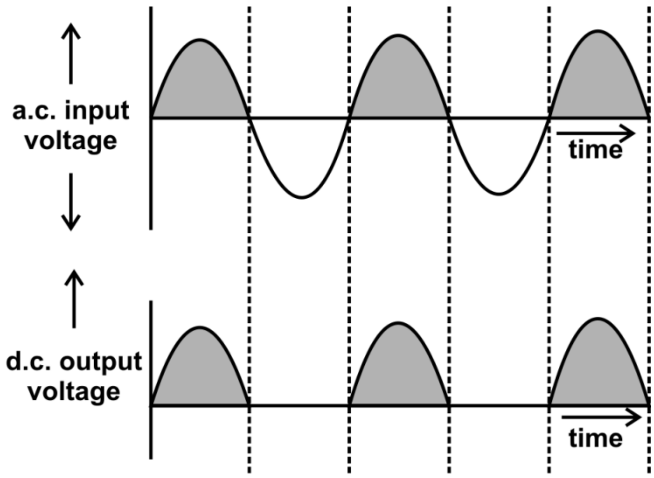


Fig. 1.11B Half Wave rectifier

- During the second half cycle, junction diode will get reverse biased and hence no output will be obtained across R_L.
- Critically, a small current will flow due to minority carriers and a negligible output will be obtained during this half cycle also.
- During the next half cycle, output is again obtained as the junction diode gets forward biased.
- Thus a half wave rectifier gives discontinuous and pulsating output across the load resistance as shown in Fig. Hence half wave rectification involves a lot of wastage of energy and hence it is not preferred.

<u>1.7.2. FULL WAVE RECTIFIER :</u>

I. ***A rectifier which rectifies both halves of each A.C. input cycle is called a full wave rectifier*.**
To make use of both the halves of input cycle, two junction diodes are used.

II. **Principle.** It also works on the principle that a junction diode offers low resistance during forward bias and high resistance, when reverse biased. Here, two junction diodes are connected in such a way that if one diode gets forward biased during first half cycle of A.C. input, the other gets reverse biased but when the next opposite half cycle comes, the first diode gets reverse biased and the second forward biased. Thus, output is obtained during both the half cycles of the A.C. input.

III. **Arrangement:** The a.c. supply is fed across the primary coil P of a step-down transformer. The two ends of the secondary coil S of the transformer are connected to the p-sections of the junction diodes D1 and D2. A load resistance R_L is connected across the n-sections of the diodes and the central tapping of the secondary coil. The d.c. output will be obtained across load resistance R_L.

IV. **Theory:** Suppose that during first half of the input cycle upper end of coil S is at positive potential and the lower end is at negative potential, the junction diode D1 will get forward biased, while the diode D2 reverse biased. The conventional current due to the diode D1 will flow along the path of full arrows.

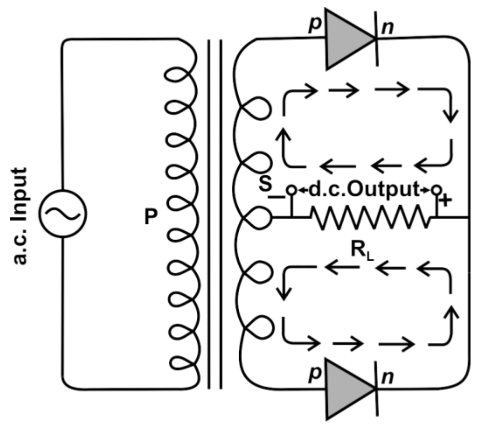


Fig 1.12A input and output waveform of full wave rectifier

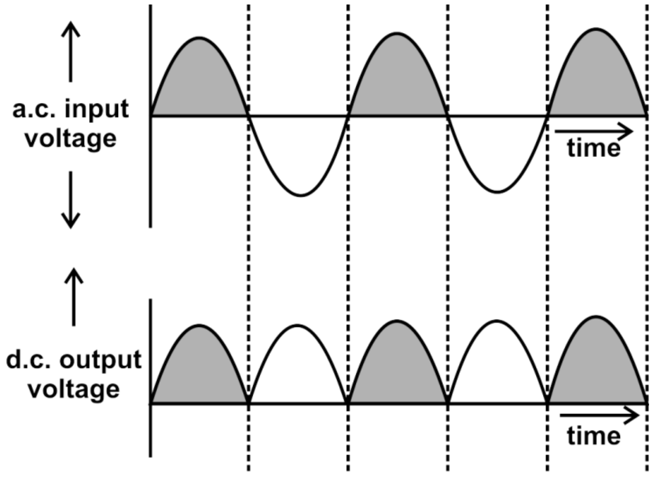


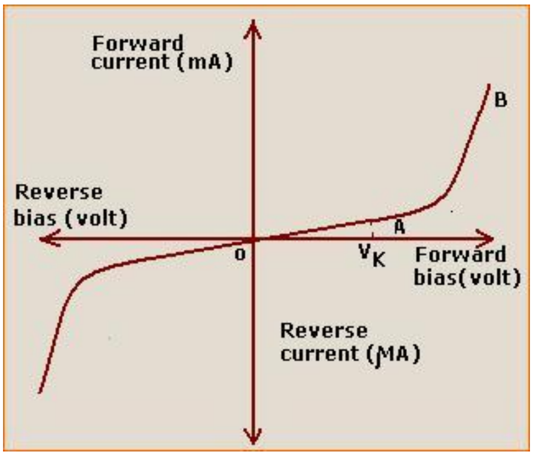
Fig.1.12B Full Wave rectifier

V. When the second half of the input cycle comes, the situation will be exactly reverse. Now, the junction diode D2 will conduct and the conventional current will flow along the path of the dotted arrows.

VI. Since current during both the half cycles flows from right to left through the load resistance R_L, the output during both the half cycles will be of the same nature. The right end load resistance R_L will be at positive potential w.r.t. its left end.

VII. Thus, in a full wave rectifier, the output is continuous but pulsating in nature. However, it can be made smooth by using a filter circuit.

1.8 VOLT AMPERE CHARACTERISTICS OF A PN JUNCTION:



The generalized voltage-current characteristic for a p-n junction in Figure above shows both the reverse-bias and forward-bias regions.

a. **At zero voltage:** The barrier does not permit any current to flow through it.
b. **Forward-bias:** Current rises rapidly as the voltage is increased and is quite high.
c. **Reverse-bias:** The junction offers a very high resistance called **reverse resistance**.
   Some amount (very small) of free holes and electrons still manage to cross the junction and constitute a **reverse current**

Other Important terms:

1. If the reverse bias is made very high, a large number of electron-hole pairs are created and the reverse current increases to a relatively high value. The maximum reverse potential difference, which a diode can tolerate without breakdown is called **reverse break down voltage or zener voltage**.
   In other words, the minimum reverse voltage at which a pn junction breaks down is called the breakdown voltage.
2. **Knee Voltage:** The voltage at which the pn junction begins to conduct current and shows rapid rise in the current.
3. **Maximum forward voltage:** The highest forward current that the pn junction can conduct without any damage to the junction
4. **Peak Inverse Voltage(PIV):** It is the maximum reverse voltage that can be applied to a pn junction without any damage to the junction
   Beyond PIV, the junction diode is destroyed due to excessive heat.
5. **Maximum power rating:** It is the maximum power that can be dissipated through the junction without damaging it.

It is equal to the product of junction current and voltage across the junction.

1.9 ZENER DIODE :

I. A conventional diode does not permit large current to flow when in reverse biased. When a p-n junction is reverse biased, the current through the junction is very small. However if the magnitude of reverse bias reaches a critical value, avalanche breakdown may take place. Thus a rapid avalanche breakdown occurs and the diode conducts a large current in the reverse biased mode and may get damaged permanently.

II. However, diodes may be specially built to operate in the breakdown region. By varying the degree of doping, diodes the specific breakdown voltages (ranging from about one to several hundred volts) can be fabricated. If the junction is well-designed, the breakdown will be very sharp and the current after the breakdown will be independent of voltages; such diodes designed for a specific breakdown voltage are shown as Zener diodes.

III. They are useful in voltage regular circuits. As the load current or supply voltage changes, the current through a Zener diode will accommodate itself to these changes to maintain a constant load voltage. The upper limit on the diode current is determined by the power dissipation rating of the diode.

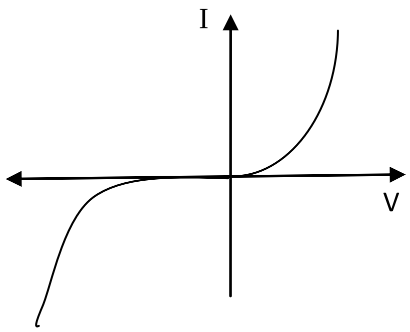


Fig 1.13 characteristic curve of Zener diode

IV. The characteristics curve of Fig. 1.13 has three regions viz. forward, leakage and breakdown. In the diode forward region, it starts conducting at 0.7 V as any other silicon diode. The region between zero and breakdown is the leakage region and only a small reverse current flows in this region. The breakdown region is very sharp. When the voltage reaches - 15V, the characteristics becomes almost vertical and the voltage becomes constant at - 15V.

V. The minus sign in the specification of the breakdown voltage does not have any significance. It only indicates that the Zener diode is reverse biased. It is preferable to say that the Zener diode has a breakdown voltage of (say) 15 V.

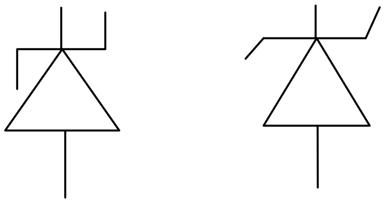


Fig. 1.14 Symbols of zener diodes.

VI. Figure 1.14 shows two symbols for a Zener diode. - the symbols the lines resemble the alphabet 'z' which stands for Zener.

1.9.1 Zener Regulator :

I. Fig 1.15 (a) shows a simple circuit using zener regulator. The series resistance R_S limits the current to less than the maximum current rating of zener. It is seen that

II. $$I_s = \frac{V_S - V_Z}{R_S}$$

III. Fig 1.15 (b) shows a rectifier circuit along with a zener and series resistor. As the rectifier output changes the current through the circuit changes but the output voltage remains constant equal to V_Z.

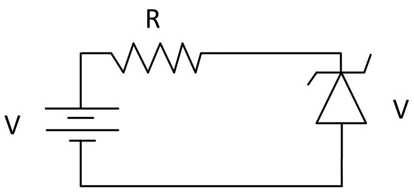


Fig 1.15 (a) Zener regulator

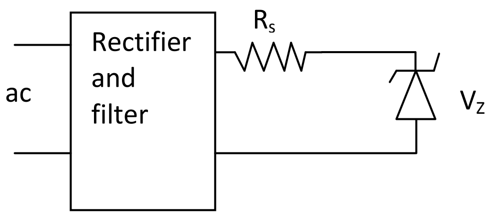


Fig. 1.15 (b) Rectifier with zener regulator

IV. An ideal Zener diode has zero resistance and the voltage across it remains absolutely constant. In actual practice the
Zener has some resistance and small voltage drop occurs across it. For most of calculations this small voltage drop can be neglected.

V. It is necessary to ensure that Zener operates in the breakdown region. It is only in this region that the voltage across the Zener is constant.

1.10 SUMMARY :

- The energy levels of an atom get modified due to the presence of other surrounding atoms and the energy levels in the outermost shells of all the atoms form valence band and the conduction band separated by a forbidden energy gap.
- Depending upon the energy gap between valence band and the conduction band, the solids behave as conductors, insulators and semiconductors

- Conductors *(e.g.* copper, aluminium) are those substances which easily allow the passage of electric current through them.
- Insulators are those substances which do not the passage of electric current through them.
- A semiconductor has :
  - I) almost full valence band
  - II) almost empty conduction band
  - III)small energy gap(1 eV) between valence and conduction bands.
- The chemically pure semiconductor is called intrinsic semiconductor and chemically impure is called as extrinsic semiconductor.
- The process of adding impurity to a pure semi conductor crystal (Si or Ge-crystal) so as to improve its conductivity is called doping.
- When a p-type crystal is placed in contact with n-type crystal so as to form one piece, the assembly so obtained is called **p-n junction or junction diode or crystal diode**. The surface of contact of p and n- type crystals is called **junction.**
- *When an external d.c. source is connected to the junction diode with p-section connected to positive pole and n-section to the negative pole, the junction diode is said to be forward biased*
- *When a battery is connected to junction diode with p-section section connected to negative pole and n-section connected to the positive pole, the junction diode is said to be reverse biased*.
- An electronic device which converts a.c. power into d.c. power is called a rectifier.

1.11 UNIT END EXERCISE:

1. Distinguish between metals, insulators and semiconductors on the basis of energy band structures.
2. Explain how intrinsic semiconductors are converted into n-type and p-type. Give example.
3. Distinguish between p-type and n-type semiconductors.
4. Explain the formation of p-n junction and the terms depletion region and potential barrier for p-n junction.
5. With the help of diagram explain the use of junction as reverse biased.
6. Describe the working of the forward biased p-n junction diode.
7. State any four / five advantages of semiconductors
8. What is rectification? With the help of diagram explain (1) half wave rectifier and (2) full wave rectifier.
9. What is Zener diode/ explain Zener diode as regulator?

❖❖❖❖

2

SEMICONDUCTOR ELECTRONICS PART (II) (Introduction to transistors)

Unit structure :

2.0 OBJECTIVES :

In this lesson we are focussing on particular semiconducting device transistor, its working and few basic features like input and output characteristics, amplifier and its use as digital switch.

2.1 INTRODUCTION:

When a third doped element is added to a semiconductor diode in such a way that two p-n junction are formed, the resulting device is called a transistor. Transistors are smaller than the vacuum tubes, have no filament in it and may be operated in any position. They are mechanically strong have practically unlimited life and is capable of achieving amplification of weak signals.

2.2 TRANSISTORS :

- A junction diode cannot be used for amplifying a signal. For amplification, another type of semiconductor device called transistor is used. It is a three-section semiconductor.

- The three sections are combined, so that the two at extreme ends have the same type of majority carriers; while the section that separates them, has the majority carriers of opposite nature. Therefore, a transistor can be n-p-n or p-n-p type. In other words, in an n-p-n transistor, the p-section is sandwiched between two n-sections .On the other hand, in a p-n-p transistor, the n-section is sandwiched between two p-sections.

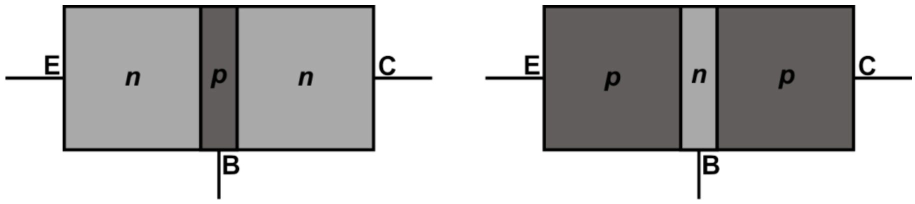


Fig 2.1 Schematic diagram of n-p-n and p-n-p transistors

- The three sections of the transistor are called **emitter (E), base (B) and collector (C).**
- The base of a transistor is made thin. Further, in comparison to emitter and collector, the base of a transistor is lightly doped i.e. the number density of majority carriers in the base is always lesser as compared to that in emitter or collector.
- The emitter supplies the majority carriers for current flow and the collector collects them. The base provides the junctions for proper interaction between the emitter and the collector.
- Symbol for transistors, In the symbol for a transistor, the arrow points hole current i.e. conventional current. Therefore, the emitter in n-p-n transistor is represented by an arrow pointing away from the base, while the emitter in p-n-p transistor is represented by an arrow pointing towards the base. The symbols for n-p-n and p-n-p transistors are respectively shown in Fig. 2.2 (i) and Fig. 2.2 (ii)

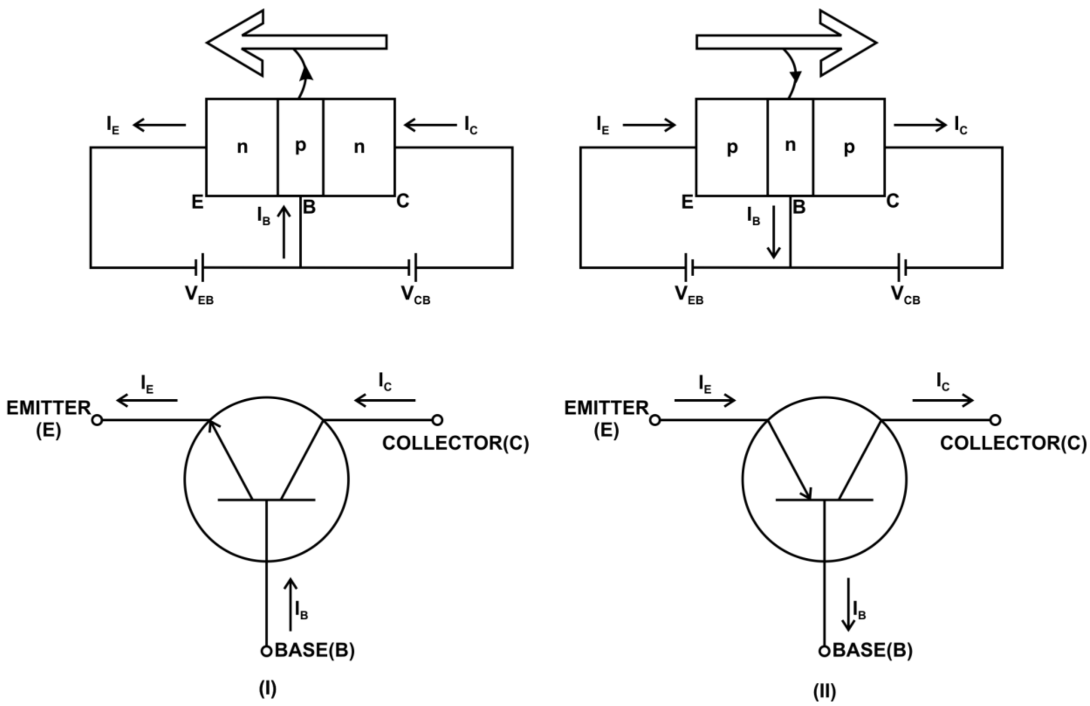


Fig 2.2

- When a transistor is used in a circuit, the base-emitter junction is always forward biased and the base-collector junction is always reverse biased.

2.3. ACTION OFTRANSISTOR

The action of both the types of transistors i.e. n-p-n and p-n-p is similar, except that the majority and minority carriers in the two cases are of opposite nature.

2.3.1 Action of n-p-n transistor:

- Fig. 2.3 shows the proper biasing of an n-p-n transistor. The n-type emitter is forward biased by connecting it to negative pole of the battery Vee (emitter-base battery) and n-type collector is reverse biased by connecting it to the positive pole of the battery Vcc (collector-base battery).

- The majority carriers (which are electrons) in the emitter are repelled towards the base due to the forward bias. The base contains holes as majority carriers but their number density is small as it is doped- lightly as compared to emitter or collector. Due to this, the probability of electron-hole combination in base region is very small (5%). Most of the electrons (95%) cross into collector region, where they are swept away by the positive terminal of the battery Vcc connected to the collector. Fig. 2.3

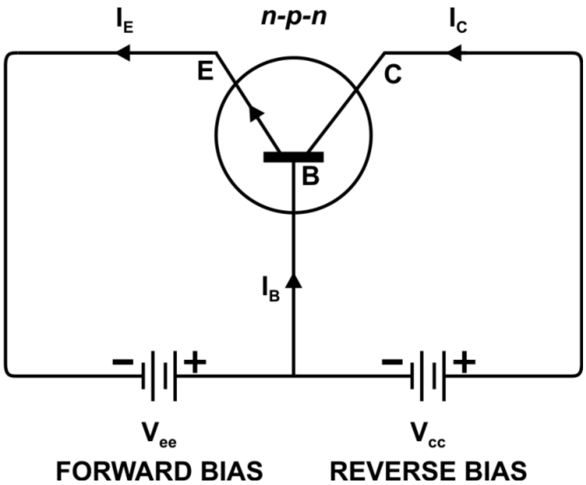


Fig. 2.3 Action of n-p-n transistor

- Corresponding to each electron that is swept by the collector and that enters the positive pole of the collector-base battery Vcc an electron enters the emitter from the negative pole of the emitter-base battery Vee.

- If I_e, I_b and Ic are respectively the emitter current, base current and collector current, then $\mathbf{I_e = I_b + I_c}$

2.3.2 Action of p-n-p transistor.

- The p-type emitter of a p-n-p transistor is forward biased by connecting it to positive pole of emitter-base battery Vee and the p-type collector is reverse biased by connecting it to the negative pole of the collector-base battery Vcc as shown in Fig. 2.4

- In this case, majority carriers in emitter are holes and they are repelled towards the base due to the forward bias. As base is thin and lightly doped (very small as compared to collector and emitter), it has a low number density of electrons.

When holes enter the base region, then only about 5% electron-hole combination takes place. Most of the holes (95%) reach the collector under the influence of reverse bias.

- As one hole reaches the collector, an electron leaves the negative pole of collector- base battery Vcc. At the same time, an electron is released from some covalent bond in the emitter, creating a hole in the emitter. The electron so released, enters the positive pole of the emitter base battery Vee. Thus, current in p-n-p transistor is carried by holes and at the same time their concentration is maintained as explained above.

- In this case also, $\mathbf{I_e = I_b + I_c}$

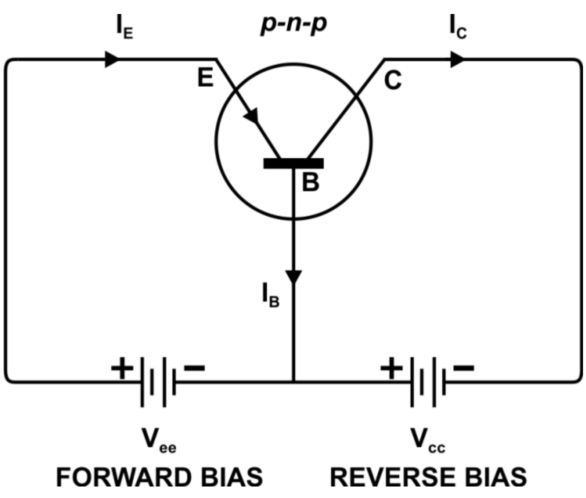


Fig. 2.4 Action of p-n-p transistor

2.4. TRANSISTOR CHARACTERISTICS :

The relationship between the current and voltages in a transistor are represented by curves called characteristics of the transistor. Transistor characteristics are classified as input and output characteristics. We are going to study common emitter characteristics of a transistor.

2.4.1 Common Emitter Characteristics of A Transistor:

- Common emitter characteristics of a transistor are graphs that are obtained between voltages and current, when emitter is earthed, base is used as input terminal and collector as the output terminal. Fig. 2.5 shows the experimental arrangement to study the common emitter characteristics of an n-p-n transistor.

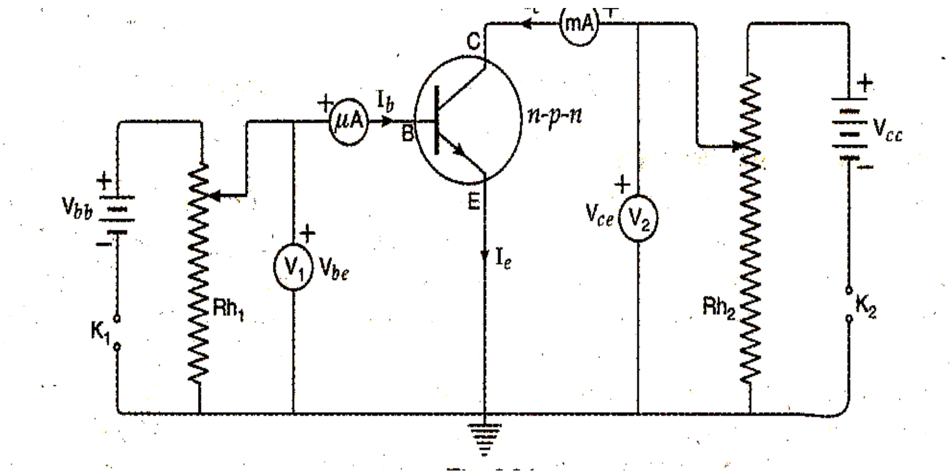


Fig. 2.5 Circuit diagram for CE characteristics of n-p-n transistor

- The base-emitter junction is forward biased by using base-emitter battery V_{bb} while the base-collector junction is reverse biased by using collector-emitter battery V_{cc}.
- The micro- ammeter and voltmeter in base-emitter circuit measure base current (I_b) and the voltage between base and emitter V_{be}
- Similarly, milli-ammeter and voltmeter connected in collector-emitter circuit measure the collector current (Ic) and voltage between collector and emitter (Vce).

(I) Input characteristics:

- **These are graphs between base voltage (V_{be}) and the base current (I_b) at different constant values of collector voltages (V_{ce}).**
- Let us first plot input characteristic at Vce = 2 V. For this, potential divider arrangement in emitter collector circuit is adjusted, till the voltmeter connected in collector-emitter circuit reads 2 V. It will mean that the voltage between collector and emitter is 2 V i.e. collector is at + 2 V w.r.t. emitter of the transistor. Now, by making use of potential divider in emitter-base circuit, make base voltage zero. It will be noted that base current is also zero.
- Now, keeping collector voltage constant at 2V, start increasing base voltage gradually. It will be found that the base current remains zero as long as the base voltage does not exceed barrier voltage (0·3 V or so). As soon as the base voltage becomes greater than the barrier voltage, the current increases slowly and then suddenly to a large value as shown in Fig. 2.6 It is input characteristic at Vce = 2 V.

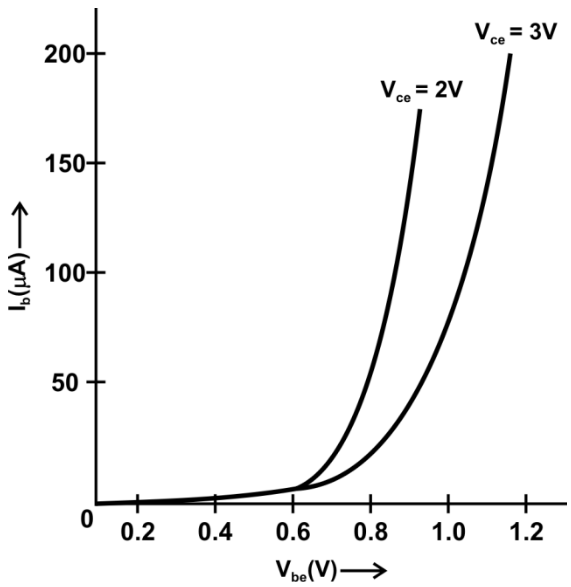


Fig. 2.6 input characteristics of CE transistor

- Similarly, we can draw input characteristic of the transistor at Vce = 3 V and so on. It may be noted that the nature of input characteristics is similar to the forward characteristics of a junction diode.

- **A.C. input resistance.** The a. c. input resistance of the transistor in common emitter configuration is defined as **the ratio of small change in base voltage to the small change produced in base current at constant collector voltage.** It is denoted by R_{in}.

  If ΔV_{be} is small change in base voltage and ΔI_b, small change produced in base current at constant collector voltage Vce, then

$$R_{in} = \left(\frac{\Delta V_{be}}{\Delta I_b}\right)_{V_{ce}}$$

(II) Output characteristics:

These are graphs between collector voltage (Vce) and the collector current (Ic) at different constant values of base current (Ib).

- Let us first plot output characteristic at say I_b = 50 µA. To do so, collector voltage is made zero and voltage between base and emitter is adjusted, till the micro-ammeter in base-emitter circuit reads 50 µA.

- Now collector voltage is increased gradually and the corresponding collector current is noted. Take care that base current always remains steady at 50 µA. Then, graph between Vce and corresponding values of Ic gives output characteristics at I_b = 50 µA.

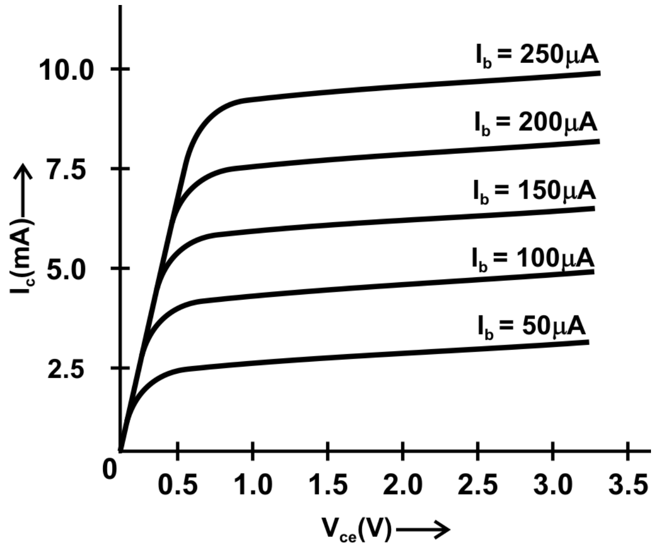


Fig. 2.7 Output characteristics of transistor

- Similarly, output characteristics can be obtained at Ib = 100 µA, 150 µA, 200 µA, Following points may be noted about output characteristics: Fig. 2.7

  (i) The collector current changes rapidly in beginning, but soon it becomes more or less independent of applied collector-emitter voltage.

  (ii) For a given value of collector voltage, larger the base current, larger is the collector current.

  (iii) In audio frequency amplifier circuits, the linear part of the output characteristics is used in order to obtain undistorted output.

- **A.C. output resistance.** The a.c. output resistance of the transistor in common emitter configuration is defined as **the ratio of small change in collector voltage to the small change produced in collector current at constant base current.** It is denoted by $R_{out.}$ Thus,

$$R_{out} = \left(\frac{\Delta V_{ce}}{\Delta I_c}\right)_{I_b}$$

(III). Transfer characteristics.

- **These are graphs between the collector current (Ic) and the base current (I_b) at different constant values of collector voltages (Vce) .**

- Let us plot transfer characteristics at say Vce = 3 V. For this, the potential divider arrangement in emitter-collector circuit is adjusted, so that the voltmeter connected in emitter-collector circuit reads 3V. Now, by making use of potential divider arrangement in emitter-base circuit, the t 8 base-emitter voltage (V_{be}) is varied. Due to this, base current will change and consequently the collector current will change. Each time, the base current (I_b) and the corresponding value of the collector current (Ic) is noted. It is ensured that the collector emitter voltage always remains 3 V.

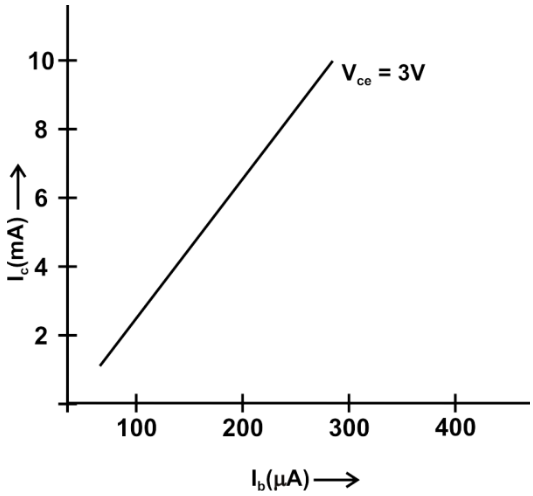


Fig. 2.8 transfer characteristics of transistor

- Then, the graph plotted between I_b and I_c is the transfer characteristic of the transistor at Vce = 3 V . It will be found that the graph will be a straight line. It implies that the output current (Ic) varies linearly with the input current (I_b).

- **A.C. current gain**. It is defined as **the ratio of change in collector current to the change in base current at constant collector voltage,** It is also called current transfer ratio and is denoted by β. Therefore,

$$\beta = \left(\frac{\Delta I_c}{\Delta I_b}\right)_{V_{ce}}$$

2.5 TRANSISTOR AS AN AMPLIFIER

A Transistor can also be used as an amplifier, oscillator, detector, etc. When a transistor is to be used as an amplifier, one may have following three types of the amplifier circuits:

1. Common base amplifier,
2. Common emitter amplifier and
3. Common collector amplifier.

We shall study only the common emitter amplifier.

2.5.1 COMMON EMITTER AMPLIFIER :

- In a common emitter amplifier, the emitter of the transistor is common to base and collector. Fig. 2.9 shows the common emitter amplifier circuit using an n-p-n transistor.

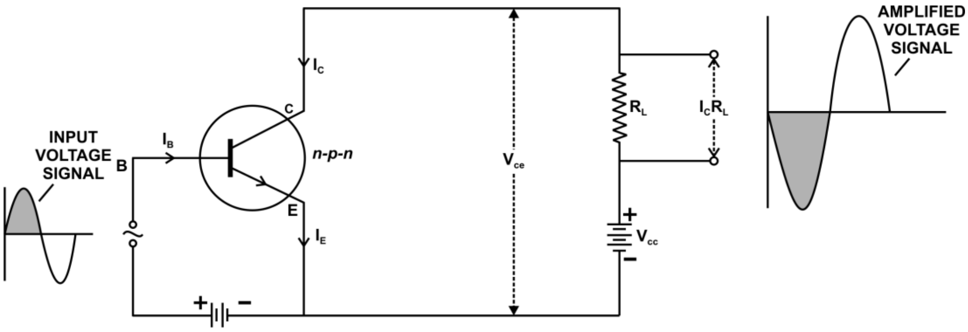


Fig. 2.9 Circuit diagram for common emitter amplifier

- The emitter is common to both the input and output circuits. The emitter is forward biased by using base-bias battery V_{bb} and due to the forward bias, the resistance of input circuit is low. The collector is reverse biased by using collector bias battery Vcc.
- The low input voltage signal is applied in base-emitter circuit (input-circuit) and the amplified output is obtained across the collector and emitter. The arrows point in the direction of the hole current or the conventional current.
- The emitter current, base current and collector current are related to each other by the equation, $\mathbf{I_e = I_b + I_c}$
- Also, corresponding to collector current I_c, the collector voltage i.e. voltage across the collector and emitter will be

$$\mathbf{V_{ce} = V_{cc} - I_c R_L}$$

- The variation in input signal voltage causes variation in emitter current, which in turn produces change in collector current and hence in collector voltage. These variations in collector voltage appear as amplified output voltage.
- Phase relation between input and output signals: The input voltage signal and output voltage signal obtained across collector and emitter are out of phase with each other in common emitter amplifier as explained below.
- Suppose that first half cycle of a.c. input voltage is positive. From fig 2.9, as base is connected to the positive pole of the battery V_{bb} it will make base more positive. Thus, negative forward bias of emitter will increase. It will increase the emitter current and hence the collector current. The increase in collector current will increase potential drop across R_L and then accordingly the collector voltage V_{ce} will decrease. As collector is connected to positive pole of the battery V_{ce} the decrease in collector voltage means that collector will become less positive i.e. negative output signal will be obtained. Thus, corresponding to the positive half cycle of a.c. input, negative output half cycle will be obtained.
- Similarly, we can prove that corresponding to negative half cycle of a.c. input, positive output half cycle will be obtained. Therefore, input and output voltage signals are out of phase with each other as shown in Fig. 3.30. From the operation of common emitter amplifier, it follows that a large collector current (output current) flows corresponding to a small base current (input current) in the circuit.
- **A.C. current gain**. It is defined as the ratio of the change in collector current to the change in base current at constant collector voltage. It is denoted by β_{ac}

Therefore, $$\beta_{ac} = \left(\frac{\Delta I_c}{\Delta I_b}\right)_{V_{ce}}$$

- **Transconductance.** It is defined as the ratio of change in collector current to the change in base voltage at constant collector voltage. It is denoted by g_m

  Therefore, $g_m = \left(\frac{\Delta I_c}{\Delta V_{be}}\right)_{V_{ce}}$

- **a.c. voltage gain**. It is the ratio of the change in output voltage to the change in input voltage. It is denoted byA_v Suppose that on applying an a.c. input signal, the input current changes by ΔI_b and correspondingly, the output current changes by ΔIc' Then

$$A_v = \frac{change\ in\ output\ voltage}{change\ in\ input\ voltage} = \frac{\Delta I_c \times R_{out}}{\Delta I_b \times R_{in}}$$

$$A_v = \beta_{ac} \times resistance\ gain$$

- **A.C. power gain**. It is the ratio of the change in output power to the change in input power.

$$a.c. power\ gain = \frac{change\ in\ output\ power}{change\ in\ input\ power} = \frac{(\Delta I_c)^2 R_{out}}{(\Delta I_b)^2 \times R_{in}}$$

$$a.c. power\ gain = \ \beta_{ac}^2 \times resistance\ gain$$

2.6 TRANSISTOR AS SWITCH:

- Transistor can be operated either in saturation or in cut-off mode.

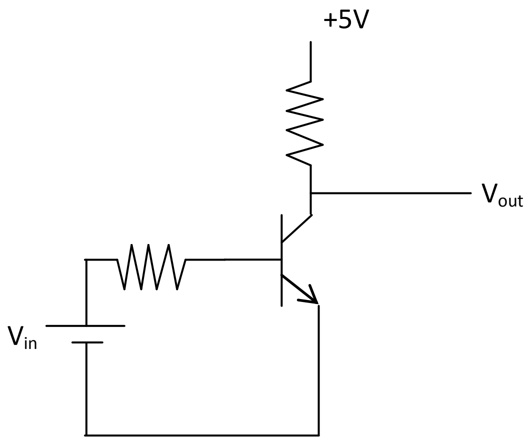


Fig. 2.10 transistor as switch

- When transistor is in saturation region the output voltage is zero and when in cut-off, output voltage it is 5V.

- When the switch is open the transistor goes into the cut-off mode and output voltage drops to zero volts.

- When switch is closed, there is a flow of base current. Hence collector current drops to zero. If there is no current through collector resistor then collector voltage rises to maximum value (+5V) thus output has two values either low (O V) or high (+5V) .

2.7 SUMMARY:

- Transistors are classified in two types n-p-n and p-n-p.
- BJT is a device which consists of two p-n junction diodes formed by sandwitching a p-type semiconductor between two n- type semiconductor or vice-a – versa.
- A BJT is used in three modes CE, CB, and CC.
- Transistor can be used as an amplifier ti amplify weak signal to desired valu.
- Transistor is used as a digital switch.

2.8 UNIT END EXERCISE.:

1. Write a short note on transistor.
2. With the help of a circuit diagram explain the action of n-p-n transistor.
3. Draw circuit symbol of n-p-n and p-n-p transistors.
4. Draw circuit diagram to obtain the characteristics of a n-p-n transistor in common emitter configuration.
5. Describe how will you obtain input and output characteristics. Give shapes of the curves.
6. Draw a neat labelled circuit diagram of n-p-n transistor as amplifier.
7. Explain with circuit diagram how transistor works as switch.

❖❖❖❖

3

AMPLIFIERS

Unit structure :

3.0 OBJECTIVES

After studying this chapter you should be able to

- Understand the concept of amplification.
- Show how to calculate voltage gain and current gain.
- Show how to calculate input impedance and output impedance.
- Draw a circuit diagram of single stage amplifier.
- Discuss the frequency response of single stage amplifier.

3.1 INTRODUCTION

This chapter discusses the concept of amplification. The invention of transistor was crucial to the evolution of electronics. Without amplification, there would be no radio, no television and no computers. This chapter also discusses voltage gain, current gain, input impedance and output impedance

3.2 CONCEPT OF AMPLIFICATION

- Amplification is a process of adding strength to input signal without changing its shape.
- The circuit which amplifies a small input signal is called as an “Amplifier”.
- It is important that the magnified output signal must have the same shape as that of the input signal; i.e. If the input signal is a sine wave then the magnified output signal also be a sine wave.

- In the **Common Emitter** or grounded emitter configuration, the input signal is applied between the base, while the output is taken from between the collector and the emitter as shown. This type of configuration is the most commonly used circuit for transistor based amplifiers and which represents the "normal" method of bipolar transistor connection.
- The common emitter amplifier configuration produces the highest current and power gain of all the three bipolar transistor configurations. This is mainly because the input impedance is LOW as it is connected to a forward-biased PN-junction, while the output impedance is HIGH as it is taken from a reverse-biased PN-junction.

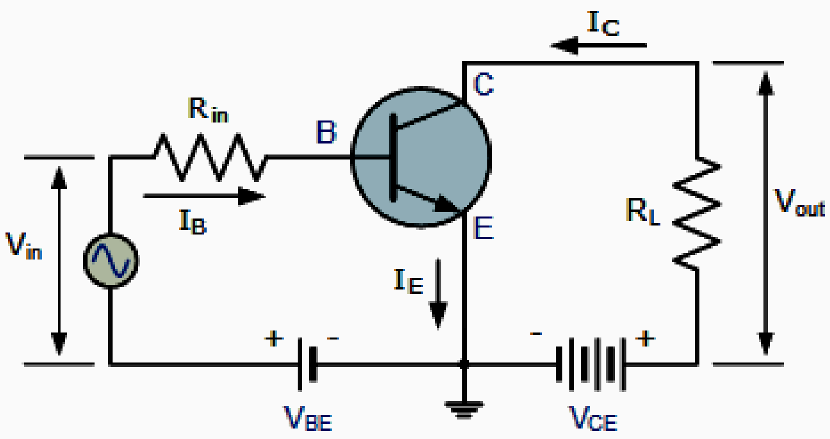


Fig 3.1

- In this type of configuration, the current flowing out of the transistor must be equal to the currents flowing into the transistor as the emitter current is given as $I_e = I_c + I_b$. Also, as the load resistance (R_L) is connected in series with the collector, the current gain of the common emitter transistor configuration is quite large as it is the ratio of Ic/Ib and is given the Greek symbol of Beta, (β).

- Any small change in the base current (I_b), will result in a much larger change in the collector current (I_c). Then, small changes in current flowing in the base will thus control the current in the emitter-collector circuit. Typically, Beta has a value between 20 and 200 for most general purpose transistors.

- This type of bipolar transistor configuration has greater input impedance, current and power gain than that of the common base configuration but its voltage gain is much lower. The common emitter configuration is an inverting amplifier circuit resulting in the output signal being 180^oout-of-phase with the input voltage signal.

3.3 AMPLIFIER NOTATIONS

3.3.1 Gain :

Because amplifiers have the ability to increase the magnitude of an input signal, it is useful to be able to rate an amplifier's amplifying ability in terms of an output/input ratio. The technical term for an amplifier's output/input magnitude ratio is gain. As a ratio of equal units (power out / power in, voltage out / voltage in, or current out / current in), gain is naturally a unitless measurement. Mathematically, gain is symbolized by the capital letter “A”.

Voltage Gain: (A_v)

It is ration of output Voltage (V_o) to the Input Voltage (V_i)

$$\text{Voltage Gain } A_v = \frac{\text{Output Voltage}}{\text{Input Current}} = \frac{Vo}{Io}$$

Current Gain: (A_i)

It is ratio of Output current to input current.

$$\text{Current Gain } A_i = \frac{\text{Output Current}}{\text{Input Current}} = \frac{Io}{Ii}$$

Power Gain: (A_p)

It is ratio of Output power to input power

Power Gain A_p =

	DC gains	AC gains
Voltage	$A_V = \frac{V_{output}}{V_{input}}$	$A_V = \frac{\Delta V_{output}}{\Delta V_{input}}$
Current	$A_I = \frac{I_{output}}{I_{input}}$	$A_I = \frac{\Delta I_{output}}{\Delta I_{input}}$
Power	$A_P = \frac{P_{output}}{P_{input}}$	$A_P = \frac{(\Delta V_{output})(\Delta I_{output})}{(\Delta V_{input})(\Delta I_{input})}$
	$A_P = (A_V)(A_I)$	

Δ = *"change in . . ."*

Fig 3.2

3.3.2 Impedance

Impedance (symbol Z) is a measure of the overall opposition of a circuit to current, in other words: how much the circuit **impedes** the flow of current. It is like resistance, but it also takes into account the effects of capacitance and inductance. Impedance is measured in ohms, symbol Ω.

Input Impedance (Z_i):

It is the resistance seen looking into the input terminal of an amplifier.

$$\text{Input Impedance } Z_i = \frac{\text{Input Voltage}}{\text{Input Current}} = \frac{Vi}{Ii}$$

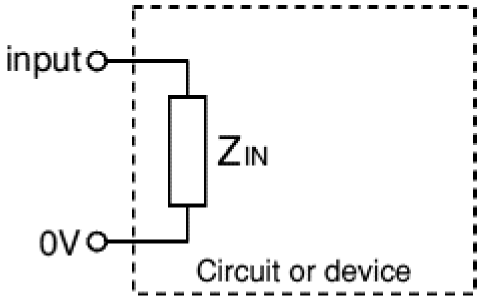


Fig 3.3

Output Impedance (Z_o) :

The output impedance of an amplifier is determined at the output terminals looking back into the system with the applied signal set to zero.

$$\text{Output Impedance } Z_o = \frac{\text{Output Voltage}}{\text{Input Current}} = \frac{Vo}{Io}$$

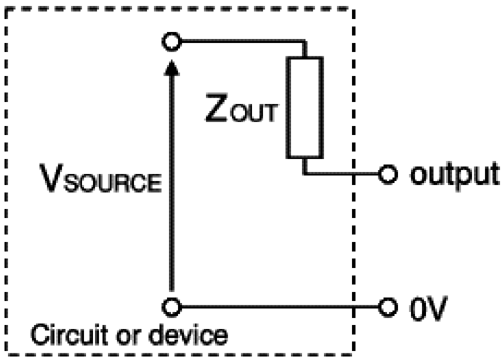


Fig 3.4

3.4 APPLICATION OF BJT AS SINGLE STAGE AMPLIFIER

- When only one transistor is used for amplifying a signal, the circuit is known as **single stage amplifier.**
- For using transistor as an amplifier the transistor should bias in such away it provides maximum gain. The Voltage divider bias circuit provides good stabilization of the operating point. A single stage common emitter with a voltage divider bias and coupling and bypass capacitor is shown in figure.

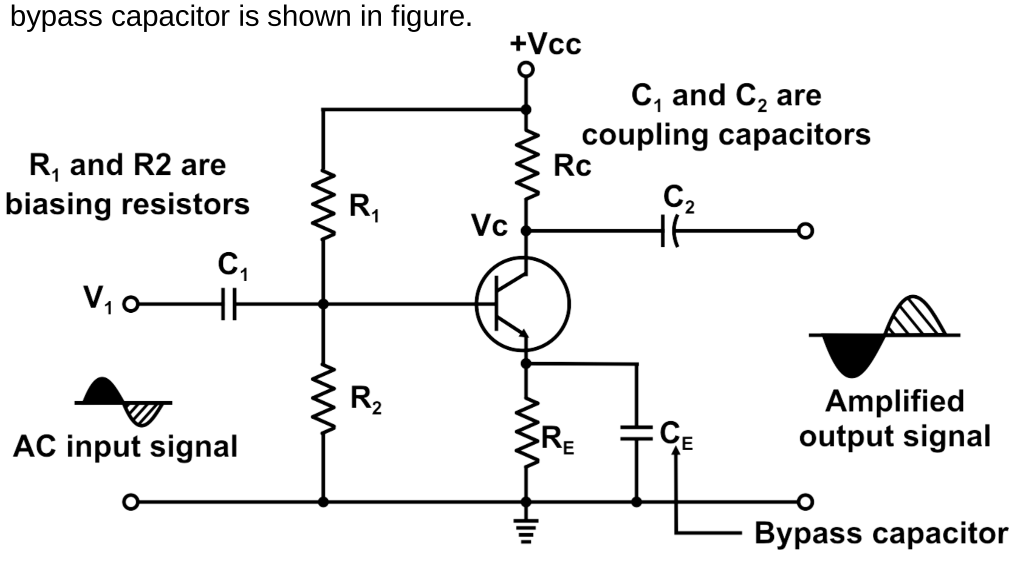


Fig 3.5 Single stage RC Coupled CE amplifier

- A coupling capacitor passes the AC signal form one side to other side and blocks the DC signal. The bypass capacitor bypasses the all AC current form emitter to ground.

- The signal Voltage Vi applied between base and emitter of the transistor gets amplified by the device and the output can be obtained across R_L. A coupling capacitor passes an AC signal form one point to another. At the some time it will not allow the DC Voltage to pass through it. The output across load resistance is free from the collector's DC Voltage here.

3.4.1 How the circuit works?

- When a weak a.c. signal is applied to the base of transistor, a small base current starts flowing. Due to transistor action, a much larger collector current flows through the R_c. This collector current I_c is β times I_b. As the value of R_c is high, a large voltage appears across R_c. Thus weak signal applied to the base appears in amplified form in the collector circuit.

- In common emitter configuration, as the input voltage increases in the positive direction, the output voltage increases in negative direction and vice versa. The input signal voltage and output voltage are always 180^o out of phase. This is known **as phase reversal.**

3.4.2 Important terms :-

- **<u>Voltage divider bias</u>:** In this of biasing two resistances R_1 and R_2 are connected across the supply V_{cc}. The emitter resistance R_E is used to provide stabilization. This biasing method is named as voltage divider bias because the voltage divides among the resistor R_1 and R_2.

- **<u>CE amplifier</u>:** The transistor is connected in the common emitter configuration. Hence it is known as CE amplifier.

- **<u>Common emitter configuration</u>:** In this configuration, input signal is applied between base and emitter and output is measured across collector and emitter. Emitter is common to both input and output, hence the configuration is known as common emitter configuration.

- **<u>Capacitive Reactance</u>:** $X_C = {}^{1}\!/_{2}\ \pi f c$
    - For direct current frequency will be zero therefore X_C will infinite for all d.c. signal. It means that a capacitor offers infinite reactance to d.c. and blocks it completely whereas it allows a.c. to pass through it.

- **<u>Input coupling capacitor C_{in}:</u>** The input capacitor C_{in} is used for coupling the a.c. input voltage to the base of the transistor. It blocks any d.c. component present in the input signal.

- **<u>Output coupling capacitor C_o:</u>** The output capacitor C_o is used to couple the amplifier output to the load resistance. It allows only a.c. part of amplified output to the load resistance.

- **Bypass capacitor C_E:** An emitter bypass capacitor C_E is used in parallel with R_E to provide low reactance path to the amplified a.c. signal. If it is not used, then amplified a.c. signal flowing through R_E will cause voltage drop across it, thereby reducing output voltage.

3.5 FREQUENCY RESPONSE OF SINGLE STAGE AMPLIFIER

- The performance of an amplifier can be judged by observing all frequency components of the input signal are amplified equally well. This can be studied by frequency response curve.
- The frequency response curve is a graph of gain of amplifier versus frequency of input signal. This curve shows how the voltage gain changes with respect to the frequency of an amplifier.

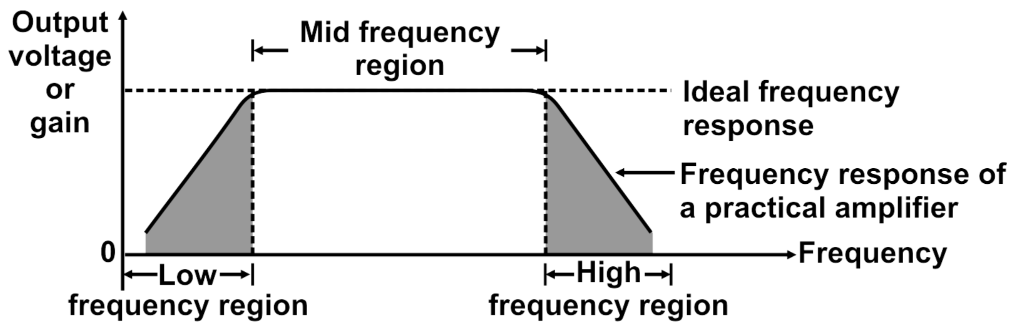


Fig 3.6 Frequency response of an amplifier

- Practically the frequency response of an amplifier Is not flat over the entire operating frequency region. The practical frequency response can be divided into three regions as follows:

    1. Low frequency region
    2. Mid frequency region
    3. High frequency region

1. Low frequency region :

In the low frequency region, the gain or output voltage decreases due to the increased reactance of the coupling and bypass capacitors.

2. Mid frequency region :

In this region, gain and output Voltage remains constant.

3. High frequency region :

In the region, the output Voltage and gain will decrease due to transistor internal capacitances and stray capacitance.

3.5.1 Bandwidth of an amplifier:

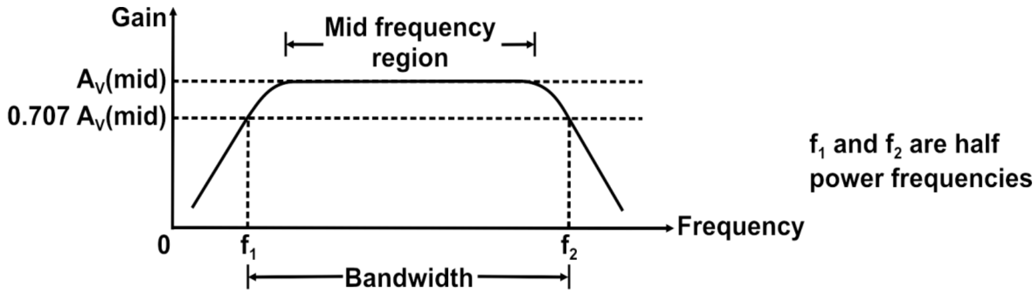


Fig 3.7 Frequency response, half power frequencies and band width.

- In the mid frequency region, the gain of an amplifier is constant. It is denoted as $A_{v\ (mid)}$

- To fix the frequency range in which the amplifier gain is relatively large, a cut-off level of 0.707 $A_{v\ (mid)}$ Or 0.707 $V_{o\ (mid)}$ is chosen. The corresponding frequencies f_1 and f_2 are known as the half power frequencies. At frequencies f_1 and f_2, the power reduces to 50% of power in mid frequency region. Therefore frequencies f_1 and f_2 are known as the half power frequencies.

- The range of frequency over which the voltage gain is equal to or greater than 70.7% of the maximum gain is known as bandwidth. From above graph, it is clear that for any frequency lying between f_1 and f_2, the gain is equal to or greater than 70.7% of the maximum gain.

- The bandwidth of the amplifier is defined as the difference between the half power frequencies.

$$\text{Bandwidth} = f_2 - f_1 \ H_z$$

3.5.2 Cut off frequencies:

- **Lower cut off frequency (f_1) or f_L :**
  It is the frequency of the input sign at which the amplifier gain or output voltage reduce to 70.7% of their mid frequency range value. f_1 is always less than f_2. And f_1 is few tens of $H_{Z.}$

- **Upper cut off frequency (f_2) or f_H :**
  It is the frequency of the input signal at which the amplifier gain or output voltage reduce to 70.7% of their mid frequency range value. f_2 is always higher than f_1 and f_2 is few KHz.

3.6 SUMMARY

- Amplification is a process of adding strength to input signal without changing its shape.

- **Voltage Gain: (A_v)**
  It is ration of output Voltage (Vo) to the Input Voltage (Vi)

$$\text{Voltage Gain } A_v = \frac{\text{Output Voltage}}{\text{Input Current}} = \frac{Vo}{Io}$$

➢ **Current Gain: (A_i)**
It is ratio of Output current to input current.

$$\text{Current Gain } A_i = \frac{\text{Output Current}}{\text{Input Current}} = \frac{Io}{Ii}$$

➢ **Power Gain: (A_p)**
It is ratio of Output power to input power

Power Gain A_p =

➢ When only one transistor is used for amplifying a signal, the circuit is known as single stage amplifier.

➢ Capacitor offers infinite reactance to d.c. and blocks it completely whereas it allows a.c. to pass through it.

➢ The frequency response of an amplifier Is not flat over the entire operating frequency region.

➢ The practical frequency response can be divided into three regions : i) Low frequency region ii) Mid frequency region

iii) High frequency region

➢ At frequencies f_1 and f_2, the power reduces to 50% of power in mid frequency region. Therefore frequencies f_1 and f_2 are known as the half power frequencies.

➢ The bandwidth of the amplifier is defined as the difference between the half power frequencies.

Bandwidth $= f_2 – f_1$ H_z

3.7 REVIEW QUESTIONS

1. Explain the concept of amplification.
2. Define the terms: A_v, A_i, A_p, Z_i and Z_o
3. How the transistor is used as an amplifier?
4. What do you understand by single stage amplifiers?
5. Draw the circuit diagram of single stage amplifier. Explain the function of each component.
6. Explain the function of coupling capacitors and bypass capacitors in single stage amplifiers.
7. Explain the frequency response of an amplifier.
8. Define the term half power frequencies.

3.8 REFERENCE

- ‘Electronic Principles’ 7th edition by Albert Malvino and David J Bates, Tata McGraw Hill.
- ‘Principles of electronics’ 11th edition by V.K.Mehta and Rohit Mehta, S. Chand.

❖❖❖❖

4

MULTISTAGE AMPLIFIERS

Unit Structure

4.0 OBJECTIVES

After studying this chapter you should be able to

- Understand the need of cascading no. of amplifier.
- Know the advantages and disadvantage of RC coupled amplifier.
- Understand the application of RC coupled amplifier.
- Understand the working of DC amplifiers and its frequency response.
- State the advantages of Darlington pair.

4.1 INTRODUCTION

This chapter discusses various multistage transistor amplifiers and their practical application. It may be noted that the practical amplifier is always a multistage amplifier. In a transistor radio receiver, the number of amplification stages may be six or more. This chapter also discusses the working and application of Darlington pair in electronics.

4.2 BASIC CONCEPTS

- The output from a single stage amplifier is usually insufficient to drive an output device. Practically, there is limitation on the maximum value of voltage gain that can be obtained from a single stage amplifier. Therefore additional amplification over two or three stages is necessary. To achieve this, the output of each amplifier stage is coupled to the input of the next stage. The resulting amplifier is known as multistage amplifier.

- There is a limit to how much gain can be achieved from a single stage amplifier. Single stage amplifiers also have limits on input and output impedance. Multistage amplifiers are used to achieve higher gain and to provide better control of input and output impedances.

- Two significant advantages that multistage amplifiers have over single stage amplifiers are flexibility in input and output impedance and much higher gain.

4.3 CASCADING

- The meaning of the word "**Cascading"** is to connect a number of amplifier stages to each other with the output of the previous stage to the input of next stage as shown in fig. below. Thus a multistage amplifier is obtained by cascading a number of amplifiers.

- The most important parameter of amplifier are its input impedance, voltage gain, band width and output resistance. Many times a single stage amplifier is not capable of providing input and output impedances of correct magnitudes.

- The required values of these parameter are dependent on the particular application.

- It is generally not possible for a single stage amplifier to fulfill all the requirements. Hence we have to use a multistage amplifier, which deals with these requirements.

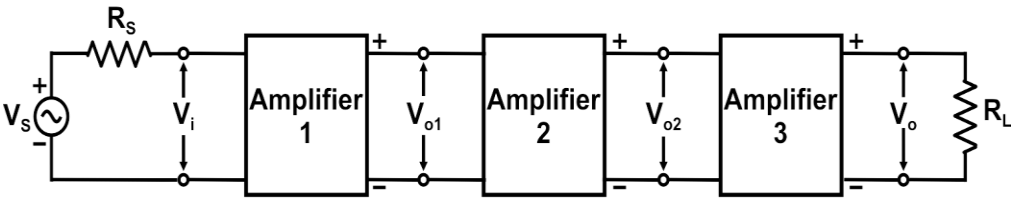


Fig .4.1 multistage amplifier is obtained by cascading three amplifier

Good to Know:

An audio amplifier, refers to an amplifier that operates in the frequency range of 20 Hz to 20 KHz. Radio frequency (RF) amplifier is one that amplifies the frequencies above 20 KHz.

4.4 RC COUPLED AMPLIFIER

- The output of first single stage is applied to the input of second stage amplifier using the coupling circuit. If the RC (Resistance Capacitance) coupling is used, then resulting multistage amplifier is known as RC coupled amplifier.

- A typical R-C coupled transistor amplifiers is as shown in fig. below.

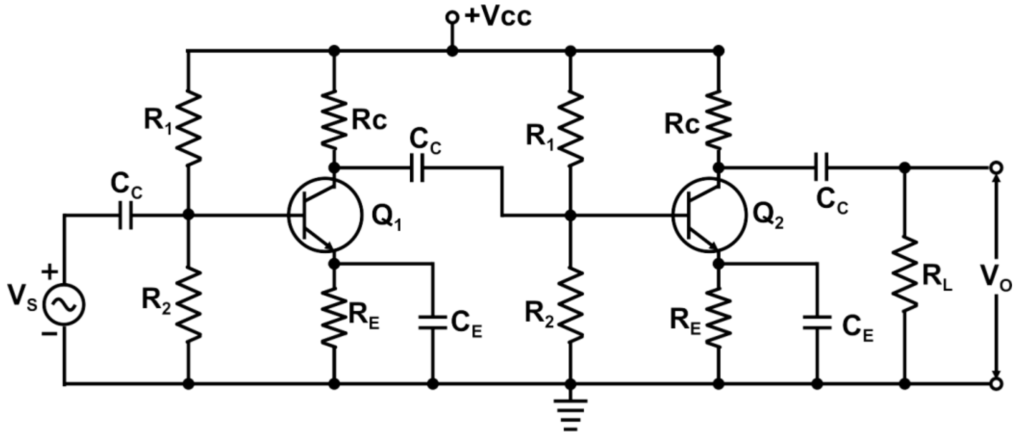


Fig.4.2 Two stage R-C Coupled transistor amplifier

- When a.c. signal is applied to the base of the first transistor, it is amplified and developed across the out of the first stage. This amplified voltage is applied to the base of next stage through the coupling capacitor C_c where it is further amplified and reappears across the out put of the second stage. Thus the successive stages amplify the signal and the overall gain is raised to the desired level. Much higher gains can be obtained by connecting a number of amplifier stages in succession (one after the other).

- **Resistance-capacitance (RC)** coupling is most widely used to connect the output of first stage to the input (base) of the second stage and so on. It is the most popular type of coupling because it is cheap and provides a constant amplification over a wide range of frequencies. Fig. shows the circuit arrangement of a two stage RC coupled CE mode transistor amplifier where resistor R is used as a load and the capacitor C is used as a coupling element between the two stages of the amplifier.
- R_1, R_2 and R_E are the biasing resistors used separately for the two stages. Voltage divider biasing is being used.

- Due to the used of coupling capacitors the dc voltages will not be coupled from one stage to the next stage will not be affected due to coupling. Thus due to RC coupling the dc operating conditions in any stage remain unaffected.

- The R-C network gives a wideband frequency response without introduction of peaks at any frequencies. Therefore R-C coupling can be used for the AF amplifiers.

4.4.1 Frequency response curve

- The curve representing the variation of gain of an amplifier with frequency is known as frequency response curve. It is shown in Fig. The voltage gain of the amplifier increases with the frequency, f and attains a maximum value. The maximum value of the gain remains constant over a certain frequency range and afterwards the gain starts decreasing with the increase of the frequency. It may be seen to be divided into three regions:

    1) Low frequency range

    2) Mid frequency range

    3) High frequency range

- Frequency response drops off at low frequencies due to the coupling capacitors and at high frequencies due to the shunting effects of the internal capacitance of the transistor and stray capacitances.

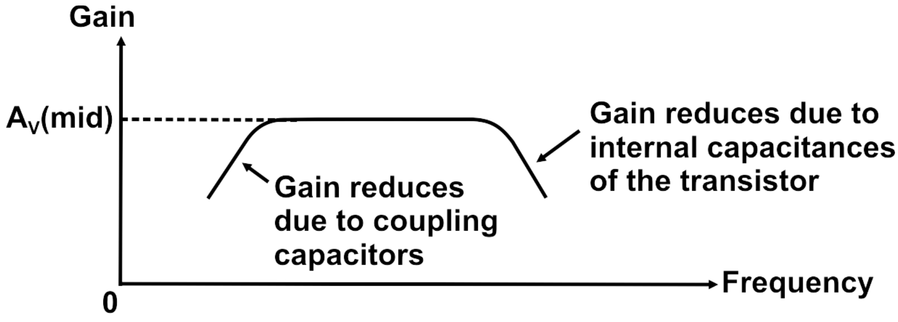


Fig.4.3 Frequency response of on RC Coupled amplifier.

4.4.2 Advantage of RC coupled amplifiers

1) It has excellent frequency response. The gain is constant over the audio frequency range.

2) It has lower cost since it employs resistors and capacitors which are cheap.

3) The circuit is very compact.

4) Due to the capacitor, the d.c. biasing conditions of individual stages will remain unchanged even after cascading.

5) The distortion in the output is low.

4.4.3 Disadvantages

1) Impedance matching is poor.

2) Overall Voltage gain is less.

3) They have the tendency to become noisy with age.

4) Gain reduces at low frequencies due to coupling Capacitors.

4.4.4 Applications

The RC coupled amplifiers have excellent audio fidelity over a wide range of frequency. Therefore, they are widely used as voltage amplifiers. For example,

1) In public address amplifier system

2) Tape recorder

3) TV, VCR and CD Player

4) Stereo amplifiers

4.5 DIRECT COUPLED AMPLIFIERS

- There are many applications in which extremely low frequency signal (<10 Hz) are to be amplified e.g. amplifying photoelectric current, thermo electric current etc. The coupling devices such as capacitances and transformers cannot be used because the electrical sizes of these components become very large at extremely low frequencies. Under such conditions, one stage is directly connected to the next stage without any coupling circuit. This type of coupling is known as direct coupling. The resulting multistage amplifier is known as Direct Coupled (DC) amplifier.

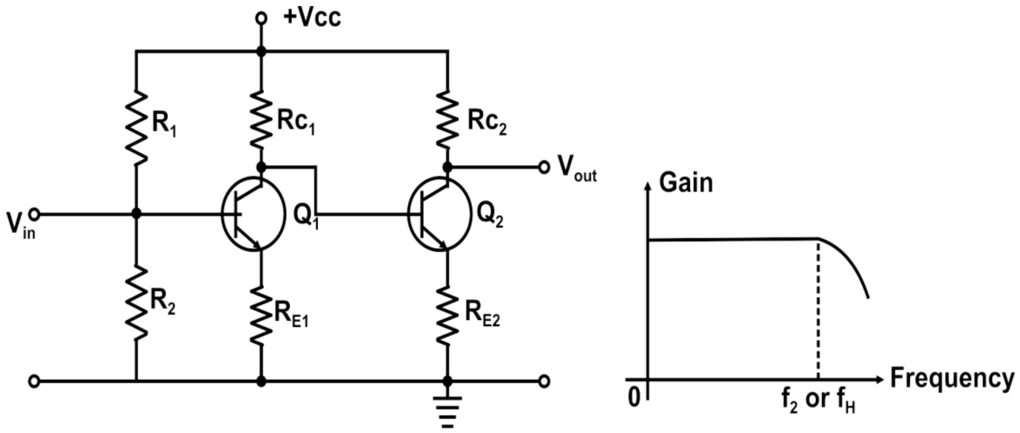


Fig.4.4 A two stage direct coupled amplifier

- Due to direct coupling, both dc and ac voltages are coupled from first stage to the second stage. Therefore the Q point of the second stage will change depending on the coupled signal.

- The low frequency response is better than due to the absence of the coupling capacitors in the circuit.

- Due to changes in temperature, the temperature dependent parameters such as V_{BE} Or β of transistor Q_1 will change. Due to these changes the collector current and voltage of Q_1 will change. The direct coupling will couple these changes to the second stage and to the output. This problem is known as "Drift".

4.5.1 Advantage of direct Coupled Amplifier.

1) Due to absence of coupling capacitors, the gain does not reduce on the lower frequency side.

2) This amplifier can amplify even the dc signal.

3) Wide frequency response.

4) Reduced cost and complexity.

4.5.2 Disadvantages:

1) The dc biasing conditions of the individual stages do not remain same after cascading.

2) The output waveform has a dc shift.

3) Poor frequency response at higher frequencies.

4) Poor temperature stability.

4.5.3 Applications:

1) In the operational amplifiers (OP-AMPs)

2) In the analog computation

3) In the linear power supplies (Voltage regulators)

4.6 DARLINGTON PAIR

- The Darlington connection of two BJTs is as shown in fig. The two transistors Q_1 and Q_2 are connected.

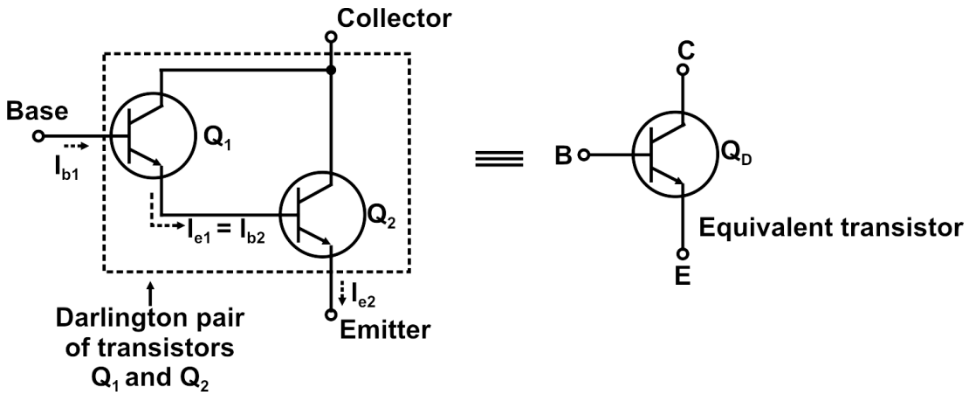


Fig. 5 a) Darlington connection of transistors b)Equivalent Darlington transistor.

- As seen, the collectors of the two transistors are connected together, emitter of Q_1 is connected to the base of Q_2 and emitter of Q_2 acts as the emitter of the Darlington connection.

- The important feature of the Darlington connection is that the composite transistor will act as a single unit, the current gain of which is equal to the product of individual current gains β_1 and β_2.

  Then overall current gain of Darlington pair is given by$\beta = \beta_1 \beta_2$

- It is possible to connect two different transistor readily available in market.

- Since its current gain is much higher, a Darlington connection can have a very high input impedance and can produce large output currents.

- Darlington connections are often used with voltage regulators, power amplifiers and high current switching applications.

4.7 SUMMARY

- There is limitation on the maximum value of voltage gain that can be obtained from a single stage amplifier.
- The meaning of the word "Cascading" is to connect a number of amplifier stages to each other with the output of the previous stage to the input of next stage.
- The advantages of multistage amplifiers have over single stage amplifiers are flexibility in input and output impedance and much higher gain.
- The output of first single stage is applied to the input of second stage amplifier using the coupling circuit. If the RC (Resistance Capacitance) coupling is used, then resulting multistage amplifier is known as RC coupled amplifier.
- Frequency response of RC coupled amplifier drops off at low frequencies due to the coupling capacitors and at high frequencies due to the shunting effects of the internal capacitance of the transistor and stray capacitances.
- Output of single stage amplifier is directly connected to the next stage without any coupling circuit. This type of coupling is known as direct coupling. The resulting multistage amplifier is known as Direct Coupled (DC) amplifier.
- Due to direct coupling, both dc and ac voltages are coupled from first stage to the second stage.
- The important feature of the Darlington connection is that the composite transistor will act as a single unit, the current gain of which is equal to the product of individual current gains β_1 and β_2.

4.8 REVIEW QUESTIONS

1. Derive the expression for voltage gain of a multistage amplifier having n amplifiers cascaded.
2. Draw a labelled circuit diagram of a two stage RC coupled amplifier.
3. Write a short note on DC amplifiers.
4. What do you understand by multistage amplifier? Mention its need.
5. What are the advantages and disadvantages of RC coupled amplifier?
6. What are the advantages and disadvantages of DC coupled amplifier?
7. With the help of neat circuit diagram explain Darlington pair.
8. What are the advantages of Darlington pair?

4.9 REFERENCE

- 'Electronic Principles' 7th edition by Albert Malvino and David J Bates, Tata McGraw Hill.
- 'Principles of electronics' 11th edition by V.K.Mehta and Rohit Mehta, S. Chand.

❖❖❖❖

5

FEEDBACK AND OSCILLATORS

Unit Structure

5.0 OBJECTIVES

After studying this chapter you should be able to

- Define positive and negative feedback.
- Understand the advantages and disadvantages of negative feedback.
- Understand the advantages and disadvantages of negative feedback.
- Understand the basic concept of oscillators.
- Describe the operation of RC phase shift oscillator.
- Describe the operation of LC oscillator.

5.1 INTRODUCTION

- This chapter discusses the positive and negative types of feedback. It also discusses the use of the feedback in amplifiers and oscillators.

- At frequencies under 1 MHz, we can use RC oscillators to produce almost perfect sinewave. These low frequency oscillators use transistor and RC resonant circuits to determine the frequency of oscillation. Above 1MHz, LC oscillators are used. These high frequency oscillators use transistors and LC resonant circuits.

5.2 CONCEPT OF FEEDBACK

- Feedback is defined as the process in which a part of output signal (voltage or current) is returned back to the input. The amplifier that operates on the principle of feedback is known as feedback amplifies. In the feedback process a part of output is sampled and fed back to the input.

- Thus at the input of an amplifier using feedback two signals will be simultaneously present. One of them is the original input signal itself and the other one is the fed back signal.

- Depending upon whether feedback energy aids or opposes the input signal, there are two basic types of feedback:
  - 1) Positive feedback 2) Negative feedback.

- If the original input signal and the feedback signal are in phase, the feedback is called as positive feedback. However if these two signals are out of phase then the feedback is called as negative feed back.
- Positive feedback is used in oscillator and negative feedback is used in amplifier.

5.3 NEGATIVE FEEDBACK

When the feedback energy (voltage or current) is out of phase with the input signal and thus opposes it, it is called negative feedback. This is illustrated in below fig. The amplifier introduces a phase shift of 180^o into the circuit while the feedback network is so designed that it introduces no phase shift. (ie 0^o phase shift) The result is that the feedback voltage V_f is 180^o out of phase with the input signal V_{in}.

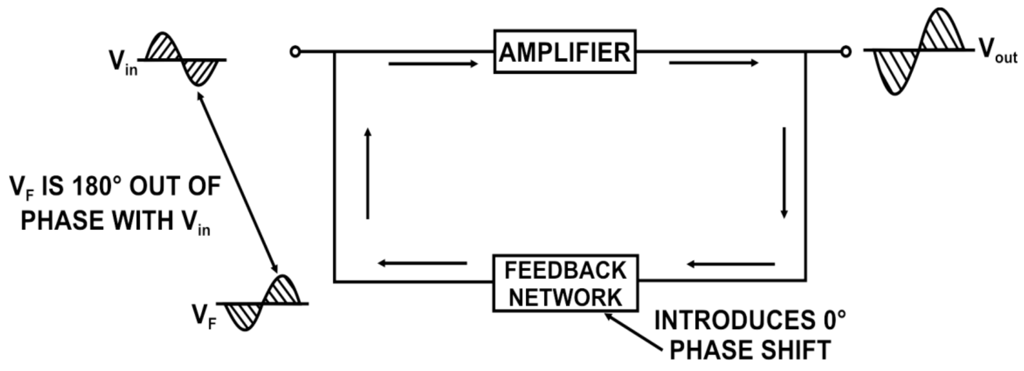


Fig.5.1 negative feedback

5.3.1 Disadvantages of Negative Feedback

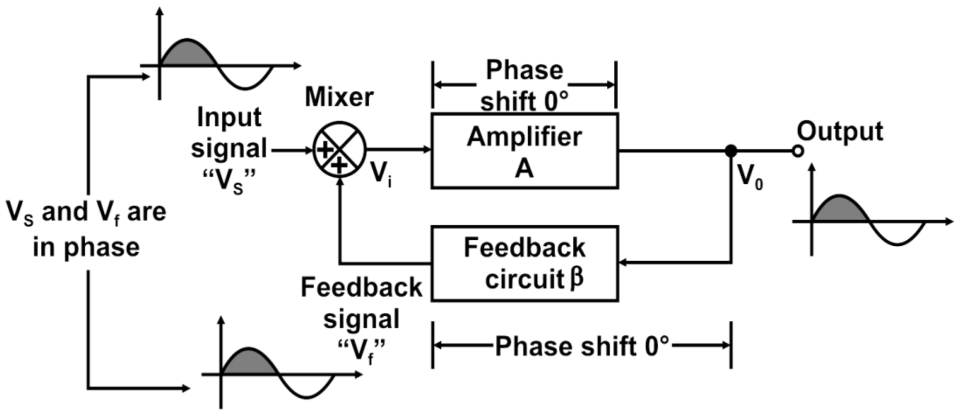
1. Negative feedback reduces the gain of the amplifier.
2. Reduction in input resistance in case of voltage shunt and current shunt type amplifier.
3. Increase in output resistance in case of current shunt and current series feedback amplifiers.

5.3.2 Advantages of Negative Feedback.

1. Negative feedback stabilizes the gain of the amplifier.
2. There is a significant increase in the bandwidth of the amplifier.
3. Distortion in the amplifier output are reduced.
4. Input resistance increases for certain feedback configurations.
5. Output resistance decreases for certain feedback configurations.
6. Operating point is stabilized.

5.4 POSITIVE FEEDBACK

- The positive feedback is used in oscillators. The concept of positive feedback can be explained with the help of fig 5.2



- **Fig. 5.2 Positive feedback**

- **An oscillator is an amplifier with positive feedback**. A part of the output is fed back through the feedback network and mixer to the amplifier input. The feedback energy (voltage or current) is "in phase" with the original input signal as shown in fig. 2

- As the phase shift introduced by the feedback network is 0^o. The positive feedback increases the gain of the amplifier.

5.4 OSCILLATOR

- Oscillators are circuits that produce specific, periodic waveforms such as square, triangular, sawtooth, and sinusoidal. There are two main classes of oscillator: relaxation and sinusoidal. Relaxation oscillators generate the triangular, sawtooth and other nonsinuoidal waveforms and are not discussed in this chapter. The focus here is on sine wave oscillators, created using transistor amplifier.

- Many electronic devices require a source of energy at a specific frequency which may range from few Hz to several MHz. This can be achieved by an electronic device called an oscillator.

- Oscillators are basically ac signal generators which you use in our laboratories. Oscillators generate alternating voltage of desired shape (sine, square, triangular etc), at desired frequency. The output voltage and frequency of an oscillator can be variable
- Block diagram of a basic oscillator is shown in Fig. below.

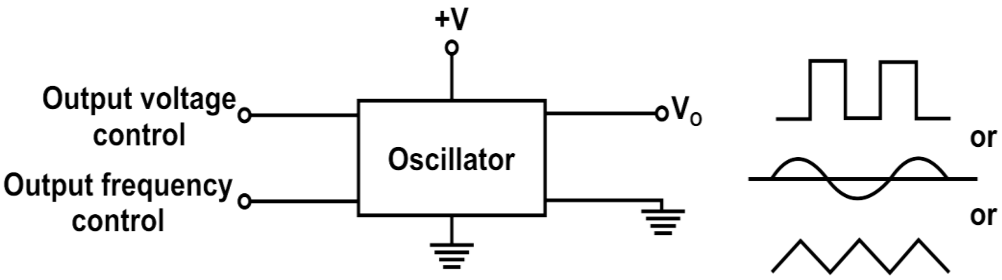


Fig.5. 3 Oscillator

- As seen from fig. the oscillator operates on a dc power supply +V volts and more importantly it produces an alternative output voltage without any alternating input voltage. Oscillators operates on the principle positive Feedback.

- An amplifier will work as an oscillator and only if it satisfies a set of conditions called the "Barkhausen criterion".

5.4.1 Barkhausen criterion:

The Barkhausen criterion states that :-

1. **The loop phase shift must be 0° or 360°**
   A oscillator will operate at that frequency for which the total phase shift introduced, as the signal proceeds from input terminals, through the amplifier and feedback network and back again to the input is precisely 0^o or 360^o or integral multiple of 360^o.

2. **The loop gain must be slightly greater than unity.**
   At the Oscillator Frequency, the magnitude of the product of open loop again of the amplifier A and the feedback factor *B* is equal to or greater than unity.

$$\therefore |AB| \geq 1$$

The product is called as the "loop gain".

There conditions are diagrammatically illustrated in fig.

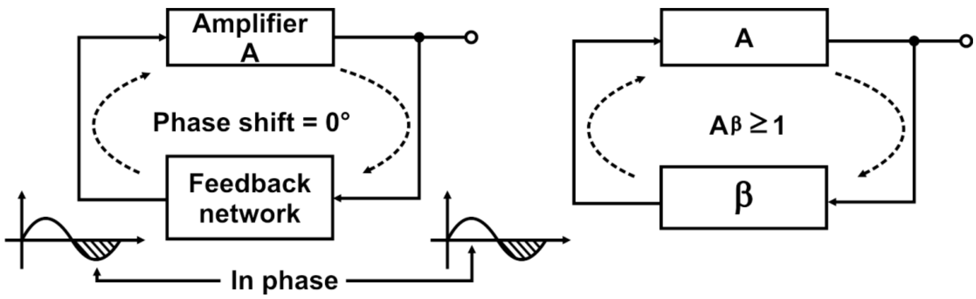


Fig. 5.4 a) The phase shift around the loop is 0°. b) Loop gain $\therefore |AB| \geq 1$

5.6 RC PHASE SHIFT OSCILLATORS

As the name suggests this is an RC Oscillators. Basically, it consists of an amplifier and a phase shifting network made of resistors and capacitors.

5.6.1 Phase shifting network

- An RC phase shifting, network can be cascaded as shown in fig. below (a).
- Each RC network is designed to introduce a phase shift of 60° precisely. Thus the total phase shift introduced by the total phase shift introduced by the 3 stage RC network is 180°. That means output of the network leads its input by 180°.
- RC network shown in Fig. 5 (a) is sometime called as the ladder network.

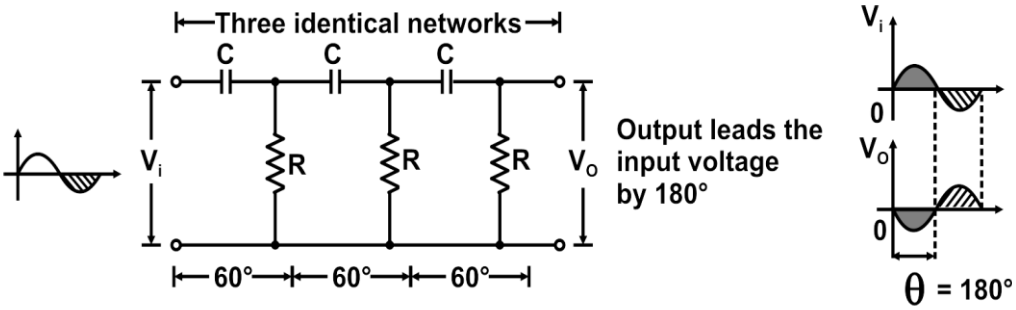


Fig 5.5 (a) RC Network

5.6.2 Circuit Description

- The amplifier used for RC phase shift oscillator can use a transistor as active device. A typical RC phase shift oscillator using transistor as an active device is shown in fig. (b)

- The circuit consists of a single stage amplifier in C.E. configuration and the RC phase shifting network. The resistors R_1, R_2 and R_E are connected for transistor biasing CE is the emitter bypass capacitor.

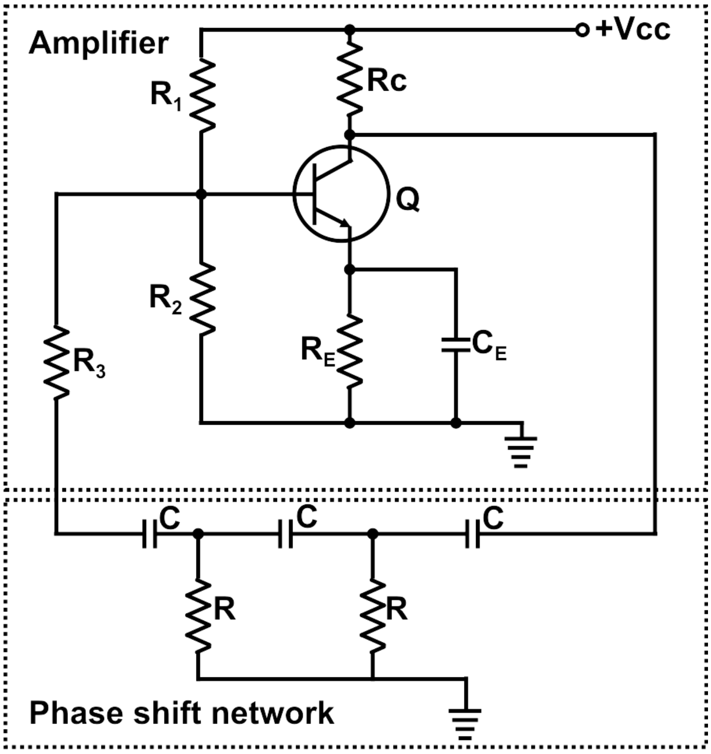


Fig 5.5 (b) : RC phase shift oscillator

5.6.3 Operation of circuit

- The circuit is set into oscillations by any random or variation caused in the base current, that may be either due to noise inherent in the transistor or minor variation in voltage of dc power supply. This variation in base current is amplified in collector circuit.
- The output of the amplifier is supplied to an R-C feedback network. The R-C network produces a phase shift of 180° between output and input voltages. Since CE amplifier produces a phase reversal of the input signal, total phase shift becomes 360° or 0° which is essential for regeneration or for sustained oscillations.
- The output of this network is thus in the same phase as the originally assumed input to the amplifier and is applied to the base terminal of the transistor. Thus sustained variation in collector current between saturation and cut-off values are obtained. R-C phase shift network is the frequency determining network.
- The phase shift around the loop will be precisely equal to 0^o only at one frequency "f" which is the frequency of operation. If the gain of the amplifier and feedback factor B is adjusted properly to have a loop gain $|AB| \geq 1$ then sustained sinusoidal oscillations will be obtained at the oscillator output.

5.6.4 Frequency of oscillations

The frequency of RC phase shift oscillator using transistor is given by

$$f = \frac{1}{2\pi RC} \frac{1}{\sqrt{6+4k}}$$

Where $K = R_C/R$

5.6.5 Advantage of Phase Shift Oscillator :

1) It does not require transformers or inductors.
2) Good quality sine wave output can be obtained.
3) It can be used to produce sine wave of very low frequency.
4) Easy to design.

5.6.6 Disadvantage of Phase shift Oscillator

1) The circuit gives small output
2) It is difficult to vary the output frequency.
3) It can not operate at very high frequencies.

5.6.7 Applications

1) The phase shift oscillator is well suited to the range of frequencies from several hertz to several hundred kilohertz (20Hz to 200 KHz), and so includes the audio frequency range (upto 20 KHz).
2) It is used in signal generators.

5.7 LC OSCILLATORS

- Resistive-capacitive sine wave oscillators can generate signals from a few hertz up to several megahertz, but inductive-capacitive (LC) oscillators can generate sinewave outputs from 20 or 30 KHz up to UHF frequencies.
- An LC oscillator includes an LC network that provides the frequency-selective feedback between the output of the amplifier and its input terminals.

5.7.1 Tank circuit

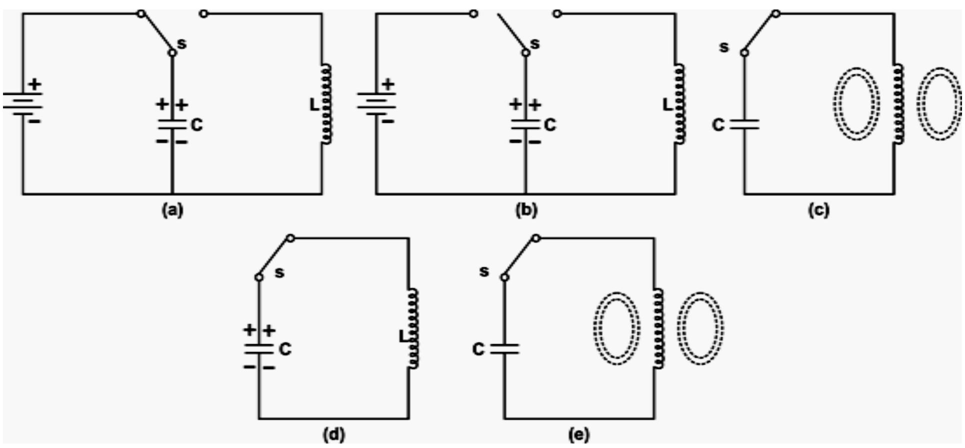


Fig. 5.6 Tank circuit

- The oscillatory circuit, also called the L-C circuit or tank circuit, consists of an inductive coil of inductance L connected in parallel with a capacitor of capacitance C. The values of L and C determine the frequency of oscillations produced by the circuit.
- The most important point is that both the capacitor and inductor are capable of storing energy—the capacitor stores energy in its dielectric field whenever a potential difference exists across its plates while the inductor stores energy in its magnetic field whenever current flows through it.
- For understanding the operation of an oscillatory circuit, let the capacitor be charged from a dc source with the polarity as shown in figure 6 (a).
- A potential difference will be across the plates of the capacitor because of the accumulation of electrons in the lower plate of the capacitor. The electrons get accumulated in the lower plate due to the supply from the negative terminal of the battery.
- Thus, a potential energy will be formed in the capacitor. Now when the capacitor is fully charged and the switch S is opened, as shown in figure 6 (b), the capacitor cannot discharge through L.
- Suppose the switch S is kept in position 'b'. The current starts flowing in the circuit but the self induced e.m.f. in the coil opposes the current flow. Thus the rate of rise of current is slow. Maximum current flows in the circuit when the capacitor is fully discharged. Due to flow of current, magnetic field is set up which stores the energy given by the electric field, as shown in figure 6 (c). Thus, at the instant the capacitor gets completely discharged, the electrostatic energy stored in the capacitor gets converted into the magnetic field energy associated with the inductor L.
- When the capacitor is completely discharged, the magnetic field begins to collapse and a counter or back e.m.f is developed which, according to Lenz's law, keeps the current flowing in the same direction. The capacitor now starts getting charge but with opposite polarity, as shown in fig. 6 (d). In this case, the energy associated with the magnetic field is again converted into electrostatic energy. In an ideal case (that is, both the L and C are loss-free), the capacitor is charged to the value it had initially while the magnetic field energy reduces to zero.
- After the collapsing field has recharged the capacitor, the capacitor now begins to discharge with a current flow in the opposite direction. The electric field starts collapsing whereas magnetic field starts building up again but in opposite direction. Fig.6 (e) shows the condition when the capacitor gets fully discharged. The sequence of charging and discharging continues, that is, the process of transformation of dielectric energy into magnetic energy and vice-versa is repeated again and again. This situation is similar to an oscillating pendulum, in which the energy keeps on interchanging between potential and kinetic energy. Thus the charge and discharge of a capacitor through inductor results in oscillating current and hence electrical oscillations are set up in the L-C or tank circuit.
- The frequency of oscillation is the same as the resonant frequency of the tank circuit and it is given as $f = 1/2\pi\sqrt{LC}$
- The interchange of energy between L and C would continue indefinitely if there were no losses in the tank circuit. But since there are losses, the indefinite

interchange of energy cannot be proved practically. The losses include the energy that is lost in the form of heat generated in the coil resistance, capacitor leakage resistance and connecting wires. As a result, while the total energy is consumed in overcoming the losses, the oscillating current goes on decreasing with time and eventually becomes zero.

5.7.2 Block diagram of LC oscillator

- The general block diagram of an LC Oscillators is as shown in fig. below. The gain of the amplifier is A and X_1, X_2, X_3 are reactance.

- The Barkhausen criterion is satisfied if and only if

$$\mathbf{X_3 = -(X_1 + X_2)}$$

- Hence X_3 must be an opposite type reactance as compared to X_1 and X_2. That means X_3 should be inductive if X_1 and X_2 are capacitive and it should be capacitive if X_1 and X_2 are inductive.

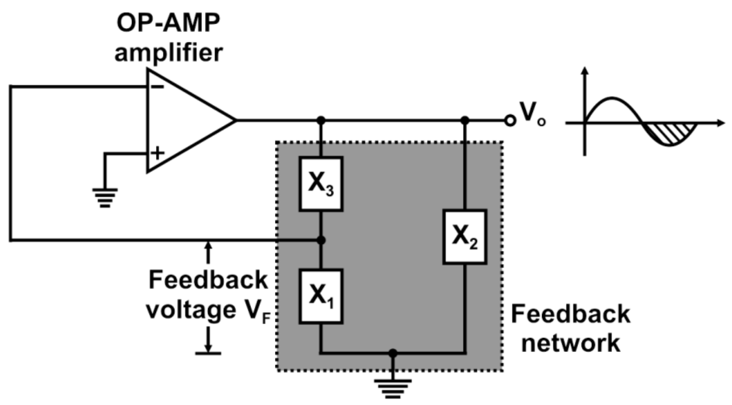


Fig.5.7 LC Oscillator

Depending on the components used in place of X_1, X_2 and X_3 we can obtain two types of LC oscillators namely:

1) Hartley Oscillator.

2) Colpitt's Oscillator

The type of reactance used in place of X_1, X_2 and X_3 for these Oscillations is given in Table.

Name of Oscillator	Components used in the feedback network		
	X_1	X_2	X_3
Hartley	L	L	C
Colpitt's	C	C	L

5.7.3 Hartley Oscillator

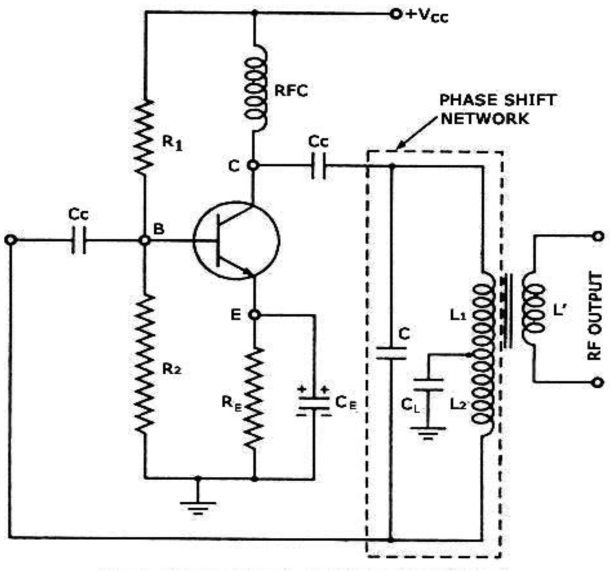


Fig.5. 8 Hartley Oscillator

- When the above circuit is turned on, the capacitor is charged. When this capacitor is fully charged, it discharges through coils L_1 and L_2 setting up oscillations of frequency determined by

$$f = \frac{1}{2\pi\sqrt{CL_T}}$$

- Where $L_T = L_1 + L_2 + 2M$
  Here M = mutual inductance between L_1 and L_2

- The output voltage of the amplifier appears across L_1 and feedback voltage across L_2. The voltage across L_2 is 180^o out of phase with the voltage developed across $L_1(V_{out})$. The voltage across L_2 provide positive feedback. . A phase shift of 180^o is produced by the transistor and a further phase shift of 180^o is produced by $L_1 - L_2$ voltage divider. In this way, feedback is properly phased to produce continuous oscillation.

5.7.4 Colpitt's Oscillator

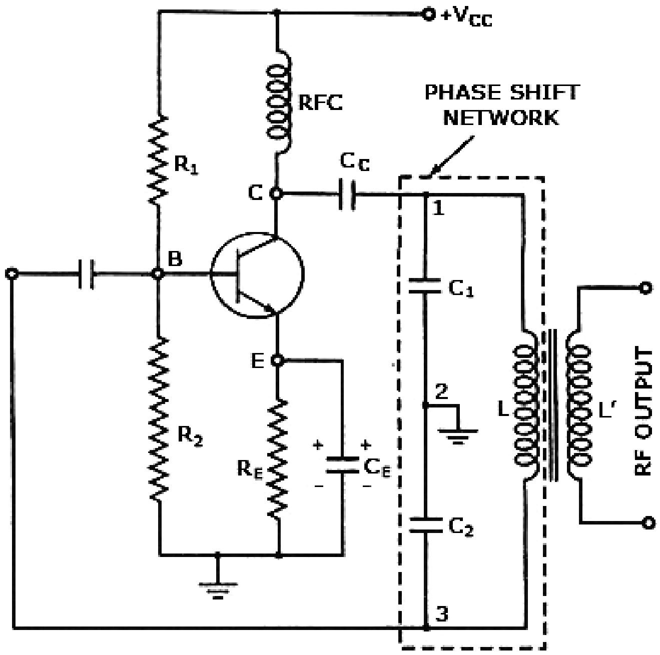


Fig.5. 9 Colpitt's Oscillator

- When the circuit is turned on, the capacitors C_1 and C_2 are charged, the capacitors discharge through L, setting up oscillations of frequency determined by

$$f = \frac{1}{2\pi\sqrt{LC_T}} \qquad \text{where } C_T = \frac{C_1 C_2}{C_1 + C_2}$$

- The output voltage of the amplifier appears across C_1 and feedback voltage is developed across C_2. The voltage across it is 180^o out of phase with the voltage developed across $C_1(V_{out})$. The voltage across C_2 provide positive feedback. A phase shift of 180^o is produced by the transistor and a further phase shift of 180^o is produced by

- $C_1 - C_2$ voltage divider. In this way, feedback is properly phased to produce continuous oscillation.

5.8 SUMMARY

- Feedback is defined as the process in which a part of output signal (voltage or current) is returned back to the input.
- Depending upon whether feedback energy aids or opposes the input signal, there are two basic types of feedback:

  1) Positive feedback 2) Negative feedback.
- Positive feedback is used in oscillator and negative feedback is used in amplifier.
- Oscillators are circuits that produce specific, periodic waveforms such as square, triangular, sawtooth and sinusoidal.
- The Barkhausen criterion states that :-
  1. The loop phase shift must be 0° or 360°
  2. The loop gain must be slightly greater than unity.

- RC phase shift Oscillators consists of an amplifier and a phase shifting network made of resistors and capacitors.
- The phase shift oscillator is well suited to the range of frequencies from several hertz to several hundred kilohertz (20Hz to 200 KHz)
- inductive-capacitive (LC) oscillators can generate sinewave outputs from 20 or 30 KHz up to UHF frequencies.
- Frequency of oscillations of Colpitt's Oscillator

$$f = \frac{1}{2\pi\sqrt{LC_T}}$$

- Frequency of Oscillations of Hartley Oscillator

$$f = \frac{1}{2\pi\sqrt{CL_T}}$$

5.9 REVIEW QUESTIONS

1. Explain the concept of feedback.
2. What do you mean by positive feedback?
3. Write a short note on negative feedback.
4. What are the advantages and disadvantages of negative feedback?
5. Explain RC phase shift oscillator.
6. What are the advantages and disadvantages of RC phase shift oscillator.
7. Explain Barkhausen criteria for oscillators.
8. Write a short note on LC oscillators.
9. Explain Colpitt oscillator in brief.
10. Explain Hartley oscillator in brief.

5.10 REFERENCE :

- 'Electronic Principles' 7th edition by Albert Malvino and David J Bates, Tata McGraw Hill.
- 'Principles of electronics' 11th edition by V.K.Mehta and Rohit Mehta, S. Chand.

❖❖❖❖

6

MULTIVIBRATORS

Unit Structure

6.0 OBJECTIVES

After studying this chapter you should be able to

- Understand the concept of multivibrators.
- Describe the working of IC 555 timer.
- Understand the working of monostable multivibrator using IC 555 timer.
- Understand the working of monostable multivibrator using IC 555 timer.
- Solve the problems based on monostable and astable multivibrators.

6.1 INTRODUCTION

This chapter discusses a popular chip called 555 timer. It is used in many applications to produce time delays, voltage controlled oscillators. Monostable and astable multivibrators can be constructed using 555 timer.

This chapter also emphasized on problems based on monostable and astable multivibrators.

6.2 MULTIVIBRATORS

- The astable multivibrator has two half stable states. It switches back and forth from one state to another. It remains in each state for a specific time depending upon the discharging of the capacitance circuit. It is an oscillatory circuit, since it requires no external triggering.

- The monostable multivibrator is also known as one shot or mono shot multivibrator. It has one stable state and one half stable state to which it returns to after application of a trigger input. It gives a single pulse of desired duration for every trigger input pulse.

- The bistable multivibrator has two stable states in which it can come to rest after some input trigger is applied. It does not oscillate and it is used as a digital memory device.

6.3 IC 555 TIMER

One of the most popular timer integrated circuit is IC 555. This timer IC was introduced by signetics corporation of U.S. in the year 1972.

IC 555 timer is widely used in a circuit that can run in either of two modes: monostable (one stable state) or astable (no stable state). In monostable mode, it can produce a accurate time delays from microseconds to hours. In astable mode, it can produce rectangular waves with a variable duty cycle.

6.3.1 Features of IC 555

- Supply voltage range : 5 to 18 volt.
- Current sinking and soursing capacity : 200 mA
- High temperature stability.
- Timing can be adjusted from microseconds to hours.
- Duty cycle of the output is adjustable.
- Output is compatible with CMOS and TTL
- Good timing stability against supply voltage variations.
- LOW Cost
- Versatile in operation

6.3.2 Functional block diagram

- The 555 Timer consists of two comparators, one R-S Flip-Flop and a discharge transistor. The simplified block diagram of the NE 555 timer, an 8-pin IC introduced by the signetics corporation.

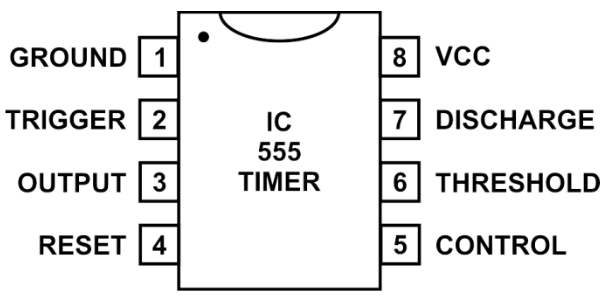


Fig. 6.1 a : The pin connection of IC 555 Timer

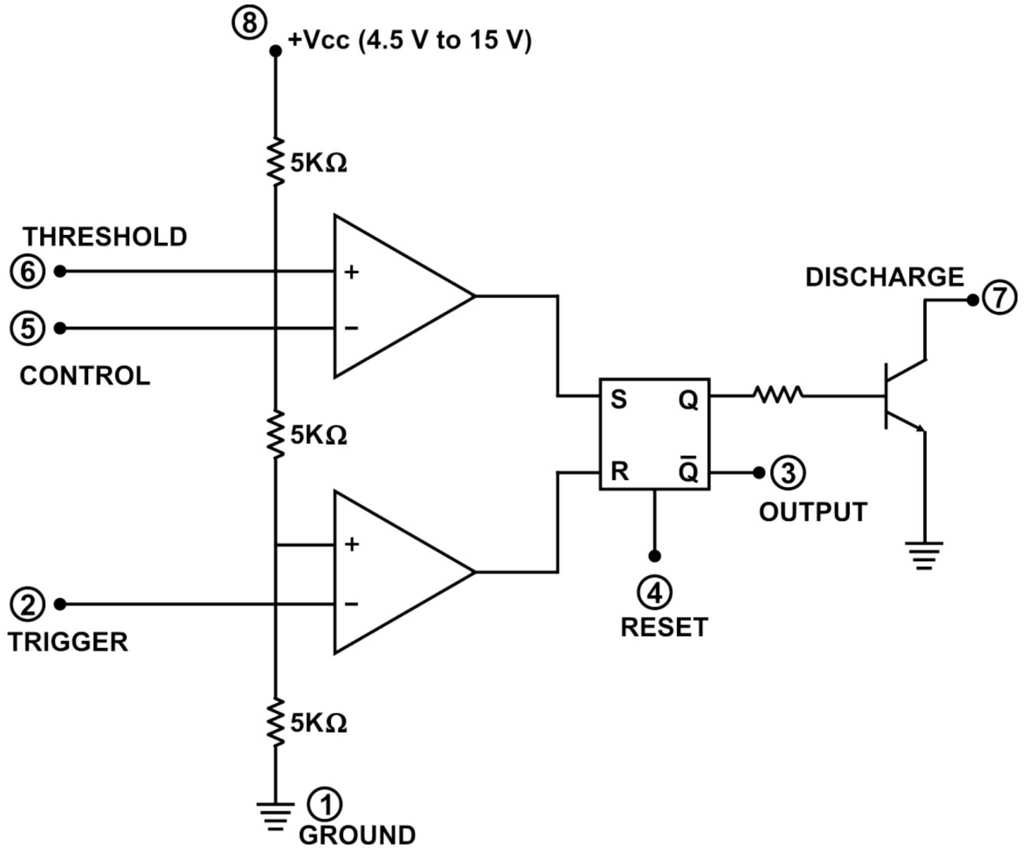


Fig. 6.1 b : Internal diagram of IC 555 timer

- Referring to Figure la Pin No. 1 is the IC's ground, While Pin No. 8 is the supply pin.

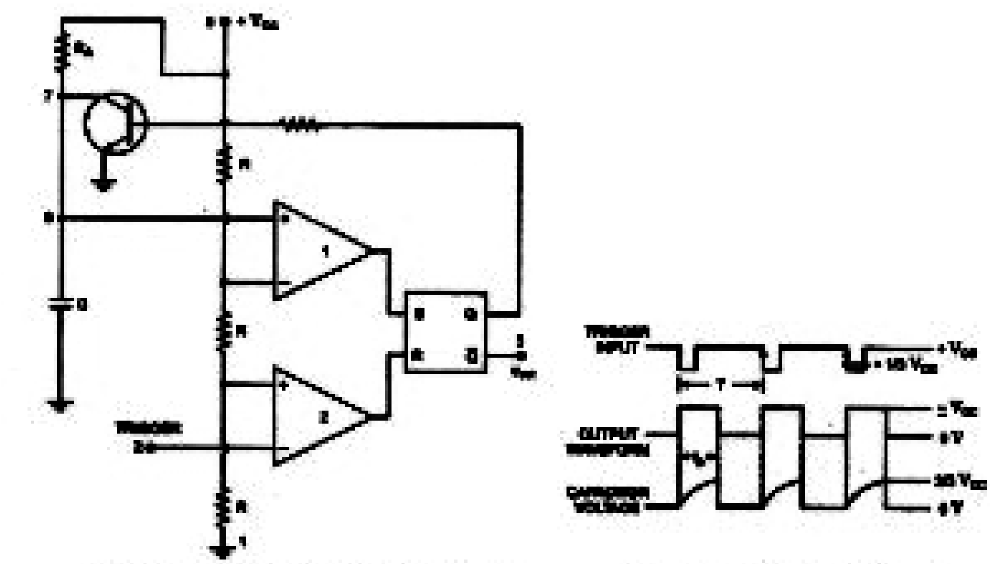
- The upper comparator has a threshold input Pin No. 6 and a control input Pin No. 5. The voltage at the control input with respect to ground is $+\frac{2}{3}$ VCC. So whenever the threshold voltage exceeds the control voltage the output of the upper comparator goes high thereby making the Flip-Flop Set.
- The Collector of the discharge transistor is connected to Pin No. 7. When this pin is connected to an external timing capacitor, a high Q-output from the Flip-Flop will saturate the transistor thereby making the capacitor to discharge. When the Flip-Flop is Reset ie. When Q is low, the transistor does not conduct and the capacitor recharges.

- The lower comparator has a Trigger input Pin No. 2 connected to its Inverting terminal. The voltage at its Non-inverting terminal with respect to ground is $+\frac{1}{3}$ VCC. When the trigger input voltage is slightly less than $+\frac{1}{3}$ VCC, the output of the lower comparator goes high which Resets the Flip-Flop.

- The complementary signal out of the Flip-Flop is connected to Pin No. 3 which is the output Pin of the IC 555 Timer.

- Pin No. 4 is the external Reset of the IC. When this pin is grounded it prevents the Flip-Flop from working. This ON/OFF. Feature is sometimes useful. In most applications, however, the external Reset is not used, but Pin No. 4 is tied directly to the supply voltage.

6.4 MONOSTABLE MULTIVIBRATORS USING AN IC 555 TIMER



Internal Circuitry With External Connections

Trigger Input, Output and Capacitor Voltage Waveforms

Monostable Operation

- The monostable type has one stable state and a quasi (half) stable state and a quasi stable state to which it returns to after the application of a trigger input. It gives a single pulse of desired duration for every trigger input pulse. In this type of multivibrator, out of the two coupling links, one link is capacitive, while the other is resistive.

- The monostable multivibrator using an IC 555 timer is shown in Figure 2. This multivibrator is also called as a one shot multivibrator.

- Referring Fig. 2 when the applied trigger input goes slightly less than + $\frac{1}{3}$ VCC, the lower comparator has a high output which resets the Flip-Flop. This cuts off the transistor, allowing the capacitor to charge.

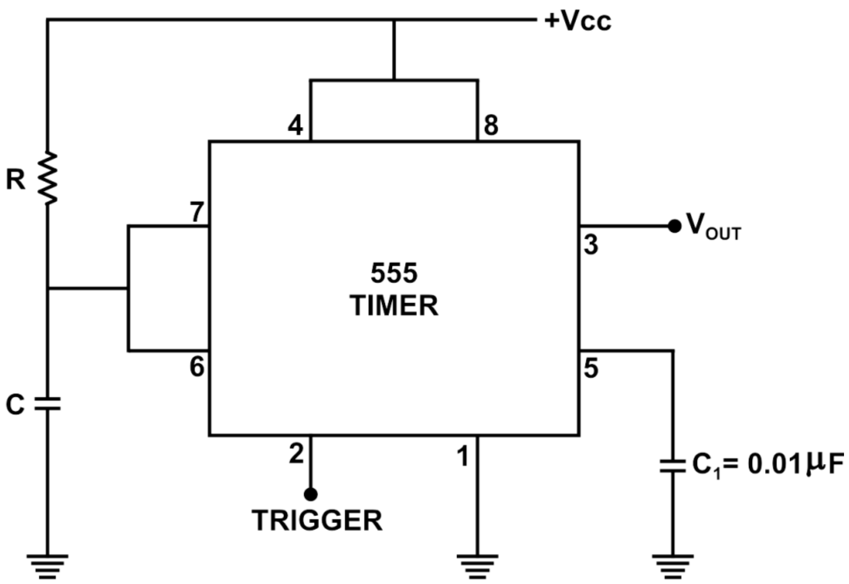


Fig. 6.2 shows the Schematic diagram of 555 Timer as an monostable multivibrator.

- As the capacitor is charging, the capacitor voltage or the threshold voltage goes on increasing. When the capacitor voltage is slightly greater than $+\frac{2}{3}$ VCC, the upper comparator has a high output, which sets the Flip-Flop. As soon as the Q-output goes high, it turns on the transistor, this quickly discharges the capacitor.

- Fig. 3 shows the waveforms at different points of the circuit. The trigger input is a narrow pulse with a quiescent value of +VCC. The pluse must drop below $+\frac{1}{3}$ VCC, to reset the Flip-Flop and allow the capacitor to charge.
- When the threshold voltage slightly exceeds $+\frac{2}{3}$ VCC, the Flip-flop sets, this saturates the transistor and discharges the capacitor. As a result we get one rectangular output pluse.

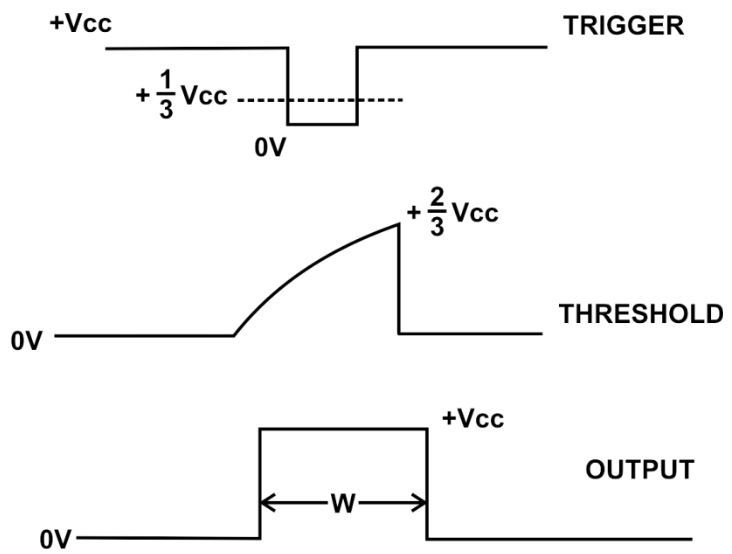


Fig.6. 3 shows the Trigger, Threshold and output Waveforms

6.4.1 Expression For Pulse Width

The instantaneous capacitor voltage is given by the equation.

$$V = V_i + (V_t - V_i)\,(1 - e^{-t/RC}) \quad ①$$

Where V = instantaneous capacitor voltage

V_i= target capacitor voltage

V_t= target capacitor voltage

t = charging time

RC= time constant

In the case of 555 – monostable multivibrator, we have

$$V = +\tfrac{2}{3} V_{CC}, \qquad V_i = 0, \quad V_t = +V_{CC},\ t = W$$

Therefore equation ① can now be written as,

$$\frac{2}{3} V_{CC} = 0 + (V_{cc} - 0)\,(1 - e^{-t/RC})$$

$$\frac{2}{3} V_{CC} = V_{cc}\ (1 - e^{-W/RC})$$

$$\frac{2}{3} = 1 - e^{-W/RC}$$

i.e. $\quad e^{-W/RC} = \tfrac{1}{3}$

or $\quad e^{-W/RC} = 3$

Taking $\log_e$ on both sides

$$\log_e e^{W/RC} = \log_e 3$$

i.e. $\quad \dfrac{W}{RC} = 1.0986$

$\therefore \quad W = 1.0986\ RC$

$\therefore \quad W \cong 1.1\ RC$

Therefore for an 555 – monostable Multivibrator, the pulse-width 'W' depends on RC.

6.5 PROBLEMS BASED ON MONOSTABLE MULTIVIBRATOR

1) For a monostable Multivibrator, the external componets are R = 15 K Ω and C = 0.1 μ F. Calculate ON time of the load voltage waveform.

Sol^n :

Given : R = 15K Ω

C = 0.1 μ F

$\therefore$ The on time,

$T_{ON} = W = 1.1\ RC$

$= 1.1 \times 15 \times 10^3 \times 0.1 \times 10^{-6}$

= 1.65 msec

2) In monostable multivibrator R = 15KΩ, the output pulse width is W= 5ms. Determine the value of C?

Sol^n :

Given : R = 10K Ω

W = 5 ms.

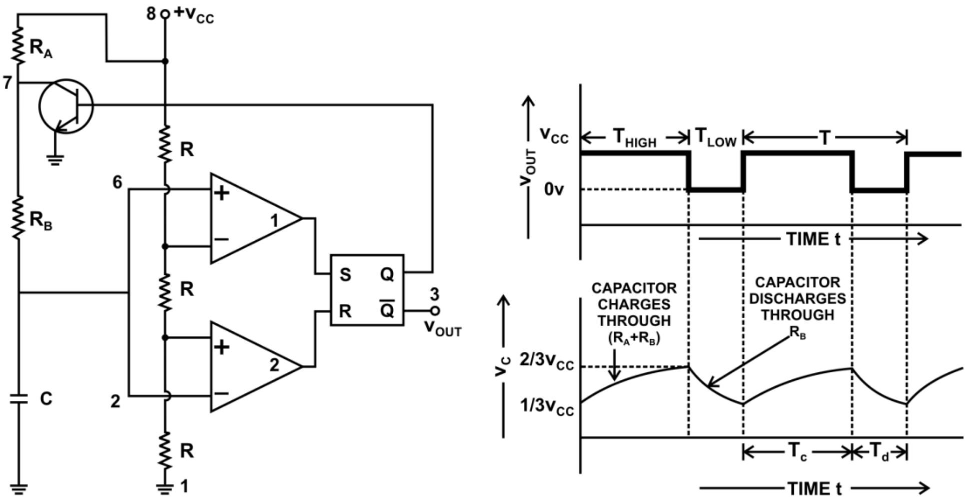
W = 1.1 RC

$$C = \frac{W}{1.1R}$$

$$C = \frac{5 \times 10^{-3}}{1.1 \times 10^{4}}$$

C = 4.5454 x 10^{-7} F.

6.6 ASTABLE MULTIVIBRATOR USING AN IC 555 TIMER



Internal Circuitry with External Connections

Capacitor and Output Voltage Waveforms

Astable Operation

- The astable Multivibrator has two quasi (half) stable states. It switches back and forth from one state to another. In this type of multivibrator both the coupling links are capacitive. It remains in each state for a specific time depending upon the discharging of the capacitive circuit. It is an oscillatory circuit, since it requires no external triggering.

- The IC 555 Timer connected as an Astable multivibrator is shown in fig. 4.a

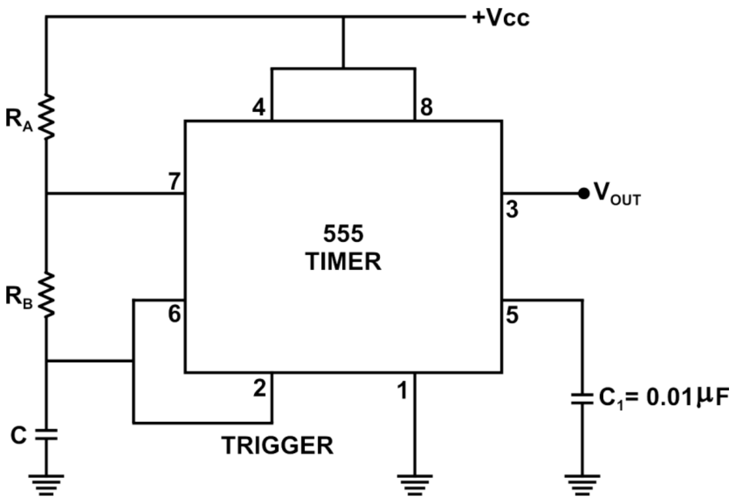


Fig. 6.4.a schematic diagram of IC 555 Timer as an astable Multivibrator

- Initially, let us assume that the Q-output of the Filp-Flop is low, there by the transistor is cut – off, hence the capacitor starts charging through a total resistance of $R_A + R_B$. Because of this, the charging time-constant is $(R_A + R_B)$ C

- As the capacitor charges, the threshold voltage increases, When the threshold voltage exceeds $+\frac{2}{3}$Vcc, then the upper comparator has a high output and this sets the Flip-Flop.

- With Q high, the transistor saturates thereby making the capacitor to discharge through R_B therefore, the discharging time constant is R_B C.

- As the capacitor is discharging, therefore when the capacitor voltage drops slightly below $+\frac{1}{3}$Vcc, the lower comparator has a high output and this resets the flip-flop. Therefore the Q output of the Flip-Flop goes low, hence the transistor is again cut-off and the capacitor starts to recharge through $R_A + R_B$. Hence the cycle continues.

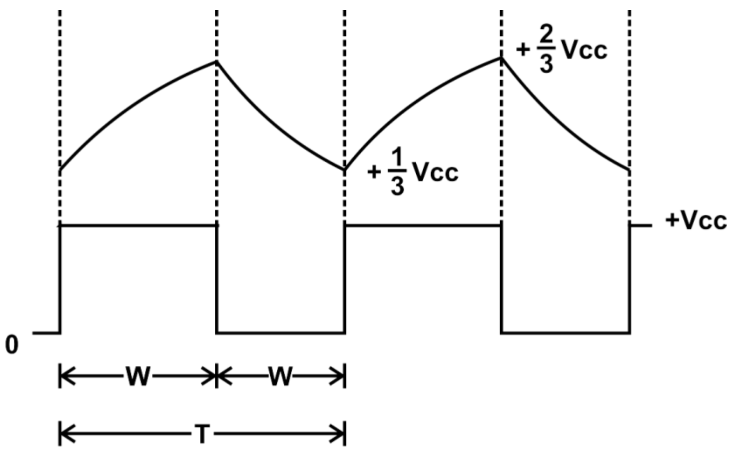


Fig.6. 4 b. capacitor and output waveforms.

- From the fig. 4.b it can been seen that the timing capacitor has an exponentially rising and falling voltage. The output is an rectangular wave.

- Since the charging time constant is more than the discharging time constant, therefore the output is not symmetrical, i.e. The high output state lasts longer than the low output state.

- To specify how unsymmetrical the output is found out by using the duty cycle equation ie.

$$\text{DUTY CYCLE D} = \frac{W}{T} \times 100\%$$

- Depending on the values of R_A and R_B, the duty cycle can be between 50% and 100%

6.6.1 Expression For the one cycle period and Frequency

The instantaneous capacitor voltage is given by equation:

$$V = V_i + (V_t - V_i)\,(I - e^{-t/RC}) \qquad ①$$

Where V = instantaneous capacitor voltage

V_i= target capacitor voltage

V_t= target capacitor voltage

t = charging time

RC = time constant

In the case of 555 – Astable multivibrator, we have

$V = +\frac{2}{3}V_{CC}$, $V_i = +\frac{1}{3}V_{CC}$, $V_t = +V_{CC}$, $t = W$, $R = (R_A + R_B)$

Therefore equation ① can now be written as,

$$\frac{2}{3}V_{CC} = \frac{1}{3}V_{CC} + (V_{CC} - \frac{1}{3}V_{CC})\left(I - e^{\frac{-W}{(R_A + R_B)C}}\right)$$

i.e. $$\frac{1}{3}V_{CC} = (V_{CC} - \frac{1}{3}V_{CC})\left(I - e^{\frac{-W}{(R_A + R_B)C}}\right)$$

$$\frac{1}{3}V_{CC} = \frac{2}{3}V_{CC})\left(I - e^{\frac{-W}{(R_A + R_B)C}}\right)$$

i.e. $$1 = 2\left(I - e^{\frac{-W}{(R_A + R_B)C}} \right)$$

i.e. $$\frac{1}{2} = I - e^{\frac{-W}{(R_A + R_B)C}}$$

$$-\frac{1}{2} = e^{\frac{-W}{(R_A + R_B)C}}$$

$$\frac{1}{2} = e^{\frac{-W}{(R_A + R_B)C}}$$

i.e. $$e^{\frac{W}{(R_A + R_B)C}} = 2$$

Taking $\log_e$ on both sides

$$\log_e e^{\frac{W}{(R_A + R_B)C}} = \log_e 2$$

i.e. $$\frac{W}{(R_A + R_B)C} = 0.693$$

$$\therefore\ W = 0.693\,(R_A + R_B)\,C \qquad ②$$

Eqn. ② gives the equation during the charging

When the capacitor discharges, it discharges only through resistor R_B. Therefore by similarly solving for the discharging equation, the discharge time (T-W) will be given as.

$$T - W = 0.693\,R_BC \qquad ③$$

Therefore the Total Time Period `T' is

$$T = 0.693\,R_BC + W$$

Substituting the value of W from eq^n ② in the above equation we get

$$T = 0.693\,R_BC + 0.693\,(R_A + R_B)\,C$$

i.e. $$T = 0.693\,C\,[R_A + 2R_B] \qquad ④$$

Equation ④ gives the total time period `T' for the 555 astable multivibrator.

As Frequency F = $\frac{1}{T}$

$$\therefore \qquad F = \frac{1}{0.693\,C\,(R_A + 2R_B)}$$

$$\therefore \qquad F = \frac{1.44}{(R_A + 2R_B)\,C} \qquad ⑤$$

Equation ⑤ gives the frequency equation for on 555 – astable multivibrator.

Similarly the Duty cycle can also be found out as follows.

Since D = $\frac{W}{T} \times 100\,\%$

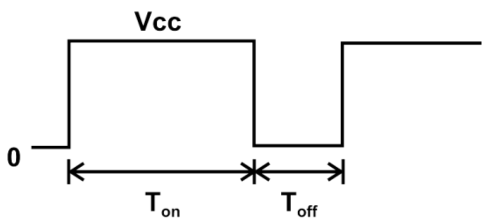
Therefore substituting eqn. ② and eqn. ④ in the above equation we get.

$$D = \frac{0.693\,(R_A + R_B)\,C}{0.693\,(R_A + 2R_B)\,C} \times 100\%$$

$$D = \frac{R_A + R_B}{R_A + 2R_B} \times 100\%$$

6.7 EXAMPLE ON ASTABLE MULTIVIBRATOR

1) In an astable circuit, R_A = 25KΩ, R_B = 33KΩ and C = 0.5 μF. Calculate the ON and OFF times of the load voltage waveform shown in below fig.



Sol^n :

The ON time corresponds to the charging time of the capacitor ie. Ton and the charging takes place through R_A and R_B

T_{ON} = 0.693 (R_A + R_B) C

= 0.693 (25 + 33) x 10^3 x 0.5 x 10^{-6}

T_{ON} = 20 msec.

The OFF time corresponds to the discharging time of the capacitor ie. T_{off} and the discharging takes place through R_B only.

T_{off} = 0.693 R_B C

$= 0.693 \times 33 \times 10^3 \times 0.5 \times 10^{-6}$

$= 11.43$ msec.

$T_{OFF} = 11.43$ msec.

2) In astable multivibrator $R_A = 2K\Omega$, $R_B = 3K\Omega$ and $C = 0.1\ \mu F$. determine pulse width and free running frequency F?

Sol^n :

$T_{ON} = 0.693\ C\ (R_A + R_B)$

$T_{OFF} = 0.693\ R_B\ C$

Given :

$R_A = 2K\Omega$

$R_B = 3K\Omega$

$C = 0.1\ \mu F$

$T_{ON} = 0.693\ C\ (R_A + R_B)$

$= 0.693 \times 0.1 \times 10^{-6}\ (2+3) \times 10^3$

$= 0.693 \times 0.1 \times 5 \times 10^{-3}$

$= 0.345$ ms

$T_{OFF} = 0.693\ R_B C$

$= 0.693 \times 3 \times 10^3 \times 0.1 \times 10^{-6}$

$= 0.207$ ms

$T = T_{ON} + T_{OFF}$

$= 0.345 + 0.207$

$= 0.552$ ms

$F = 1 / T$

$= 1 / 0.552$ ms

$= 1.811\ KH_Z$.

6.8 SUMMARY

- The astable multivibrator has two half stable states. It switches back and forth from one state to another. It remains in each state for a specific time depending upon the discharging of the capacitance circuit. It is an oscillatory circuit, since it requires no external triggering.
- The monostable multivibrator is also known as one shot or mono

shot multivibrator. It has one stable state and one half stable state to which it returns to after application of a trigger input. It gives a single pulse of desired duration for every trigger input pulse.

- The 555 Timer consists of two comparators, one R-S Flip-Flop and a discharge transistor.
- For 555 – monostable multivibrator, the pulse-width `W' depends on RC

  $$W \cong 1.1\ RC$$

- For 555 – astable multivibrator, the time period is given by

  $$T = 0.693\ C\ [R_A + 2R_B]$$

- For 555 – astable multivibrator, the duty cycle is given by

  $$D = \frac{R_A + R_B}{R_A + 2R_B} \times 100\%$$

6.9 REVIEW QUESTIONS

1. What do you mean by multivibrators?
2. What is the difference between monostable and astable multivibrators?
3. Write a short note on IC 555 timer.
4. With the help of neat circuit diagram explain monostable multivibrator.
5. Explain internal functional diagram of IC 555 timer.
6. With the help of neat circuit diagram explain astable multivibrator.
7. Design a monostable multivibrator using IC555 timer for V_{CC}=15V and pulse width of 10 ms.
8. Calculate ON time of a pulse in 555 timer monostable multivibrator if R_A= 10KΩ and C=0.1 μF.
9. In astable multivibrator R_A=20 K and R_B=22KΩ and C=0.5 μF. Calculate ON time and OFF time of output waveform.

6.10 REFERENCE

- 'Electronic Principles' 7th edition by Albert Malvino and David J Bates, Tata McGraw Hill.
- 'Principles of electronics' 11th edition by V.K.Mehta and Rohit Mehta, S. Chand.

❖❖❖❖

Unit: IV Modulation techniques

7

MODULATION

Unit structure :

7.0 OBJECTIVE :

After going through this you will be focusing on

- Communication
- Modulation
- Amplitude modulation
- Spectrum of AM and modulation index and power relationship.

7.1 INTRODUCTION:

In its basic sense, the term communications refers to the sending, receiving and processing of information by electric means. As such, it started with wire telegraphy in early stages, developing with telephony later and radio at the beginning of previous century. Radio communication made possible by the invention of the triode tube. It subsequently became even more widely used and refined through the invention and use of the transistor, integrated circuits and other semiconductor devices. More recently, the use of satellites and fiber optics has made communications even more widespread, with an increasing emphasis on computer and other data communications. A modern communications system is first concerned with the storing, processing and storing of information before its transmission. The actual transmission then follows with further processing and the filtering of noise. Finally we have reception, which may include processing steps such as decoding, storage and interpretation. In this context, forms of communications include radio communication, broadcasting, point-to-point and mobile communications, etc.

In order to become familiar with these it is necessary to know about the electronic processes and equipments and methods. Before investigating individual systems we have to define and discuss important terms such as information, noise, modulation and demodulation. In this chapter we will be discussing communication, modulation and amplitude modulation in detail.

7.2 COMMUNICATION:

Communication from its fundamental sense involve the transmission of information from one point to other point (destination) through a succession of process as follows :

1. The generation of pattern or image
2. The description of that pattern or image with certain precision by a set of symbols.
3. The encoding of these symbols in a form which is suitable for transmission over the channel.
4. The transmission of encoded symbols to desired destination.
5. The decoding and reproduction of the original symbols.
6. The recreation of original pattern or image.

The purpose of communication systems is to transmit information bearing signals from source located at one point to a user destination at some distance away.

The block diagram of a communication system consisting of transmitter, channel and receiver is shown on fig. 8.1.

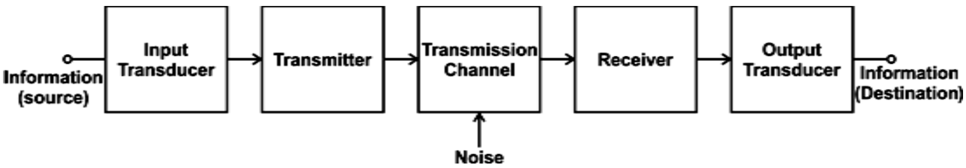


Fig.7.1 Communication system

1. **Channel:** The function of the channel is to provide a physical connection between the transmitter and receiver. It may be wired or wireless.
2. **Input Transducer:** It converts the information to be transmitted into its electrical equivalent. Ex. The microphone converts spoken words into audio signals.

3. **Transmitter :**
   - To transmit the message signal over a communication channel we need to modify it into a suitable form for efficient transmission over the channel.
   - Modification of message signal is achieved be means of a process called modulation.
   - This process involves varying some parameters of carrier wave in accordance with signal in such a way that the spectrum of modulated wave matches the assigned bandwidth.

4. **Receiver:**
   - The function of the receiver is to process the received signal so as to produce an estimate of original message signal.
   - The receiver is required to re-create the original message signal from a degraded version of the transmitted signal after propagation through the channel.

- The re-creation is accomplished by using a process called demodulation, which is inverse of modulation used in transmitter process.

5. **Output Transducer: A**fter the received signal is restored to its electrical equivalent as it was before it was modulated, it is converted back to its original form using output transducer. Ex. The loudspeaker converts the electrical signals back to audio signals.

7.3 MODULATION

- Modulation is an operation performed at the transmitter to achieve efficient and reliable information transmission.
- **Modulation** is a process of mixing a signal with a sinusoid of high frequency to produce a new signal. This new signal will have certain benefits of an un-modulated signal, especially during transmission. If we look at a general function for a sinusoid:

  $$f(t) = A \sin(\omega t + \varphi)$$

  We can see that this sinusoid has 3 parameters that affect the shape of the graph.
  A: the **magnitude** or **amplitud**e of the sinusoid,
  ω: the **frequency**, and the last term,
  φ: the **phase angle**.
  All 3 parameters can be altered to transmit data.
- The high frequency sinusoidal signal that is used in the modulation is known as the **carrier signal**, or simply "the carrier". The signal that is used in modulating the carrier signal (or sinusoidal signal) is known as the **data signal** or the **message signal**.

- A simple sinusoidal carrier contains no information of its own. In other words we can say that modulation is used because the some data signals are not always suitable for direct transmission, but the modulated signal may be more suitable.

- There are 3 basic types of modulation: Amplitude modulation, Frequency modulation, and Phase modulation.

  1. **Amplitude modulation** a type of modulation where the amplitude of the carrier signal is modulated (changed) in proportion to the message signal while the frequency and phase are kept constant.

  2. **Frequency modulation** a type of modulation where the frequency of the carrier signal is modulated (changed) in proportion to the message signal while the amplitude and phase are kept constant.

  3. **Phase modulation** a type of modulation where the phase of the carrier signal is modulated (changed) in proportion to the message signal while the amplitude and frequency are kept constant.

- Modulation involves two waveforms; a modulating signal that represents the message, and a carrier wave that suits the particular application.

- A modulator systematically alters the carrier wave in correspondence with the variations of the modulating signal. The resulting modulated wave thereby "carries" the message information. We generally require that modulation be a reversible operation, so the message can be retrieved by the complementary process of demodulation.

7.3.1 NEED FOR MODULATION

1) **Modulation for efficient Transmission**:
   - Signal transmission over appreciable distance always involves a traveling electromagnetic wave, with or without a guiding medium.
   - The efficiency of any particular transmission method depends upon the frequency of the signal being transmitted.

   - Typically, efficient line-of-sight radio propagation requires antennas whose physical dimensions are at least 1/10 of the signal's wavelength. Unmodulated transmission of an audio signal containing frequency components down to 100 Hz would thus require for antennas 300 *km long*. Modulated transmission at 100 MHz, as in FM broadcasting, allows a practical antenna size of about one meter. At frequencies below 100 MHz, other propagation modes have better efficiency with reasonable antenna sizes.

2) **Modulation to Overcome Hardware limitations**: The cost and availability of hardware is the constrained of communication system which is frequency dependent. Modulation permits the designer to place a signal in some frequency range that avoids hardware limitations. Hardware costs and complications are minimized by keeping the fractional bandwidth within 1-10 %.

3) **Modulation for Frequency Assignment:** When you tune a radio or television set to a particular station, you are selecting one of the many signals being received at that time. Since each station has a different assigned carrier frequency, the desired signal can be separated from the others by filtering. Due to modulation it is possible to modulate different sound signals with different frequency carriers thereby create the modulated signals that occupy different slots of the frequency spectrum and avoid a jumble of signals.

4) **Modulation for Multiplexing:** Multiplexing is the process of combining several signals for simultaneous transmission on one channel. Ex. Frequency-division multiplexing uses modulation to put each signal on a different carrier frequency and several filters separates the signals at the destination.

5) **Increases the range of communication:** The frequency of base band signal is low. Therefore it cannot travel over a long distance. When such signals are transmitted, they get heavily attenuated. The attenuation of the signal reduces with increase in frequency of transmitted signal and they travel larger distance. The modulation process up shifts the frequency of the signal to be transmitted. Therefore it increases the range of the communication.

7.3.2 CONCEPT OF MODULATION :

- Modulation is the fundamental requirement for a communication system. Modulation plays a very important role in transmission of signals over long distances.
- The characteristics of analog signal namely amplitude, frequency and phase are commonly used (altered) for transmission.
- Amplitude is the distance from the central axis, frequency is the number of complete wave cycles per unit time, and phase is the relative position within one complete cycle.
- To transmit signal transmitter and receiver need a common medium signaling mechanism. This connection between transmitter and receiver is provided by electronic or electromechanical equipment by using a carrier signal sent continually from the source to the destination.
- The amplitude, frequency, or phase characteristics of the carrier are then varied or modulated in such a way that the information to be transmitted is super imposed upon and carried along with the basic signal.
- Two signals are necessary to implement modulation, one low frequency signal and one high frequency signal. The message is in the low frequency band. The high frequency signal is called the “carrier” because it carries low frequency message signal. The low frequency signal changes one or more characteristics of the high frequency carrier signal to induce modulation.
- Modulation can be defined as the process in which some characteristics of a high frequency carrier wave is varied in accordance with the instantaneous value of the message, generally called modulating signal.

7.4 AMPLITUDE MODULATION (AM) :

- Amplitude Modulation is the process of varying the amplitude of high frequency carrier signal in accordance with the instantaneous amplitude of the modulating signal to obtain the modulated signal.

- An **Amplitude Modulated (AM)** signal can be produced by using the instantaneous amplitude of the information signal to vary the peak amplitude of a higher frequency signal. The high frequency signal that is combined with the information signal to produce the modulated waveform is called carrier; in practice ratio between carrier frequencies and modulating frequency is high.
- Consider high frequency carrier signal, modulating signal and the modulated signal as shown in fig. 7.2a, 7.2b &7.2c respectively.

- Upto point A, modulating signal has zero amplitude therefore the amplitude of carrier signal **V_c at point A is 0(zero).**

- As the amplitude of the modulating signal increases the amplitude of the AM also increases (above V_c). **At point B**, amplitude of the AM waveform is maximum **$V_c + V_m$.**

- As the amplitude modulating signal decreases the amplitude of the Am waveform also decreases and **at the point C the amplitude of modulating signal is zero**. Hence amplitude of AM waveform is V_c.

- **At point D,** amplitude of the AM waveform is minimum and it is **$V_c - V_m$**. This shows that modulating signal is superimposed and added with carrier signal.

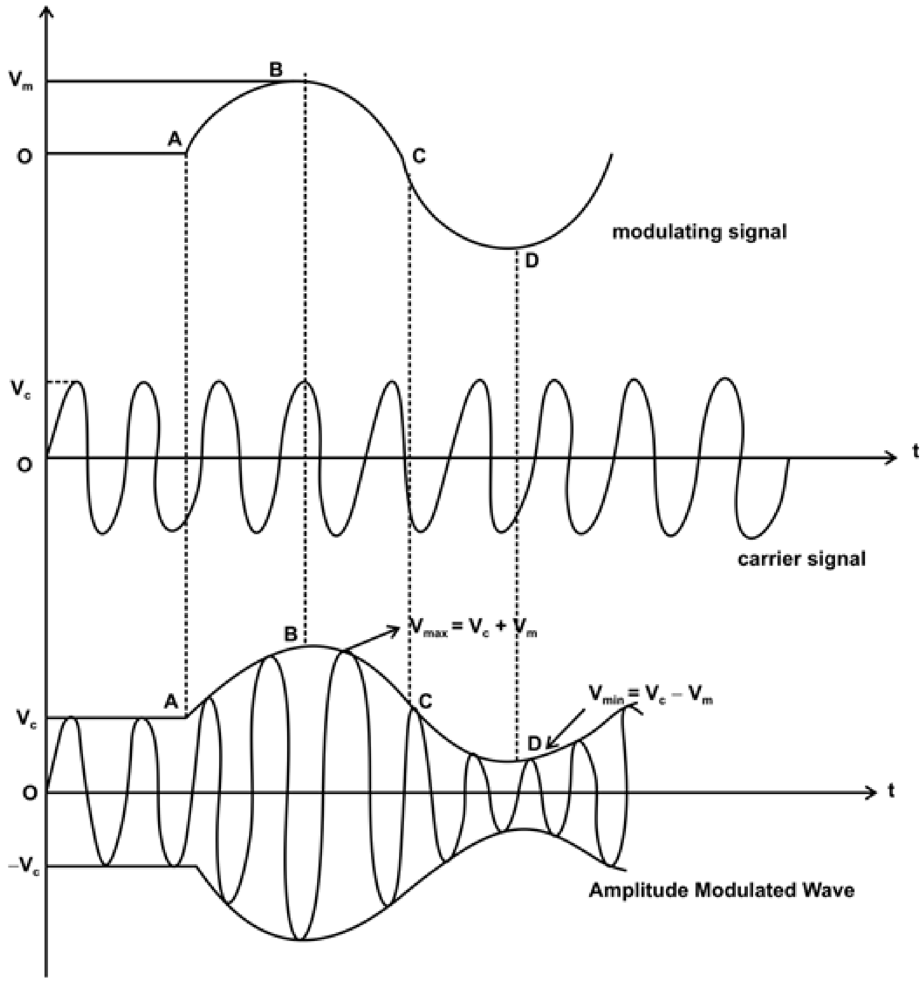


Fig 7.2 a) modulating signal b) carrier signal c) Amplitude modulated wave.

- For proper modulation the amplitude and frequency of carrier should be greater than that of the modulating signal.

- If the peak of the individual waveform of the modulated signal are connected the resulting envelop resembles the original modulating signal. The envelop repeats at the modulating frequency and the shape of each half (positive or negative) is the same as that of the modulating signal.

Summary of AM:

1) The amplitude of the carrier wave varies in accordance with the information signal.

2) Frequency domain shows that the signal component at the carrier frequency remains intact, with the same amplitude and frequency as before.

3) The amplitude of the entire signal does not change with the modulation.

4) AM is not the simple linear addition of the two signals. AM is essentially a nonlinear process. As in any nonlinear interaction between signals sum and difference frequencies are produced that contain the information to be transmitted.

7.5 ANALYSIS OF AM WAVE:

- Amplitude modulation is created by using the instantaneous modulating signal voltage to vary the amplitude of the modulated signal. The carrier is almost a sine wave. The modulating signal can be sine wave but it is more often an arbitrary wave form like audio signal. The relationship is given by ;

$$V_t = (E_C + e_m) \sin\omega_c t \quad \text{----------------------------------} (1)$$

Where ; V_t = instantaneous amplitude of the modulated signal in volts.

E_C = peak amplitude of the carrier signal in volts.

e_m = instantaneous amplitude of the modulated signal in volts

ω_c = frequency of the carrier in radians

t = time in seconds

The addition of E_C and e_m is algebraic.

Amplitude modulation involves the addition of the instantaneous amplitude of modulating signal(e_m) to the peak carrier amplitude (E_C) of the modulating signal

Equation (1) has the form ;

$$\mathbf{V_t = (E_C + E_m \sin \omega_m t) \sin\omega_c} t \quad \text{----------------------}(2)$$

Where, E_m = peak amplitude of the modulating signal in volts

ω_m = frequency of modulating signal in radian.

Another/ alternative method for determination of amplitude of the AM wave is as follows;

The unmodified carrier wave and modulating signal is ;

$$\left.\begin{aligned} e_c &= E_c \cos \omega_c t \\ e_s &= E_s \cos \omega_s t \end{aligned}\right\} \quad \text{--------------------------------} (I)$$

where, $\mathbf{e_c}$ = instantaneous amplitude of carrier signal in volts

$\mathbf{e_s}$ = instantaneous amplitude of modulating signal in volts

$\mathbf{E_c}$ =Peak amplitude of carrier signal in volts

$\mathbf{E_s}$ = Peak amplitude of modulating signal in volts

and $\boldsymbol{\omega_c >> \omega_s}$

New modulated wave will be represented as ;

$$e_m = E_m \cos \omega_c t \quad \text{-------------------------------}(II)$$

$$= (E_c + Ke_s) \cos \omega_c t ,$$

K is proportionality constant

$$= E_c \cos \omega_c t + Ke_s \cos \omega_c t$$

$$= E_c \cos \omega_c t + KE_s \cos \omega_s t \cos \omega_c t$$

$= E_c \cos \omega_c t + K\left(\frac{E_s}{E_c}\right) E_c \cos \omega_c t \cos \omega_s t$

$= E_c \left[1 + K\left(\frac{E_s}{E_c}\right) cos\omega_s t\right] \cos \omega_c t$

Let $K\left(\frac{E_s}{E_c}\right) = m$, modulation constant or index.

$\therefore \quad e_m = E_c\,[1 + m\,cos\omega_s\,t]\,cos\omega_c\,t$ ----------------------(III)

The instantaneous amplitude of the AM wave is ,

$$e_m = E_c[1 + m\,cos\omega_s\,t]\,cos\omega_c\,t$$

$$= E_c\,cos\omega_c t + mE_c cos\omega_c cos\omega_s t$$

$$= E_c\,cos\omega_c t + \frac{mEc}{2}(2cos\omega_c cos\omega_s t)$$

$$= E_c\,cos\omega_c t + \frac{mEc}{2}[cos(\omega_c + \omega_s)t + [cos(\omega_c - \omega_s)t]$$

$$= E_c\,cos\omega_c t + \frac{mEc}{2}cos(\omega_c + \omega_s)t + \frac{mEc}{2}cos(\omega_c - \omega_s)t$$

---------------------------(IV)

7.6 MODULATION INDEX

- Modulation index is defined as the ratio between the amplitudes of modulating signal and the carrier signal. It is expressed by ‘m. Mathematically it is expressed as; $m = \frac{E_m}{E_c}$

- The modulating signal is ,

$$V(t) = (E_c + E_m sin\omega_m t)sin\omega_c t$$

$$= E_c\left(1 + \frac{E_m}{E_c}\,sin\omega_m t\right)sin\omega_c t$$

$$= E_c(1 + m\,sin\omega_m t)sin\omega_c t$$

  - m varies between 0 to 1. To start with m = 0, E_m = 0 we get un-modulated carrier.

  - Following figures shows the variation of ‘m’ and corresponding E_m values.

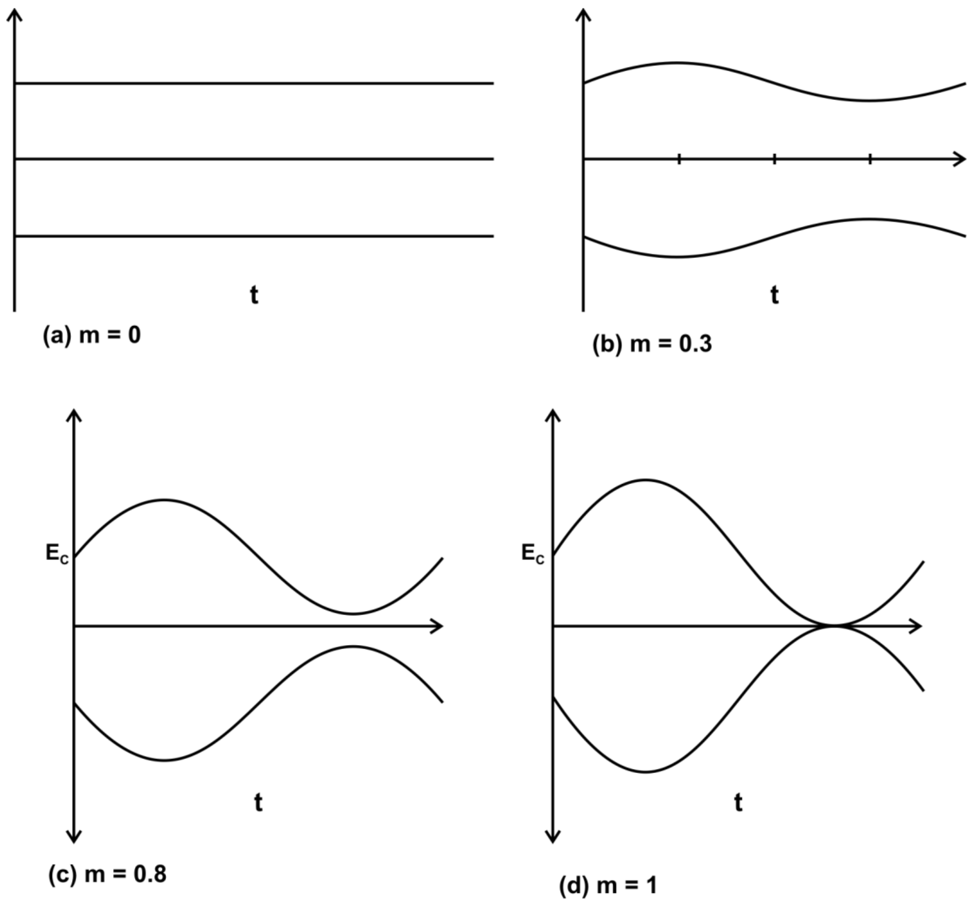


Fig. 7.3 AM wave when modulation index
(a) m =0 (b) m= 0.3 (c) m = 0.8 (d) m = 1

Modulation index can also be expressed in % ;

(a) If m = 0 and amplitude of carrier wave = A

i.e. E_m = 0 and E_c = A

% modulation index = $\frac{0}{A} \times 100\% = 0\ \%$

(b) If E_c = A and E_m = 0.5A then m = 0.5

% modulation index =0.5 x 100 = 50%

(c) If E_c = A and E_m = 0.8A then m = 0.8

% modulation index =0.8 x 100 = 80%

(d) if E_c = A and E_m = A then m = 1

% modulation index =1 x 100 = 100%

7.6.1 MEASUREMENT OF MODULATION INDEX

- If E_m and Ec are the peak modulation and carrier voltages then the maximum envelop voltage is given by ;

$$E_{max} = E_c + E_m$$
$$E_{max} = E_c\left(1 + \frac{E_m}{E_c}\right)$$
$$E_{max} = E_c(1 + m) \quad \ldots\ldots\ldots\ldots \; as\ m = \frac{E_m}{E_c}$$

- The minimum envelop voltage is

$$E_{max} = E_c - E_m$$
$$E_{max} = E_c(1-m)$$

- Using above equations we can show that ;

$$m = \frac{E_{max} - E_{min}}{E_{max} + E_{min}}$$

- Thus for m = 0 the peak voltage is E_c and for m = 1 the envelop voltage varies from 2E to zero.

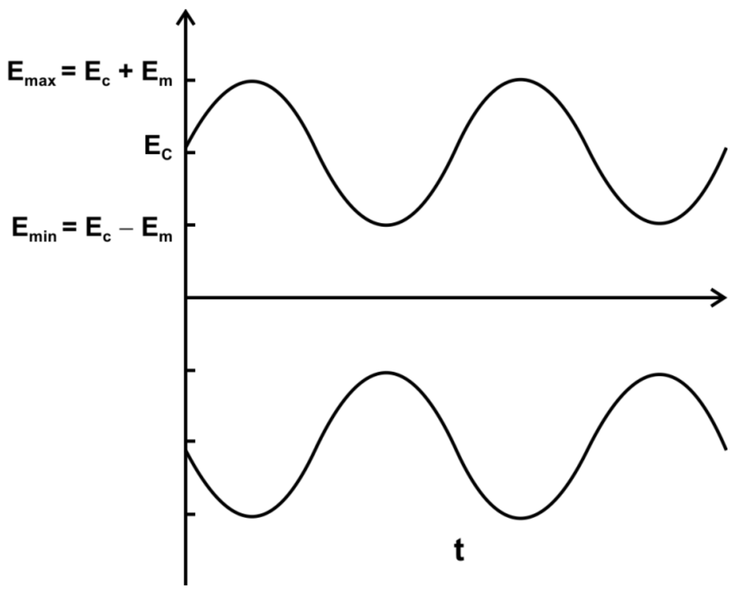


Fig.7.4 voltage relationship in AM signal.

7.6.2 MODULATION INDEX FOR MULTIPLE MODULATING FREQUENCIES:

- In most practical cases the information signal is audio signal which contains multiple frequencies. All the more the signal is not periodic.
- When there are more than two sine waves of different frequencies i.e. frequencies which are not the multiples of each other, modulating a single carrier ***m*** can be calculated by using following formula;

$$m_T = \sqrt{m_1^2 + m_2^2 + \cdots \ldots}$$

Where, m_T = total resultant modulation index
m_1, m_2, …. = modulation indices of different modulating components.

7.7 SPECTRUM OF AM AND MODULATION INDEX:

- In AM both carrier and modulating signal may be sine waves but modulated wave is not a sine wave.

- The instantaneous amplitude of the AM wave is given by:

$$e_m = E_c\, cos\omega_c t + \frac{mEc}{2} cos(\omega_c + \omega_s)t + \frac{mEc}{2} cos(\omega_c - \omega_s)t$$

- Thus besides the original signal there are two additional sinusoidal waves one above the carrier and the other below the carrier frequency. The complete signal consists of carrier wave and two additional frequencies one on each side, which are called **side frequencies.**

- The separation of the each frequency from the carrier is equal to the modulating frequency and the relative amplitude of the side frequency compared with that of the carrier, is proportional to the modulation index 'm', becoming half of the carrier voltage for m = 1.

- Each modulating frequency produces two side frequencies. Those above the carrier are called **upper side band** and the other is **lower side band**. Fig.7.5 shows AM frequency domain. (carrier, usb , lsb)

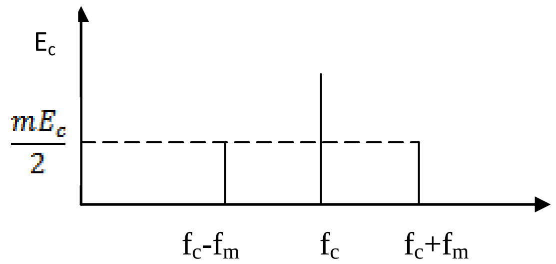


Fig. 7.5 AM frequency domain

From the above explanation;

$$f_{usb} = f_c + f_m \quad \text{and} \quad f_{lsb} = f_c - f_m$$

$$\therefore E_{lsb} = E_{usb} = \frac{mE_c}{2}$$

Where f_{usb} = frequency of upper side band
f_{lsb} = frequency of lower side band
E_{usb} = peak voltage of upper side band
E_{lsb} = peak voltage of lower side band

7.8 POWER RELATIONSHIP IN AM WAVE:

- Since AM wave contains carrier along with two side bands, total power in AM wave will be

$$P_T = P_{carr} + P_{lsb} + P_{usb}$$

$$P_T = \frac{V_c^2}{R} + \frac{V_{lsb}^2}{R} + \frac{V_{usb}^2}{R} \quad \text{------------------------(1)}$$

- In the above expression are in RMS value.

  For peak value the above equation becomes;

$$P_T = \frac{\left(\frac{V_c}{\sqrt{2}}\right)^2}{R} + \frac{\left(\frac{mV_c}{2\sqrt{2}}\right)^2}{R} + \frac{\left(\frac{mV_c}{2\sqrt{2}}\right)^2}{R}$$

$$P_T = \frac{V_c^2}{2R} + \frac{m^2V_c^2}{8R} + \frac{m^2V_c^2}{8R}$$

$$\therefore P_T = \frac{V_c^2}{2R} + \frac{m^2V_c^2}{4R}$$

$$\therefore P_T = \frac{V_c^2}{2R}\left(1 + \frac{m^2}{2}\right)$$

$$P_T = P_c\left(1 + \frac{m^2}{2}\right)$$

- From the above equation it is seen that as the % modulation increases the power required for AM transmission also increases. For m = 1, $P_T = 1.5\ P_c$

- Therefore maximum power transmitted by AM wave is 1.5 times unmodulated carrier power.

For m = 1, $P_T = \frac{3}{2}P_c$

$$\therefore P_c = \frac{3}{2}P_T$$

$\therefore$ P_c = 67 % of P_T

- **Conclusion:**
  Out of total power transmitted 67% is carrier power but it does not contain information. Hence it is wastage of power the useful information in side bands. There are two sidebands which contain same information. Therefore by transmitting one side band also we can recover original signal at the receiver. Therefore in Am carrier power and half the side band power is waste.

7.9 SUMMARY:

- Communication from its fundamental sense involve the transmission of information from one point to other point (destination) through a succession of processes

- The transmission has a function of processing the message signal into a form suitable for transmission over the channel , such an operation is called **modulation**

- Two signals are necessary to implement modulation, one low frequency signal and one high frequency signal. The message is in the low frequency band. The high frequency signal is called the "carrier" because it carries low frequency message signal.

- Modulation index is defined as the ratio between the amplitudes of modulating signal and the carrier signal. It is expressed by 'm. Mathematically it is expressed as ; $m = \frac{E_m}{E_c}$
- Total producer in AM signal increases with modulation. For 100% modulated signal total power is 50% greater than the unmodulated carrier.

- AM transmission is more efficient when m = 1

7.10 UNIT END EXERCISE

1. Explain with the help of block diagram the term communication system.
2. What is meant by Modulation? What is the need for modulation /
3. What is AM? Obtain the equation for the instantaneous amplitude of the AM wave?
4. Define modulation index. Obtain the expression for modulation index.
5. Show that maximum power transmitted by AM wave is 1.5 times unmodulated carrier power.
6. Explain what do you mean by spectrum of AM wave?

❖❖❖❖

8

GENERATION OF AM

Unit structure :

8.0 OBJECTIVES :

While reading this chapter you will understand the amplitude modulated wave, generation of AM wave using various methods, demodulation of Am and its detection.

8.1 INTRODUCTION :

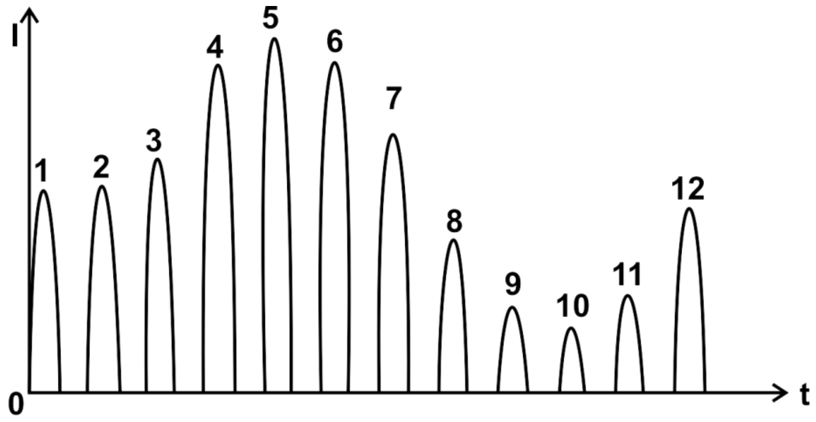
There are two types of-devices in which it is necessary to generate amplitude modulation. The first of these, the AM transmitter which generates such high powers that its prime requirement is efficiency. Hence quite complex methods of AM generation may be employed. The other device is the laboratory AM generator. Here AM is produced at such a low power level that simplicity is a more important requirement than efficiency.

8.2 BASIC REQUIREMENTS-COMPARISON OF LEVELS

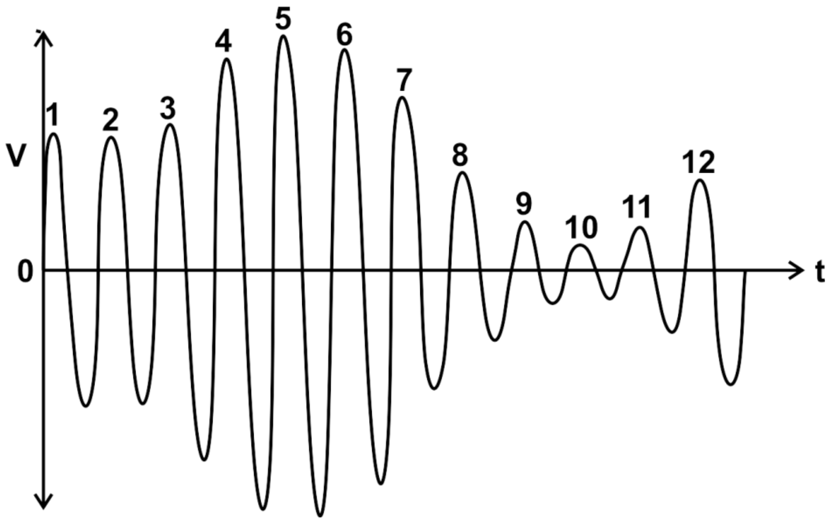
- In order to generate the AM wave it is necessary to apply the series of current pulses as shown in Fig. 8.1 to a tuned circuit. Each pulse would initiate a damped oscillation in the tuned circuit. The oscillation would have initial amplitude proportional to the size of the current pulse and a decay rate dependent on the time constant of the circuit.

- Since a train of pulses is fed to the tank circuit, it will cause a complete sine wave proportional in amplitude to the size of this pulse. This will be followed by the next sine wave, proportional to the size of the next applied pulse, and so on.

- AM wave will result if the original current pulses are made proportional to the modulating voltage. The process is known as the flywheel effect of the tuned circuit



(a) Current pulses to tuned circuit



(b) Tuned circuit AM voltage

Fig 8.1 Current requirements in AM

- In an AM transmitter, amplitude modulation can be generated at any point after the radio frequency source. If the output stage in a transmitter is collector modulated in a lower-power transmitter the system is called high-level modulation. If modulation is applied at any other point, including some other electrode of the output amplifier, then so-called low-level modulation is produced.

- **LOW LEVEL VS HIGH LEVEL MODULATION**

  Low level modulation

  I. Modulation occurs prior to the stage that drives the output through the transmitter.
  II. Less modulating signal power is required to achieve high percentage modulation

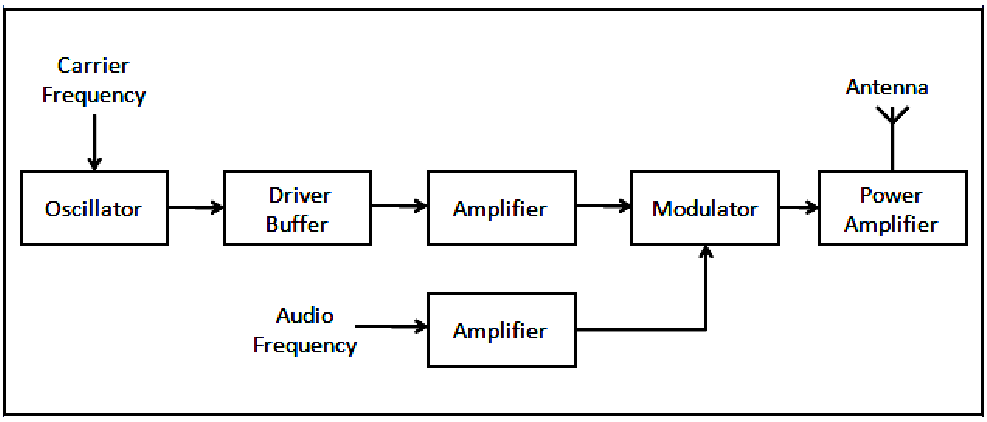


Fig 8.2 Low level modulation

High level modulation

I. Modulation occurs in the stage that drives the output through the transmitter

II. High modulating signal power is required to achieve high percentage modulation, since the modulating signal is applied to the carrier that is already amplified.

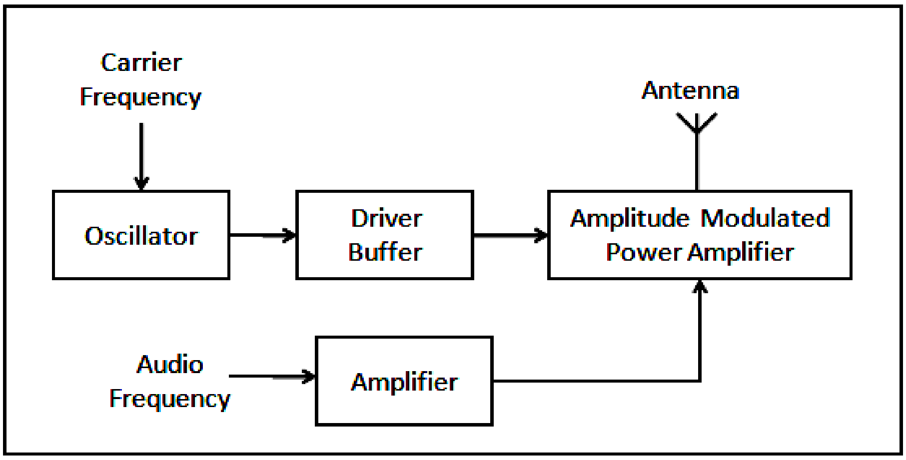


Fig 8.3 High level modulation

Note: The above diagrams are over simplified for better understanding

8.3 LOW LEVEL MODULATOR CIRCUIT

Emitter Modulator

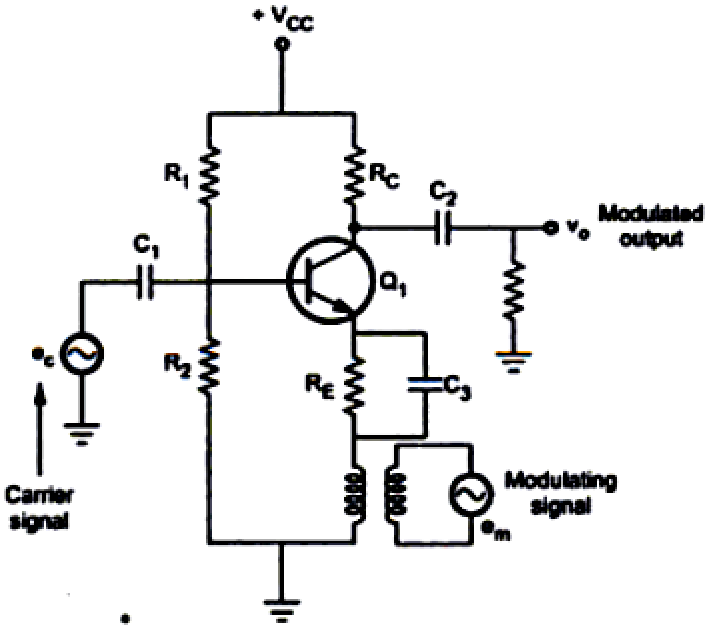


Fig 8.4 Emitter Modulator

- Carrier Signal is given to the base of the transistor and modulating signal is given to the emitter.
- In the absence of modulating signal, the fixed voltage V_{CC} will be applied to the collector generates current pulses of constant amplitude and the output of tuned circuit will be steady sine wave.
- When a modulating signal is applied to the emitter, the amplifier produces a gain that varies according to the voltage of the modulating signal.
- The amplification in the carrier signal depends upon the variation in the gain obtained. Thus the amplitude of the carrier signal is modulated by the modulating signal.

8.4 HIGH LEVEL MODULATOR CIRCUIT

Collector modulator

- Modern AM transmitter uses transistor for Modulation. Both collector and base modulation of transistor is possible but collector modulation is generally preferred. Circuit is shown in fig 8.8.

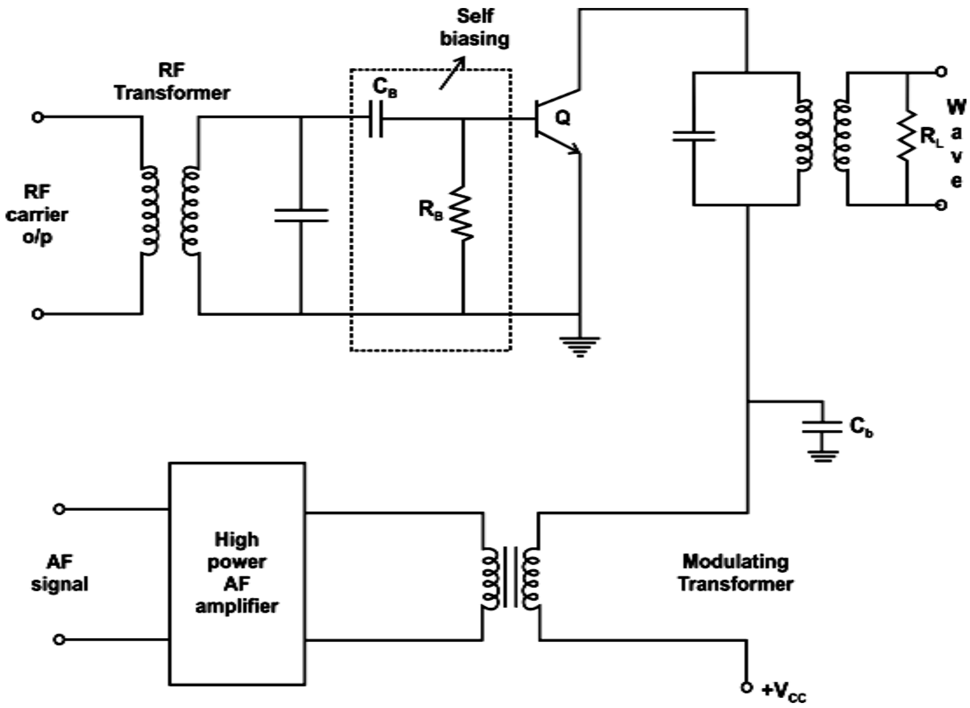


Fig. 8.5 Collector modulator

- RF carrier signal is applied to the base of the transistor through RF transformer, C_B and R_B are used to provide self –ve bias to the base of the transistor, so that transistor operates in class C mode.
- Since the collector modulator is high level modulator the power level of AF signal is increased by high AF amplifier. This high power AF signal is given in series with V_{CC} to the collector of class c amplifier.
- In the absence of modulating signal, the fixed voltage V_{CC} will be applied to the collector generates current pulses of constant amplitude and the output of tuned circuit will be steady sine wave.
- When modulating signal is applied, AC voltage across the secondary of modulating transformer will be added to and subtracted from $V_{CC.}$ Such a varying supply voltage is applied to the collector of class C amplifier which results in variation of current pulses at the collector of the transistor. Then we get current pulses whose amplitude varies with the collector bias which in turn is proportional to the modulating signal. Such current pulses given to a tank circuit, will oscillate once for each pulse and produce a desired AM wave.

8.5 DEMODULATION OF AM:

- Demodulation is the reverse process of modulation.
- It is the process of recovering original modulating signal from modulated waveform.
- The circuit used for this process is called **Demodulator.**The process of Demodulation is also called as Detection
- There are two types of AM detectors:
    A. Square Law Detector
    B. Envelope Detector

Envelop detector :

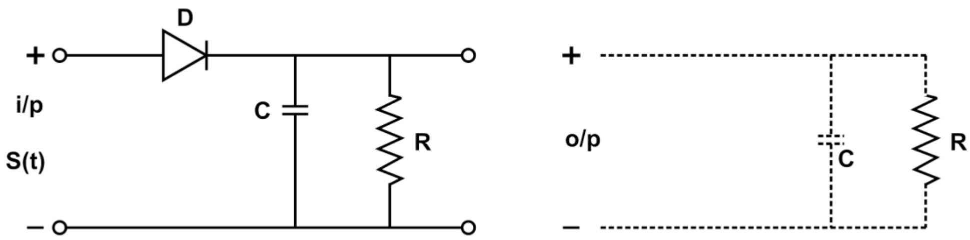


Fig 8.6 simple envelop detector

- An envelope detector is a simple and yet highly effective device that is well-suited for the demodulation of a narrow-band AM wave (i.e. the carrier frequency is large compared with the message bandwidth), for which the percentage modulation is less than 100%.
- Ideally, an envelope detector produces an output signal that follows the envelope of the input signal waveform exactly. Figure 8.6 shows the circuit diagram of an envelope detector that consists of a diode and a resistor-capacitor filter.

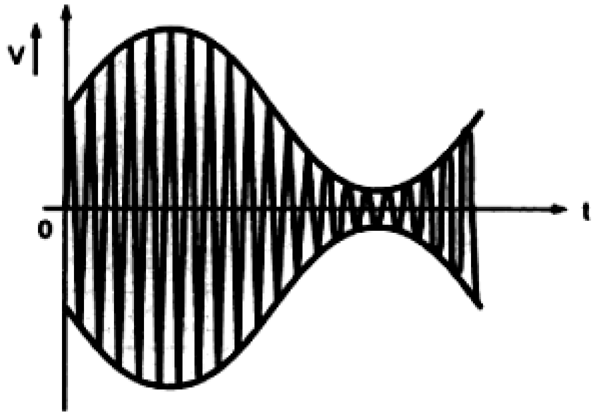
Operation:

- On the positive half-cycle of the input signal, the diode is forward-biased and the capacitor C charges up rapidly to the peak value of the input signal.
- When the input signal falls below this value, the diode becomes reverse-biased and the capacitor C discharges slowly through the load resistor R_l
- The discharging process continues until the next positive half-cycle. When the input signal becomes greater than the voltage across the capacitor, the diode conducts again and the process is repeated.
- It is assumed that,
    1. The diode is ideal, presenting zero impedance to current flow in the forward-biased region, and infinite impedance in the reverse-biased region.
    2. The AM wave applied to the envelope detector is supplied by a voltage source of internal impedance R_s The charging time constant R_sC must be short compared with the carrier period $\frac{1}{f_c}$ that is, $R_sC < \frac{1}{f_c}$.
- Hence, the capacitor C charges rapidly and thereby follows the applied voltage up to the positive peak when the diode is conducting.
- On the other hand, the discharging time constant R_lC must be long enough to ensure that the capacitor discharges slowly through the load resistor R_l between positive peaks of the carrier wave, but not so long that the capacitor voltage will not discharge at the maximum rate of change of the modulating wave,

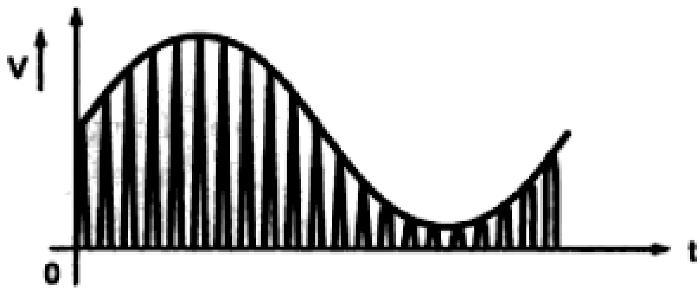
$$\frac{1}{f_c} < R_l C < \frac{1}{\omega}$$

  Where ω is the message bandwidth.

- The result is that the capacitor voltage or detector output is very nearly the same as the envelope of the AM wave, shows in Figs. 8.7 (a) and (b).



(a) AM Signal



(b) Current pulses through diode D

(c) Demodulating Signal

8.6 SUMMARY :

- AM wave will result if the original current pulses are made proportional to the modulating voltage. The process is known as the flywheel effect of the tuned circuit

- A class C amplifier may be modulated by the introduction of the modulating voltage in series with the grid bias.

- Plate current pulse for each RF i/p cycle whose amplitude is proportional to the modulating signal and whose frequency is carrier frequency. Such pulses are applied to the LC tank circuit. For each pulse tank circuit oscillates once and produces desired AM wave.

- Modern AM transmitter uses transistor instead of triodes or vacuum tubes. Both collector and base of transistor can modulated but collector modulation is generally preferred.

- Demodulation is the process of recovering original modulating signal from modulated waveform

8.7 UNIT END EXERCISE

1. Draw block diagram of Am transmitter. Explain current requirements in AM?
2. Explain with the help of circuit diagram the Grid-Modulated Class C Amplifier **?**
3. Explain the working of plate modulated class C amplifier circuit for the generation of AM.
4. Draw input and output wave forms of plate modulated AM wave and discuss them.
5. Explain with suitable diagram collector modulator method for generation of AM.
6. What do you mean by demodulation? Discuss analytically.
7. What is envelop detector explain with circuit diagram.

❖❖❖❖

9

SINGLE SIDE BAND (SSB) TECHNIQUES

Unit structure :

9.0 OBJECTIVES :

In this chapter we will be studying the SSB (single sideband) techniques in the generation of AM. In this chapter we will show the various methods of generation of SSB modulated wave and its demodulation.

9.1 INTRODUCTION:

- In the study of AM wave, we have discussed that a carrier and two sidebands are produced in AM generation.

- To reconstruct the original signal it is not necessary to transmit all the signals to receiver with enough information. The carrier as well as one of the two side bands may be removed or attenuated.

- The resulting signal will acquire less transmitted power, will occupy less bandwidth and acceptable communication is possible.

9.2 Description Of SSB And Power Requirements:

- In case of AM generation, when a carrier is amplitude modulated by a single sine wave the resulting signal consists of three frequencies: the original carrier frequency(fc), the upper side band frequency (fc + fm) and lower side band frequency (fc - fm).

- Most of the power in an AM signal is in the carrier, which contains no information. It would be a better use of bandwidth and power to send just one of the sidebands, without the carrier.

- Such an AM signal, consisting of only one sideband, is known as a single sideband suppressed carrier signal, which is often shortened to SSB.

9.3 POWER REQUIREMENTS :

- It is fact that the carrier of double sideband full carrier (DSBFC) AM conveys no information as the carrier component remain constant in amplitude and frequency irrespective of modulating voltage. The two side bands are images of each other since each is affected by changes in modulating voltage amplitude through exponent $\frac{m\,V_c}{2}$. Hence all the information can be conveyed by the use of one side band only. The carrier is superfluous and the sideband is redundant.

- The AM power equation states that the ratio of total power to carrier power is ;

$$\frac{P_T}{P_C} = \left(1 + \frac{m^2}{2}\right)$$

- Thus if carrier is suppressed then only sideband power remain is $\frac{P_C m^2}{2}$. If one of the sideband is removed the remaining power is $\frac{P_C m^2}{4}$.

- Thus 2/3 saving is effected at 100% modulation and 50% is over carrier suppressed AM.

- **Advantages of SSB** :
  1. Less transmitting power is required to produce same quality signal in the receiver as it is achieved in DSB-FC. The power saving in SSB = $\frac{power\ saved}{total\ power}$

$$= \frac{P_C + \left(\frac{m^2 P_C}{4}\right)}{P_C\left(1+\frac{m^2}{2}\right)}$$

$$= \frac{1 + \left(\frac{m^2}{4}\right)}{\left(1+\frac{m^2}{2}\right)}$$

If m = 1, them % power saving in SSB = 83.33%

  2. Bandwidth reduces to half and bandwidth of SSB is f_m.
  3. The bandwidth reduces to half the number of channels that can be transmitted over a given frequency range becomes double.
  4. As the bandwidth decreases, noise picked up by the R_x at its output decreases. Therefore signal to noise ratio increases.

9.4 CONCEPT OF BALANCED MODULATOR:

- The side bands of AM or DSB signal contains new frequencies that were not present in the carrier or message. The modulator must be therefore be a time-varying or nonlinear system because low transmission information system never produce new frequency components.

9.4.1 Balanced modulator :

The circuit of balanced modulator is shown in the following figure.

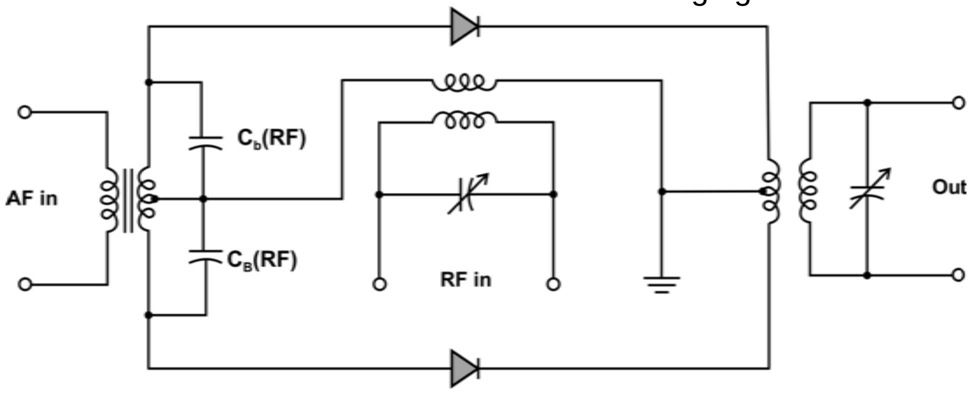


Fig.9.1 Balanced modulator

- The circuit can be constructed by using pair of identical diodes.

- The modulating voltage V_2 is fed in push pull and the carrier voltage V_1 in parallel to a pair of identical diodes. The modulated output of the two diodes is combined in the centre tapped primary of the push pull output transformer.

- If the system is perfectly symmetric the carrier frequency will be completely canceled. But in practice no system is perfectly symmetrical so that the carrier will be heavily suppressed rather than completely removed.

- The output of the balanced modulator thus contains two sidebands and some of the miscellaneous components which are taken out by the tuning output transformer's secondary winding.

- Thus the final output consists only of sidebands. The output V_0 is given by the equation;

$$V_0 = P\, sin\omega_C t \quad + Q\ \cos(\omega_c - \omega_m)\, t - Qcos(\omega_c - \omega_m)\, t$$

- Modulating signal, lower sideband & upper sideband frequency in the above equation shows that under ideally symmetrical conditions the carrier has been canceled out, leaving only the two sidebands and the modulating frequency.

- The tuning of the output transformer will remove the modulating frequencies from the output.

9.5 SSB GENERATION :

- Three main systems are employed for the generation of the SSB :
  1. The Filter Method
  2. The Phase Cancelation Method And
  3. The Weaver Method.

- They differ from one another in the way of suppressing the unwanted sideband, but all use some form of balanced modulator to suppress the carrier.

9.5.1 Filter method :

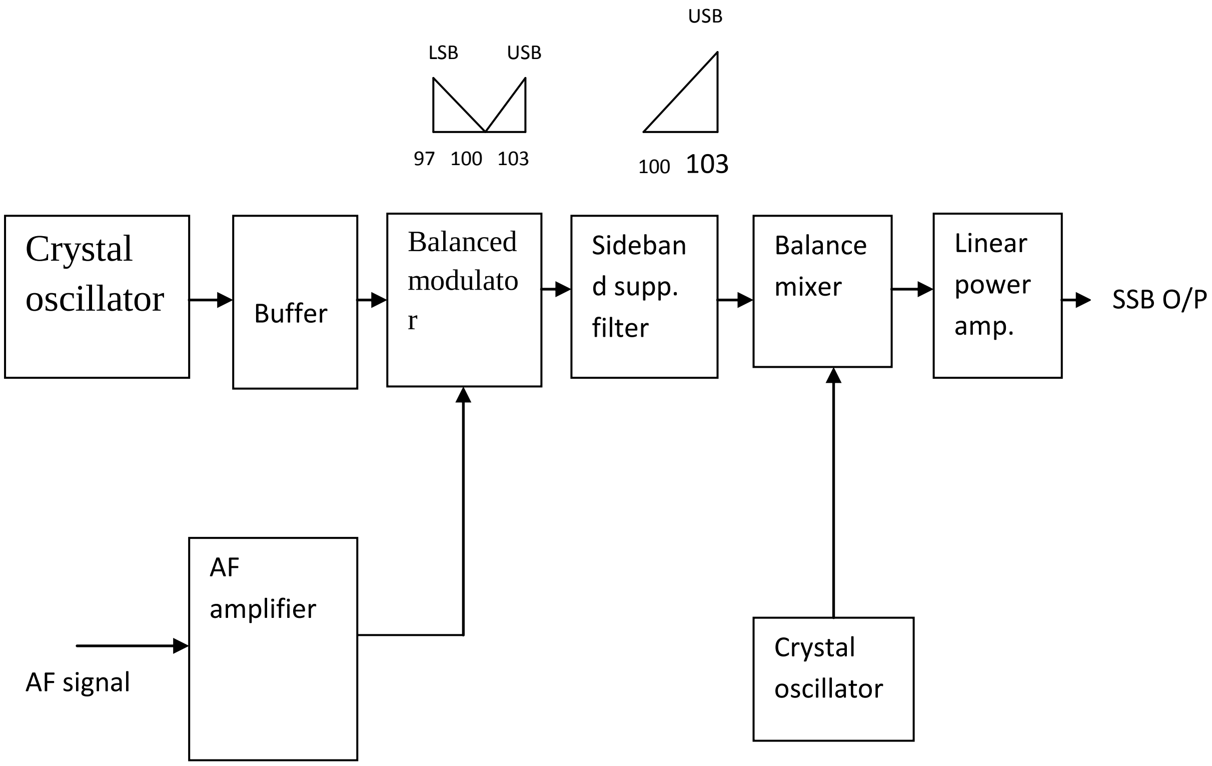


Fig. 9.2 The Filter method for SSB generation

- Crystal oscillator generates carrier signal and the AF signal amplified by the AF amplifier are given as input to Balanced Modulator. The o/p of the balanced modulator contains USB and LSB only. The sideband suppression filter is needed to remove unwanted sideband. The low frequency unwanted sideband can be easily suppressed by the band suppression filter.

- The output of the filter produces a low frequency signal. A process called Frequency Up Conversion is used to boost the frequency of this signal. It is done by giving this signal to a balance mixer along with a signal generated by a second crystal oscillator. The filter must have a flat pass-band and extremely high attenuation outside the pass band

- The initial modulation takes place at a low carrier frequency in balanced modulator since at high frequency the attenuation to be offered by the filter is practically impossible to achieve.

- This sideband is then amplified by the linear amplifier and then transmitted. The Amplification process prevents signal distortion i.e. prevent changes to the shape of the signal.

Advantages:-

1. It gives adequate side band suppression
2. Bandwidth is sufficiently flat and wide
3. The bulk of LC filter is overcome by mechanical or crystal filter. Hence it is practically used

Disadvantages:-

1. It cannot be used to generate SSB at very high frequency because repeated mixing networks are required with stable crystal oscillators.
2. Very low audio frequencies cannot be used because it is difficult to design filter having very small bandwidth or pass band
3. Expensive filters are required to remove one side of the side band

Even with the above disadvantages, it is used in most commonly sued commercial systems.

9.5.2 Phase shift method:

- It uses phase shifting techniques that causes one of the side band to be cancelled out.
- The block diagram of phase shift method to generate SSB is as shown:

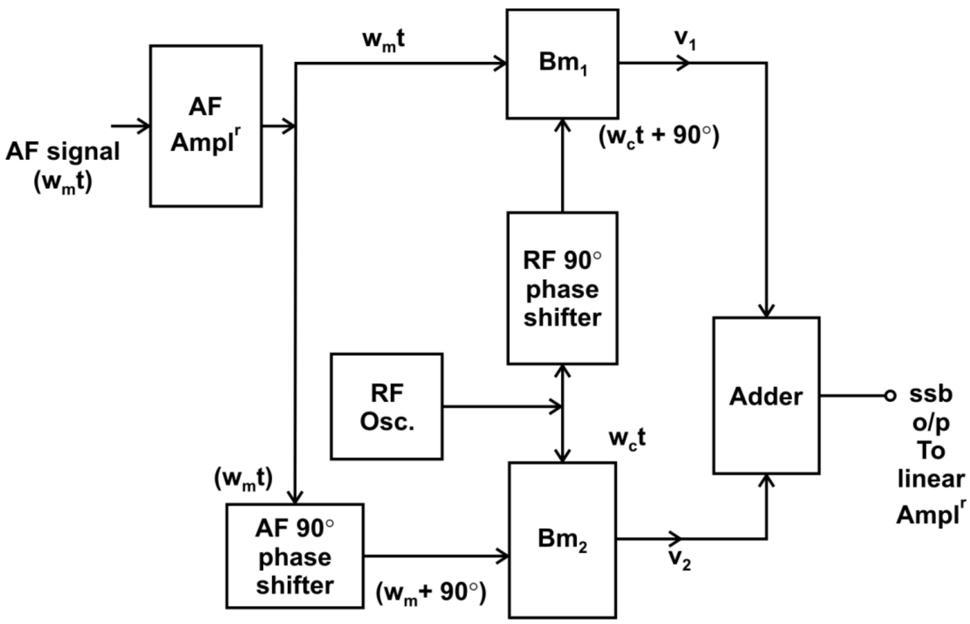


Fig. 9.3 the phase shift method.

- This method uses two Balance Modulator and two phase shifting networks. Here inputs to BM1 are the carrier signal shifted by 90^o and AF signal. The inputs to BM2 are the carrier signal and the AF signal shifted by 90^o. Both the modulators produce an output contains only side bands. The output of both modulators are added to cancel out one of the side band width that SSB output can be obtained.

Mathematical analysis:

- Input to BM1 are Sinω_mt and Sin (ω_ct+ 90^o).
  The Output of BM1 is the product of these two inputs
  V_1 = Sin (ω_ct+ 90^o). Sinω_mt
  $V_1 = \frac{1}{2}\left[\cos(\omega_c t + 90^o - \omega_m t) - \cos(\omega_c t + \omega_m t + 90^o)\right]$
- Similarly, to Bm_2, we apply sin ω_ct and sin (ω_mt + 90^o)
  Therefore, the output of Bm_2 = product of two inputs
  V_2 = Sinω_ct . Sin (ω_mt+ 90^o)
  $V_2 = \frac{1}{2}\left[\cos(\omega_c t - 90^o - \omega_m t) - \cos(\omega_c t + \omega_m t + 90^o)\right]$
- Therefore, the output of the adder is
  SSB output = $v_1 + v_2$ = - cos [ω_ct + ω_mt + 90^o]
  Therefore SSB output = sin (ω_ct + ω_mt)
- Thus, the first term in v_1 and v_2 are cancelled out as they are 180^o out of phase. Therefore, the output of the phase shift method contains only USB.

LSB Generation:-

- If the input to Bm_1 are sin ω_mt and sin ω_ct and input to BM are sin ($\omega_m t + 90^o$) and sin ($\omega_c t + 90^o$), then we get SSB and LS Band.

- **Advantages:-**
  1. It can generate SSB at any frequency.
  2. The audio frequency may be used for modulation.
  3. It is very easy to switch from one side band to the other.
  4. To generate SSB, at high radio frequencies up convergenbs and hence repetitive mixing is mot read.

- **Disadvantages:-**
  1. It is difficult to achieve 90^o phase shift at all modulating frequencies.
  2. The output of two balanced modulators must be exactly the same. Otherwise cancellation will be incomplete.

9.5.3 Third or Weavers method:-

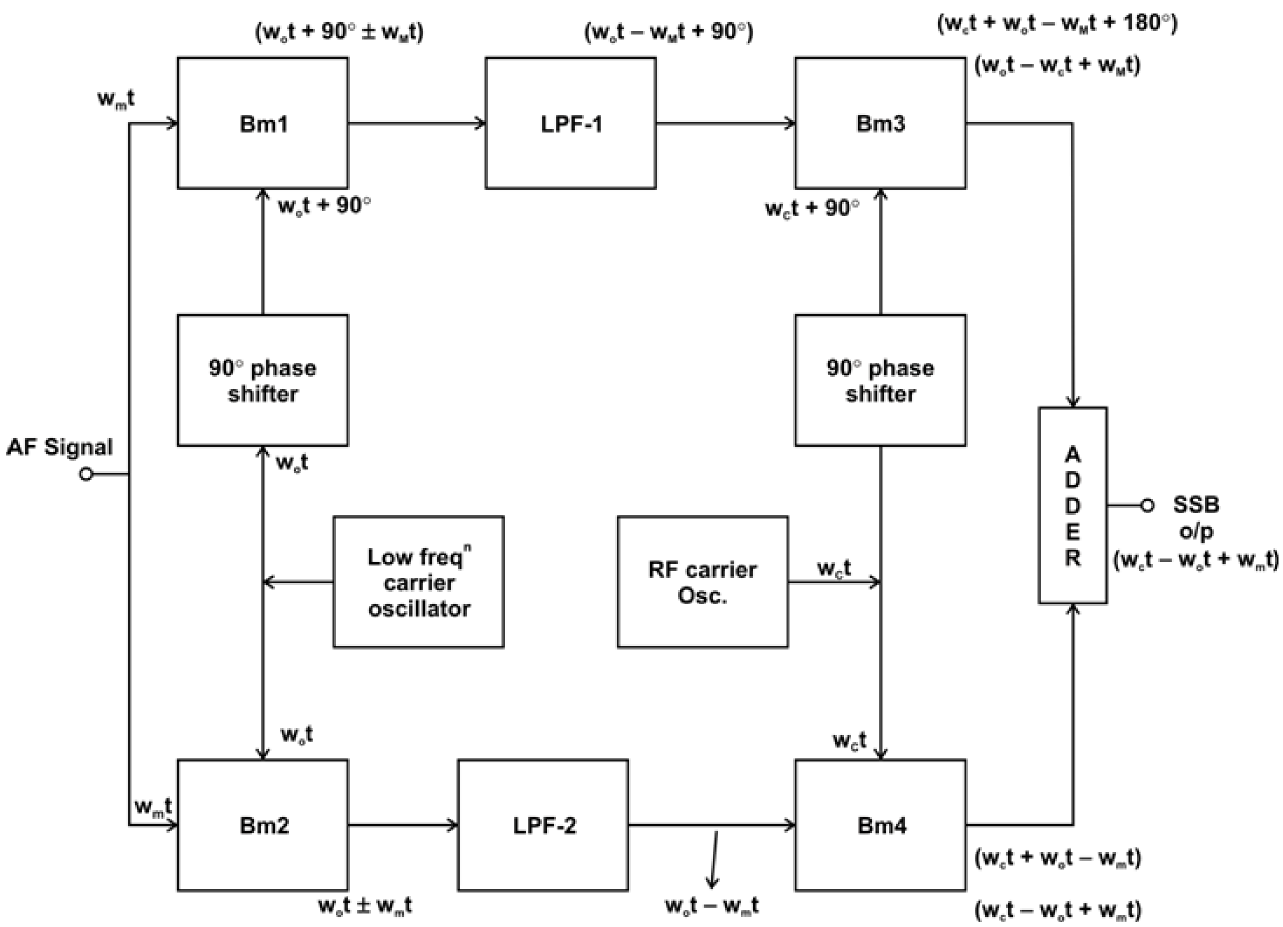


Fig. 9.4 Third method of SSB generation

- The input of Bm_1 is $\omega_m t$ and $\omega_c t + 90^o$. Therefore, its output contains LSB and USB i.e. $(\omega_o t + 90^o + \omega_m t)$ or $(\omega_o t + 90^o - \omega_m t)$. the input to Bm_2 is $\omega_m t$ and $\omega_o t$. Therefore, its output contain LSB and USB i.e. $(\omega_o t + \omega_m t)$ or $(\omega_o t - \omega_m t)$.

- The LSB pass filter 1 i.e. LPF – 1 will pass only LSB. Therefore, the output of the LPF-1 is $\omega_o t - \omega_m t + 90^o$. The output of LPF – 2 is $\omega_o t - \omega_m t$. The input of Bm_3 is $\omega_c + 90^o$ and $(\omega_o t - \omega_m t + 90^o)$. Therefore the output of Bm_3 contains LSB and USB.

$$USB = \omega_c t + \omega_o t - \omega_m t + 180^o$$
$$LSB = \omega_c t - \omega_o t + \omega_m t$$

- The input of Bm_4 is $\omega_c t$ and $\omega_o t - \omega_m t$ frequency signals. Therefore, its output contains

$$USB = \omega_c t + \omega_o t - \omega_m t$$
$$LSB = \omega_c t - \omega_o t + \omega_m t$$

- The output of Bm_3 and Bm_4 are added in the adder circuit. Therefore, the output of the added is the LSB of both.

$$SSB\ o/p = \omega_c t - \omega_o t + \omega_m t$$

- The USB signal in the output of Bm_3 and Bm_4 are 180^o out of phase. Therefore, they cancel each other. Thus at output, we get only LSB.

- **Advantages:-**
  The third method has all advantages of phase shift method. The additional advantage is that AF signal need not to be phase shifted. Therefore, the design of 90^o phase shifter becomes very simple because this circuit has to provide the 90^o phase shift and only one frequency i.e. carrier frequency.

- **Disadvantages:-**

  This is complicated and therefore rarely used.

9.6 COMPARISION BETWEEN SSB GENERATION METHODS:-

Parameters	Filter method	Phase shift method	Third method
1. Method to cancel unwanted SB.	1. Using filter.	1. By shifting RF and AF signal to Bm by 90^o.	1.By shifting RF signal to Bm by 90^o.
2. Design of 90^o phase shifting network.	2. Not applicable	2.Design is critical.	2. Design nis easy.
3. Possibility to generate SSB at high frequency.	3. Not possible	3.Possible	3. Possible.
4. Need for frequency up conversion.	4. Needed.	4. Not needed.	4.Not needed.
5. Use of law modulating frequency.	5. Not possible	5. Possible.	5.Possible.
6. Critical point ion design.	6. Filter characteristic, its size, weight and cut off frequency.	6. Design of 90^o phase shifter at AF and symmetry of Bm.	6.Symmetry of Bm for proper carrier cancellation.

9.7 PILOT CARRIER SYSTEM :

- A single sideband suppressed carrier requires frequency stability on the part of both transmitter and receiver, because any frequency shift, anywhere along the chain of events through which the information passes, will cause an equal frequency shift to the received signals. Imagine a 50-Hz frequency shift in a system through which three signals are being transmitted at 250, 500 and 700 Hz. Not only will they all be shifted in frequency to (say) 200, 450 and 650 Hz, respectively, but their relation to one another will also stop being harmonic. The result is that good-quality music is obviously difficult to transmit via single sideband suppressed carrier. Speech will also be impaired unless long-term stabilities are attained.

- Such frequency stability is presently available for fixed-frequency transmitters, but problem for receivers, since they must be tunable. The technique that was used to solve this problem is to transmit a pilot carrier with the wanted sideband.

- The carrier is normally reinserted and it provides a reference signal to help demodulation in the receiver. The receiver can then use automatic frequency control (AFC).Since the frequency stability obtainable over long-term periods with single sideband suppressed carrier is of higher order.

- Such systems are widely employed. They are particularly found in transmarine point-to-point radiotelephony and in maritime mobile communications, especially at the distress frequencies.

9.8 INDEPENDENT SIDEBAND (ISB) SYSTEMS:

- Multiplexing techniques are used for high-density point-to-point communication. For low or medium-density traffic, ISB transmission is often employed.

- The block diagram of ISB transmitter is shown below in the figure 10.5 ISB essentially consists of single sideband reduced carrier with two SSB channels added to form two sidebands around the reduced carrier. However, each sideband is quite independent of the other. It can simultaneously convey a totally different transmission, to the extent that the upper sideband could, for example, be used for telephony while the lower sideband carries telegraphy. Each 6-kHz channel is fed to its own balanced modulator, each balanced modulator also receiving the output of the 100-kHz crystal oscillator. The carrier is suppressed in the balanced modulator and the following filter, the main function of the filter still being the suppression of the unwanted sideband.

- The difference here is that while one filter suppresses the lower sideband, the other suppresses the upper sideband. Both outputs are then combined in the adder with the 26-dB carrier, so that a low-frequency ISB signal exists at this point, with a pilot carrier also present. Through mixing with the output of another crystal oscillator, the frequency is then raised to the standard value of 3.1 MHz.

- The use of balanced mixers is to permit easier removal of unwanted frequencies by the output filter. The signal now leaves the drive unit and enters the main transmitter. Its frequency is raised yet again, through mixing with the output of another crystal oscillator, or frequency synthesizer. This is done because the frequency range for such transmissions lies in the HF band, from 3 to 30 MHz. The resulting RF ISB signal is then amplified by linear amplifiers, as might be expected, until it reaches the ultimate level, at which point it is fed to a directional antenna for transmission. The typical power level at this point is generally between 10 and 60 kW peak.

- Since each sideband can carry two voice circuits, so that a total of four conversations may be transmitted simultaneously.

- Demodulation of ISB in the receiver follows a path similar to that of the modulation process.

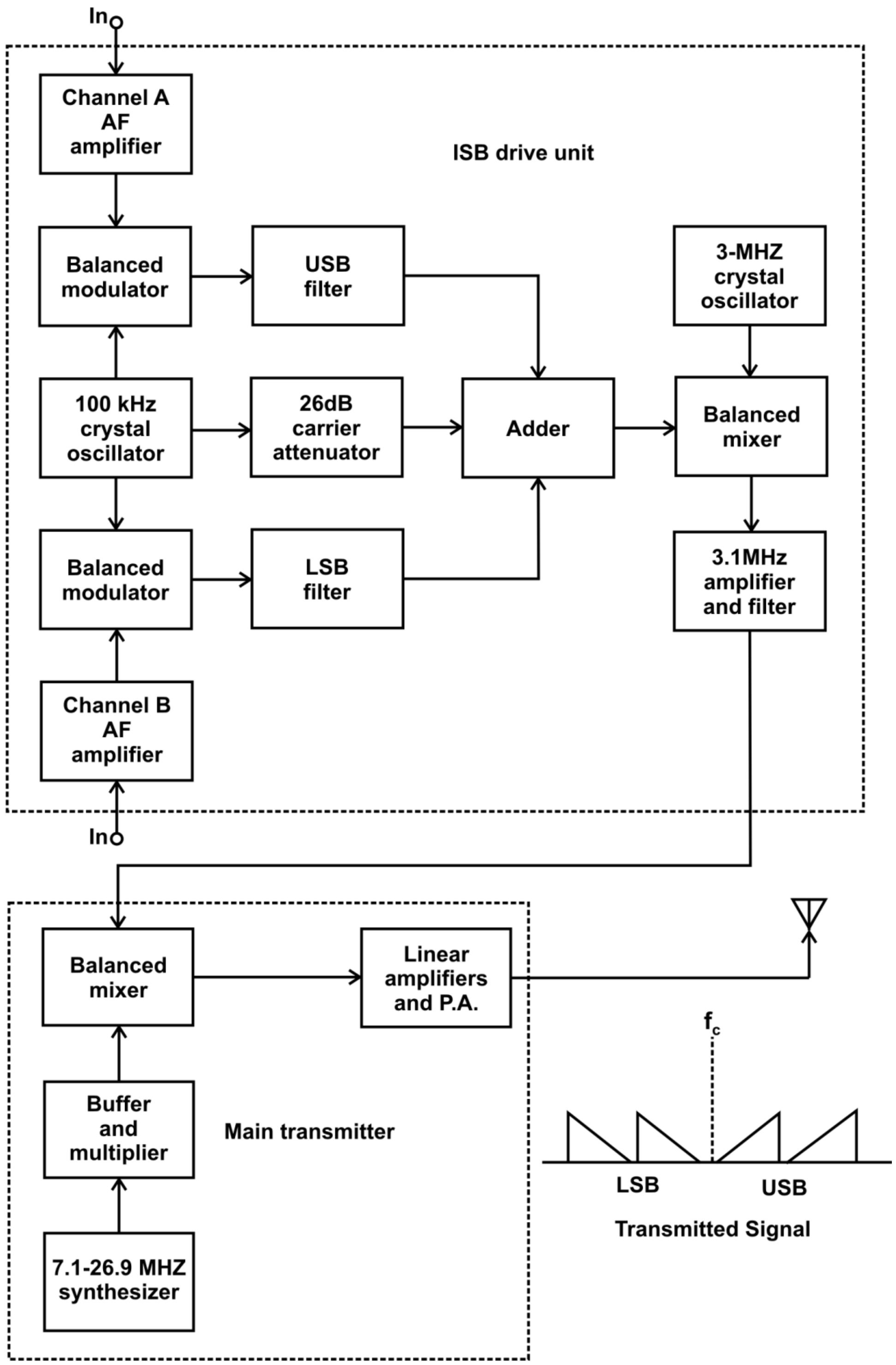


Fig .9.5 ISB transmitter.

9.9 VESTIGIAL SIDEBAND TRANSMISSION :

- The major advantage of single sideband is the bandwidth saving that occurs from its use, as well as the power saving.

- On the other hand some demodulation complications arise from the use of single sideband suppressed carrier, as opposed to AM systems in which a carrier is sent.

- Also, the greater the bandwidth occupied by a signal, the greater is the spectrum space that can be saved by sending one sideband instead of both. The more information that must be sent in a given time (i.e., per second), the larger the bandwidth required to send it. For proper transmission and reception of television signal, the bandwidth occupied by such signals is at least 4 MHz. knowing filter characteristics, a transmitted bandwidth of 9 MHz would be the minimum required for video transmissions.

- The use of some form of SSB is clearly indicated here to ensure spectrum conservation. So as to simplify video demodulation in the receiver, the carrier is, in practice, sent undiminished. Because the phase response of filters, near the edges of the flat pass band, would have a harmful effect on the received video signals in a TV receiver, a portion of the unwanted (lower) sideband must also be transmitted. The result is vestigial sideband transmission, or C3F, as shown in **Fig. 9.6(a).**

- By sending the first 1.25 MHz of the lower sideband (the first 0.75 MHz of it undiminished), it is possible to make sure that the lowest frequencies in the wanted upper sideband are not distorted in phase by the vestigial-sideband filter. Because only the first 1.25 MHz of the lower sideband is transmitted, 3 MHz of spectrum is saved for every TV channel. Since the total bandwidth requirement of a television channel is now 6 MHz instead of 9 MHz, clearly a great saving has been made, and more channels consequently can be accommodated.

- For completeness, **Fig. 9.6 (a)** shows also the location, in frequency, of the frequency-modulated sound transmissions that accompany the video.

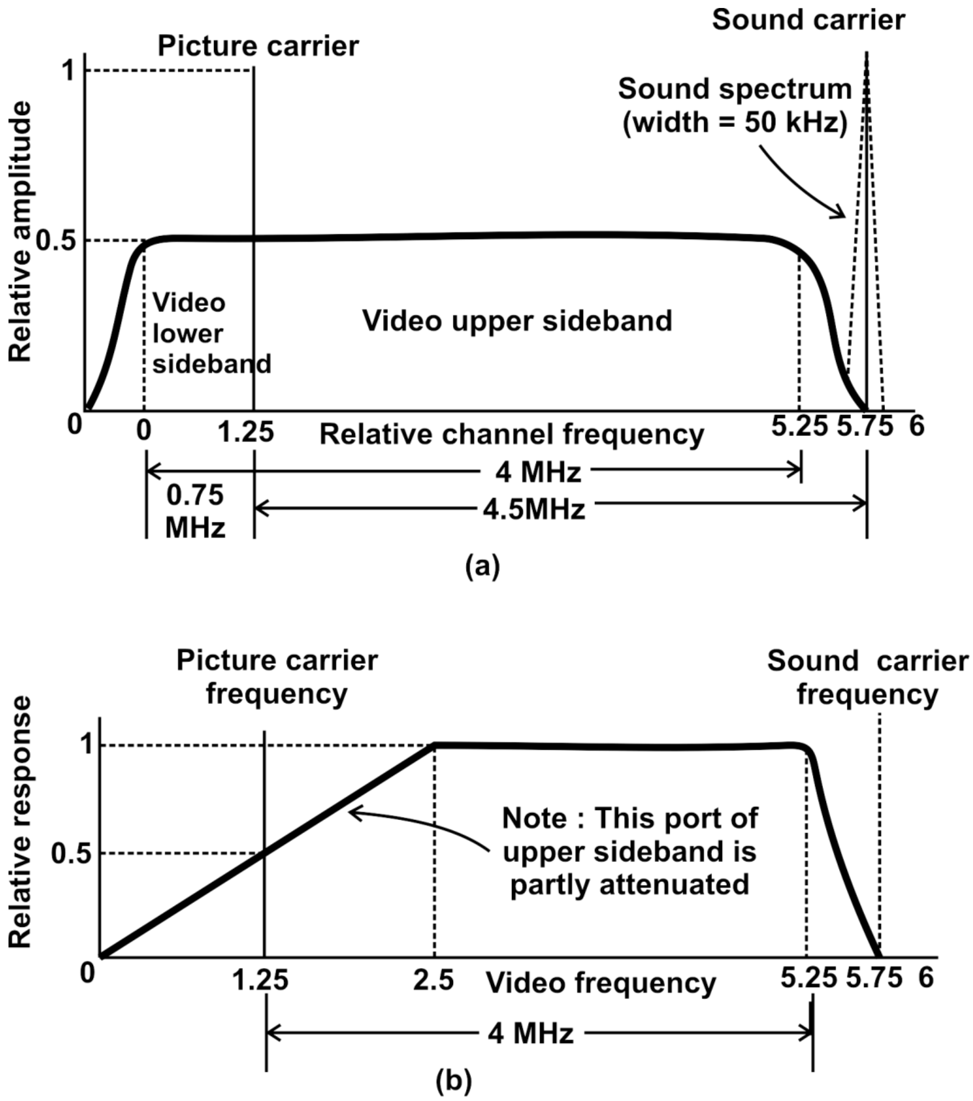


Fig. 9.6 Vestigial sideband for TV video transmission (a) Spectrum of transmitted signals;
(b) Corresponding receiver video amplifier response.

- It should be noted that these transmissions have nothing to do with the fact that the modulation system for video is vestigial sideband transmission and would have been there regardless of the video modulation system. All these signals occupy frequencies near the video transmissions simply because sound is required with the pictures, and it would not be very practical to have a completely separate receiver for the sound, operating at some frequency remote from the video transmitted frequencies.

- Fig. 9.6 (b) shows the video frequency response of the television receiver. Attenuation is provided for the video frequencies from 0 to 1.25 MHz. The reason is quite simple extra power is transmitted at these frequencies (since they are sent in both sidebands, whereas the remaining video frequencies are not only in the upper sideband). Accordingly, these frequencies would be unduly emphasized in the video output of the receiver if they were not attenuated appropriately

9.10 SUMMARY :

- In case of AM generation, when a carrier is amplitude modulated by a single sine wave the resulting signal consists of three frequencies : the original carrier frequency , the upper side band frequency ($f_c + f_m$) and lower side band frequency ($f_c - f_m$).
- The AM power equation states that the ratio of total power to carrier power is ;
  $\frac{P_T}{P_C} = \left(1 + \frac{m^2}{2}\right)$
- Three main systems are employed for the generation of the SSB : the filter method, the phase cancelation method and the third method. They differ from one another in the way of suppressing the unwanted sideband, but all use some form of balanced modulator to suppress the carrier.

9.11 UNIT END EXERCISE :

1) What is SSB modulation?
2) Show that minimum 50% power is saved when carrier is suppressed or removed.
3) What is balanced modulator? With the help of diagram explain the its working.
4) Explain the following in the generation of SSB signal

   (I) The filter method (ii) The phasing method (iii) The third method

5) Explain pilot carrier system.
6) Explain ISB system
7) Write a note on vestigial sideband transmission.

❖❖❖❖

UNIT V – Frequency Modulation

10

FREQUENCY MODULATION

Unit structure :

10.0 OBJECTIVE :

In this lesson we will be learning frequency and phase modulation. The mathematical analysis will help us to understand the concept.

10.1 INTRODUCTION :

Angle Modulation : It is the process of varying total phase angle or frequency of a carrier wave in accordance with the instantaneous value of modulating signal keeping amplitude of carrier constant. The unmodulated carrier is

$$E_c = A\sin(w_c t + \varphi)$$

Where φ instantaneous phase. If this phase angle varies with modulating signal we get angle modulation. φ can vary either by varying or by phase of the carrier. There are two types of angle modulation

- Frequency modulation
- Phase modulation.

10.2 FREQUENCY MODULATION:

Definition: *<u>If frequency of carrier signal is varied in accordance with the instantaneous amplitude of the modulating signal, then it is called frequency modulation.</u>*

In FM, amplitude of the carrier remains constant. The variation in the carrier frequency from the unmodulated carrier frequency is called as frequency deviation **(δ)**. In FM, the information is contained in the frequency deviation of the FM wave.

FM is a non – linear process. Here, the frequency deviation is proportional to the amplitude of the modulating signal.

10.2.1 Mathematical expression of FM:

The instantaneous frequency of the frequency modulated wave is

$$\omega_i = \omega_c + \delta$$

Where,

W_i is the instantaneous frequency of frequency modulated wave,
W_c instantaneous frequency of carrier wave,
d is deviation in carrier frequency

The frequency deviation above or below ω_c, depends on the amplitude of modulating signal i.e.

$\delta \propto V_m \cos \omega_m t$ (directly proportional)

$\delta = K V_m \cos \omega_m t$

Where *'K' is the frequency deviation constant of FM and it is defined as the ratio of change in output frequency of the FM modulator with respect to change in amplitude of modulating signal.* Its unit is Hz/v.

The maximum value of frequency deviation is

$$\boxed{\delta_{max} = KV_m}$$

Instantaneous frequency is

$$\omega_i = \omega_c + \delta$$
$$\omega_i = \omega_c + KV_m Cos\omega_m t$$

The frequency modulated signal is represented mathematically as

$V_m = V_c + \sin \theta$

Where θ is the instantaneous phase and it can be determined from instantaneous frequency as

$$\omega_i = \frac{d\theta}{dt}$$
$$\int d\theta = \int \omega_i \, dt$$
$$\delta \, d\theta = ϛ w_i dt$$

$$\therefore \Psi = \theta = \int\{\omega_c + K \, Vm \cos \omega_m t]dt$$

$$\therefore \theta = w_c t + \frac{KV_m}{w_m} \sin w_m t$$

$$= w_c t + \frac{\delta \max}{w_m} \sin w_m t$$

$$\therefore \theta = w_c t + M_F \sin w_m t$$

Where $M_F = \frac{\delta\text{max}}{w_m} = \frac{KV_m}{w_m}$

M_F is the modulation index of FM. ; K is in rad/v.

∴ the equation of frequency modulation is

$$V_{FM} = Vc\ sin\,(w_c t + MF\ sin\, w_m t)$$

If the modulating signal and the carrier signal are 'cosine' signals then

$V_{FM} = V_c \cos (w_c t + M_F \sin w_m t)$

If both the siganls are 'sine' signals, then

$V_{FM} = V_c \cos (w_c t - M_F \sin w_m t)$

10.3 BANDWIDTH OF FM:

The effective bandwidth of FM wave can be defined in two ways:

1. As per **Carson's rule** is:

$$BW = 2(\delta + f_m)$$

2. And as per Bessel Table, the bandwidth of FM is

$$BW = 2 * n * f_m$$

Where n = number of significant sidebands.

10.3.1 Deviation Ratio:

It is the worst case modulation index and is the ratio of maximum frequency deviation to the maximum modulating frequency.

$$\text{Deviation Ratio} = \frac{\delta_{max}}{fm_{(max)}}$$

10.3.2 Power in FM:

Power transmitted by FM wave is constant. It does not change due to change in modulation index. The total power in FM is equal to the sum of power of modulated carrier component and sideband components.

$$P_T = P_0 + P_1 + ... + P_n$$

An FM wave contains infinite sidebands and the amplitude of each sideband is given by $J_n M_f V$, n indicates the number of particular sideband.

As the modulation index increases, $J_0(M_f)$ decreases. Therefore, the power of carrier component in FM decreases, but the significant sideband increases. However, the total power transmitted by FM waves remains constant and it is equal to the unmodulated carrier power.

$$P_T = \frac{V_c^2}{2R}$$

Where V_c = amplitude of unmodulated carrier.

$P_T = P_0 + P_1 + ... + P_n$

$$= \frac{V_c^2}{2R} + \frac{2V_1^2}{2R} + ... + \frac{V_n^2}{2R}$$

Where $V_c = J_o (M_F)V_c$ = amplitude of the carrier component in FM
$V_1 = J_o (M_F)V_c$ = amplitude of the first pair of the sideband.
$V_2 = J_o (M_F)V_c$ = amplitude of the second pair of the sideband.

$$\therefore P_T = \frac{J_O(MF)V_c^2}{2R} + \frac{J_1(MF)V_c^2}{2R} + \frac{J_2(MF)V_c^2}{2R} + ... + \frac{J_n(MF)V_c^2}{2R}$$

$$= \frac{V_c^2}{2R}\,[\mathrm{Jo\ (MF)} + \mathrm{J1\ (MF)} + ... + \mathrm{Jn\ (MF)}]$$

$\therefore P_T = P_c\,[J_o^2(M_F) + 2\,\Sigma\,J_n^2\,(M_F)]$

For any value of M_F, the bracketed term is approximately equal to unity.

$\therefore P_T = P_c = \frac{V_c^2}{2R}$

10.4 NOISE IN FM:

- Noise signal produces amplitude modulation as well as phase modulation in FM. Let us consider the noise signal and carrier signal vectorially.

- When noise signal superimposed on carrier, the amplitude of noise is added with carrier amplitude.

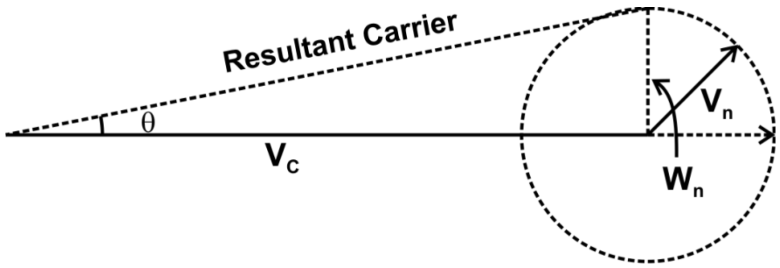


Fig10.1

- Therefore, the amplitude of the resultant carrier changes due to noise. Therefore, noise voltage produces amplitude modulation in FM.
  Maximum deviation of amplitude carrier = V_n.

- At the same time, noise vector is constantly changing phase angle w.r.t carrier signal V_c which will change the phase deviation (θ) of FM wave i.e. phase deviation θ of FM changes due to noise. Thus, noise modulates the carrier in terms of amplitude as well as phase.

- The amplitude variation of the carrier due to noise can be removed by using amplitude limiter circuits.

.

<u>10.4.1 Noise Triangle:</u>

- The effect of noise in FM does not remain constant but it increases with the increase in frequency of modulating signal.

- To explain this, consider noise amplitude constant $V_n = \frac{1}{4} v_c$. Therefore, amplitude modulation due to noise in FM is content as

$$M = \frac{V_n}{V_c} = \frac{1}{4} \text{ and}$$

the maximum phase deviation produced when noise vector is perpendicular to the resultant carrier and it is

$$\theta = sin^{-1}\left(\frac{V_n}{V_c}\right).$$

$$\therefore \quad \theta = 14.5^o. \text{ in FM,}$$

- Therefore, V_c is constant and single frequency noise voltage, we assume constant. Therefore, V_n/V_c = constant and θ = constant. Thus, frequency deviation due to noise is constant. Therefore, modulation index due to noise is a constant as $M_{fn} = \frac{\delta}{fn}$, fn = single noise frequency. But as the frequency of modulating signal increases, the modulation index due to signal decreases as $M_{fn} = \frac{\delta}{fn}$. Therefore, the signal to noise ratio in FM is $\frac{S}{N} = \frac{M_{fs}}{M_{fn}}$

Thus, as the frequency of modulating signal increases, Mf_S decreases. Therefore, signal to noise ratio decreases i.e. the effect of noise in FM increases as shown in noise triangle.

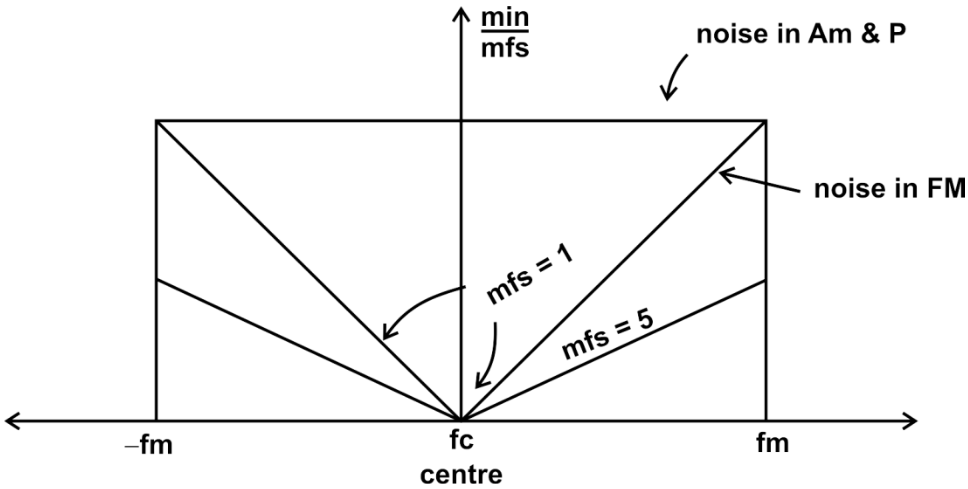


Fig.10.2 Noise Triangle in FM

- From noise triangle, it is seen that noise in AM and PM remains constant for entire audio range because the modulation index due to signal are independent on modulating frequency but in FM the effect of noise will be increased with the increase in modulating frequency. Thus, noise has more effect on higher frequencies in FM. *The triangular distribution of noise in FM is called as FM noise triangle.*
- Noise triangle shows that noise has greater effect on higher modulating frequencies than the lower ones. Hence, if the higher modulating frequencies are artificially boosted up at the transmitters and correspondingly cut at the receivers then the effect of noise on high modulating frequency can be minimized and we can maintain constant S/N over the entire audio range like AM and PM. The boosting up of high frequencies modulating signals is done by '**Pre – emphasis circuit**' and attenuation at the receiver is done by '**De – emphasis circuit**'.

10.5 PRE – EMPHASIS AND DE – EMPHASIS:

- From noise triangle, it is seen that noise has more effect on higher modulating frequencies than the lower ones. Therefore, for modulating signal with uniform signal level, the non – uniform S/N ratio is produced and higher modulating frequency have low S/N ratios than the lower ones as shown in Fig.10.3(a).

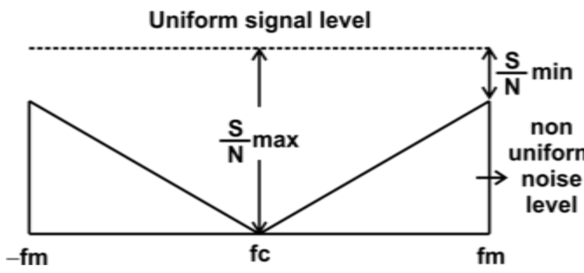


Fig.10.3(a)

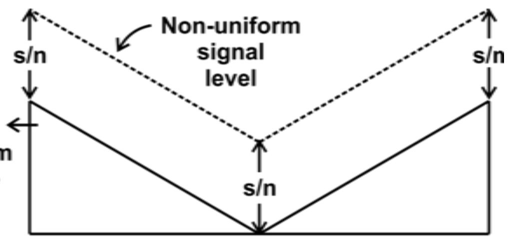


Fig.10.3(b)

- To compensate this, a high frequency modulating signal is emphasised or boosted in amplitude in transmitter before forming modulation. To compensate for this boost, the high frequencies are attenuated or de – emphasised in the receiver after the demodulation has been performed. Due to pre – emphasis and de – emphasis, the S/N ratio at the output of receiver is maintained constant as shown in Fig.10.3(b).

- A pre – emphasis network is a high pass filter i.e. differentiator and the de – emphasis network is a low pass filter i.e. integrator.

10.6 PHASE MODULATION:

If the phase of carrier signal is varied in accordance with the amplitude of modulating signal, then the resulting wave is said to be phase modulated. Here, the phase deviation is proportional to amplitude of modulating signal.

$\therefore$ phase deviation = θ

$\theta \propto$ modulating signal.

$\therefore \theta = k_p V_m \cos w_m t$

Where k_p = phase deviation sensitivity.

The maximum value of phase deviation is

$$\theta_{max} = k_p V_m$$

The instantaneous phase of PM wave is

$$\Psi = w_c t + \theta$$
$$= w_c t + k_p V_m \cos w_m t$$

The equation of PM wave is

$$e(t) = A \sin \Psi$$
$$= A \sin (w_c t + k_p V_m \cos w_m t)$$
$$\therefore e(t) = A \sin (w_c t + M_P \cos w_m t)$$

Since, $A = V_c$

Where, $M_P = k_p V_m$

M_P is modulation index of phase modulation and it is proportional to the amplitude of modulating signal and it is independent of frequency of modulating signal.

10.6.1 Bandwidth of PM:

The instantaneous phase of PM is

$$\Psi = w_c t + k_p V_m \cos w_m t,$$

the instantaneous frequency is

$$w_i = \frac{d\Psi}{dt} = \frac{d}{dt}\left[w_c t + k_p V_m \cos w_m t\right]$$

$$= w_c - k_p V_m w_m \sin w_m t$$

$$w_i = w_c - \delta$$

so the maximum value of frequency deviation of PM is

$$\delta_{max} = k_p V_m w_m$$

Frequency deviation in PM is α to the amplitude as well as frequency of modulating signal as per Carson's rule bandwidth required for PM is

$$BW = 2(\delta + f_m)$$

10.7 COMPARISON BETWEEN FM AND PM

Frequency modulation	Phase modulation
1. The frequency of the carrier signal varied according to instantaneous value of modulating signal.	1. The phase angle of carrier is varied according to instantaneous value of modulating signal.
2. The equation of FM signal is $V_{FM} = V_c \sin(\omega_c t + M_F sin\omega_m t)$	2. the equation of PM signal is $V_{PM} = V_c \sin(\omega_c t + M_P sin\omega_m t)$
3. The modulation index changes with modulating frequency	3. the modulation index is independent of modulating frequency.
4. Frequency deviation is proportional to amplitude of modulating signal	4. Frequency deviation is proportional to amplitude as well as frequency of modulating signal

5. BW changes slightly with f_m but the M_f changes considerably	5. BW changes considerably with f_m but M_f remain unchanged.
6. Fm can be obtained from phase modulator	6. PM can be obtained from frequency modulator.

10. 8 RECEIVER TOPOLOGY :

- There are two popular types of receivers for radio signals: A Tuned Receiver & Super Heterodyne Receiver.

- Almost all modern receiver designs use the super heterodyne principle, which have many advantages and overcome the shortcomings of tuned receivers.

- The simplest receiver would be a demodulator connected directly to an antenna, as in Figure 10.10(a). Any signal arriving at the antenna would be demodulated, and the detector output would be connected to sensitive headphones. Only strong signals received by a good antenna could be heard at all. In addition, this receiver would have no ability to discriminate against unwanted signals and noise and so would receive all local stations at once. Obviously the results would not be at all satisfactory.

- This receiver could be improved by adding a tuned circuit at the input, as shown in Figure 10.10 (b). This would provide some selectivity, that is, the receiver could be tuned to a particular station. Signals at the resonant frequency of the tuned circuit would be passed to the detector, and those at other frequencies would be attenuated. However, there is still no gain.

- The addition of an audio amplifier, as shown in Figure 10.10(c), could provide enough output power to operate a speaker. However, the selectivity would remain poor because of the single tuned circuit, and the receiver would not be sensitive enough to receive weak signals because a demodulator needs a relatively large input voltage to operate efficiently, with low noise and distortion. More sensitive detectors can be devised, but a better solution is to provide gain before the detector.

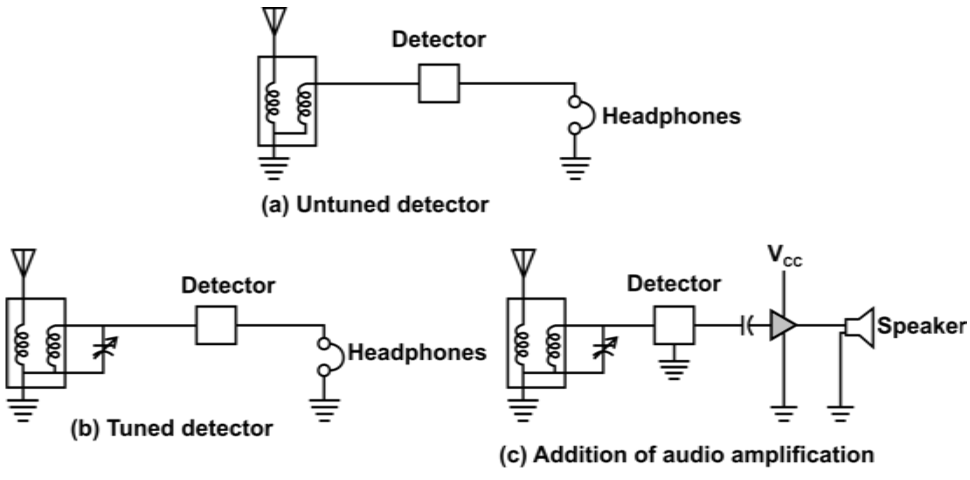


Fig. 10.10 Simple Receivers

10.9 TUNED RADIO FREQUENCY RECEIVER :

- Figure 10.11 shows a block diagram for a tuned-radio-frequency (TRF) receiver.
- Several RF amplifiers, each tuned to the signal frequency, provide gain and selectivity before the detector. An audio amplifier after the detector supplies the necessary power amplification to drive the speaker.
- The problems with this receiver are in the RF stages. To achieve satisfactory gain and selectivity, several stages will probably be needed. All of their tuned circuits must tune together to the same frequency (or track very closely); this tends to cause both electrical and mechanical problems. Having several high-gain tuned amplifier stages in close physical proximity is likely to lead to oscillation, since it is difficult to prevent feedback. In addition, some means must be provided to tune all the circuits simultaneously. The usual way, when TRF receivers were popular, was to use several variable capacitors connected together mechanically either ganged (mounted on the same shaft) or linked by belts or gears. However, component tolerances cause the circuits to tune to slightly different frequencies when their capacitors are at the same position, unless the frequencies are corrected with small adjustable capacitors called *trimmers and padders*. Even with these adjustments, the tracking is never perfect.
- Another problem arises from the fact that the bandwidth of a tuned circuit does not remain constant when its resonant frequency is changed.
- At higher frequencies, internal magnetic fields in the wire cause the current to flow mainly in the region near the surface of the conductor. This decreases the effective cross-sectional area of the conductor and increases its resistance. The resistance varies with the square root of frequency. Therefore, the bandwidth of a tuned circuit increases approximately with the square root of frequency.
- Thus a receiver with the correct selectivity at the low end of its tuning range will have a wide bandwidth at the high end.

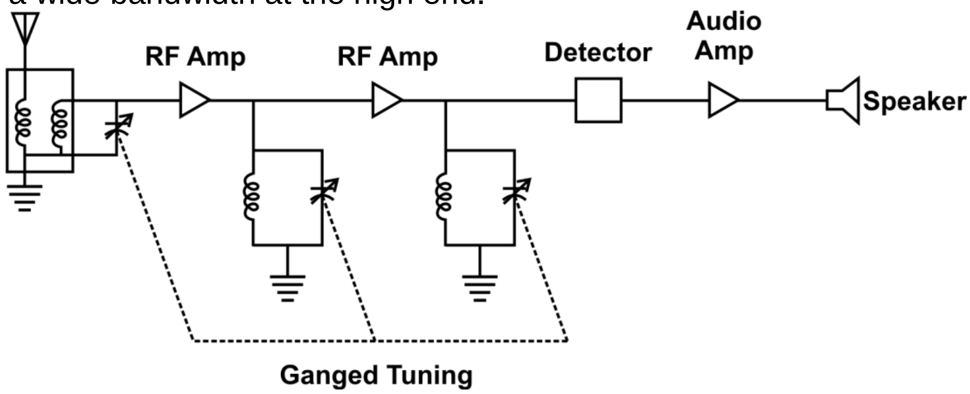


Fig 10.11 TRF receiver

- For high signal frequencies and narrow bandwidths, it may not be possible to obtain satisfactory results with a reasonable number of RF stages, because of limitations on the Q of conventional tuned circuits. Circuits with higher Q, such as crystal filters, are not practical because of the difficulty in operating them over a wide range of frequencies. In addition, most active devices show a reduction of gain with increasing frequency.
- The TRF system is now used only for simple, fixed-frequency receivers, where most of its disadvantages do not apply.

10.9.1 The Superheterodyne Receiver :

- The superheterodyne receiver or superhet was invented by Edwin H. Armstrong in 1918 and is still almost universally used, in many variations. Figure 10.13 shows its basic layout.
- It may contain one or more stages of RF amplification, and the RF stage can be either tuned, as in the TRF receiver, or broad banded. This stage should have a good noise figure as being the first stage in the receiver; it is largely responsible for the noise performance of the entire system. Low-cost receivers sometimes omit the RF amplifier, but they still include some -sort of input filter, such as a tuned circuit. The input filter and RF stage (or mixer, if there is no RF stage) are sometimes referred to as the front end of a receiver.

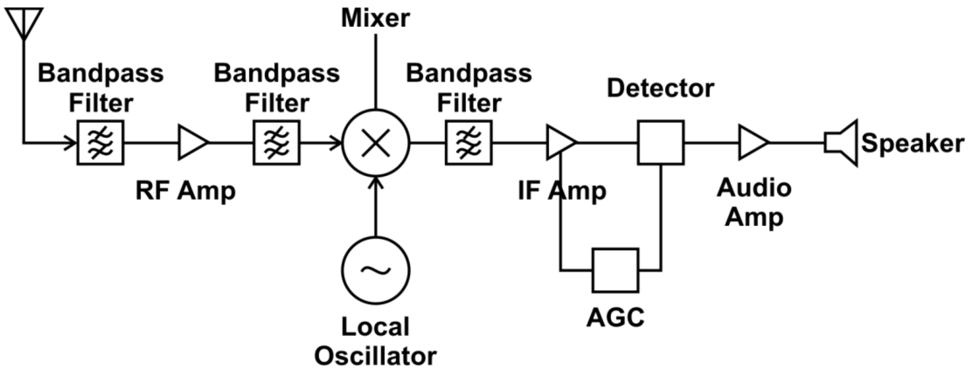


Fig 10.11 The Superheterodyne Receiver

- The next stage is a mixer. The signal frequency is mixed with a sine-wave signal generated by an associated stage called the local oscillator, creating a difference frequency called the **intermediate frequency (IF)**. The local oscillator is tunable, so the IF is fixed regardless of the signal frequency. The combination of mixer and local oscillator is known as a converter.

- Receivers with conventional variable-capacitor tuning usually use a two- or three-gang tuning capacitor. One section tunes the local oscillator, and the other section or sections tune the mixer input circuit and the input circuit for the RF amplifier (if present).

- The mixer is followed by the IF amplifier, which provides most of the receiver's gain and selectivity. Generally there are two or more IF stages, with selectivity provided either by resonant circuits or, in more advanced designs, by a crystal filter or a ceramic filter. The use of a fixed IF greatly simplifies the problem of achieving adequate gain and selectivity. The Automatic Gain Control (AGC) is used to adjust the gain of the IF. The remainder of the receiver is straightforward and resembles the TRF design. There is a detector to demodulate the signal and an audio amplifier to increase the signal power to the level required to operate a loudspeaker.

10.10 SUMMARY :

- <u>If frequency of carrier signal is varied in accordance with the instantaneous amplitude of the modulating signal, then it is called frequency modulation.</u>

- The frequency deviation sensitivity of FM is defined as the ratio of change in output frequency of the FM modulator w.r.t change in amplitude modulating signal. Its unit is Hz/v.

- Noise signal produces amplitude modulation as well as phase modulation in FM. The effect of noise in FM does not remain constant but it increases with the increase in frequency modulating signal.

- The boosting up of high frequencies modulating signals is done by '**Pre – emphasis circuit**' and attenuation at the receiver is done by '**De – emphasis circuit**'.

- Due to pre – emphasis and de – emphasis, the S/N ratio at the output of receiver is maintained constant.

- If the phase of carrier signal is varied in accordance with the amplitude of modulating signal, then the resulting wave is said to be phase modulated.

- The simplest conceivable receiver would be a demodulator connected directly to an antenna.

- The superheterodyne receiver or superhet was invented by Edwin H. Armstrong in 1918 and is still almost universally used, in many variations.

10.11 UNIT END EXERCISE :

1. Explain the term FM
2. Obtain expression for FM , BW in FM and power in FM.
3. What is the relation between noise and FM. What is noise triangle?
4. Explain in details (i) Pre-emphasis (ii) De- emphasis.
5. What do you mean by PM?
6. Distinguish between FM and PM
7. Write short note on (i) Receiver topology , (ii) TRF receiver And (iii) The super heterodyne receiver.

❖❖❖❖

UNIT V

11

ANALOGUE PULSE MODULATION

Unit structure :

11.0 OBJECTIVE ;

In this chapter we are mainly discussing the different types of modulation techniques like PAM, PWM, PPM, PCM AND Delta modulation.

11.1 INTRODUCTION:

In pulse modulation, carrier is the train of pulses. Some parameters of these pulses (amplitude, width, position) are changed wrt the modulating signal. There are three types of analogue pulse modulation.

1. If the amplitude of pulse varies according to modulating signal, then it is called 'pulse amplitude modulation' (PAM).
2. If the width of pulse varies according to the modulating signal, then it is called as 'pulse width modulation' (PWM).

3. If the position of pulse changes according to modulating signal, then it is called 'pulse position modulation' (PPM).

11.2 ADVANTAGES OF PULSE MODULATION:

In pulse modulation, usually the pulses are quite short as compared to the time between the two pulses of the same signal. Hence, pulse modulated wave is off most of the time. Due to this property, pulse modulation offers some advantages over continuous wave modulation.

1. In pulse modulation, power is required to transmit pulses only rather than being delivered continuously. Hence, transmitting power is saved.

2. In pulse modulation, the time of pulse is very short in width and there is a very large off time between two pulses of the same signal. This off time can be utilized for samples for other signal. This permits transmission of many messages on single communication channel. This is called division multiplexing.

11.3 SAMPLING PROCESS:

1. In a sampling process, a continuous time signal (AF signal) is converted to a discrete time signal (PAM signal).
2. This conversion can be done by switch. Switch position is controlled by the sampling signal (carrier pulses).
3. The sampling signal is a periodic train of pulses with unit amplitude and of period T_s. This time is known as sampling time and during this time; switch is closed so that sampled signal is equal to the input signal.
4. In the remaining time, switch is open as no input signal appears at the output.

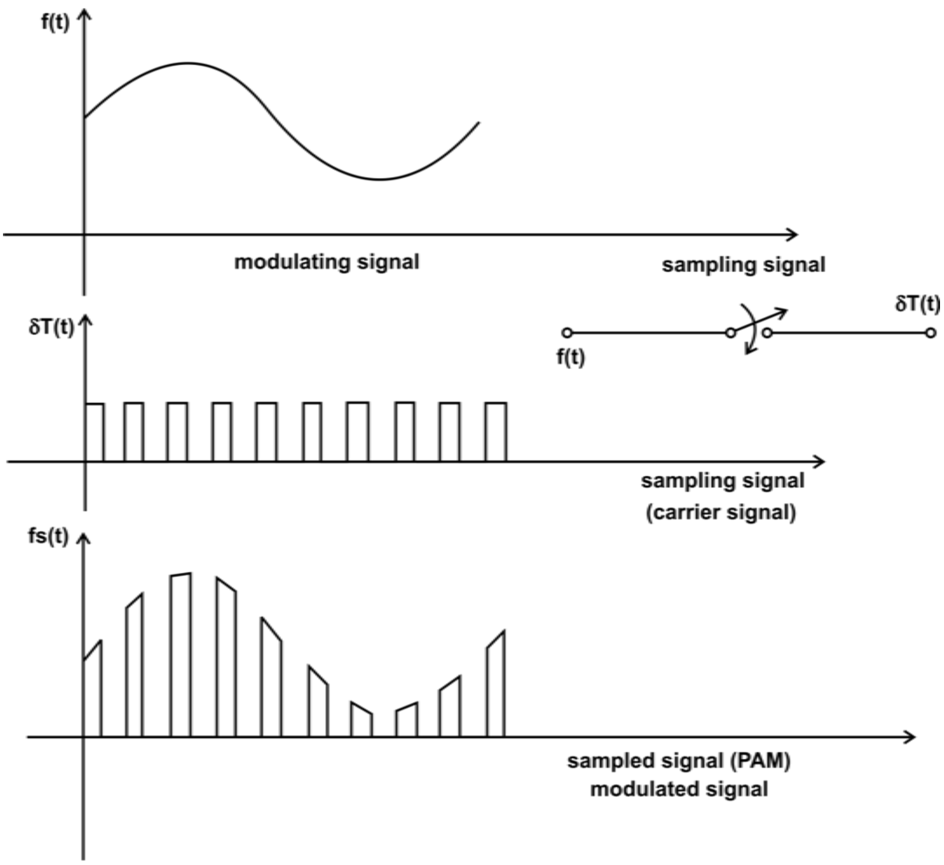


Fig. 11.1 Sampling Process

11.3.1 Sampling Theorem:

For proper reproduction of the modulating signal at the receiver, the modulating signal should be sampled at the rate of $f_s \geq 2f_m$ where f_m = highest modulating frequency.

Sampling Techniques:

There are two types of sampling techniques:

1. Natural sampling.
2. Flat top sampling.

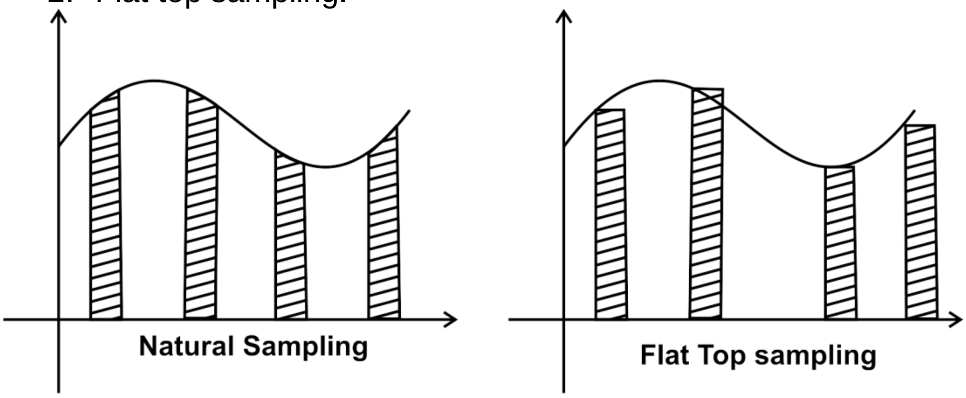


Fig. 11.2

In natural sampling, the top of the pulse is kept same as of analogue signal in that pulse duration. In flat – top sampling, the signal is sampled and held for pulse duration. This is done using sample and hold circuit.

11.4 PULSE AMPLITUDE MODULATION (PAM):

In PAM, the amplitude of carrier pulses is varied in accordance with the instantaneous amplitude of modulating signal. In this system, signal is sampled at regular intervals and each sample is made proportional to the amplitude of signal at the instant of sampling.

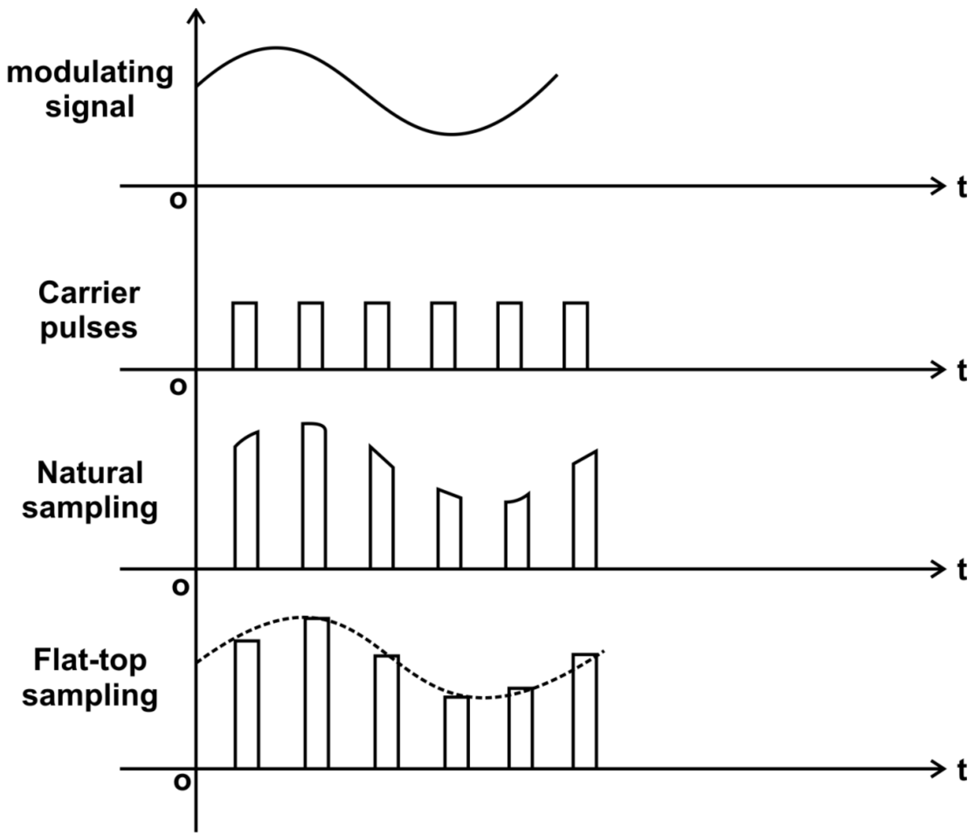


Fig. 11.3

Disadvantages of PAM:

1. Since noise affects the amplitude of waveform and in PAM the information contains amplitude variation like AM; PAM is less immune to noise.
2. Due to changes in amplitude of PAM pulses, the transmitted power is not constant.
3. Large bandwidth is required to transmit PAM.

11.5 PULSE WIDTH MODULATION (PWM):

- In this system, the width of carrier pulses is varied in accordance with the instantaneous amplitude of modulating signal. The amplitude of pulses is fixed.

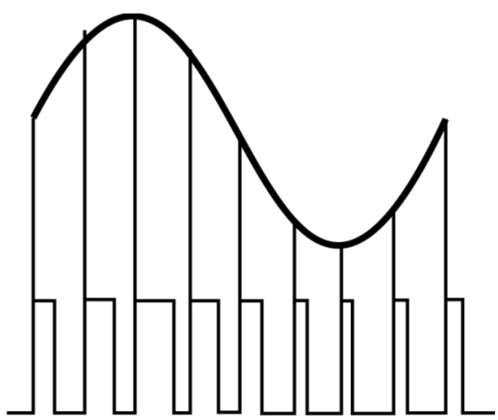


Fig. 11.4

- From the diagram it is seem that the width of pulses goes on increasing as the modulating signal goes on increasing and decreases from decreasing modulating signal.

Advantages of PWM:

1. Less effect of noise i.e. very good noise immunity.
2. Synchronisation between the transmitter and the receiver is not essential.

Disadvantages of PWM:

1. Due to variable pulse width, the pulses have variable power constant so the transmitter must be powerful enough to handle maximum width pulse.
2. Requires large bandwidth.

11.6 PULSE POSITION MODULATION (PPM):

The position of pulses is varied in accordance with the modulating signal. For increasing AF signal, pulses shift to right and for decreasing AF signal, pulses shift to left.

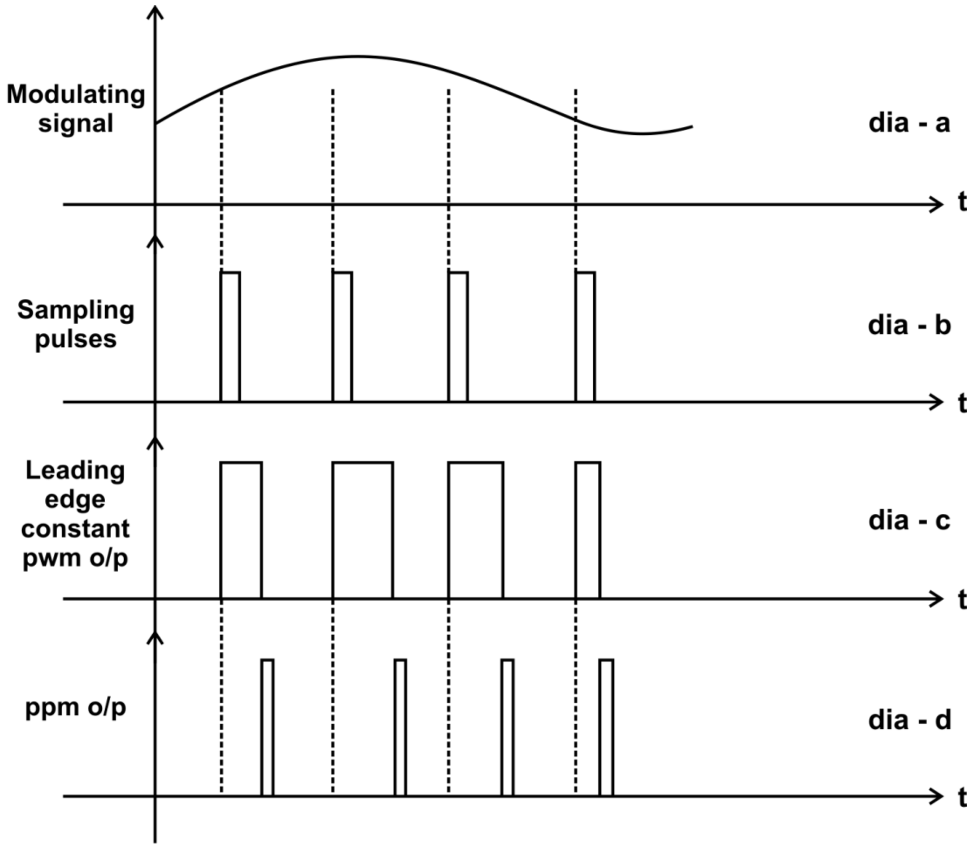


Fig. 11.5 PPM output

Advantages of PPM:

1. It has good noise immunity.
2. Due to constant amplitude and width pulses, the transmitted power always remains constant.

Disadvantages of PPM:

1. Large bandwidth required.
2. Synchronization between T_X and R_X required.

11.7 COMPARISON BETWEEN PAM, PWM AND PPM

Sr. No.	Parameter	PAM	PWM	PPM
1.	Information contained.	Amplitude variation.	Width variation.	Position variation.
2.	Bandwidth requirement	Low.	High.	High.
3.	Noise immunity	Low	High	High
4.	Transmitted power	Varies with amplitude and pulses.	Varies with variation in width.	Remains constant.
5.	Need to transmit synchronisation pulses	Not needed	Not needed	Necessary
6.	Complexity of generation and detection	Complex	Easy	Complex

11.8 PULSE CODE MODULATION (PCM):

In digital transmission instead of transmitting analogue signal, the analogue signal is sampled and converted into binary bits and pulses corresponding to binary bits are transmitted.

11.8.1 Advantages of Digital Transmission (PCM):

1. Primary advantage of digital transmission over analogue transmission is noise immunity.

   With digital transmission it is not necessary to evaluate the amplitude, frequency and phase as precisely as it is with analogue transmission. Instead, the received pulses are evaluated and a simple determination is made weather the pulse is above or below a certain threshold level. The exact amplitude, frequency or phase of the received signal is not important. Therefore, it has better noise immunity.

2. Digital signals are better suited to processing and multiplexing than analogue signal.

3. Regeneration of digital signal along transmission path is possible. Therefore, the digital transmission systems are more noise resistant than their analogue counterparts. Digital systems use signal regeneration rather than signal amplification.

4. Communication can be kept private and secured through the use of encryption.

5. Digital signals are simpler to measure and evaluate. Therefore, it is easier to compare the performance of alternate digital systems with unequal signalling and information capacities than it is with comparable analogue systems

6. Digital systems are better suited to evaluate error performance. Transmission errors in signals can be detected and corrected more easily and accurately than it is possible with analogue systems.

11.8.2 Disadvantages of Digital Transmission:

1. Transmission of digital signal requires more bandwidth as compared to analogue transmission.

2. Analogue signal must be converted to digital codes prior to transmission and converted back to analogue form at the receiver. Thus, additional encoding and decoding circuitry needed.

3. Digital transmission requires precise time synchronization between transmitter and receiver clock. Therefore, digital systems require expensive clock recovery circuit in all receivers.

11.9 QUANTIZATION:

- In analogue pulse modulation the signal is sampled and transmitted in analogue pulse form. In PCM also, sampling process is used but instead of transmitting analogue pulses, the PAM pulses are converted into a series of binary bits and the number of binary bits form a word and it represents the amplitude of the corresponding PAM pulse. Since series of digits are transmitted, this modulation is known as digital modulation.

- The basic principle of PCM is quantization. In PCM, the train of carrier pulses first amplitude is modulated by sampling process, so we get the PAM. Thus, PAM is quantized and encoded into binary code and then transmitted.

- *Quantization is a process in which PAM signal or actual signal is compared with fixed amplitude level called quantization level. Therefore, PAM signal is quantized i.e. rounded off to a particular quantization level.*

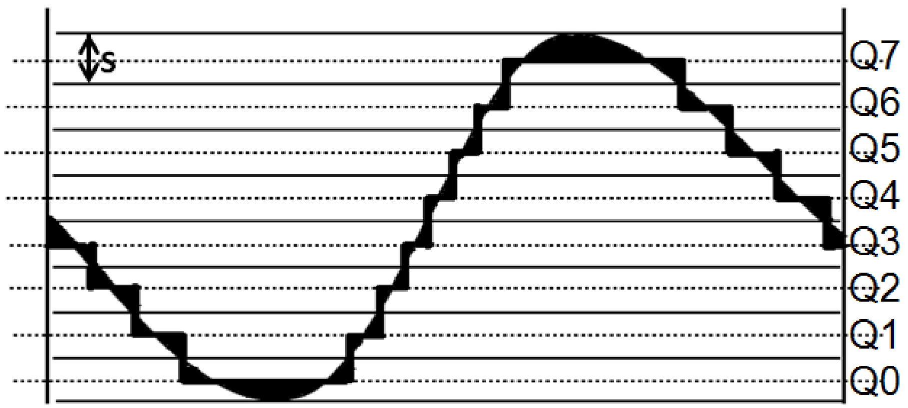


Fig. Quantization (a)

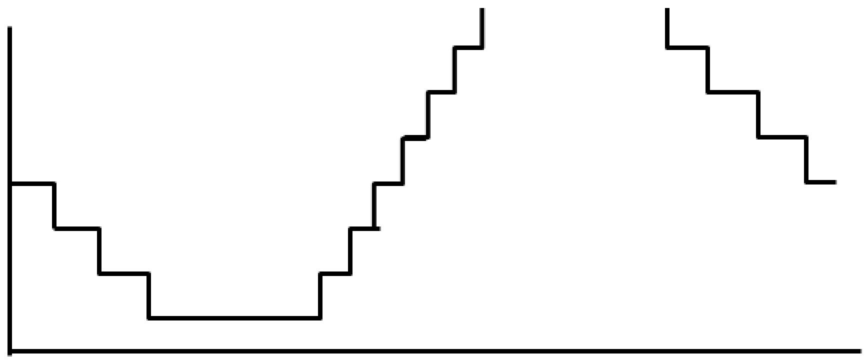


Fig. 11.6 Quantization (b) Output Digital signal

- The difference between two quantisation levels is called step size and it is

  $S = \frac{V_H - V_L}{2^N}$

  where N is the number of bits per sample.
- The number of quantization level required for quantization depends on the number of bits used to transmit a pulse and it is

  $Q = 2^N$ where

  Where Q = number of quantization levels

  N = number of bits per sample.

- As Q level increases, the number of bits per sample (per pulse) increases therefore, the bandwidth of PCM increases.

11.9.1 Quantization Error or Quantization Noise:

1. By comparing the quantized signal with actual signal, it can be seen that quantization process introduced distortion.
2. The difference between quantized signal and actual signal is called as **quantization error or quantization noise.**
3. The maximum value of quantization error is $\frac{S}{2}$ where S is the step size.
4. It can be reduced by increasing the number of Q levels but increment in Q level increases the number of bits to be transmitted per sample therefore, the bandwidth of PCM increases which is undesirable.
5. In practical system 128 Q level are used for speech signal and it is quite sufficient.

11.9.2 Types of Quantization Processes:

Quantization process can be classified into two types.

1. Uniform quantization (linear).
2. Non – uniform quantization (non – linear).

A quantizer is called as uniform quantizer if the step size remains constant throughout the input range. However, if the step size varies depending on the input then the quantizer is known is non – uniform quantizer.

Need of Non – Uniform Quantization or Taperal Quantization:

1. For uniform quantization the step size is fixed throughout quantization process. Therefore, the maximum quantization error is fixed irrespective of signal amplitude and it is $\pm \frac{S}{2}$. Thus for large amplitude input signal, the S/N ratio will be large i.e. the effect of Q noise will be less but for small amplitude input signal S/N ratio will be less i.e. the effect of Q noise will be more.
2. Thus uniform quantization is not suitable for weak input signal. To avoid this or to maintain constant S/N ratio throughout the peak – peak input signal range, there is a need of variable step size or non – uniform quantization.
3. In non – uniform quantization, corresponding to large amplitude signal step size is more and corresponding to weak signal, the step size is less. Thus, we get constant S/N ratio throughout the quantization process but it is very difficult or complicated to achieve non – uniform quantization in PCM. Therefore, there is another alternative to reduce the effect of Q noise on weak signal or to improve S/N ratio corresponding to weak signal. This is known as **companding.**

11.10 COMPANDING:

1. It is seen that in PCM small amplitude signals suffer from more quantization error than large amplitude signal in the process of PCM. Thus, companding is used to reduce Q error associated with small amplitude signal in PCM.

2. Companding is the process of compressing the signal at the transmitter and correspondingly expanding it at the receiver.

3. **Compressor:**
   - In this process(companding), the signal to be modulated is passed through such an amplifier (compressor) which is adjusted correctly for non – linear characteristics.
   - The low amplitude signal is amplified with higher gain whereas high amplitude signals are amplified with small gain. This process is called 'compressing'.
   - The amplifier which produces it is called as compressor and it is at the transmitter.
   - High amplitude signals are relatively compressed as compared to low amplitude signals. It can also be said that quantization error with small signals is compressed. Hence the name compressor.

4. **Expander:**
   1. At the receiver, there must be amplifier that should have exactly opposite characteristics as that of compressor.
   2. Such amplifier is called expander because it attenuates the low amplitude signal as compared to the high amplitude signal.

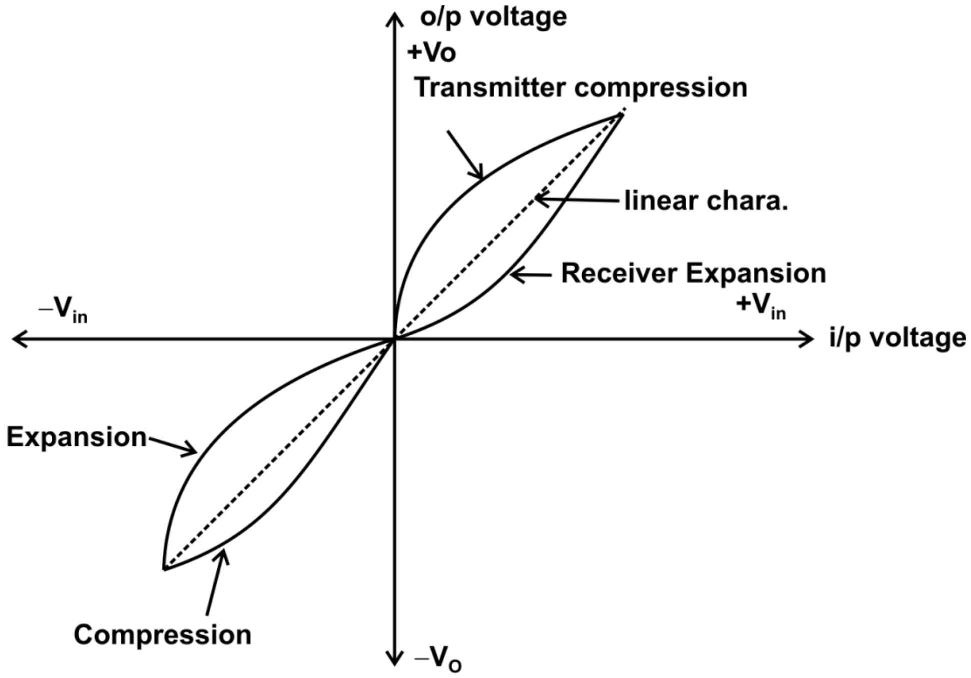


Fig. 11.7 Companding curve for PCM

Signalling Rate and Bandwidth of PCM:

- In PCM, input signal is sampled at the sampling rate f_s i.e. there are f_s number of samples per second. Each sample is converted into N bit word.

  $\therefore$ Number of bits per second
  = no. of samples per second x no. of bits per sample
  = $f_S N$
- Signalling rate is nothing but the number of bits per second.
  Signalling rate of PCM = Nf_s
- The transmission bandwidth of PCM is at least equal to half of signalling rate.
  $\therefore$ PCM bandwidth = $\frac{1}{2} N f_s$

11.11 DIFFERENTIAL PCM:

1. Instead of coding the entire sample amplitude for each sample, it is possible to code and transmit only the difference between amplitude of the current sample and that of previous sample.
2. Since successive samples often have similar amplitudes, it should be possible to use fewer bits to encode the changes.
3. The most common (and most extreme) example of this process is delta modulation.

11.12 DELTA MODULATION:

1. In PCM system, N bits per sample are transmitted.
2. Therefore, the transmission channel bandwidth of PCM system is very large.
3. This problem can be overcome by using delta modulation. Here only one bit per sample is transmitted instead of N bits as in PCM.

4. In delta modulation, the present sample value m(t) is compared with previous sample value $\hat{m}(t)$ and the result of comparison which is single bit is transmitted which simply indicates weather the present sample is larger or smaller than the previous sample.

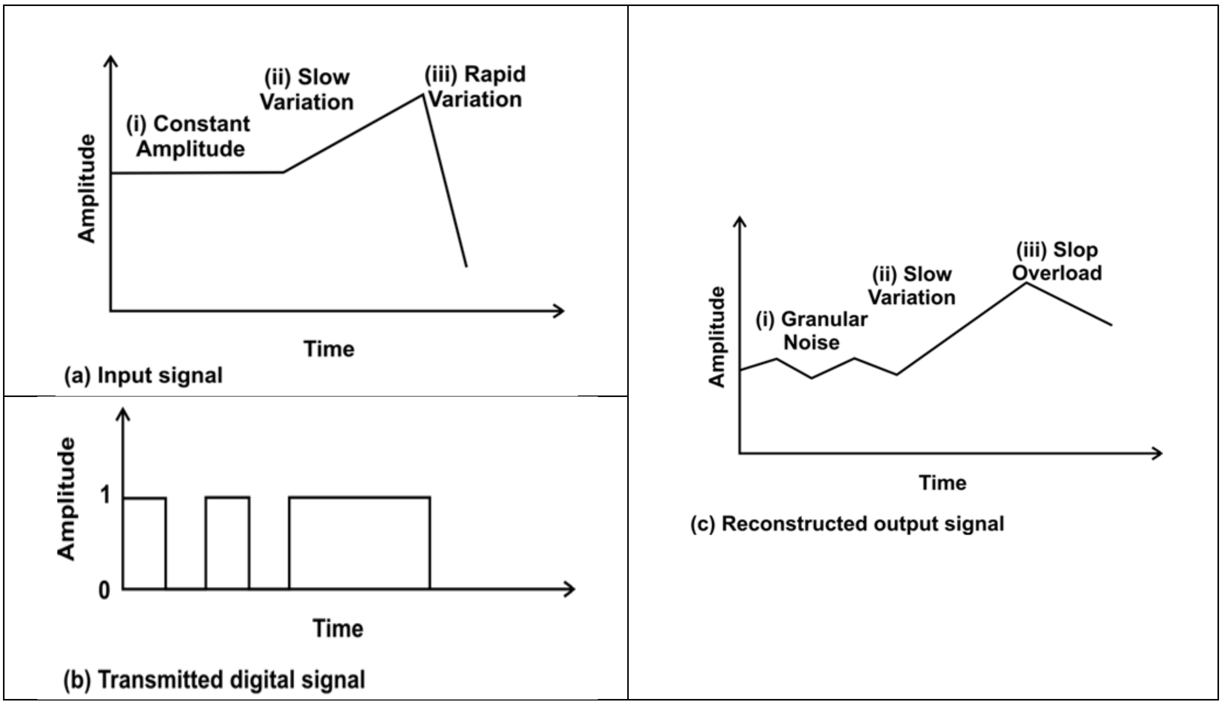


Fig.11.8 Delta modulation

1. Figure shows that how delta modulation generates error.
2. In region (i) the signal is not varying at all, transmitter can send only ones and zeros however the output waveform has a triangular shape, producing noise signal called granular noise.
3. On the other hand the signal in the region (iii) changes more rapidly than the system can follow creating an error in the output called slope over load

Problems with Delta Modulation:

There are two errors associated with DM.

1. Slope overload error.
2. Granular error.

11.13 SUMMARY:

- In pulse modulation, carrier is the train of pulses. Some parameters of these pulses (amplitude, width, position) are changed wrt the modulating signal
- The sampling signal is a periodic train of pulses with unit amplitude and of period T_s. This time is known as sampling time
- For proper reproduction of the modulating signal at the receiver, the modulating signal should be sampled at the rate of $f_s \geq 2f_m$ where f_m = highest modulating frequency
- There are two types of sampling techniques:
  - Natural sampling.
  - Flat top sampling.

- In PAM, the amplitude of carrier pulses is varied in accordance with the instantaneous amplitude of modulating signal
- In this system, the width of carrier pulses is varied in accordance with the instantaneous amplitude of modulating signal.
- The position of pulses is varied in accordance with the modulating signal
- *Quantization is a process in which PAM signal or actual signal is compared with fixed amplitude level called quantization level. Therefore, PAM signal is quantized i.e. rounded off to a particular quantization level.*
- The difference between quantized signal and actual signal is called as quantization error or quantization noise. The maximum value of quantization error is $\frac{S}{2}$ where S is the step size.
- Companding is the process of compressing the signal at the transmitter and correspondingly expanding it at the receiver
- In PCM system, N bits per sample are transmitted. Therefore, the transmission channel bandwidth of PCM system is very large. *This problem can be overcome by using delta modulation. Here only one bit per sample is transmitted instead of N bits as in PCM.*

11.14 UNIT END EXERCISE:

1. What is pulse modulation ? what are the advantages pf pulse modulation?
2. Explain Sampling process. State Sampling theorem.
3. Write shore notes on 1. PAM 2. PWM 3. PPM
4. What is PCM? hat are its advantages?
5. Explain the terms (i) Quantization , (ii) Quantization noise
6. Waht do you mean by Companding ?
7. Write in short on Delta modulation.

❖❖❖❖

UNIT V

12

MULTIPLEXING

Unit structure :

12.0 OBJECTIVE :

In this we are studying multiplexing viz FDM and TDM. We are also focussing on television system and television transmitter and television receiver for monochrome and colour television.

12.1 INTRODUCTION:

Simultaneous transmission of multiple messages over a channel is called multiplexing. There are 2 types of multiplexing.

1. Frequency division multiplexing.
2. Time division multiplexing.

FDM uses analogue modulation system, whereas TDM uses pulse modulation system.

12.2 FREQUENCY DIVISION MULTIPLEXING (FDM):

1. FDM consists simultaneous transmission of messages of different channels by shifting them in frequency domain. The block diagram of FDM transmitter(T_X) and receiver (R_X) is as shown below

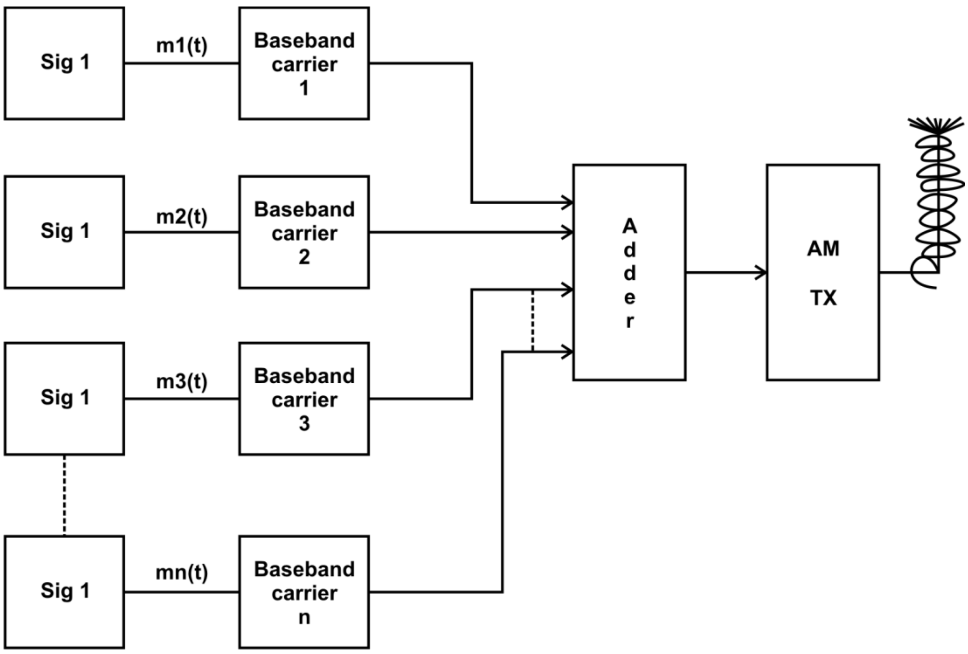


Fig. 12.1 Block diagram of FDM

2. In above diagram, all the baseband carrier blocks are the amplitude modulators.
3. All the above modulating signals have the same frequency spectrum. Therefore, all these signals are up shifted in different frequency slots using amplitude modulators.
4. For every BaseBand Carrier (BBC) block, the carrier frequency is different. Hence, the output of BC–1 will be $fc_1 \pm fm_1$. The output of BBC–2 will be $fc_2 \pm fm_2$. If all the modulating signals have the same BW, then the output adder will have the following spectrum.

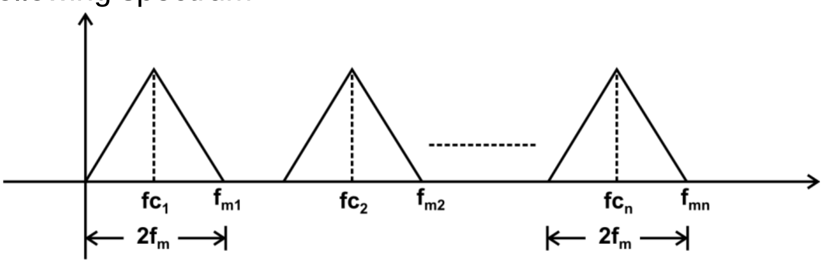


Fig. 12.2 frequency spectrum of output adder

5. This FDM signal is transmitted by Transmitter (T_X). The block diagram of FDM Receiver (R_X) is as shown in figure 12.3.

6.

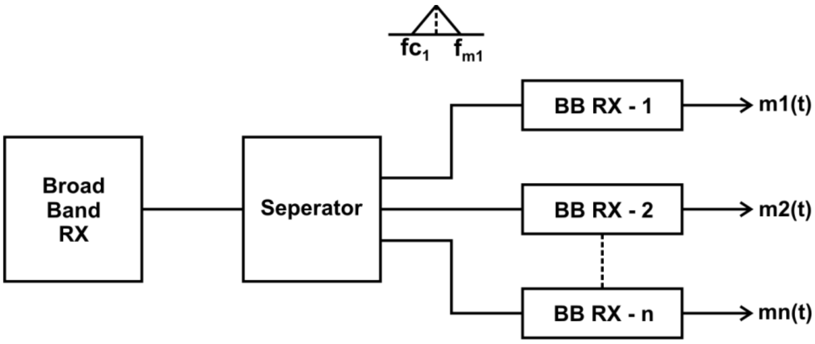


Fig. 12.3 Block diagram of FDM R_X

7. At the receiving station, the FDM signal is received by the broadband R_X. The separator circuit uses approximate mixers and filters, tuned at proper frequencies. The separator has N outputs and each o/p is the modulated signal centred around fc_1, fc_2 ... fc_n respectively. These outputs are given to Baseband Receiver (BBR_X). Hence, the output of each block is corresponding modulating signal.

12.2.1 FDM Hierarchy:

1. The Hierarchy of FDM signals in descending order is as follows:
    1. Super Jumbo Group - 43200KHz (3 Jumbo groups)
    2. Jumbo Group - 14400KHz (6 Master groups)
    3. Master Group - 2400KHz (10 Super groups)
    4. Super Group - 240KHz (05 basic groups)
    5. Basic Group - 48KHz (12 Single Voice Channels)
    6. Single Voice Channel - 4 KHz
2. The FDM using voice channel has certain standard for grouping of voice channels. Every voice channel occupies 4 KHz band including guard band from adjacent channels.
3. 12 voice channels are grouped together and form **basic group**. Therefore, basic group bandwidth is 12 X 4 KHz = 48 KHz.

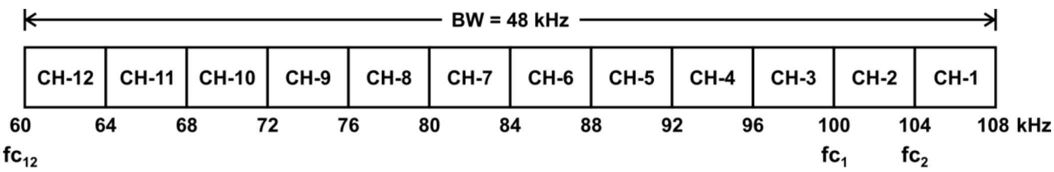


Fig. 12.4

4. Five such basic groups are multiplexed together to form **supergroup**. Thus single supergroup can carry information from 12 X 5 = 60 voice channels. Therefore, the supergroup band is 60 X 4 = 240 KHz.

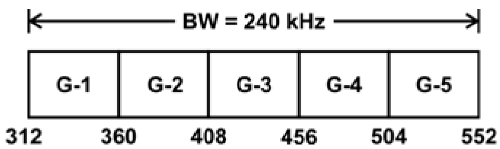


Fig. 12.5

5. 10 such supergroups multiplexed together to form **master group**. Thus single master group can carry 60 X 10 = 600 voice channels. Therefore, this bandwidth is 600 X 4 = 2400 KHz.
6. Six such master groups are multiplexed together to get **jumbo group**. Therefore a jumbo group has 600 X 6 = 3600 voice channels.
7. A **super jumbo group** has 3 jumbo groups i.e. 10, 800 voice channels.

12.2.2 Disadvantages of FDM :

1. Bandwidth is wasted because of guard band requirement between adjacent channels, to minimise channel interference i.e. cross – talk between channels.

2. The system is not flexible to BW changes e.g. if the BW of one channel is increased, the centre frequency of all other channels may need to be redefined or the number of channels should reduce.

12.3 TIME DIVISION MULTIPLEXING (TDM):

1. When a signal is sampled by narrow pulses, then the time interval between two pulses can be used to transmit samples of other signals.
2. In this technique, the signals are multiplexed in the time domain and hence called as Time Division Multiplexing (TDM).
3. It is also used to transmit number of signals on a single transmission media and hence act as alternative to FDM.

12.3.1 TDM System:

The block diagram of PAM/TDM is as shown in the figure.

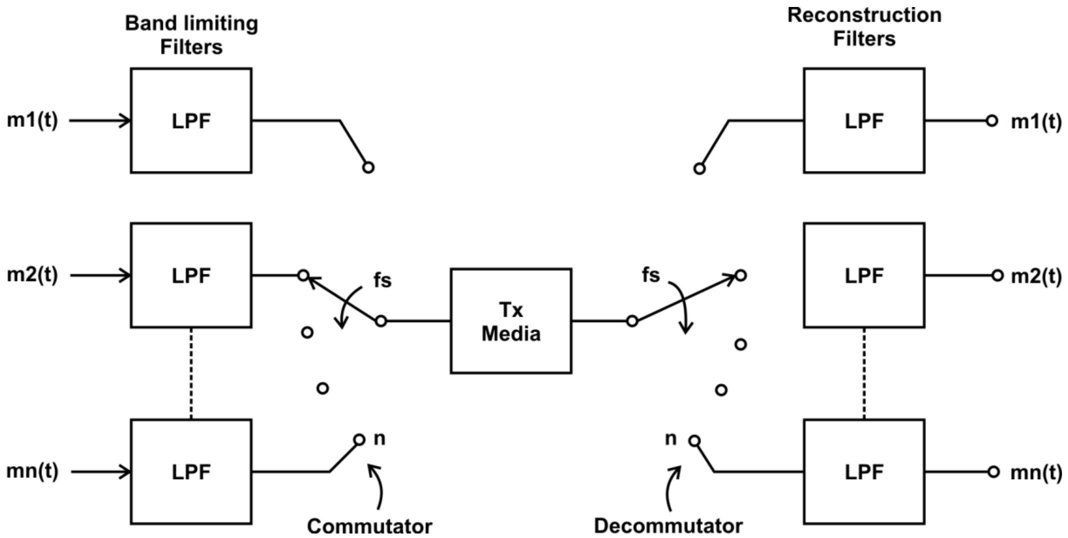


Fig. 12.6 Block diagram of TDM

1. The switch arrangement at the T_X is provided by the **commutator circuit**. In each one of its rotation, the commutator extracts of samples one sample from each message input $m_1(t)$, $m_2(t)$ ... $m_n(t)$. Thus at the output of commutator, we get PAM waveform which contains the sample of message inputs which are periodically interlaced in time. These multiplexed message samples are transmitted over communication channel.

2. At the receiving end, **decommutator** is used which distributes the pulses to different receivers. The demodulator is again a switching arrangement at the receiving end, similar to that of the transmitting end. This decommutator is used to separate various received samples and to distribute them to an assembly of LPFs. The LPF then reconstructs individual messages $m_1(t)$, $m_2(t)$ ... $m_n(t)$ at the output.

3. Here it is necessary that the rate of switching of commutator and decommutator must be same and they must be synchronised to each other. This synchronization is achieved by sending a synchronization pulse. Thus after sending (n – 1) pulses (each pulse from different channels) one synchronization pulse is sent. Thus overall, n pulses are sent in time T_S.

12.4 COMPARISON OF FDM AND TDM:

FDM	TDM
1. Multiplexing is done in frequency domain. 2. Circuitry required is complex because different carrier modulator, mixers and demodulators and filters of different frequencies are required. 3. Cross – talk problems are significant because of improper filtering. 4. Synchronization between the transmitter and receiver is not required. 5. Analogue signals are not sampled. 6. The linear sampling amplifier is used to obtain FDM. 7. No synchronizing signals. 8. It uses analogue modulation system.	1. Multiplexing is done in time domain. 2. Circuitry in TDM is not as complex as that of FDM system. Moreover, the blocks (circuits) used in TDM will be identical in design. 3. Cross – talk problems are not significant as that of in FDM. 4. Synchronization between transmitter and receiver is must. 5. Analogue signals are sampled. 6. The time division multiplexer is used to obtain TDM. 7. Synchronization bits and frame bits are added with PCM samples. 8. It uses digital modulation system.

12.5 TELEVISION SCANNING:

1. Television and video systems form very important part of the communication environment. Video systems form pictures by scanning process.

2. The image is divided into a number of horizontal lines which are traced out synchronously at the camera and the receiver. The number of lines is arbitrary, but increasing it gives a better resolution in vertical direction. Not all the lines are visible on the television screen. The number of scan lines actually used to form the image is less than the total number of lines. The remaining lines are transmitted between the intervals between images.

3. In order to make the image visible all at once, rather than a series of consecutively drawn lines, the process must be completed quickly. Also in order to provide the illusion of motion, many images must be drawn in quick succession.

4. The more quickly the images follow one another, less is the flicker present. However, this requires more information to be transmitted, increasing the required bandwidth. For this, images are sent at the rate of 30 per second.

5. Frame rates of 25 to 30 Hz cause noticeable flicker. To reduce this, most televisions use a technique known as interlaced scan, which involves transmitting alternate lines of picture, then returning and filling the missing lines. Figure 1.7 illustrates the idea. Each half of the picture thus sent is called a field, and of course the field rate is twice the frame rate i.e. 60Hz in North American system. This result is reduced flicker without increased bandwidth.

6. As figure 12.7 shows, the picture is scanned from left to right and top to bottom. The electron beam that scans the picture is blanked during the time intervals the beam retraces its path, from right to left and from bottom to top. These time intervals are called the horizontal and vertical blanking intervals respectively.

7. The ratio of width to height, or aspect ratio, is 4: 3. If the frame rate is 30 Hz, then the rate at which the lines are sent is 525 multiplied by 30 or 15750 Hz. These frequency specifications are somewhat arbitrary, but obviously the same standards must be in use at both ends of communication path.

8. Good colour reproduction can be achieved by mixing three primary colours: red, blue and yellow. Thus considerably, more information must be transmitted for colour television than monochrome. Colour television also requires a rather elaborate picture tube containing three electron guns.

.

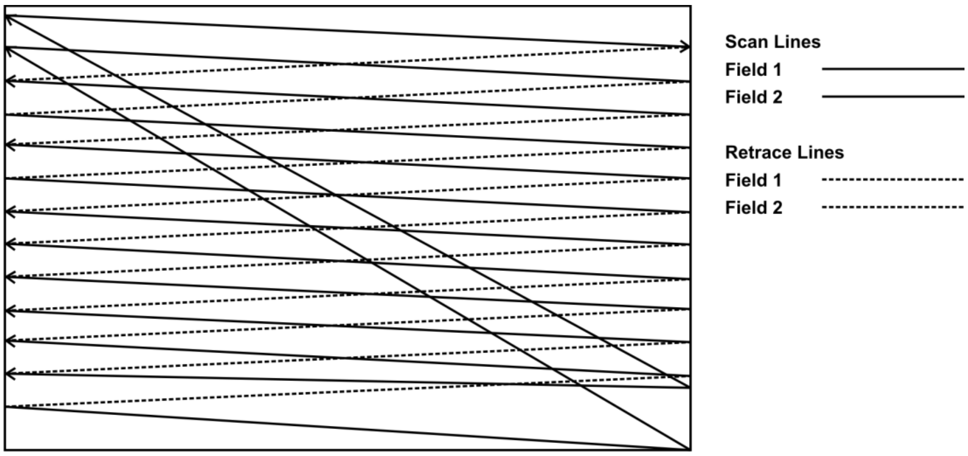


Fig. 12.7 Television Scanning

12.6 COMPOSITE VIDEO SIGNAL:

1. The signal we examine most closely is called a composite video signal, because it combines all the picture information, along with synchronization pulses.

2. With a composite signal, the whole signal can be transmitted on the same cable or the same radio channel. It is also possible to separate various components of a video signal, and this is often done when transmission signal is not great – with computer monitors, for instance.

12.6.1 Luminance Signal:

1. Figure 13.10 shows one line of monochrome signal that conforms to the North American standard. Note that the duration of line is 63.5 µs. This is simply the period corresponding to the horizontal line frequency of 15.75 KHz. About 10 µs of this is used by the horizontal synchronizing (sync) pulse. The rest of the time is occupied by analogue signal that represents the variation of luminance (brightness) level along the line.

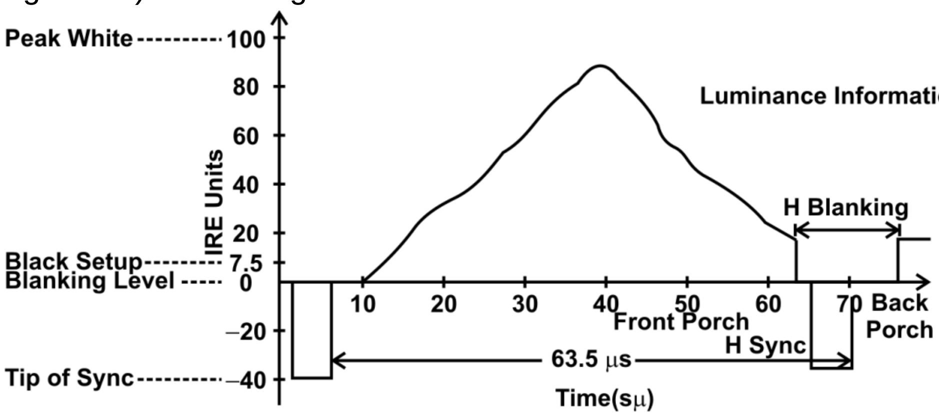


Fig. 12.8 Monochrome video signal

2. The video signal shown in figure 12.8 has negative sync, that is, the sync pulses are in negative direction. An inverted version of this signal, called positive sync, is also possible. The polarity of luminance portion depends on

that of the sync pulses, which are always in the direction of black. Both polarities and a variety of amplitudes are used in video equipment, but for the interconnection of equipment, the standard is negative sync with amplitude of 1V peak – to – peak, into a 75 Ω terminating resistance.

3. Because of the variety of possible levels and polarities for the video signal, a scale of relative amplitudes is used for setting the correct proportions between the level of synchronizing pulses and the luminance signal. It is called the IRE scale, for the institute of Radio Engineers, a precursor to the Institute of electrical and Electronic Engineers (IEEE). On this scale, which is also shown in figure 13.10, zero represents blanking level, and –40 represents the level of the sync pulses. The maximum luminance level, called peak white, is 100, and 7.5 is a level that should represent black the receiver.

4. From this, it can be seen that the blacking and the synchronizing pulses are “blacker than black” and ensure that the CRT electron beam is completely turned off while it retraces from right to left.

12.7 TELEVISION TRANMITTER:

1. An oversimplified block diagram of a monochrome TV transmitter is shown in Fig. 1.8.

2. The signal from the camera is amplified and synchronizing pulses are added before feeding it to the modulating amplifier.

3. Synchronizing pulses are transmitted to keep the camera and picture tube beams in step.

4. The allotted picture carrier frequency is generated by a crystal oscillator. The output wave is given large amplification before feeding to the power amplifier where its amplitude is made to vary (AM) in accordance with the modulating signal received from the modulating amplifier. The modulated output is combined (see Fig. 1.8) with the frequency modulated (FM) sound signal in the combining network and then fed to the transmitting antenna for broadcasting.

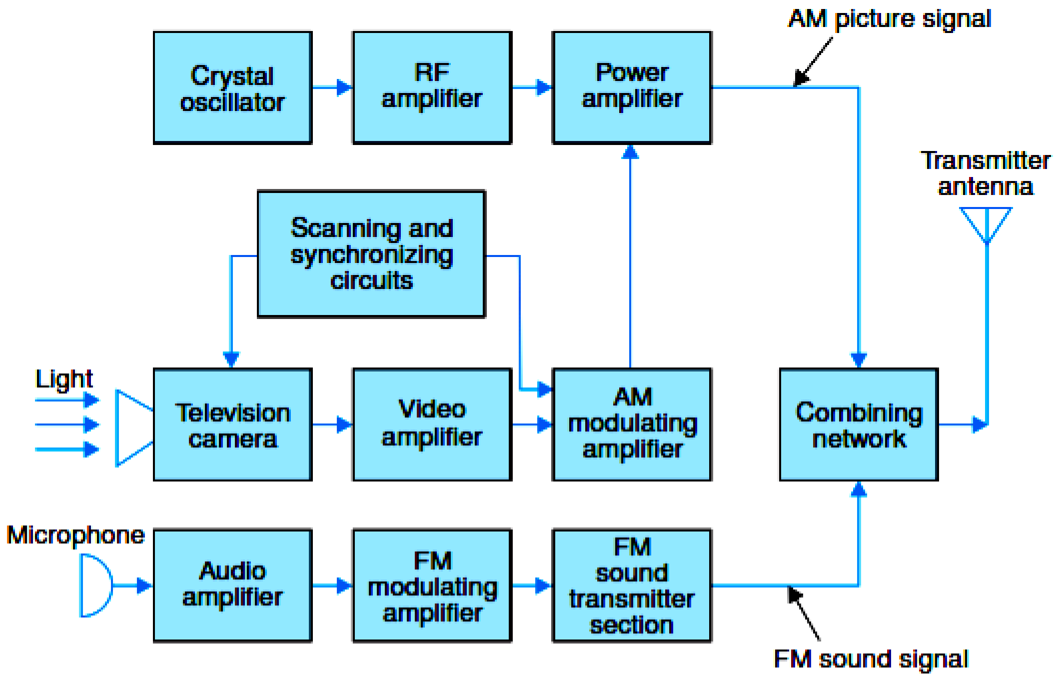


Fig. 12.9 Elementary block diagram of a monochrome television transmitter

12.7.1 Colour Transmitter

1. A colour TV transmitter is essentially the same as the monochrome transmitter except for the additional need that colour (chroma) information is also to be transmitted.

2. Any colour system is made compatible with the corresponding monochrome system. Compatibility means that the colour TV signal must produce a normal black and white picture on a monochrome receiver and a colour receiver must be able to produce a normal black and white picture from a monochrome TV signal.

3. For this, the luminance (brightness) signal is transmitted in a colour system in the same way as in the monochrome system and with the same bandwidth. However, to ensure compatibility, the colour camera outputs are modified to obtain (B-Y) and (R-Y) signals. The symbol used for luminance is "Y." The colour difference signals also contain Y and one of the primary colours -- usually B-Y and R-Y. For example, B-Y is the Blue video signal minus the luminance signal.

4. These are modulated on the colour sub-carrier, the value of which is so chosen that on combining with the luminance signal, the sidebands of the two do not interfere with each other *i.e.,* the luminance and colour signals are correctly interleaved. A colour sync signal called 'colour burst' is also transmitted for correct reproduction of colours.

12.7.2 Sound Transmission

1. There is no difference in sound transmission between monochrome and colour TV systems.

2. The microphone converts the sound associated with the picture being televised into proportionate electrical signal, which is normally a voltage. This electrical output, regardless of the complexity of its waveform, is a single valued function of time and so needs a single channel for its transmission.

3. The audio signal from the microphone after amplification is frequency modulated, employing the assigned carrier frequency.

4. In FM, the amplitude of carrier signal is held constant, whereas its frequency is varied in accordance with amplitude variations of the modulating signal. As shown in Fig. 13.8, output of the sound FM transmitter is finally combined with the AM picture transmitter output, through a combining network, and fed to a common antenna for broadcasting in the form of electromagnetic waves.

12.8 TELEVISION RECEIVER:

1. A simplified block diagram of a black and white TV receiver is shown in Fig. 12.10.

2. The receiving antenna intercepts broadcasted RF signals and the tuner selects desired channel's frequency band and converts it to the common IF band of frequencies. The receiver employs two or three stages of intermediate frequency (IF) amplifiers.

3. The output from the last IF stage is demodulated to recover the video signal. This signal that carries picture information is amplified and coupled to the picture tube which converts the electrical signal back into picture elements of the same degree of black and white.

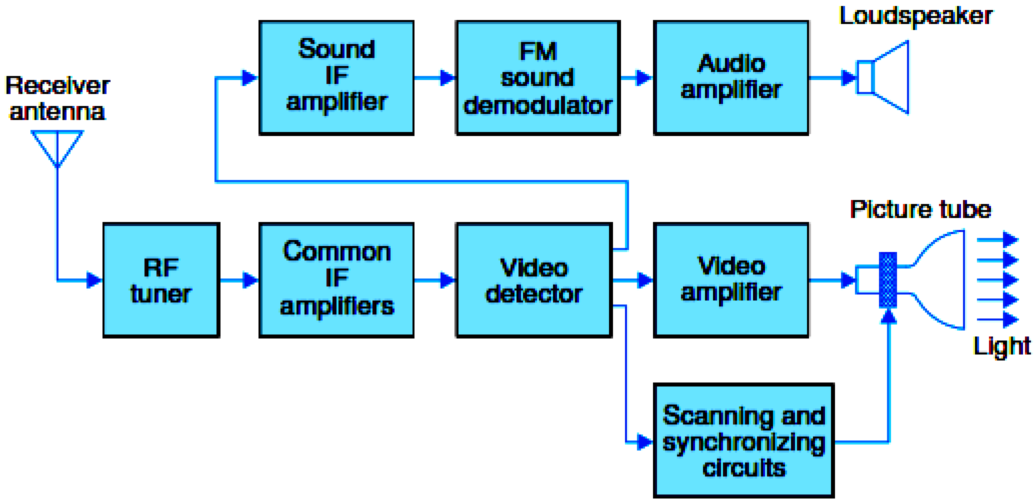


Fig. 12.10 Block diagram of a black and white TV receiver.

4. The picture tube shown in Fig. 12.11 is very similar to the cathode-ray tube used in an oscilloscope. The glass envelope contains an electron-gun structure that produces a beam of electrons aimed at the fluorescent screen. When the electron beam strikes the screen, light is emitted. The beam is deflected by a pair of

deflecting coils mounted on the neck of picture tube. The amplitudes of currents in the horizontal and vertical deflecting coils are so adjusted that the entire screen, called raster, gets illuminated because of the fast rate of scanning.

5. The video signal is fed to the grid or cathode of picture tube.
   - When the varying signal voltage makes the control grid less negative, the beam current is increased, making the spot of light on the screen brighter. More negative grid voltage reduces brightness. If the grid voltage is negative enough to cut-off the electron beam current at the picture tube, there will be no light. This state corresponds to black.

   - Thus the video signal illuminates the fluorescent screen from white to black through various shades of grey depending on its amplitude at any instant. This corresponds to brightness changes encountered by the electron beam of the camera tube while scanning picture details element by element. The rate at which the spot of light moves is so fast that the eye is unable to follow it and so a complete picture is seen because of storage capability of the human eye.

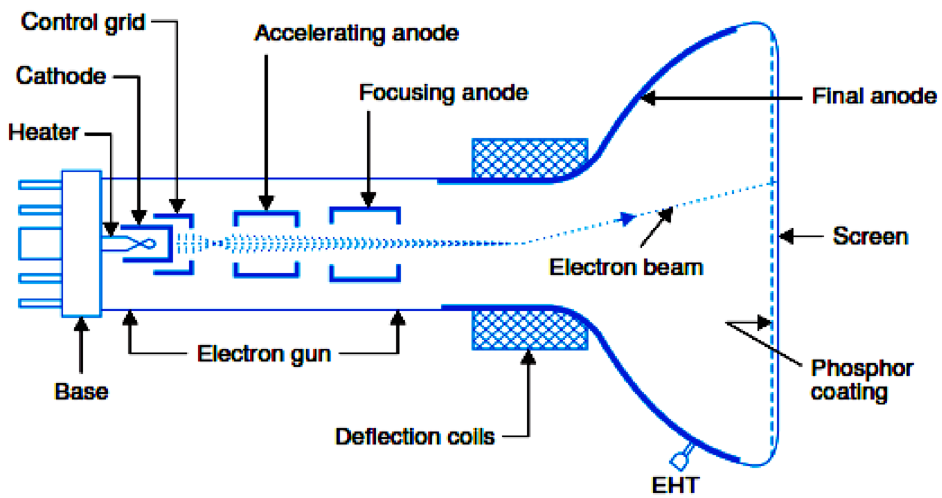


Fig. 12.11 Elements of picture tube

Sound Reception

1. The path of sound signal is common with the picture signal from antenna to video detector section of the receiver. Here the two signals are separated and fed to their respective channels.

2. The frequency modulated audio signal is demodulated after at least one stage of amplification.

3. The audio output from the FM detector is given due amplification before feeding it to the loudspeaker.

Colour Receiver

1. A colour receiver is similar to the black and white receiver as shown in Fig. 12.12.

2. The main difference between the two is the need of a colour subsystem. It accepts only the colour signal and processes it to recover (B-Y) and (R-Y) signals.
3. These are combined with the Y signal to obtain VR(video red), VG(video green) and VB(video blue) signals as developed by the camera at the transmitting end. VG becomes available as it is contained in the Y signal.
4. The three colour signals are fed after sufficient amplification to the colour picture tube to produce a colour picture on its screen.
5. As shown in Fig. 12.12, the colour picture tube has three guns corresponding to the three pick-up tubes in the colour camera.
6. The screen of this tube has red, green and blue phosphors arranged in alternate stripes. Each gun produces an electron beam to illuminate corresponding colour phosphor separately on the fluorescent screen. The eye then integrates the red, green and blue colour information and their luminance to perceive actual colour and brightness of the picture being televised.
7. The sound signal is decoded in the same way as in a monochrome receiver.

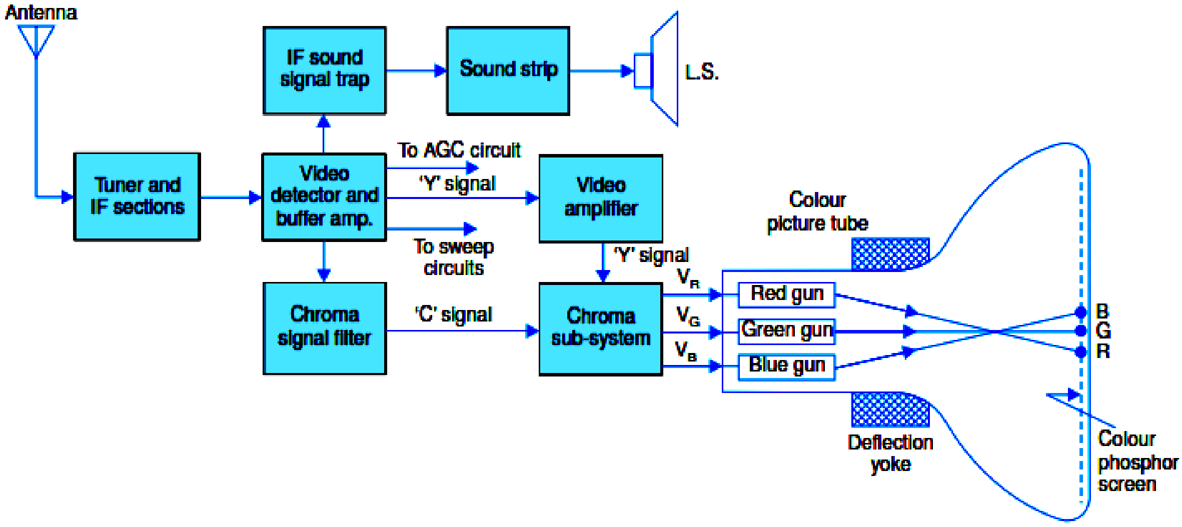


Fig. 12.12 Block diagram of a colour receiver

12.8. RECEIVER CONTROLS

1. Most black and white receivers have on their front panel (*i*) channel selector, (*ii*) fine tuning, (*iii*) brightness, (*iv*) contrast, (*v*) horizontal hold and (*vi*) volume controls besides an ON-OFF switch. Some receivers also provide a tone control.
2. The channel selector switch is used for selecting the desired channel.
3. The fine tuning control is provided for obtaining best picture details in the selected channel.
4. The hold control is used to get a steady picture in case it rolls up or down.
5. The brightness control varies beam intensity of the picture tube and is set for optimum average brightness of the picture.

6. The contrast control is actually gain control of the video amplifier. This can be varied to obtain desired contrast between white and black contents of the reproduced picture.
7. The volume and tone controls form part of the audio amplifier in sound section, and are used for setting volume and tonal quality of the sound output from the loudspeaker.

12.9 SUMMARY :

- Simultaneous transmission of multiple messages over a channel is called multiplexing.
- FDM consists simultaneous transmission of messages of different channels by shifting them in frequency domain
- When a signal is sampled by narrow pulses, then the time interval between two pulses can be used to transmit samples of other signals. In this technique, the signals are multiplexed in the time domain and hence called as Time Division Multiplexing (TDM).
- Video systems form pictures by scanning process.
- The signal we examine most closely is called a composite video signal, because it combines all the picture information, along with synchronization pulses.
- Most black and white receivers have on their front panel (*i*) channel selector, (*ii*) fine tuning, (*iii*) brightness, (*iv*) contrast, (*v*) horizontal hold and (*vi*) volume controls besides an ON-OFF switch.
- In colour receivers there is an additional control called ‘colour’ or ‘saturation’ control.

12.10 UNIT END EXERCISE :

1. What is multiplexing ? explain its types.
2. What is FDM? explain with diagram the working of FDM.
3. What is TDM? explain with diagram the working of TDM.
4. Write a note on (i) television Scanning , (ii) composite video signal.
5. With the help of block diagram explain television transmitter.
6. Explain the working of television receiver.
7. How is a colour receiver different from a black and white receiver? Explain why separate colour sync pulses (colour burst) are needed in the CTV system.
8. Describe briefly the functions of various controls provided on the front panel of a monochrome receiver. What is the purpose of saturation (colour) control in a colour receiver?

❖❖❖❖

UNIT VI – Digital communication and Fibre optics System

13

DIGITAL MODULATION

Unit Structure :

13.0 OBJECTIVES:

While going through this chapter you will be focussing on –
- digital modulation
- FSK and its parameters
- FM detectors
- PSK demodulators

13.1 INTRODUCTION:

Digital signals have become very important in both wired and wireless communication. To send data by radio it is necessary to use a higher frequency carrier wave. Since the high frequency carrier wave is a sine wave for digital as well as analogy signal the same three basic parameters are available for modulation; amplitude, frequency, phase. All three, singly and combination are used in digital systems.

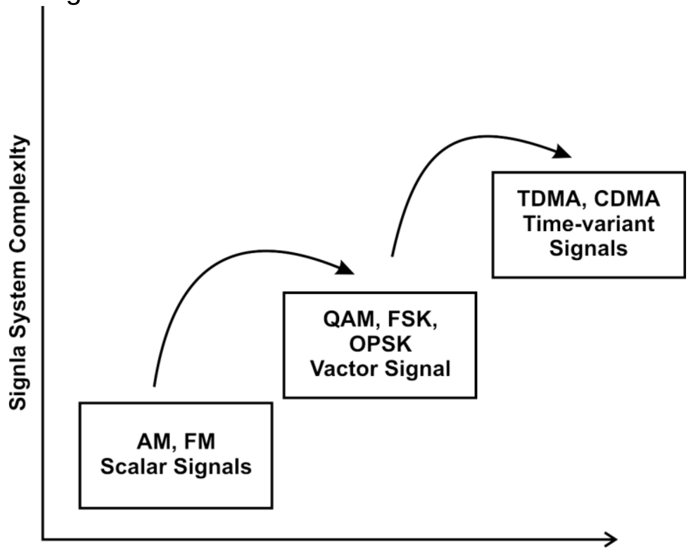
13.2 WHY DIGITAL MODULATION?

1. The move to digital modulation provides more information capacity, compatibility with digital data services, higher data security, better quality communications, and quicker system availability. Developers of communications systems face these constraints:
    - Available bandwidth
    - Permissible power
    - Inherent noise level of system.

2. The RF spectrum must be shared, yet everyday there are more users for that spectrum as demand for communications services increases. Digital modulation schemes have greater capacity to convey large amounts of information than analogue modulation schemes.

3. The transmission of the digital signals is increasing at a rapid rate. Low – frequency analogue signals are often converted to digital form before transmission. The source signals are generally referred to as the ***baseband*** signals. We can send analogue and digital signals directly over as medium. From electro – magnetic theory, for efficient radiation of electrical energy from an antenna, it must at least be in the order of magnitude of a wavelength in size; $c = \lambda f$ where c is the velocity of light, f is the signal frequency and λ is wavelength. For a 1 kHz audio signal, the wavelength is 300 km. the low – frequency signal is often frequency – translated to a higher frequency range for better transmission. This process is called ***modulation***. The use of a higher frequency range reduces antenna size.

4. In the modulation process, the baseband signals constitute the modulating signal and the high – frequency carrier signal is a sinusoidal waveform. There are three basic ways of modulating a sine wave carrier. For binary data modulation, they are called ***Binary Amplitude Shift Keying (BASK)***, ***binary frequency shift keying (BFSK)*** and ***binary phase – shift keying (BPSK)***. Modulation also leads to possibility of frequency multiplexing. In a frequency – multiplexed system, individual signals are transmitted over adjacent, non – over lapping frequency bands. They are therefore transmitted in parallel and simultaneous time. If we operate at higher carrier frequencies, more bandwidth is available for frequency – multiplexing more signals.



13.1 Recent trends in communication

5. Moreover, over a past few decades a major transition from simple analogue Amplitude modulation (AM) and Frequency/Phase Modulation (FM/PM) to new digital modulation techniques. Another layer of complexity in many new systems is multiplexing. Two principle types of multiplexing (or “multiple accesses”) are TDMA (Time Division Multiple Access) and CDMA (Code Division Multiple Access). These are the two ways to add diversity to signals allowing different signals to be separated from one another.

13.3 FREQUENCY SHIFT KEYING (FSK):-

- In Frequency shift keying, we change the frequency of the carrier wave.
- Bit 0 is represented by a specific frequency, and bit 1 is represented by a different frequency.
- In the figure below frequency used for bit 1 is higher than frequency used for bit 0

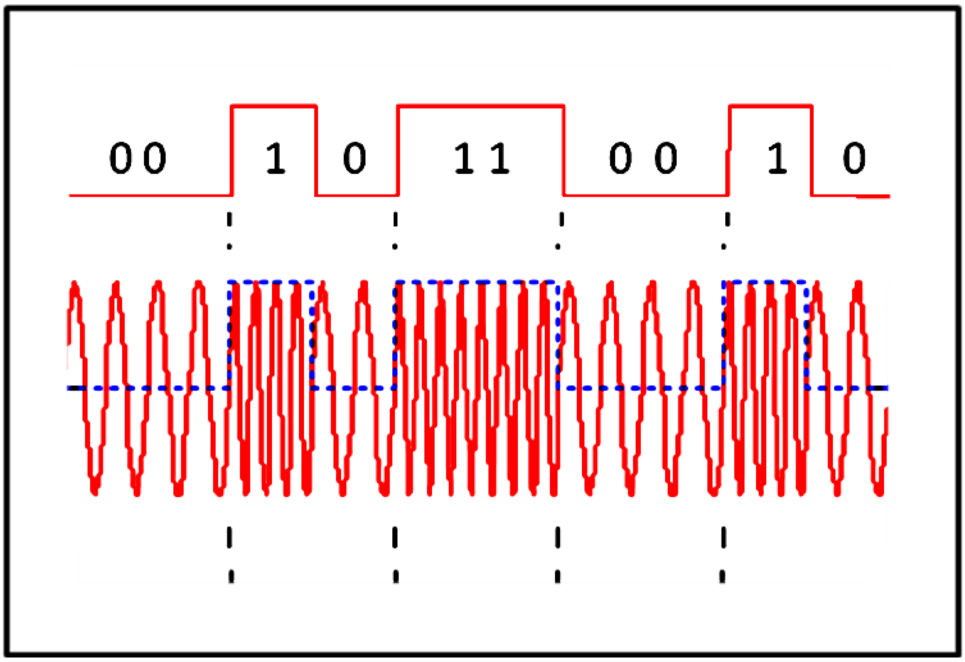


Figure :13.2 Frequency Shift Keying (FSK)

13.4. AMPLITUDE SHIFT KEYING (ASK)

- In amplitude shift keying, the amplitude of the carrier signal is varied to create signal elements.
- Both frequency and phase remain constant while the amplitude changes.
- **Binary ASK (BASK)**

  ASK is normally implemented using only two levels and is hence called binary amplitude shift keying.

  Bit 1 is transmitted by a carrier of one particular amplitude.

  To transmit Bit 0 we change the amplitude keeping the frequency is kept constant

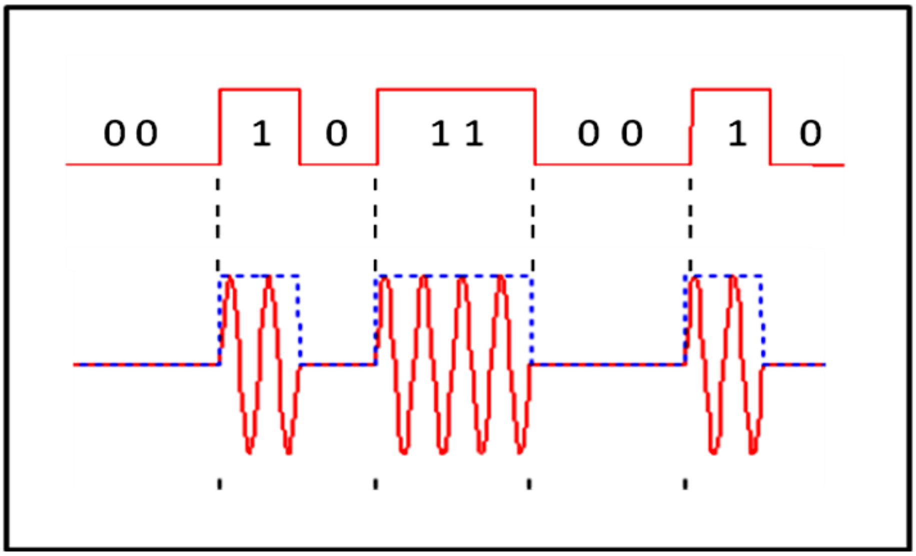


Figure :13.3 Amplitude Shift Keying (ASK

13.5 PHASE SHIFT KEYING (PSK)

- Phase shift keying (PSK) is a method of transmitting and receiving digital signals in which the phase of a transmitted signal is varied to convey information.
- Both amplitude and frequency remain constant as the phase changes.
- The simplest from of PSK has only two phases, 0 and 1.
- If the phase of the wave does not change, then the signal state stays the same (low or high).
- If the phase of the wave changes by 180 degrees, that is, if the phase reverses, then the signal state changes (from low to high or from high to low)

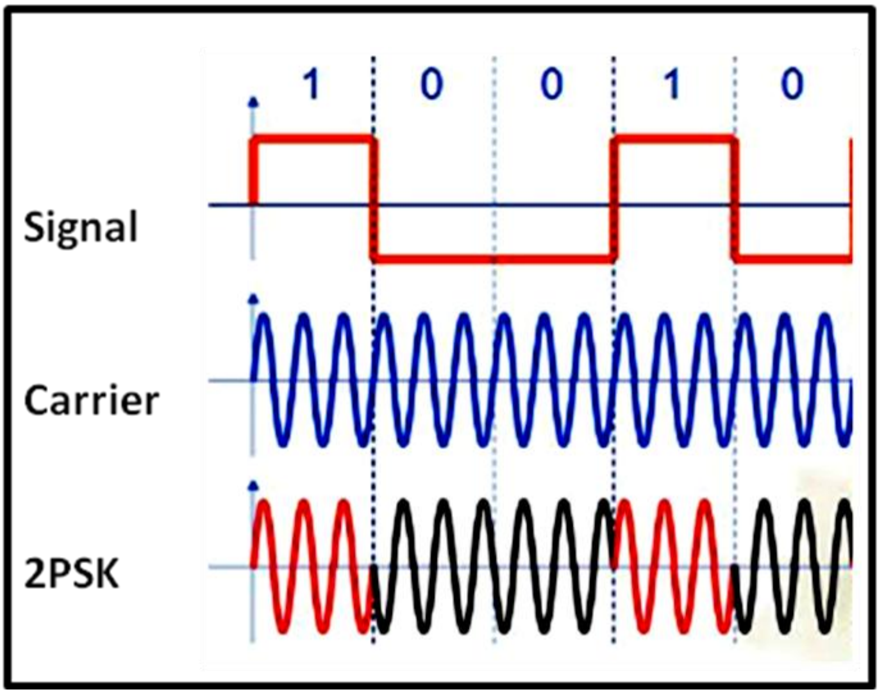


Figure: 13.4 Phase Shift Keying (PSK)

13.7 SUMMARY :

- The source signals are generally referred to as the ***baseband*** signals.
- There are three basic ways of modulating a sine wave carrier. For binary data modulation, they are called ***Binary Amplitude Shift Keying (BASK)***, ***binary frequency shift keying (BFSK)*** and ***binary phase – shift keying (BPSK)***.
- Frequency Shift Keying is the most common form of digital modulation in high frequency radio spectrum, and has most important applications in telephone circuits.

13.8 UNIT END EXERCISE :

1. What is the need of digital modulation?
2. Write Short Notes on : ASK, FSK and PSK.

❖❖❖❖

14

INTRODUCTION TO FIBER OPTIC SYSTEM

Unit Structure :

14.0 OBJECTIVES

This chapter serves as an introduction to Fiber optic system of data transmission. The fiber optic wire carries data in the form of light rather than electromagnetic signals as done by copper cables.

14.1 INTRODUCTION

In the communication system, the wired transmission lines were mostly used. But nowadays, the optic fibres are rapidly replacing the wire transmission lines. These optic fibre lines offer several important advantages over wire lines.

1. As light is the same as radio frequency radiation, but at a very high frequency (3,000,000 GHz) the information carrying capacity of a fibre is very much greater than the microwave radio systems.
2. The materials used in fibres is silicon dioxide or silica glass which are the most abundant materials on the earth, and this results in much lower material cost than with wire lines.
3. The fibres are not electrically conductive and therefore, they may be used in areas where isolation from electrical and electromagnetic interference is a problem.
4. The fibre optic lines also have a much higher information capacities and multiple channel routes which can be compressed into much smaller cables using fibre optics.

In the present technology although the fibre optic communication systems are still more expensive than wire or radio systems, the fibre optic systems have more advantage over wired line systems.

14.2 OPTICAL FIBRE CABLE

1. Optical fibres are thin hair like strands, and consist of glass or silica fibre of few micrometers called as the **core**.
2. The core is enclosed in a cover called **cladding**. The cladding has reflective index less than the core.
3. The schematic diagram of an optical fibre is shown in figure 14.1.

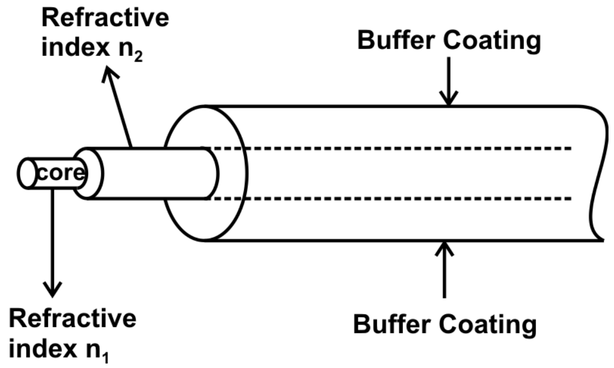


Fig. 14.1: Schematic diagram of optical fibre structure

4. The core diameter is about 10 µm to 100 µm while the cladding diameter is about 100 µm to 400 µm.
5. The fibre is enclosed in an outer protective coating called the buffer. This buffer coating provides strength to the optical fibre. The outer buffer coating also provides protection from environmental factors.

14.3 PROPAGATION OF LIGHT IN OPTICAL FIBRE

1. The optical fibre works on the principle of **Total Internal Reflection**. Whenever light enters one end of the glass fibre under right conditions, most of the right will propagate along the length of the fibre and exit from far end.
2. In this process, some part of light will escape through side walls of the fibre and some will be lost through inter absorption. But a portion of light will be guided to the far end. This fibre is called **light pipe** or **light guide.**
3. Whenever a ray of light travels from one medium to another medium of different refractive index, the ray of light may be either reflected or refracted. In the fibre optics the light stays in it because it is **totally reflected** by the surface of the fibre.
4. Light entering the end of a fibre at slight angle to the axis follows a zigzag path inside the fibre. The total internal reflection occurs only when two following conditions are satisfied:-

- The refractive index n_1 of the core should be greater than the refractive index n_2 of the cladding i.e. $n_1 > n_2$.
- The light approaching the fibre must have an angle of incidence Φ greater than the critical angle Φ_c which is defines as

$$\Phi_c = \frac{n_2}{n_1}$$

5. The reflected light will leave the fibre wall at the same angle as it had struck the wall before reflection. This can be seen from figure 14.2.

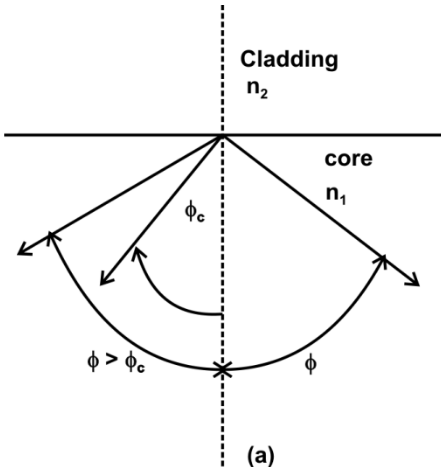


Fig. 14.2: Reflection from the inside of the core wall

6. If the angle of incidence is less than the critical angle, then the ray of light will pass through the wall (core and cladding interfacing wall) into the cladding region by refraction and it will be lost. This can be seen from figure 14.3.

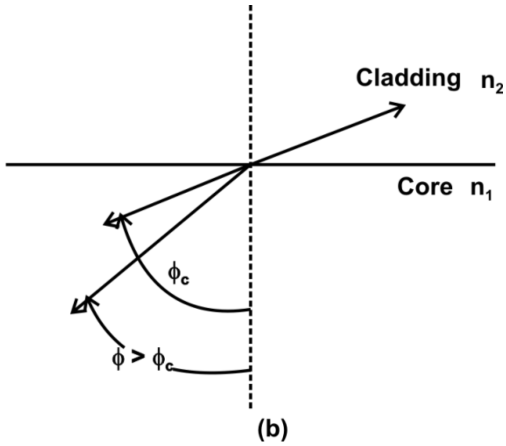


Fig. 14.3: Ray escaping through the core wall by refraction

14.4 TYPES OF FIBRES

1. Generally optical fibres are broadly categorized **into single mode and multimode.** In general, mode means the various paths light that can be taken in the fibre.
2. In the single mode fibres only one mode is allowed to propagate in the fibre. Only a single ray path is possible for single mode fibre.

3. Multimode means several paths are available. Hence the dispersion process caused due to differences in transmit times of different rays in multimode fibre would be completely absent.

14.4.1 Step index optical fibre

1. In step index optical fibre, the core has a uniform refractive index. The cladding is also of uniform refracting index.
2. If n_1 and n_2 are the refractive indices of the core and cladding respectively, such that ($n_1 < n_2$). This gives a step discontinuity in the refractive index profile also the refractive index changes in steps not gradually. Hence the name is given as step – index optical fibre.

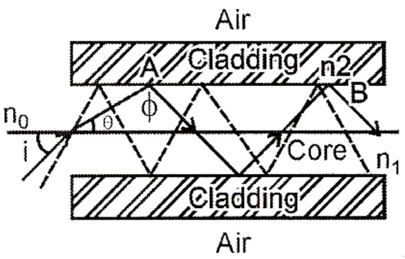


Fig. 14.4: Step – Index Optical Fibre (Propagation of light)

14.4.2 Graded Index Fibre

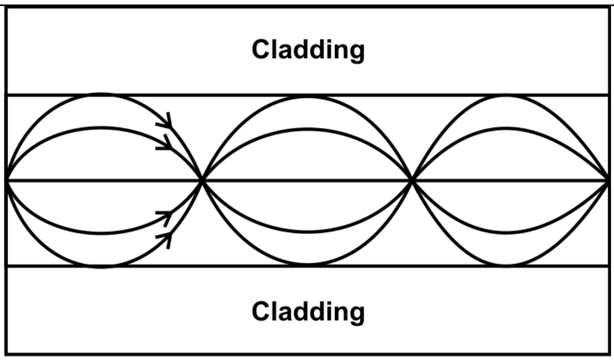


Fig. 14.5 (a) Graded index fibre

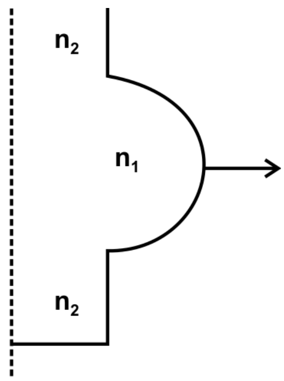


Fig. 14.5 (b) Refractive index profile

1. In the graded index fibre, the refractive index varies radially decreasing in a parabolic manner, from maximum value of n_1 at the centre of core to a constant value of n_2 at the core cladding interface as shown in figure 14.5 (b).
2. In the graded index fibre, light rays travel at different speeds in different parts of the fibre due to the refractive index variation throughout the fibre. The refractive index is lower near outer edge, this is the reason the rays near the outer edge travel faster than the centre of the core. But still all the rays arrive at the end of the fibre at proximately the same time. All rays take the same amount of time in traversing the fibre.

14.5 SIGNAL DISTORTION

Although optical fibres have great advantages over other communication lines, they also have some disadvantages. When optical signals are passed through optical fibres, they are sometimes lost in the fibres due to various reasons. The power loss in the optical fibres can be due to attenuation or dispersion.

14.5.1 Losses due to attenuation:

Losses due to attenuation take place inside the fibre cable. The attenuation loss causes the amplitude of the signal to get reduced. The signal on optical fibre attenuates due to the following mechanisms:-

1. Intrinsic loss in the fibre material.
2. Scattering due to micro irregularities inside the fibre.
3. Micro – bending losses due to micro – deformation of the fibre.
4. Bending or radiation losses on the fibre.

The first two types of losses are usually present in any fibre and next two depend on the environmental factors.

1. Material Loss

- The material loss is due to impurities present in glass used for making the fibres. Although various purification methods are used, some impurities like Fe, Ni, Al, etc. are present in the fibre material. These impurities ultimately reduce the signal greatly.

- Glass intrinsically is a good infra – red absorber. As we as we increase the wavelength, the infra – red loss increases rapidly.

- Minute quantities of water molecules trapped in the glass during manufacturing contribute OH^- ions into the material. These OH molecules contribute to the major absorption of light.

2. Scattering Loss

- The glass in the optical fibres is an amorphous (non – crystalline) solid that is formed by allowing the glass to cool from its molten state at high temperature until it freezes. While it is still plastic, the glass is drawn out into the fibre form. During this formation process sub – microscopic variation in the glass densities and doping impurities are formed into the glass creating reflecting and refracting facets which scatters the light passing through the glass creating losses.

- These losses due to scattering can be reduced by carefully manufacturing techniques but cannot be eliminated.

3. Micro – bending Losses

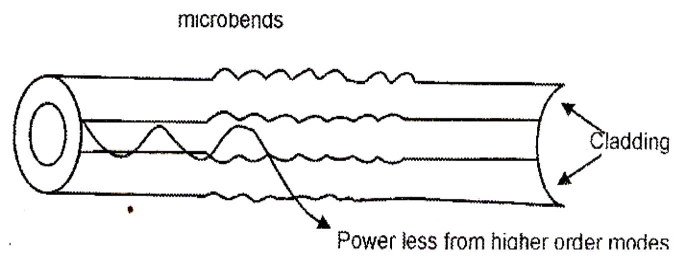


Fig. 14.6: Micro – bending in optical fibre

- In a fibre without micro – bends, the light is guided by total internal reflection (TIR) at the core – cladding boundary. The rays which are guided inside the fibres have incident angle greater than the critical angle at the core – cladding interface. In presence of micro – bends, however, the direction of the local normal to the core – cladding interface deviates and therefore the rays may not have an angle of incidence greater than the critical angle and consequently will be leaked out. Hence a part of the propagating optical energy therefore leaks out due to micro – bends. Depending upon the roughness of the surface though which the fibre passes, the micro – bending loss varies.

4. **Radiation or Bending Loss**
   - While laying the fibre, the fibre may undergo a slow bend. In micro – bend, the bending is on micron scale, whereas in slow bend, the bending is on the centimetre scale. One example of a slow bend is the formation of the optical fibre loop. Because of this bending, sometimes cracks may be developed in the core layer. The data (signal) will escape from such cracks.

14.6 DISPERSION

Theoretically, a pulse of light with a given width and amplitude transmitted into one end of the fibre would reach far end with its shape and width unchanged but only its amplitude reduced by losses.

There are different types of dispersions. This dispersion of pulse during transmission widens out the pulse which can lead to overlapping with adjacent pulses. Three separate dispersion mechanisms exist in a fibre. There are intermodal dispersion, material or chromatic dispersions and waveguide dispersions.

14.6.1 Intermodal dispersion

1. Each mode of a fibre optic has a different effective group velocity, even though the phase – velocity, in each ray path may be identical. This happens because the total path followed by the guided rays is zigzag in nature and has a different length of each mode.

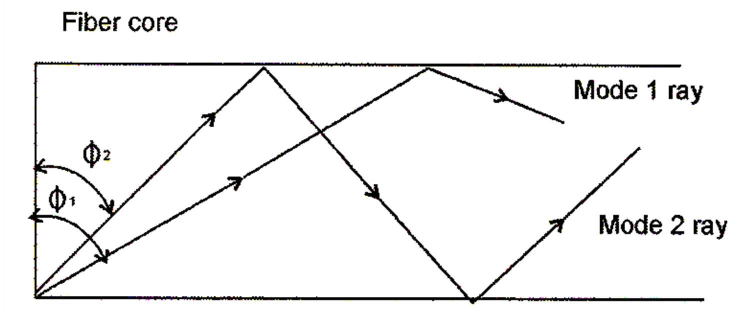


Fig. 14.7: Inter modal dispersion

2. From figure 14.7, the pulse transmitted in the fibre propagates in different paths depending upon the modes. The received pulse in the other end of the fibre is the addition of rays arriving at different time intervals. The received pulses are called the modal pulses. The received pulse therefore spreads out giving rise to dispersion.

3. In the figure 14.7, the mode 2 ray travels along the longer path than compared to mode 1. Hence, the mode 2 ray will reach the other end other end of the fibre with a small delay compared to mode 1 ray. The delay in the rays gives rise to dispersion as shown in figure 14.12.

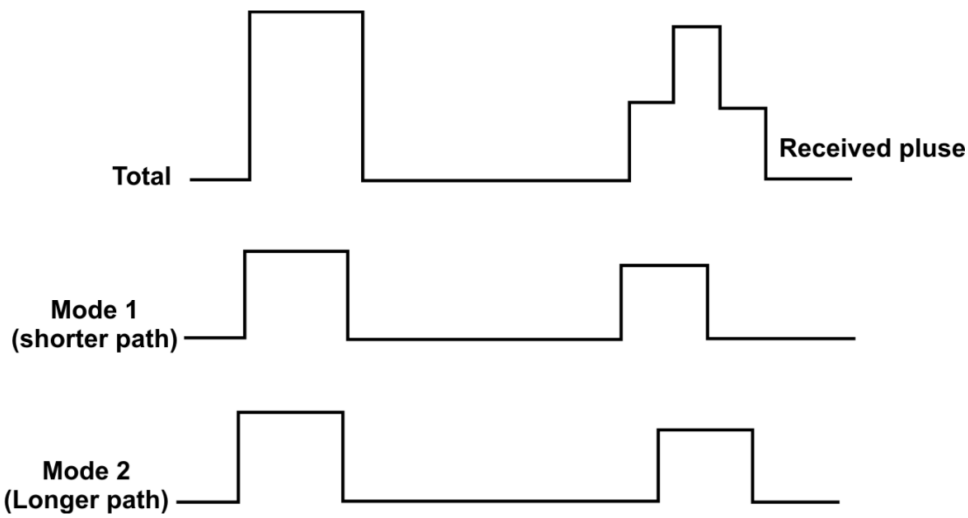


Fig. 14.8: Effect of Intermodal dispersion on a transmitted pulse

14.6.2 Material or Chromatic Dispersion

1. The refractive index of a core is not the same for the lights of different wavelengths but it varies across the spectrum. The practical light sources do not give pure monochromatic light but produce a spectrum distributed about a central wavelength λ_0 as shown in fig 14.9 with a spectrum bandwidth of 3dB. The light components of the pulse with shorter wavelength will experience more delay and the light components of same pulse with longer wavelengths will experience less delay. Therefore, the result is time dispersion at the receiver end of the fibre.

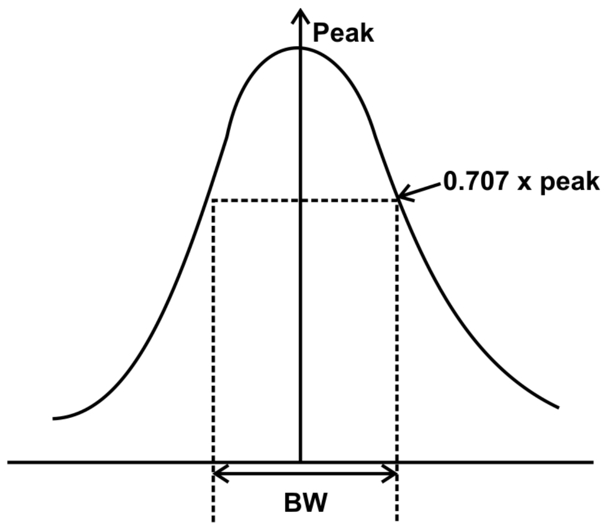


Fig. 14.9: Spectral content of light source

2. A narrow bandwidth source will produce less dispersion than a wideband source. The fig. 14.9 clearly shows that the material dispersion is proportional to the bandwidth.

14.6.3 Waveguide dispersion

- All practical light sources contain light components of different wavelengths.
- The group velocity for these wavelengths would be different for a specific mode. Hence, every wavelength has its own group velocity which will result in the dispersion of the incident light pulse known as the waveguide dispersion. T
- he waveguide dispersion can be reduced by reducing the bandwidth of the light source used.

14.7 OPTICAL SOURCES

The light sources used for fibre optics actually act as light transmitters which should meet the requirements if they are to be acceptable. They are as follows:

1. The light produced must be nearly monochromatic (single frequency) as possible.

2. The light source should have a high intensity output so that sufficient energy is transmitted to the fibre to overcome losses due to transmission.

3. The light devices should be small and should be able to easily coupled to the fibres so that excessive coupling looses are avoided.

4. They should be relatively inexpensive to manufacture.

5. The two most widely used light sources are the Light Emitting Diodes (LED) and semiconductor lasers.

14.7.1 Light Emitting Diodes (LED)

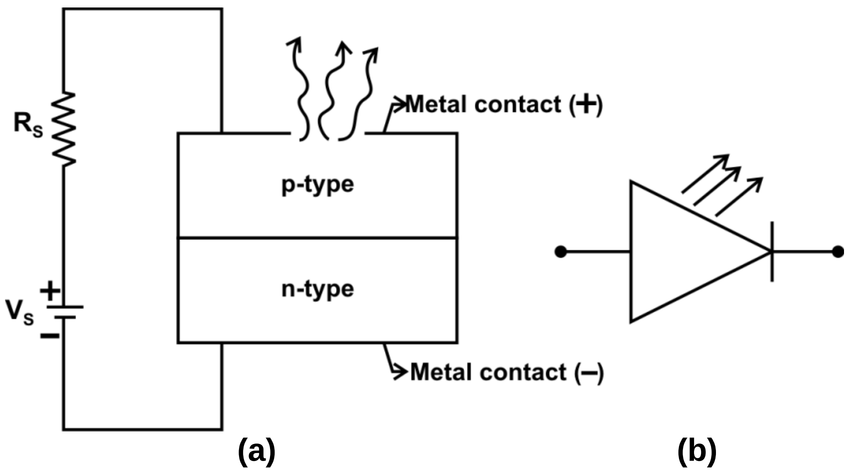


Fig. 14.10 (a) LED construction
Fig. 14.10 (b) Symbol of LED

1. A p – n junction diode which emits light when it is in forward bias is called the light emitting diode. The principle of working of an LED is **"in any p – n junction diode when electrons are combined with holes energy is released**."

2. Also for an LED the spontaneous emission principle is also applicable. An LED is like a normal semiconductor diode having a P – type material with large number of holes as majority charge carriers and n – type materials with electrons in large number as majority charge carriers.

3. When a p – n junction is forward biased, the electrons from n – side recombine with the holes on P – side. The free electrons and holes are in conduction band and valance band respectively. Therefore energy level of free electron is higher than that of holes. Thus, when electrons recombine with the holes it falls from the conducting band to the valance band by making a transaction from higher energy band to lower energy band.

4. This difference in energy levels of electrons is released in the form of visible light by LED.

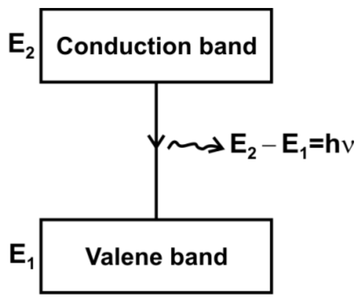


Fig. 14.11: Electron transition

Advantages of LED:

- Its power consumption is less.
- It operates at low voltages (approximately 1.5V).
- It requires a small current.
- It is cheap and easy to handle.
- It emits light of different colours (this depends upon the type of material used for doping).
- Its life is long i.e. thousands of hours.
- It can be used for various purposes i.e. in optical communications, burglar alarms, digital watches, etc.
- It is tough and can't easily break.

14.7.2 Semiconductor Laser Diodes

1. The term LASER is an acronym for **Light Amplification by Stimulated Emission of Radiation**. The laser action can be obtained by using many different gases such as carbon dioxide, Helium, Neon or materials like Ruby. The semiconductor laser uses the solid semiconductor as the lasing material.

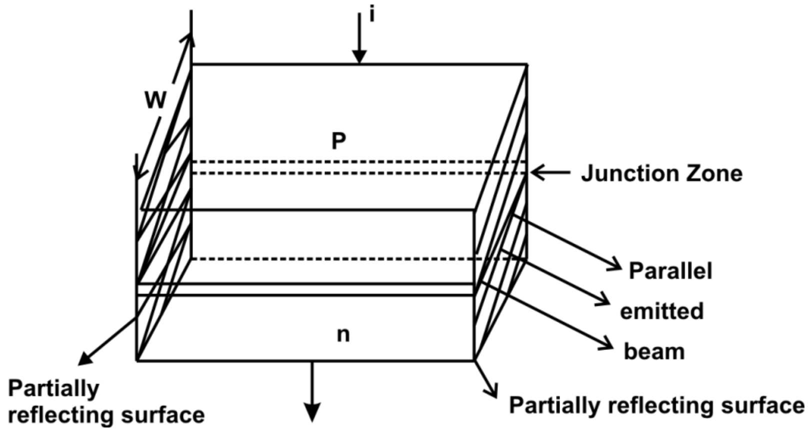


Fig. 14.12: Semiconductor Laser Diode

2. In the semiconductor LASER diode when the current is passed through the junction diode, light is emitted (as discussed in section of LED). As a critical current level is reached the minority charge carriers on both sides of the junctions' increases. The increase in the number of recombination of the electrons and holes and also the number of photons emitted. Therefore, the

density of photons increases to such a level that they collide with the exited minority carriers giving rise to a process called the stimulation emission in which two photons are liberated instead of one. Both the photons have same frequency and energy level. This laser action is further enhanced by placing the reflection surfaces on either sides of the junction as shown in figure 14.12.

One advantage of the semiconductor Laser is that the laser beam diameter is very small. This property makes them useful in the optic system.

14.8 OPTICAL DETECTORS

Different types of photosensitive devices can be used as detectors for use with fibre optics. Some of the devices include silicon photodiodes, phototransistors and photo resistors. But all of them do not have high sensitivity and speed of response. In the following section, we will study a detector which has good sensitivity and speed of response.

p – i – n photodiode

The construction of a p – i – n diode is as shown in figure 14.13.

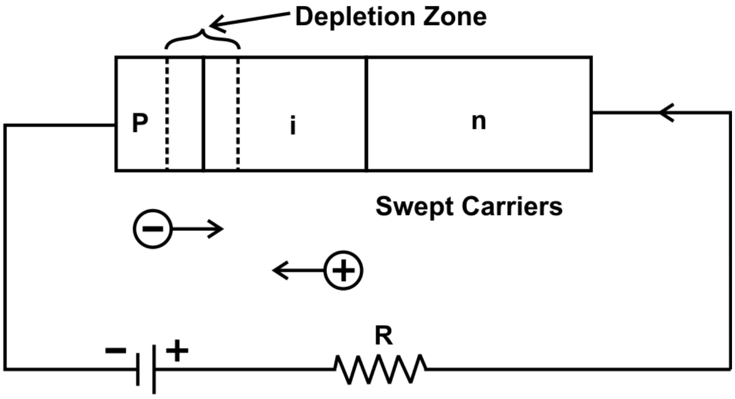


Fig. 14.13: p – i – n diode

The sensitivity of a pn photodiode can be improved by including a lightly doped (or almost intrinsic) n layer between the junction and more heavily doped n – contact region from the p – i – n diode. The intrinsic layer shown in figure 14.17 is made thick enough so that most of the photons that pass through the junction without ionizing are absorbed within this layer. The reverse bias electric field applied extends deep into this region and any holes produced by the photons are swept across the junction adding the photocurrent. Hence the incident photons are used more efficiently in pin photodiode than in the case of pn photodiode resulting in larger photocurrent and much higher sensitivity.

14.9 THE OPTICAL FIBRE COMMUNICATION LINK

The figure 14.14 shows the basic optical communication system. The major components of the optical communication system are optical transmitter, optical fibre and optical receiver.

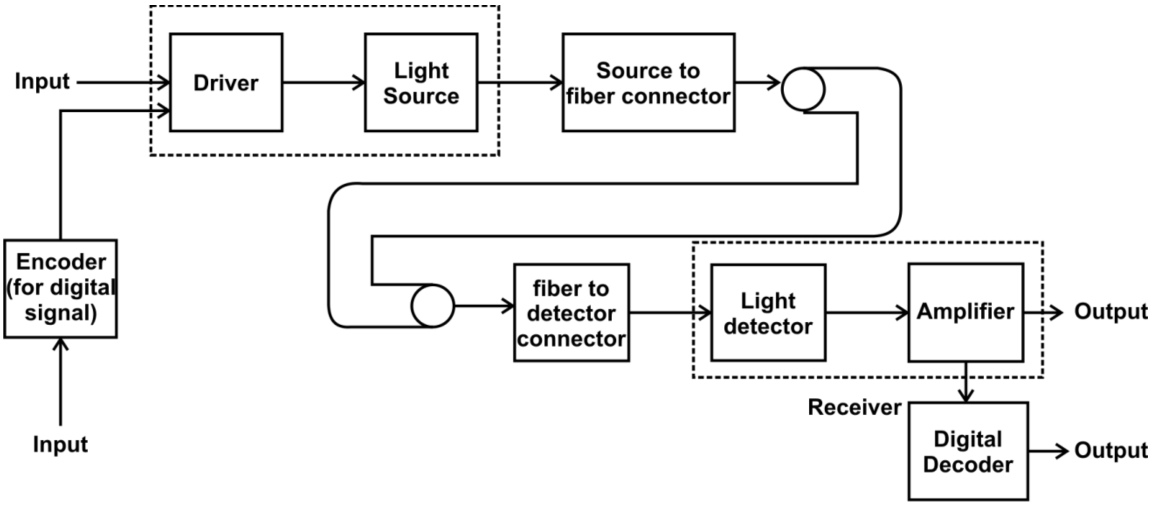


Figure 14.14: Fibre Optic Communication System

The transmitter basically consists of a laser diode or a light emitting diode (LED). The optical fibres essentially transmit the light waves. The detectors are usually positive intrinsic (PIN) diode. Fibre optic system basically converts electric signal to an infra – red light signal. This signal is transmitted by the transmitter through an optical fibre. At the end of the fibre, it is reconverted back into an electric signal. In fibre optic system, the encoder and decoder are necessary only for digital signals. Encoder and Decoder are electrical circuit in which information is encoded and decoded respectively. The information is encoded and decoded into binary sequence of zeros and ones which is then transmitted through the Optical Communication System.

14.10 UNIT END EXERCISE::-

1. Write a note on optical fibre cable.
2. Explain the propagation of light in optical fibre. Draw the necessary diagram.
3. Write short notes on the following:
    a. Step index optical fibre.
    b. Graded index optical fibre.
4. Describe the structure of a step – index optical fibre. Explain the propagation of light through it.
5. Write a note on losses in optical fibre due to attenuation.
6. Discuss in brief the losses in optical fibres due to dispersion.

7. Write a short note on photo detectors.
8. Write a short nor on optical sources.
9. Write short notes on the following:
    a. Light emitting diode (LED).
    b. Semiconductor laser diode.
10. Explain the fibre optic communication system.
11. Write a note on p – i – n photodiode.

❖❖❖❖