Communication System

- Anit Kumar

AM/FM Transmitters

Unit-I

Transmitters on the basis of modulation

- AM Transmitter
- FM Transmitter
- PM Transmitter

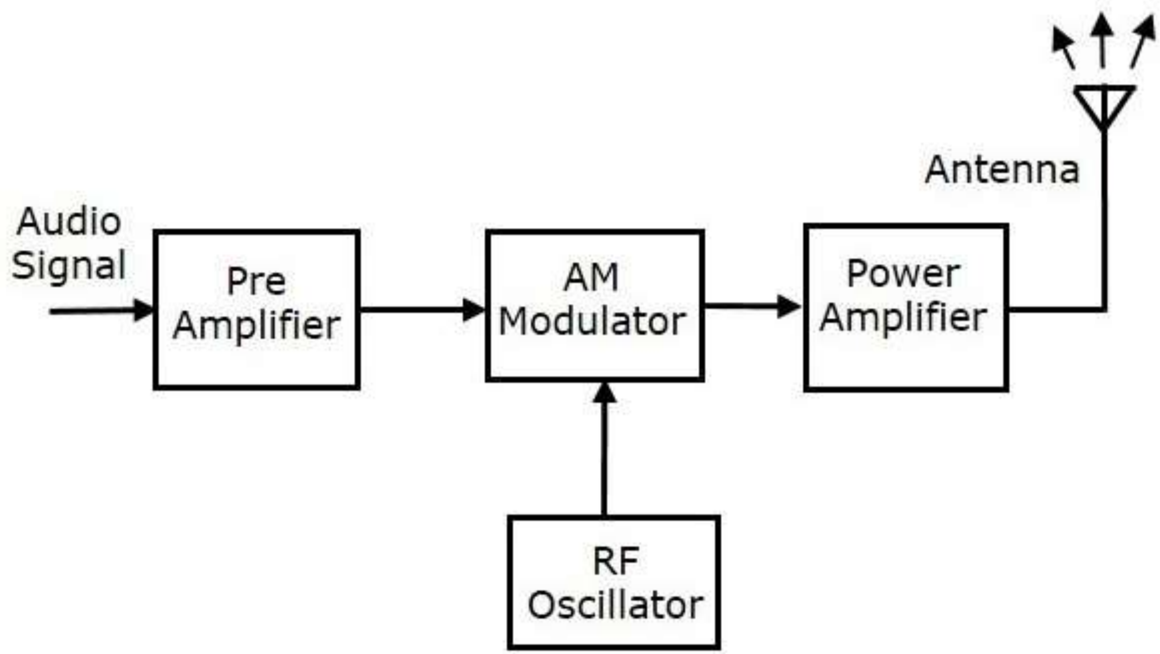
Transmitters on the basis of service

- Radio Transmitter
- TV Transmitter
- Radar Transmitter
- Communication Transmitter
- Etc.

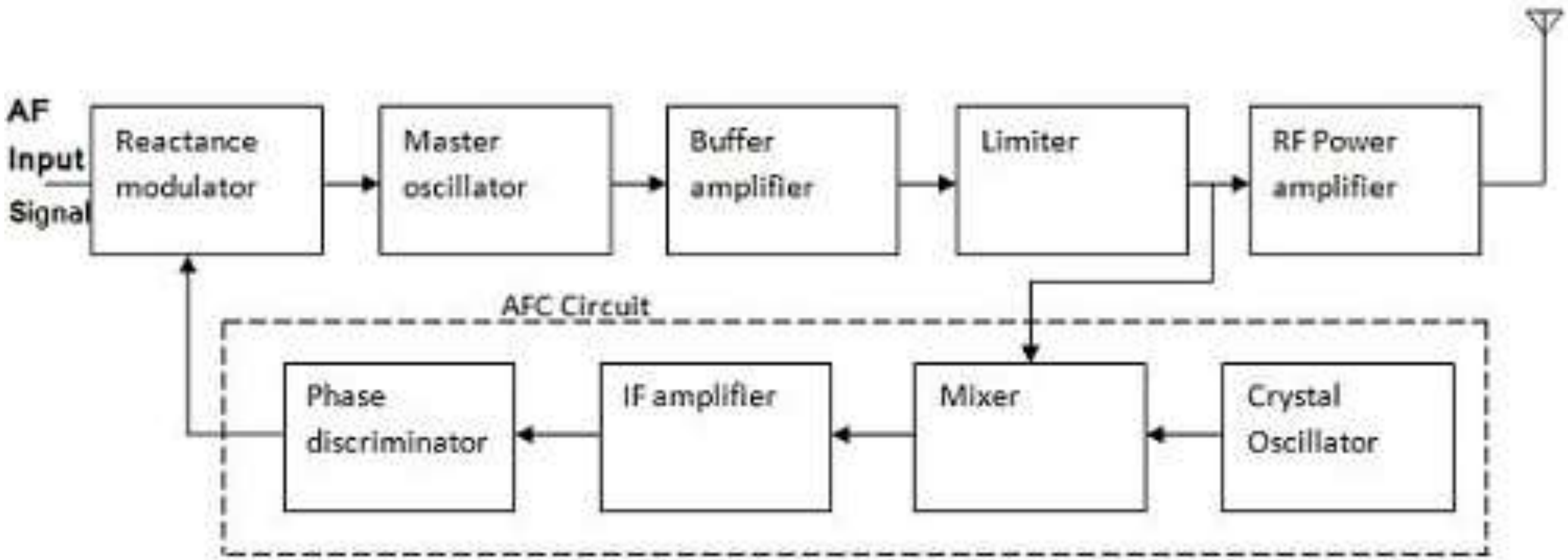
Transmitters on the basis of frequency

- MW Transmitter
- SW transmitter
- VHF & UHF Transmitter
- Microwave Transmitter

AM Transmitter

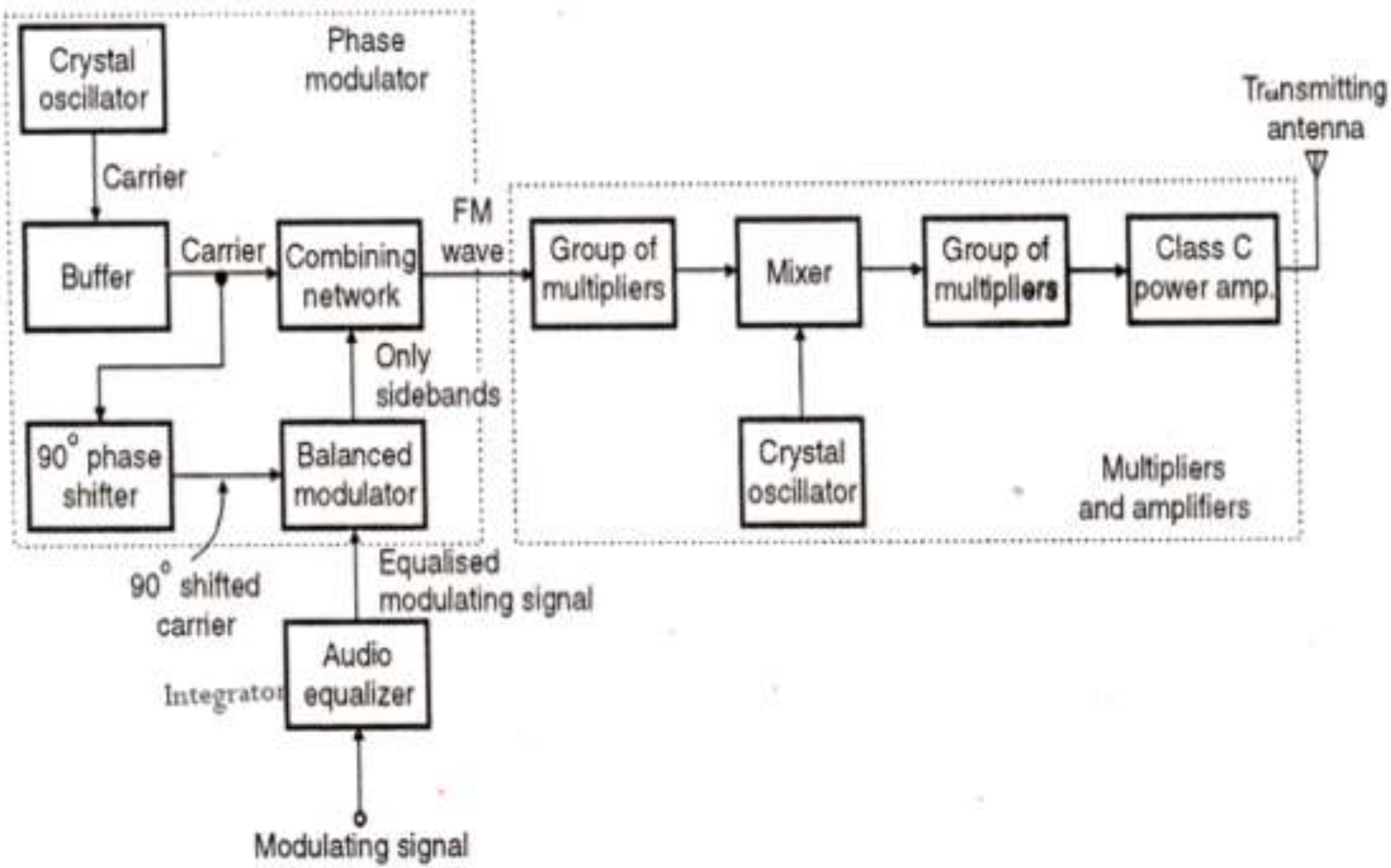


Reactance Transmitter



FM Transmitter Block Diagram

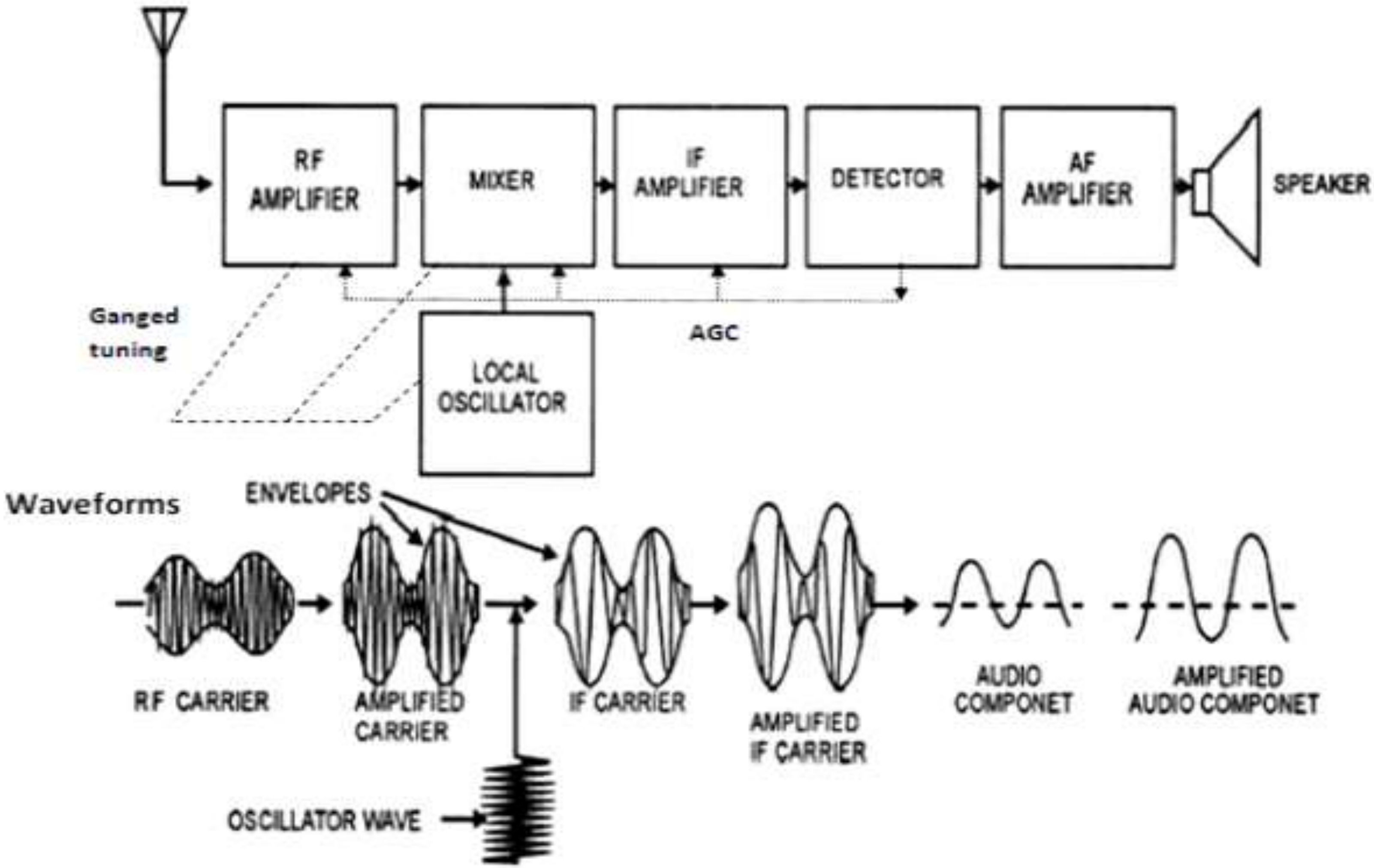
Armstrong Transmitter



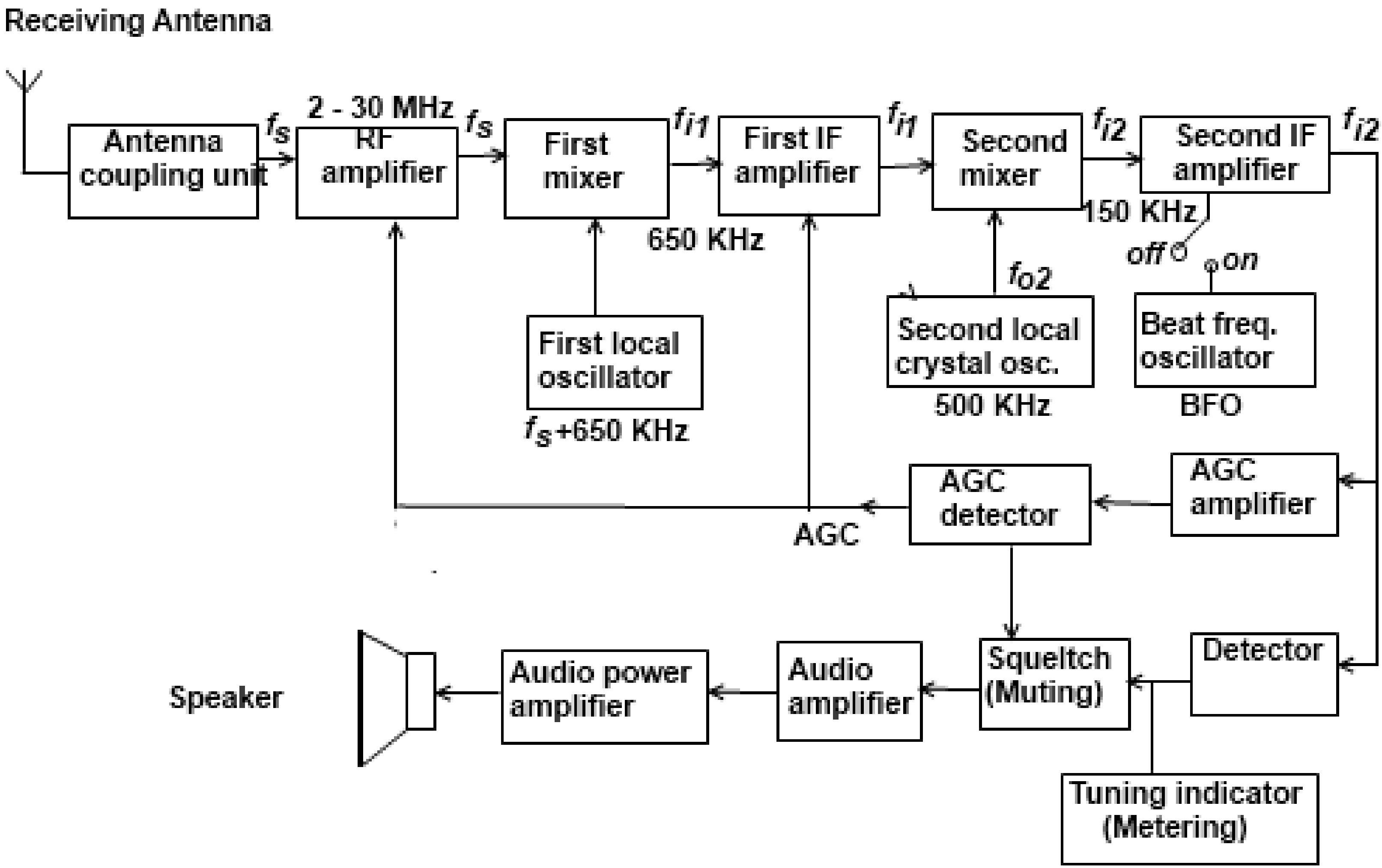
AM/FM Radio Receivers

Unit-2

Super heterodyne Receiver



Communication Receiver



Important Definitions

<u>Sensitivity:-</u>

The minimum voltage required at the receiver end which can produce standard output.

<u>Selectivity:-</u>

Ability of the receiver to reject the unwanted signal and accept the desired signal.

Important Definitions

- <u>Fidelity:-</u>

Ability of the receiver to produce replica of the transmitted signal.

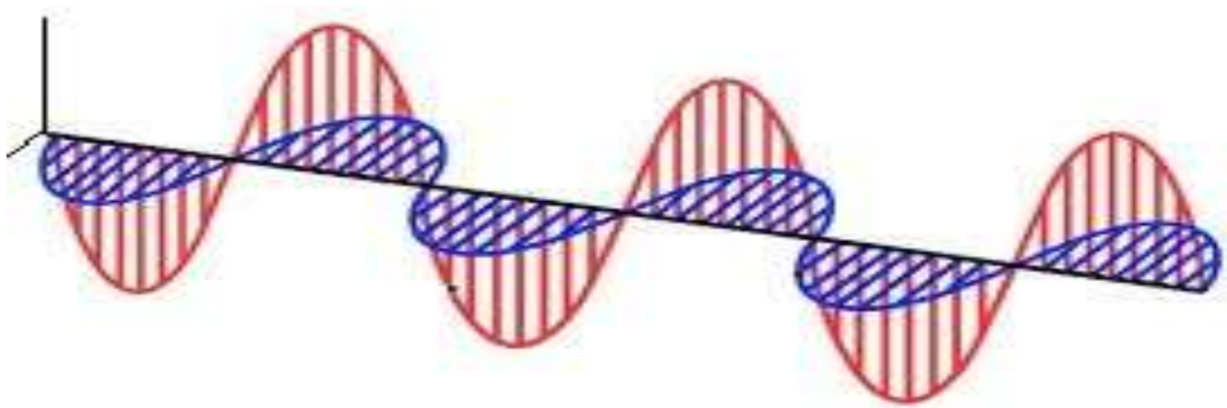
<u>S/N Ratio:-</u>

Ratio of signal power to the noise power

Antennas

Unit-3

EM Wave



Frequency Spectrm

Frequency	Class	
3Hz	ELF	Extremely Low Frequency
30Hz	SLF	Super Low Frequency
300Hz	ULF	Ultra Low Frequency
3KHz	VLF	Very Low Frequency
30KHz	LF	Low Frequency
300KHz	MF	Medium Frequency
3MHz	HF	High Frequency
30MHz	VHF	Very High Frequency
300MHz	UHF	Ultra High Frequency
3GHz	SHF	Super High Frequency
30GHz	EHF	Extremely High Frequency

Polarization of Wave:-

Polarization tells the orientation of oscillation.

Point Source:-

An antenna/radiator of a size of point.

Antenna Gain:-

Ability of antenna to radiate in a direction compared to reference antenna.

Directivity:-

Maximum radiation intensity of a test antenna divided by average radiation intensity of test antenna.

Antenna Aperture:-

Area whose orientation is normal to the direction from where EM wave reaches antenna.

Radiation pattern :-

The three dimensional sketch drawn to represent the radiation of an antenna.

Beam width :-

Angular separation in which the magnitude of the radiation pattern decrease by 50% (or -3 dB) from the peak of the main beam.

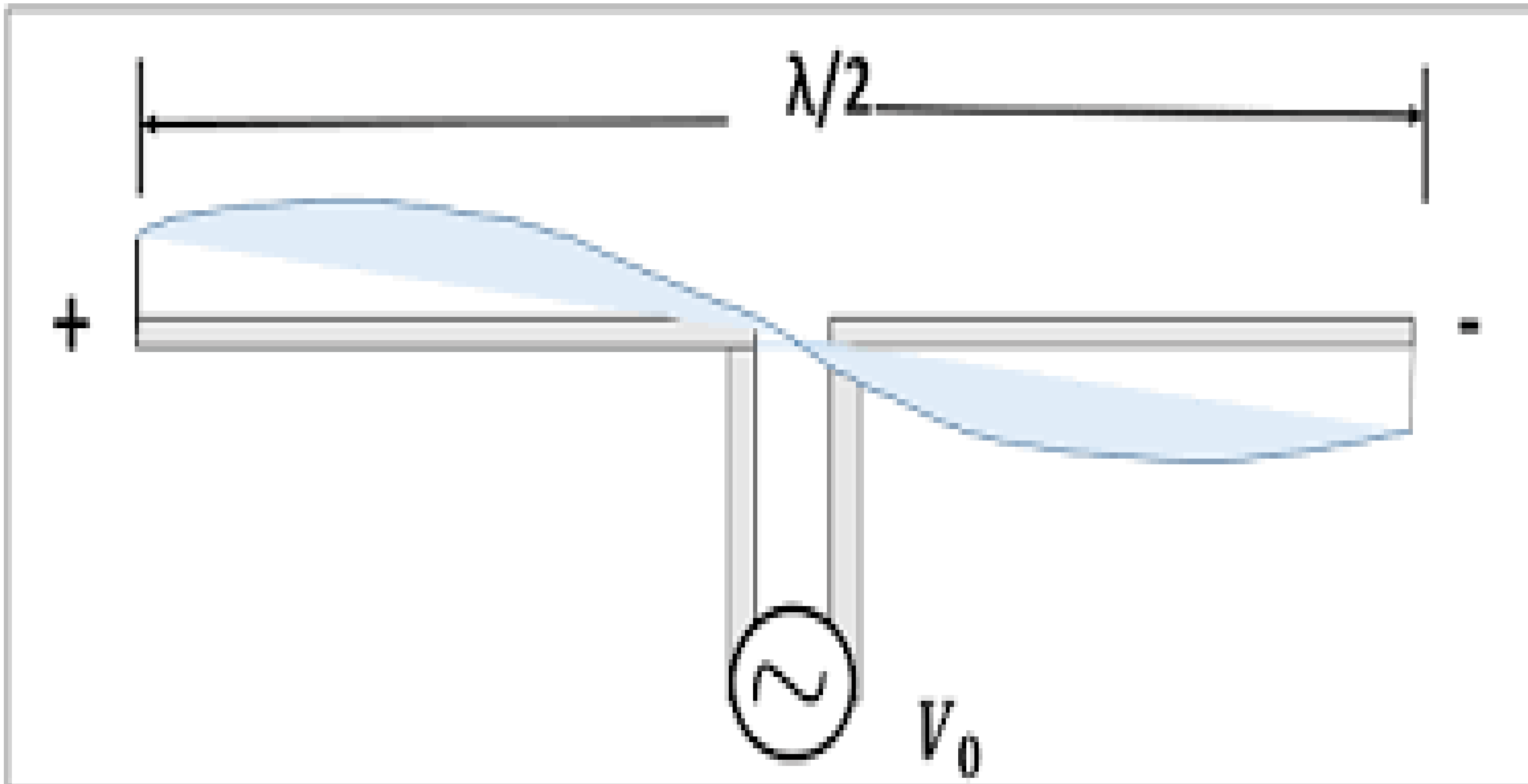
Radiation resistance :-

Radiated power in watts divided by the square of the effective current in amperes at the point of power supply

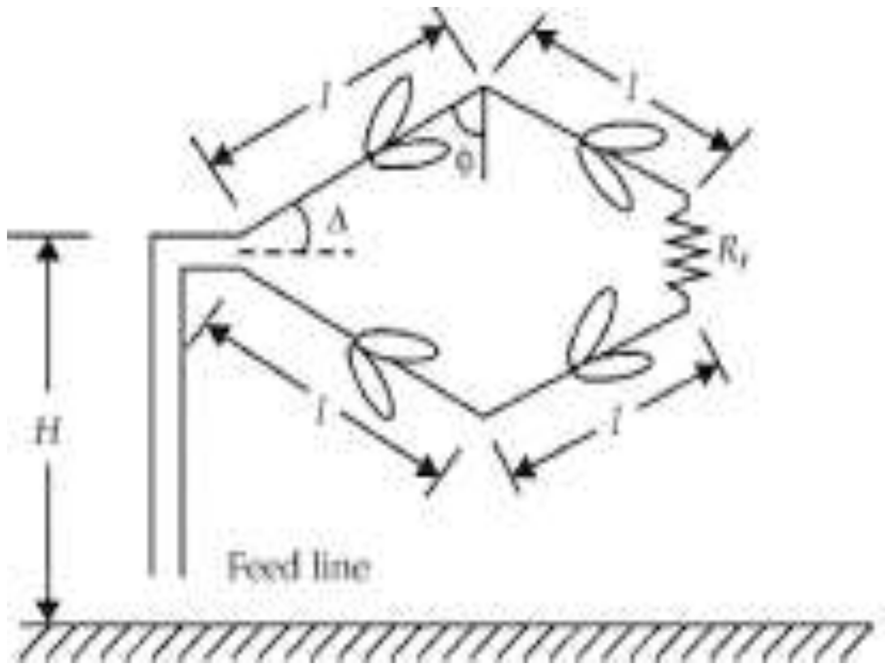
Types of antenna

1. Resonant
2. Non Resonant

Half Wave Dipole Antenna

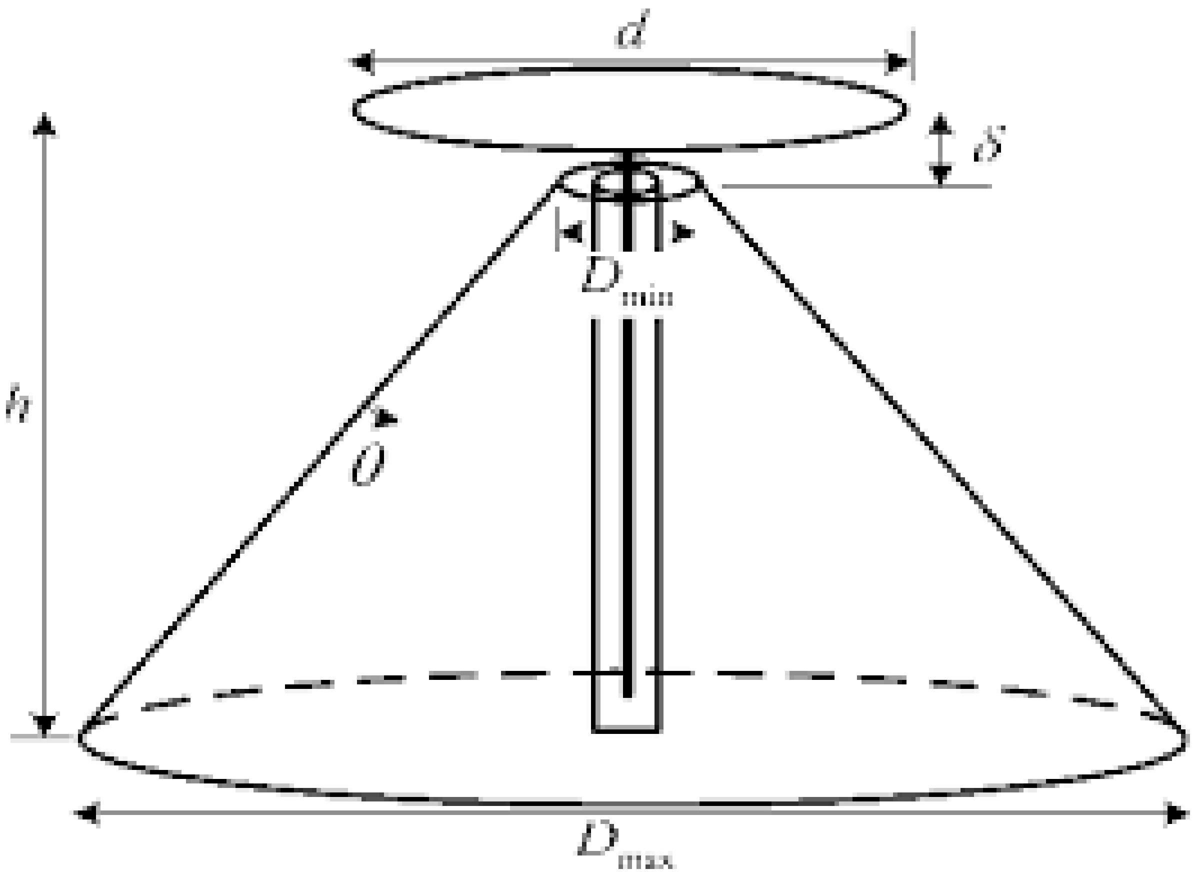


Rohambic Antenna

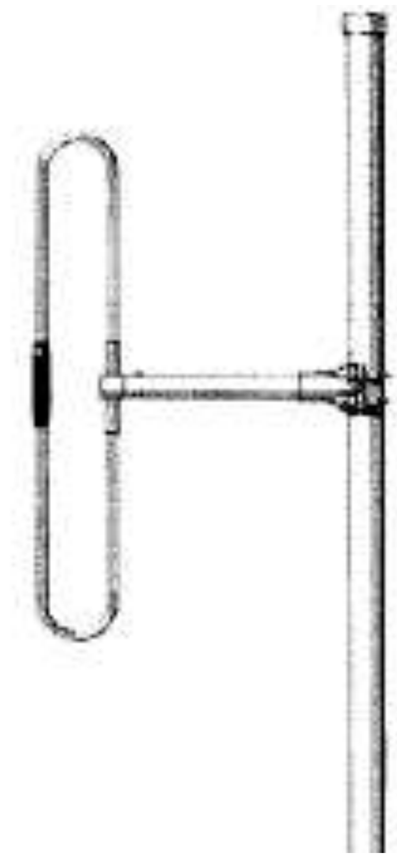



Radiation pattern

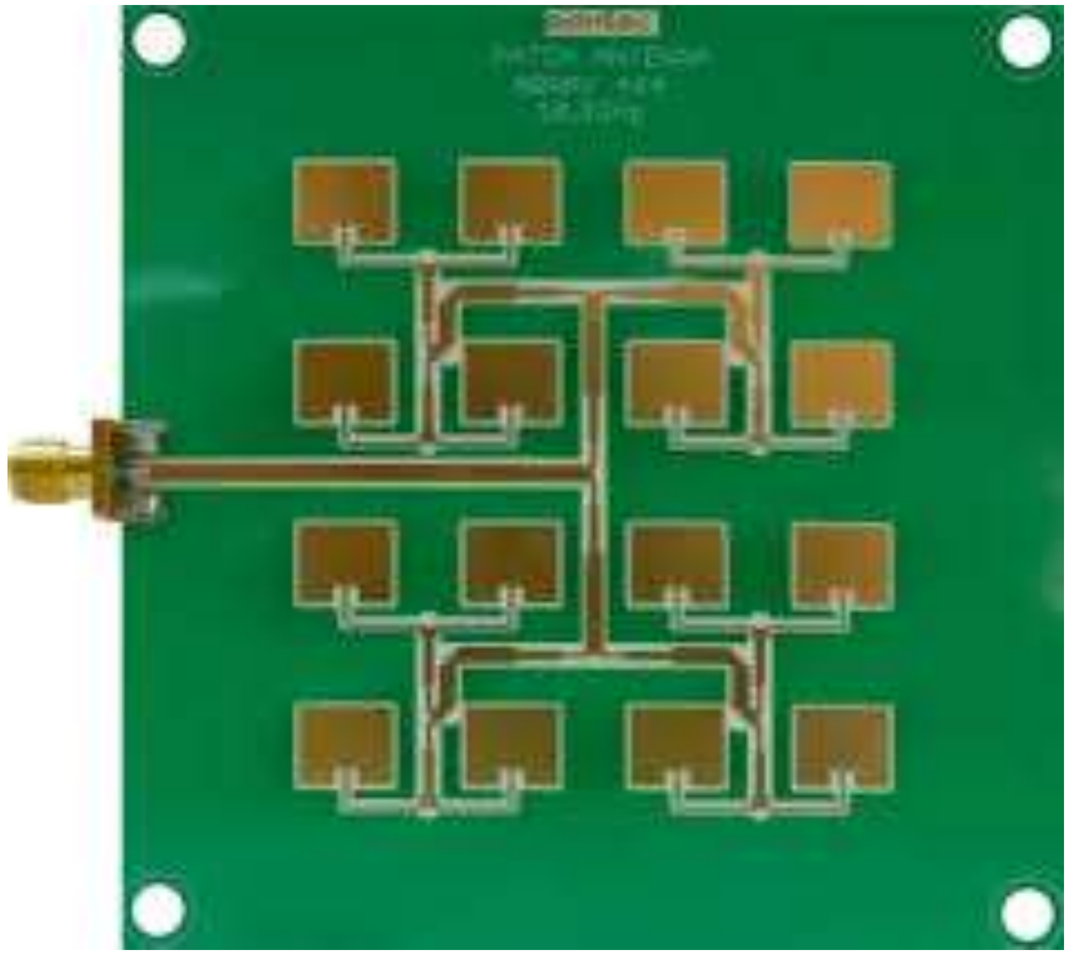
Discone Antenna



Folded Dipole

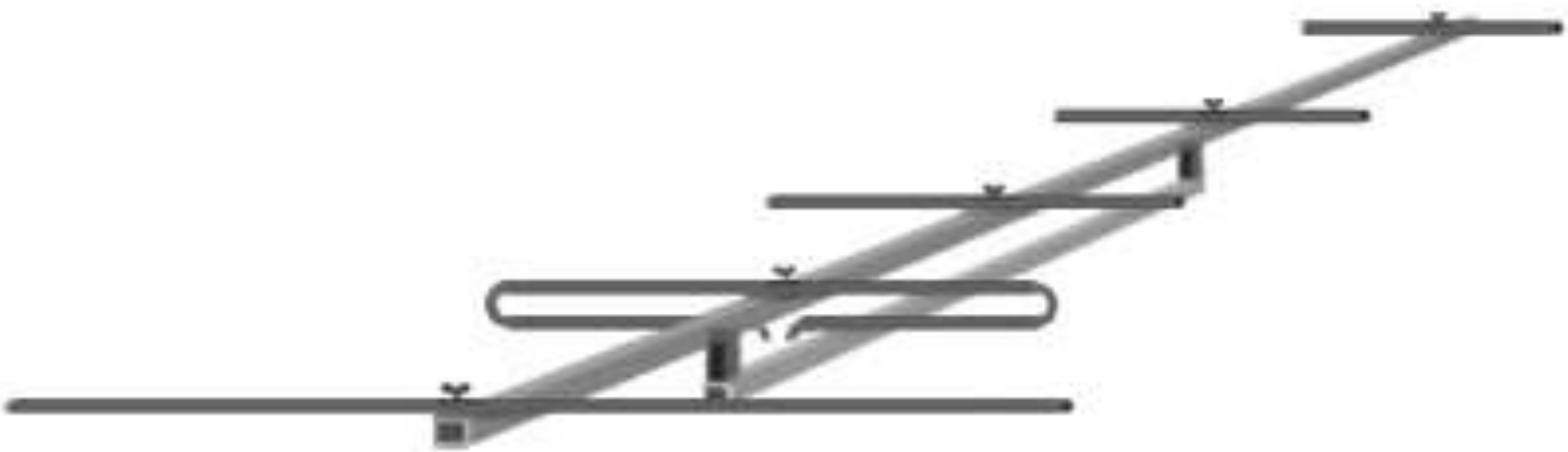


Patch Antenna



Ferrite Rod

Yagi Antenna



Dish Antenna

Propagation

Unit-4

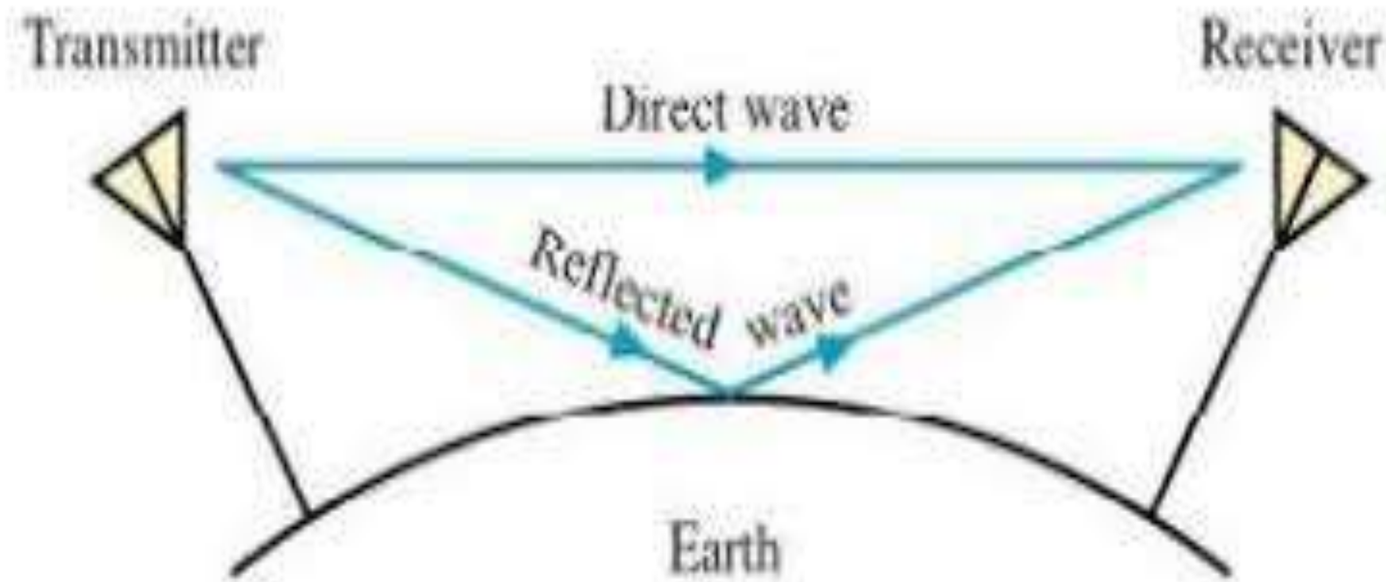
Propagation

The mechanism by which a wave travels is called propagation.

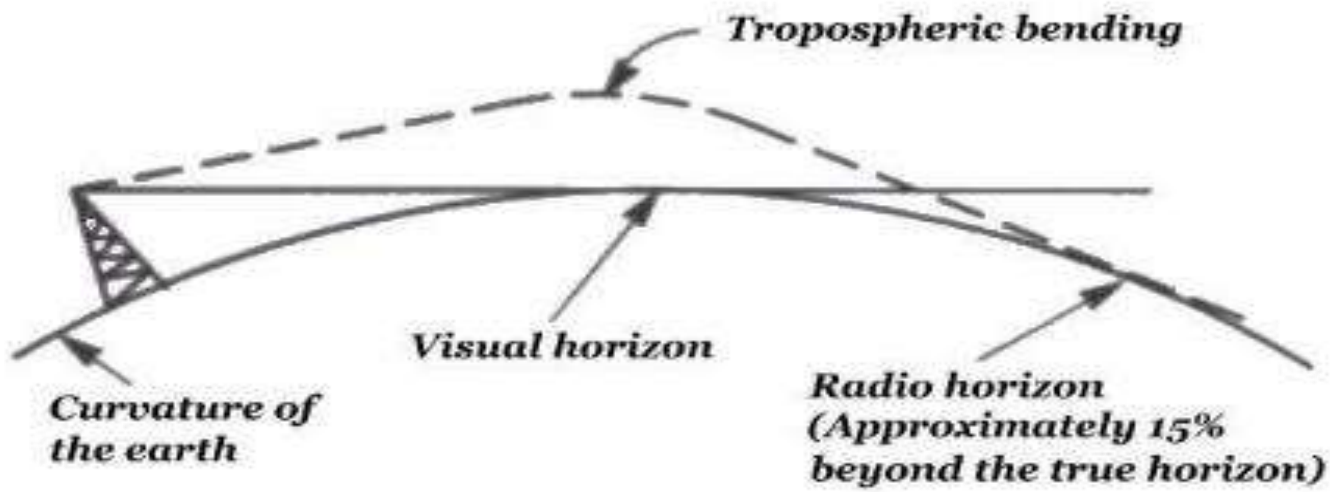
Space Wave Propagation

- It is the type of propagation in which a wave reaches the receiving antenna directly and/or after reflection from the earth.
- Frequencies used in this type of propagation are VHF(30MHz -300MHz) UHF(300MHz -3000MHz) and Microwave(More than3000MHz)

Space Wave Propagation



Space Wave Propagation



Space Wave Propagation

- Earths curvature affects the SWP it creates a path difference and hence phase difference between reflected and direct wave.

Space Wave Propagation

- Maximum distance between Tx and Rx Antenna on earth depends on antenna height and is limited to horizon.
- It may be up to the 15% more than the optical horizon as shown in figure in above slide.
- Distance between transmitter and receiver is given by the below equation.

Space Wave Propagation

$$d = \sqrt{2ht} + \sqrt{2hr}$$

d –distance

ht –height of transmitting antenna

ht –height of receiving antenna

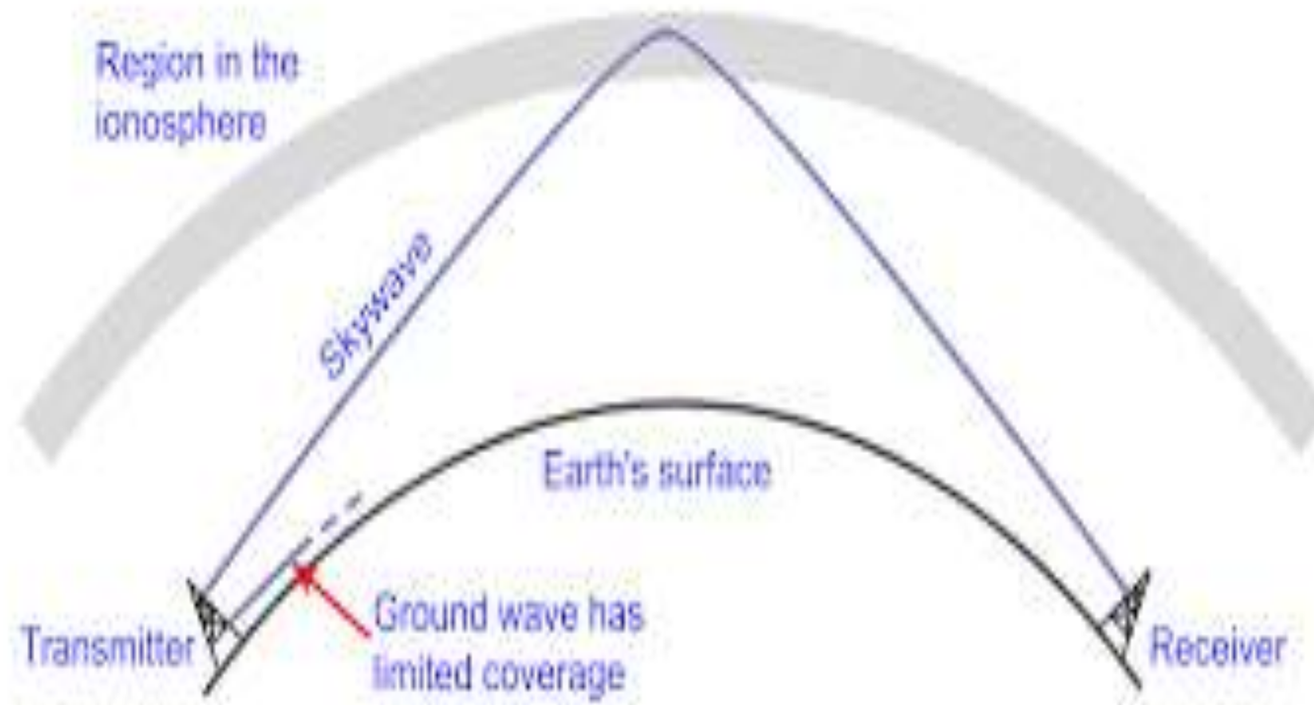
Space Wave Propagation

- Fading is also an important factor in SWP due to multipath fading and changes in atmospheric conditions field strength changes.

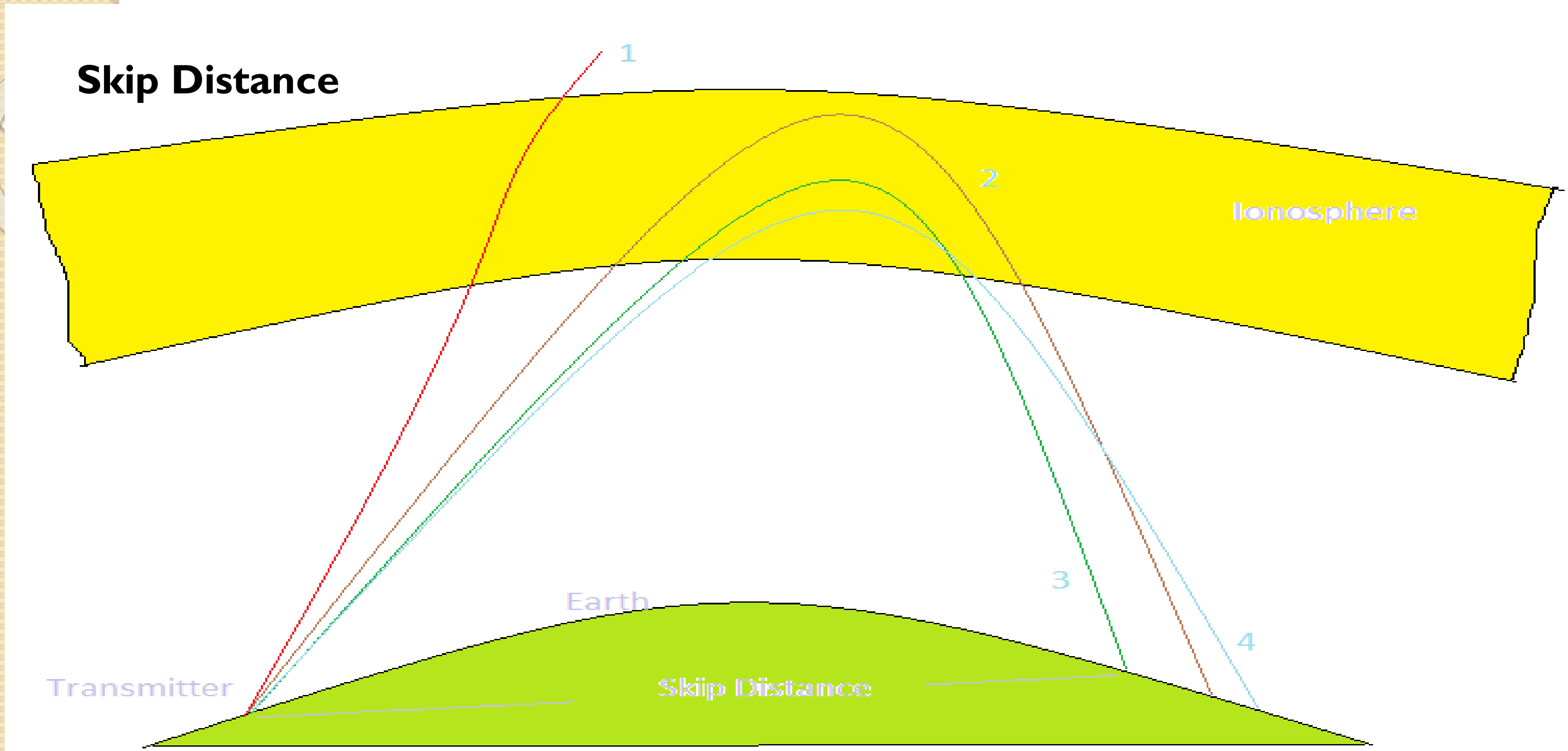
Sky Wave Propagation

- In this type of propagation wave is directed towards the sky and reflected or refracted by the ionosphere(top most layer of earth's atmosphere).
- For long distance communication on earth we use SWP.

Sky Wave Propagation



Sky Wave Propagation



In above figure there are four waves of same frequency transmitted from the same point at different angles.

First wave passes the ionosphere and doesn't come back.(RED)

When we reduce the angle wave 2 comes back.(ORANGE)

When we reduce the angle more wave 3 comes back at shorter distance.(GREEN)

When we reduce the angle more wave 4 comes back at the distance more than wave 3.(BLUE)

Sky Wave Propagation

- Skip Distance

It is the minimum distance between transmitter and the point where wave comes back after reflection/refraction from ionosphere.

Skip Distance Depends on

- 1.Frequency of wave
- 2.Critical frequency of ionosphere

Sky Wave Propagation

- Maximum Usable Frequency(MUF)

It is the highest frequency that can be used for transmission between two points via reflection from the ionosphere at a specified time, independent of transmitter power.

Sky Wave Propagation

- <u>Critical Frequency</u>:-

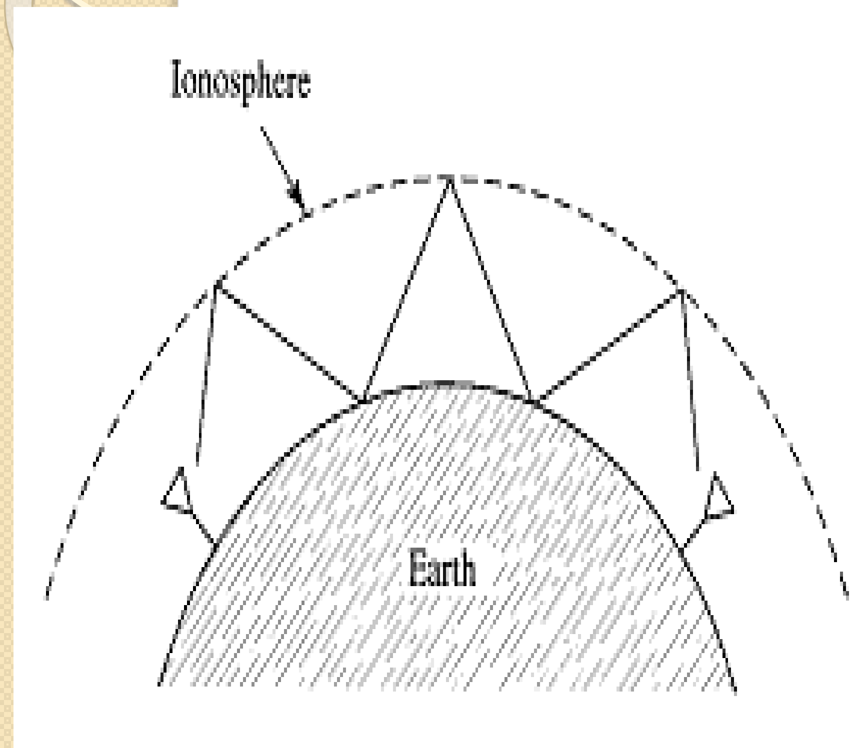
The highest frequency which can be reflected back from ionosphere at vertical incidence is called critical frequency.

Sky Wave Propagation

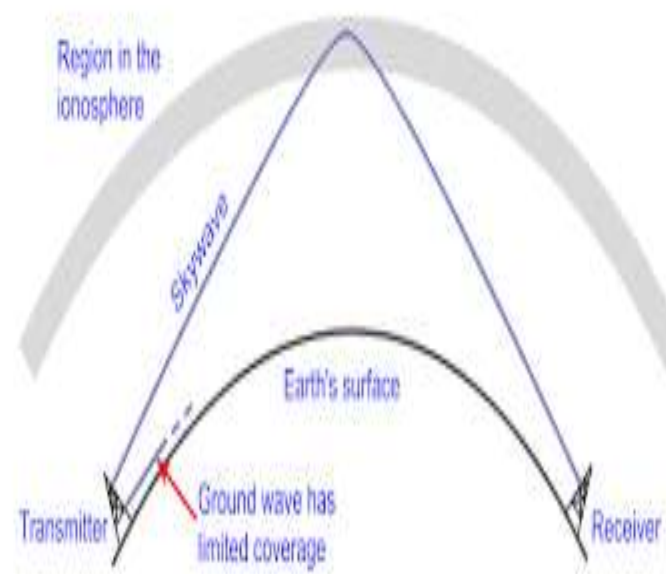
- <u>Virtual Height</u>

The apparent height of a layer in the ionosphere, determined from the time required for a radio pulse to travel to the layer and return, assuming that the pulse propagates at speed of light

Sky Wave Propagation



Multiple Hop



Single Hop

Sky Wave Propagation

- When the distance between the transmitter and receiver is so large that can not be achieved using single hop in that case we use multiple hop sky wave propagation.

Sky Wave Propagation

- **Layers in the Ionosphere**
- There are four ionized layers or regions in the ionosphere. They are called the D, E, F1 and F2 layers. All four layers are present during daytime. At night, the F1 and F2 layers thin out and tend to merge into one F layer. The D and E layers disappear at night. These layers have a lower degree of ionization. After the Sun sets, the intensity of the ultra-violet radiation decreases, recombination of the ionized particles occurs and the lower ionized layers will dissolve. Just before sunrise, ionization is at its lowest point.

Sky Wave Propagation

- **D-Layer**
- The D-layer, which is present during the day only, has the least ionization and therefore has the least effect on radio propagation. However, it almost completely absorbs radio waves of lower frequencies (up to 10 MHz). The ionization is proportional to the angle of elevation of the Sun and is the greatest at noon.

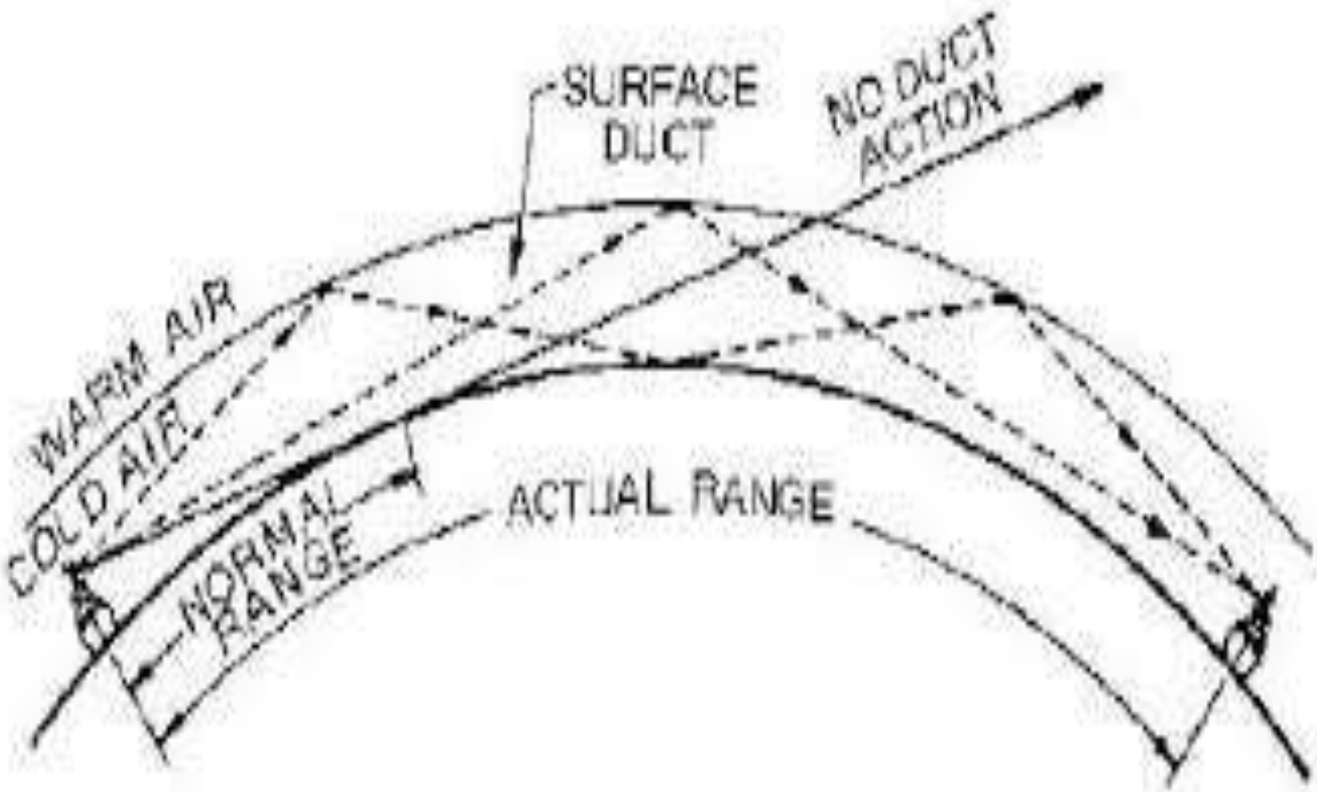
Sky Wave Propagation

- **E-Layer**
- The *Kennelly-Heaviside layer* is a layer of ionised gas occurring between roughly 90-150 km. It reflects medium-frequency radio waves (up to about 20 MHz), and because of this reflection radio waves can be propagated beyond the horizon.
  Propagation is affected by time of day. During the daytime the solar wind presses this layer closer to the Earth, thereby limiting how far it can reflect radio waves. On the night side of the Earth, the solar wind drags the ionosphere further away, thereby greatly increasing the range which radio waves can travel by reflection, called **sky wave**. The extent of the effect is further influenced by the season, and the amount of sunspot activity.
  This is the lowest useful ionized layer, and normally dissolves after sunset.

Sky Wave Propagation

- **F-Layer**
- The F region contains ionized gases at a height of around 150-800 km above sea level. It has the highest concentration of free electrons and ions anywhere in the atmosphere.
- During daytime when solar radiation is much higher than during the night time, the F layer splits into two sub-layers: the F1 and F2 layer with average virtual heights of 200 and 500 Km.
- After sunset these two layers merge again into a single F region.

Duct Propagation



Duct Propagation

- In duct propagation due to density variation of atmosphere/air formation of a duct takes place and signal face less attenuation than it would face in the absence of duct.
- Such type of propagation is not stable for a long time because of variation of temperature and density of atmosphere.

Satellite Communication

Unit-5

Satellite Communication

Satellite is any object which revolves around a planet.

There are two types of satelllite.

1. Natural satellite
2. Artificial(Man Made) Satellite

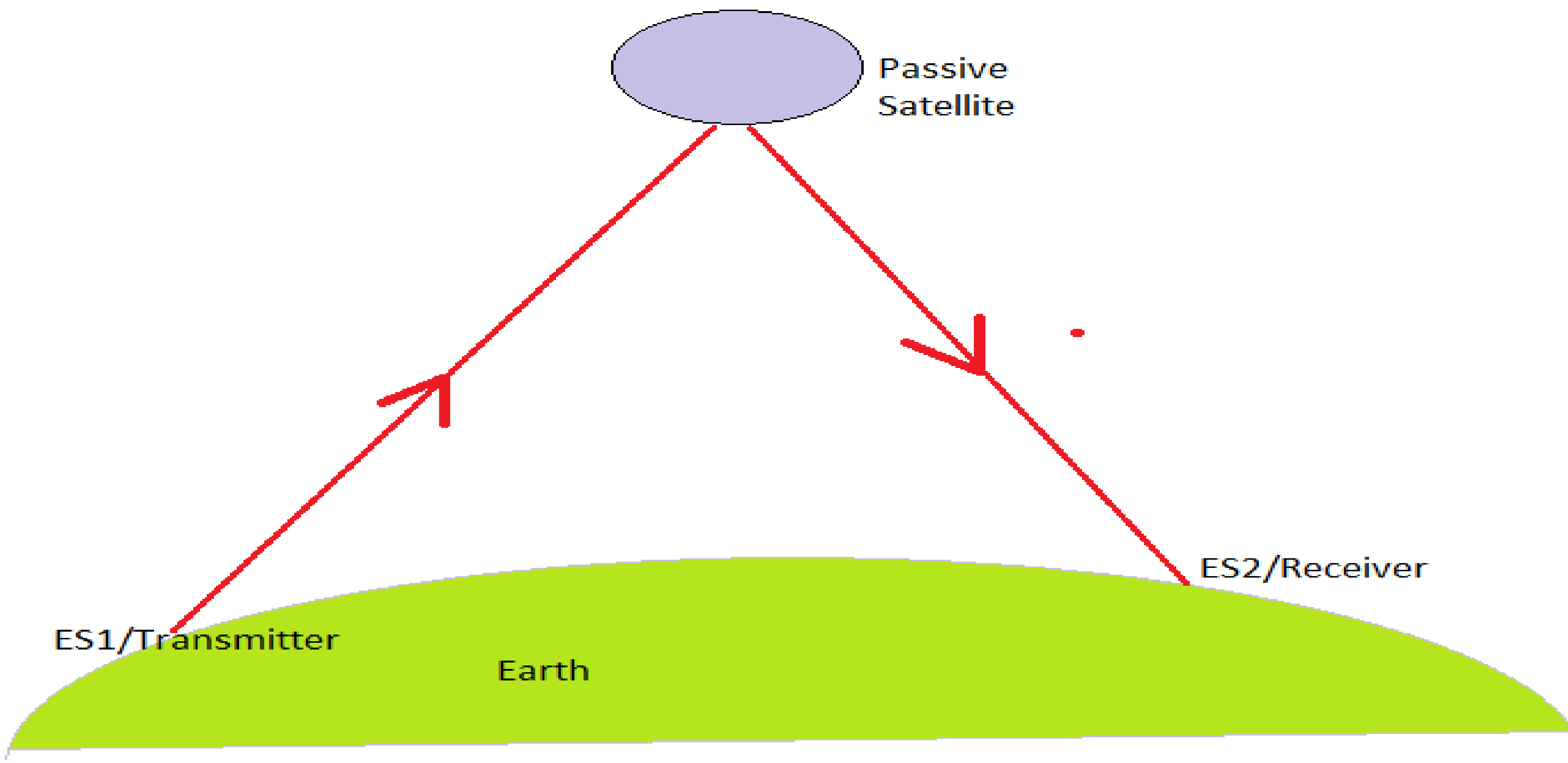
- Artificial Satellite:-

Man made satellites are called artificial satellites. These are of two types

1. Passive Satellites
2. Active Satellites

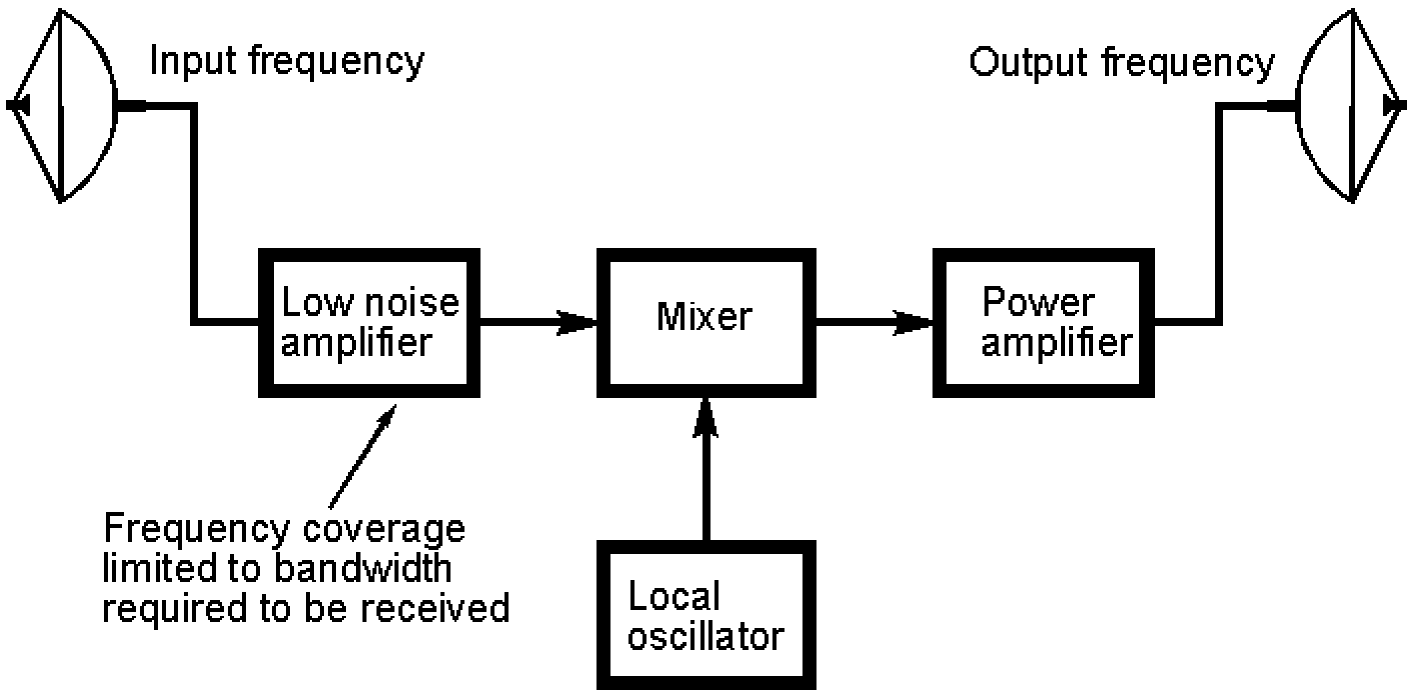
Passive Satellites

Passive Satellites are those which don't have any electronic device and simply used to reflect the signal.

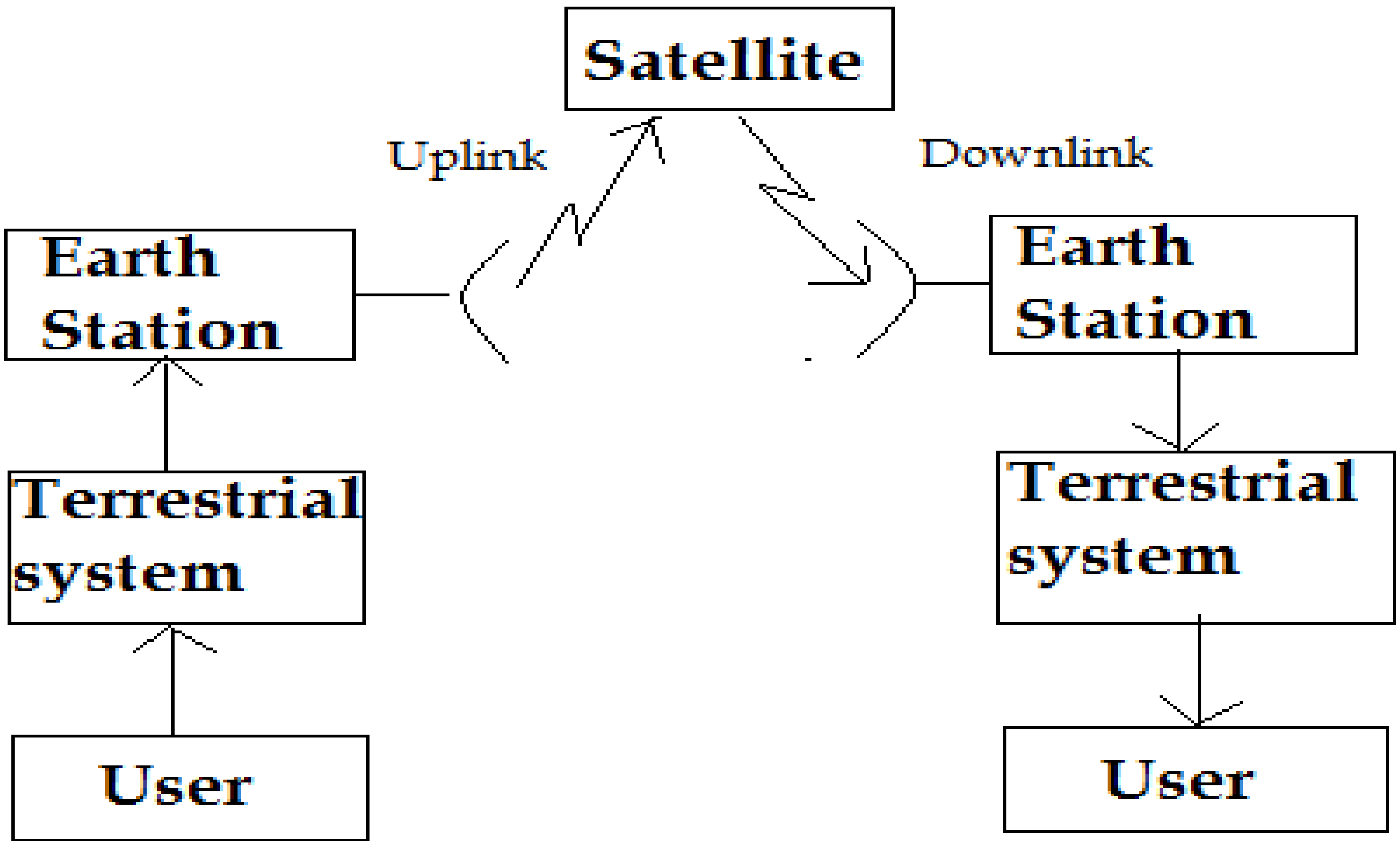


- Geostationary Satellites
- Geostationary Satellites are the satellites which have same angular velocity as earth and always have a fixed position with respect to earth.
- Satellites used for communication are generally geostationary satellites to avoid the tracking complexity.

Block Diagram of Transponder



Block Diagram of Satellite Communication



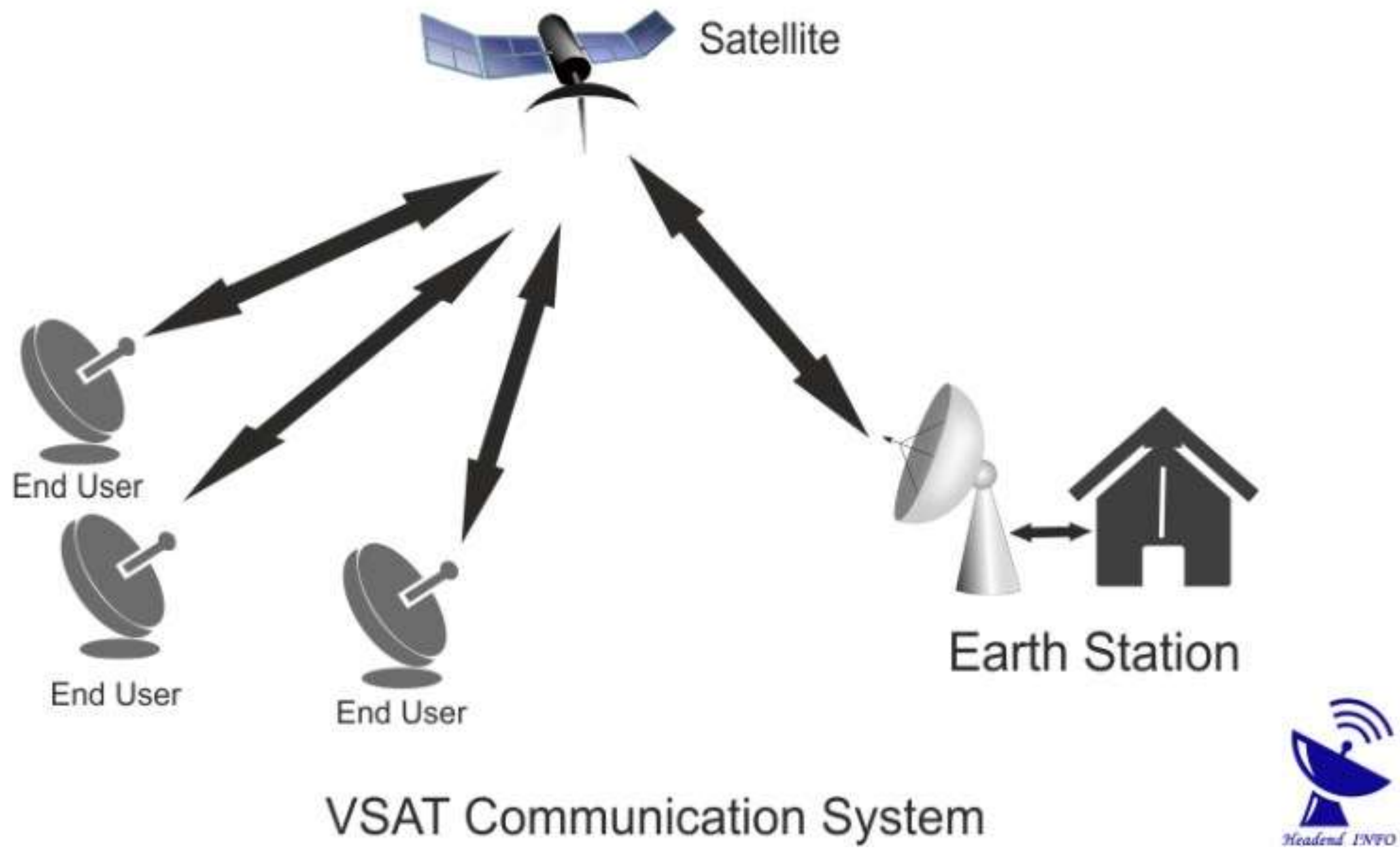
Advantages of Satellite Communication

- Large area can be covered
- Cost is independent of distance
- Communication in remote area is easy and cheaper.
- Communication can be restored easily at the time of disaster.

Disadvantages of Satellite Communication

- Time delay is high
- Repair of satellite is almost impossible
- Communication may disturbed in bad weather.

VSAT(Very Small Aperture Terminal)



- VSATs are like small earth stations but generally are the end users in satellite communication,
- Small size of antennas are used in VSATs.
- VSATs use higher frequency than C band and could be possible after the improved frequency handling capacity of satellites.

Advantages of VSATs

- Facilities like TV and ATMs and BANKs can be established in remote areas.
- Small Antennas are adequate
- No need of tower
- Size of TRx is also very small

Instrumentation

BY:- Yogender Singh

UNIT–I & II

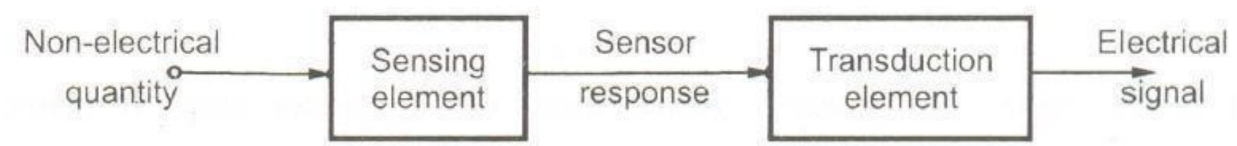
Measurement refers **to the comparison of an unknown quantity with a known quantity**. The result of the measurement is a numeric value with certain units.

Transducers

Definition of transducers – Classification of transducers – Advantages of Electrical transducers –Characteristics and choice of transducers – Principle operation of resistor, inductor, LVDT and capacitor transducers – LVDT Applications – Strain gauge and its principle of operation – Guage factor – Thermistors – Thermocouples – Synchros – Piezo electric transducers – Photo diodes.

INTODUCTION

A device which converts a physical quantity into the proportional electrical signal is called a transducer. The electrical signal produced may be a voltage, current or frequency. A transducer uses many effects toproduce such conversion. The process of transforming signal from one form to other is called transduction. A transducer is also called pick up. The transduction element transforms the output of thesensor to an electrical output, as shown in the Fig.



A transducer will have basically two main components. They are

1. Sensing Element

The physical quantity or its rate of change is sensed and responded to by this part of the transistor.

2. Transduction Element

The output of the sensing element is passed on to the transduction element. This element is responsiblefor converting the non-electrical signal into its proportional electrical signal.

There may be cases when the transduction element performs the action of both transduction and sensing. The best example of such a transducer is a thermocouple. A thermocouple is used to generate a voltage corresponding to the heat that is generated at the junction of two dissimilar metals.

Classification of Transducers

The Classification of Transducers is done in many ways. Some of the criteria for the classification are based on their area of application, Method of energy conversion, Nature of output signal, According to Electrical principles involved, Electrical parameter used, principle of operation, & Typical applications.

The transducers can be classified broadly

i. On the basis of transduction form used

ii. As primary and secondary transducers
iii. As active and passive transducers
iv. As transducers and inverse transducers.

Broadly one such generalization is concerned with energy considerations wherein they are classified asactive & Passive transducers. A component whose output energy is supplied entirely by its input signal(physical quantity under measurement) is commonly called a „passive transducer". In other words the passive transducers derive the power required for transduction from an auxiliary source. Active transducers are those which do not require an auxiliary power source to produce their output. They are also known as self generating type since they produce their own voltage or current output. Some of thepassive transducers (electrical transducers), their electrical parameter (resistance, capacitance, etc), principle of operation and applications are listed below.

Resistive Transducers

1. Resistance Strain Gauge – The change in value of resistance of metal semi-conductor due toelongation or compression is known by the measurement of torque, displacement or force.
2. Resistance Thermometer – The change in resistance of metal wire due to the change intemperature known by the measurement of temperature.
3. Resistance Hygrometer – The change in the resistance of conductive strip due to the change ofmoisture content is known by the value of its corresponding humidity.
4. Hot Wire Meter – The change in resistance of a heating element due to convection cooling of aflow of gas is known by its corresponding gas flow or pressure.
5. Photoconductive Cell – The change in resistance of a cell due to a corresponding change in light flux is known by its corresponding light intensity.
6. Thermistor – The change in resistance of a semi-conductor that has a negative co-efficient of resistance is known by its corresponding measure of temperature.
7. Potentiometer Type – The change in resistance of a potentiometer reading due to the movement of the slider as a part of an external force applied is known by its corresponding pressure or displacement.

Capacitance Transducers

1. Variable capacitance pressure gage -
Principle of operation: Distance between two parallel plates is varied by an externally appliedforce Applications: Measurement of Displacement, pressure
2. Capacitor microphone
Principle of operation: Sound pressure varies the capacitance between a fixed plate and a movablediaphragm. Applications: Speech, music, noise
3. Dielectric gauge
Principle of operation: Variation in capacitance by changes in the dielectric.
Applications: Liquidlevel, thickness

Inductance Transducers

1. Magnetic circuit transducer
Principle of operation: Self inductance or mutual inductance of ac-excited

coil is varied bychanges in the magnetic circuit. Applications: Pressure, displacement

2. Reluctance pickup
Principle of operation: Reluctance of the magnetic circuit is varied by changing the position ofthe iron core of a coil. Applications: Pressure, displacement, vibration, position
3. Differential transformer
Principle of operation: The differential voltage of two secondary windings of a transformer isvaried by positioning the magnetic core through an externally applied force. Applications: Pressure, force, displacement, position
4. Eddy current gage
Principle of operation: Inductance of a coil is varied by the proximity of an eddy current plate.Applications: Displacement, thickness
5. Magnetostriction gauge
Principle of operation: Magnetic properties are varied by pressure and stress. Applications: Force,pressure, sound

Voltage and current Transducers

1. Hall effect pickup
Principle of operation: A potential difference is generated across a semiconductor plate (germanium) when magnetic flux interacts with an applied current. Applications: Magnetic flux,current
2. Ionization chamber
Principle of operation: Electron flow induced by ionization of gas due to radioactive radiation.Applications: Particle counting, radiation
3. Photoemissive cell
Principle of operation: Electron emission due to incident radiation on photoemissive surface.Applications: Light and radiation
4. Photomultiplier tube
Principle of operation: Secondary electron emission due to incident radiation onphotosensitive cathode. Applications: Light and radiation, photo-sensitive relays

Self-Generating Transducers (No External Power) – Active Transducers

They do not require an external power, and produce an analog voltage or current when stimulated bysome physical form of energy.

1. Thermocouple and thermopile
Principle of operation: An emf is generated across the junction of two dissimilar metals or semiconductors when that junction is heated. Applications: Temperature, heat flow, radiation.
2. Moving-coil generator
Principle of operation: Motion of a coil in a magnetic field generates a voltage. Applications:Velocity. Vibration
3. Piezoelectric pickup
An emf is generated when an external force is applied to certain crystalline materials, such asquartz Sound, vibration. acceleration, pressure changes
4. Photovoltaic cell
Principle of operation: A voltage is generated in a semi-conductor junction device

when radiantenergy stimulates the cell Applications: Light meter, solar cell

Primary Transducers and Secondary Transducers- Bourden tube acting as a primary detecter senses the pressure and converts the pressure into a displacement of its free end.The displacement of the free end moves the core of a linear variable differential transformer(LVDT) which produces an output voltage.

Analog Transducers-These transducers convert the input quantity into an analog output which is acontinuous function of time. ◦ Strain Gauge ◦ LVDT ◦ Thermocouple ◦ Thermistor

Digital Transducers-These transducers convert the input quantity into an electrical output which is in the form of pulses. ◦ Glass Scale can be read optically by means of a light source,an optical system and photocells

Transducers and Inverse Transducers- -A Transducer can be broadly defined as a device which converts a non-electrical quantity into an electrical quantity. Ex:- Resistive,inductive and capacitive transducers -An inverse transducer is defined as a device which converts an electrical quantity into a non-electrical quantity. Ex:- Piezoelectric crystals

Advantages of Electrical transducers

Mostly quantities to be measured are non-electrical such as temperature, pressure, displacement, humidity, fluid flow, speed etc., but these quantities cannot be measured directly. Hence such quantities are required to be sensed and changed into some other form for easy measurement. Electrical quantities such as current, voltage, resistance, inductance and capacitance etc. can be conveniently measured, transferred and stored, and, therefore, for measurement of the non-electrical quantities these are to be converted into electrical quantities first and ten measured. The function of converting non-electrical quantity into electrical one is accomplished by a device called the electrical transducer.

Basically an electrical transducer is a sensing device by which a physical, mechanical or optical quantity to be measured is transformed directly, with a suitable mechanism, into an electrical signal (current, voltage and frequency). The production of these signals is based upon electrical effects which may be resistive, inductive, capacitive etc. in nature. The input versus output energy relationship takes a definite reproducible function. The output to input and the output to time behavior is predictable to a known degree of accuracy, sensitivity and response, within the specified environmental conditions. Electrical transducers have numerous advantages. Modern digital computers have made use of electrical transducersabsolutely essential.

Electrical transducers suffer due to some draw-backs too, such as low reliability in comparison to that of mechanical transducers due to the ageing and drift of the active components and comparative high cost of electrical transducers and associated signal conditioners. In some cases the accuracy and resolution attainable are not as high as in mechanical transducers.

Some of the advantages are:

1. Electrical amplification and attenuation can be done easily and that to with a static device.
2. The effect of friction is minimized.
3. The electric or electronic system can be controlled with a very small electric power.

4. The electric power can be easily used, transmitted and process for the purpose of measurement.

Factor to be considered while selecting transducer:

It should have high input impedance and low output impedance, to avoid loading effect. It should have good resolution over is entire selected range.
It must be highly sensitive to desired signal and insensitive to unwanted signal. Preferably small in size.
It should be able to work n corrosive environment.
It should be able to withstand pressure, shocks, vibrations etc.. It must have high degree of accuracy and repeatability.
Selected transducer must be free from errors.
The transducer circuit should have overload protection so that it will withstand overloads.

Requirements of a good transducers

• Smaller in size and weight.
• High sensitivity.
• Ability to withstand environmental conditions.
• Low cost.

RESISTIVE TRANSDUSERS

Resistance of an electrical conductor is given by,

$$R=\rho l/A$$

Where ,

R = Resistance in „Ω"

P = Resistivity of the conductor

(Ω - cm) l = Length of the

conductor in cm.

A = Cross-sectional area of the metal conductor in cm2

It is clear from the equation that, the electrical resistance can be varied by varying,
(i) Length

(ii) Cross-sectional area and

(iii) Resistivity or combination of these.

Principle:-
A change in resistance of a circuit due to the displacement of an object is the measure of displacement of that object ,method of changing the resistance and the resulting devices aresummarized in the following

Method of changing resistance-
Length - Resistance can be changed varying the length of the conductor,(linear and rotary).

Dimensions - When a metal conductor is subjected to mechanical strain, change in dimensions ofthe conductor occurs, that changes the resistance of the conductor.

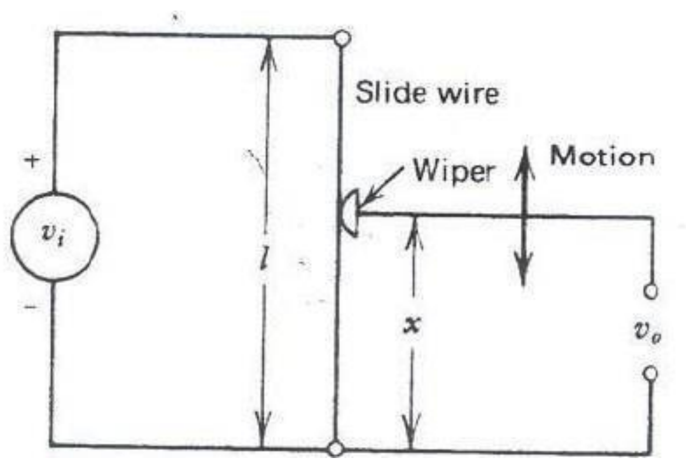
Resistivity -
When a metal conductor is subjected to a change in temperature and change in resistivity occurs whichchanges resistance of the conductor.

Resulting device:-
Resistance potentiometers or sliding contact devices displacements ,Electricalresistance straingauges.Thermistor and RTD

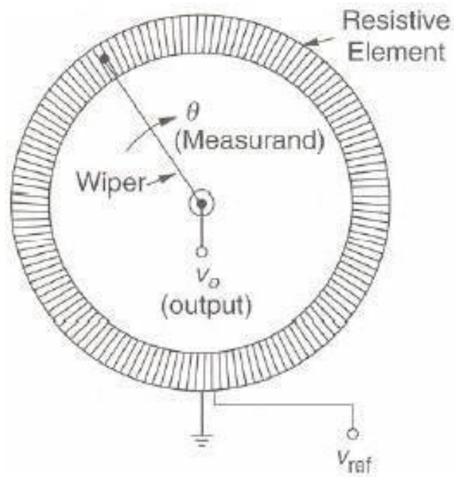
Use:-
the resistive transducer used for the measurement of linear and angular, and used for thetemperature mechanical strain measurement.

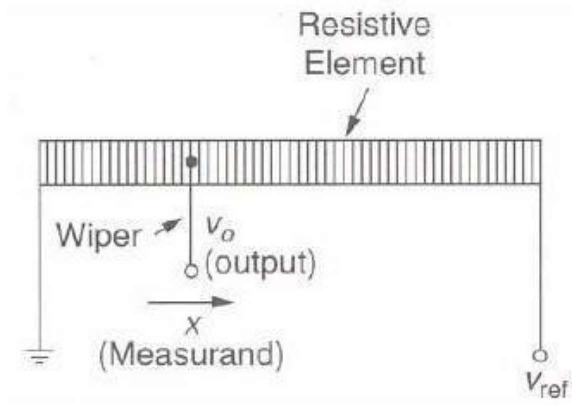
How Potentiometer works

A potentiometer is a resistive sensor used to measure linear displacements as well as rotary motion. In a potentiometer an electrically conductive wiper slides across a fixed resistive element. A voltage is appliedacross the resistive element. Thus a voltage divider circuit is formed. The output voltage(Vout) is measured as shown in the figure below. The output voltage is proportional to the distance travelled.



There are two types of potentiometer, linear and rotary potentiometer. The linear potentiometer has a slideor wiper. The rotary potentiometer can be a single turn or multi turn.





The important parameters while selecting a potentiometer are

•Operating temperature

•Shock and vibration

•Humidity

•Contamination and seals

•life cycle

•dither

Types of Potentiometer:

Wire-Wound type potentiometer

• The resistance range between 10Ω and 10M Ω
• The resistance increase in a stepwise manner.
• It is possible to construct potentiometers with 100 –200 turns per cm length (The resolution range between 0.1 to 0.05 mm).
• Linear potentiometers are available in many lengths up to 1m.
• Helical potentiometers are commercially available with 50 to 60 turns (The angular displacement is between 18000 – 21600 degree)
• Potentiometer life exceed 1 million cycles.

Thin film type potentiometer

• Higher resolution.
• Lower noise.
• Longer life (exceed 10 million cycles)
• Resistance of 50 to 100 Ω/mm can be obtained with conductive plastic film.
• Commercially available resolution is 0.001 mm.

- Power rating

$$P = \frac{V_{in}^2}{R_P}$$

$$V_{in(\max)} = \sqrt{PR_P}$$

- Sensitivity

$$V_o = \frac{V_{in}}{L} x$$

$$Sensitivity = \frac{V_{in}}{L}$$

- Linearity $V_o = V_{in}(\frac{R}{R_P})\left[\frac{R_M/R_P}{(R_M/R_P)+(R/R_P)-(R/R_P)^2}\right]$

- Optimum sensitivity (0.2 V/degree and 2 V/cm)

Some of the advantages of the potentiometer are

•Easy to use
•low cost
•High amplitude output
•Proven technology
•Easily available

Some of the disadvantages of the potentiometer are

•Since the wiper is sliding across the resistive element there is a possibility of friction and wear. Hence the number of operating cycles are limited.
•Limited bandwidth
• Inertial loading

Some of the applications of the potentiometer are

- •Linear displacement measurement
- •Rotary displacement measurement
- •Volume control
- •Brightness control
- •Liquid level measurements using float

Strain Gauge

Strain gage is one of the most popular types of transducer. It has got a wide range of applications. It canbe used for measurement of force, torque, pressure, acceleration and many other parameters. The basic principle of operation of a strain gage is simple: when strain is applied to a thin metallic wire, its dimension changes, thus changing the resistance of the wire. Let us first investigate what are the factors, responsible for the change in resistance.

Gauge Factor

Let us consider a long straight metallic wire of length *l* circular cross section with diameter *d* (fig). Whenthis wire is subjected to a force applied at the two ends, a strain will be generated and as a result, the

dimension will change (*l* changing to $l + \Delta l$, *d* changing to $d + \Delta d$ and *A* changing to $A + \Delta A$). For the time being, we are considering that all the changes are in positive direction. Nowthe resistance of the wire:

Let us consider a long straight metallic wire of length l circular cross section with diameter d (fig. 5). When this wire is subjected to a force applied at the two ends, a strain will be generated and as a result, the dimension will change (l changing to $l+\Delta l$, d changing to $d+\Delta d$ and A changing to $A+\Delta A$). For the time being, we are considering that all the changes are in positive direction. Now the resistance of the wire:

$$R = \frac{\rho l}{A}, \text{ where } \rho \text{ is the resistivity.}$$

From the above expression, the change in resistance due to strain:

$$\begin{aligned}\Delta R &= \left(\frac{\partial R}{\partial l}\right)\Delta l + \left(\frac{\partial R}{\partial A}\right)\Delta A + \left(\frac{\partial R}{\partial \rho}\right)\Delta \rho \\ &= \frac{\rho}{A}\Delta l - \frac{\rho}{A^2}\Delta A + \frac{l}{A}\Delta \rho \\ &= R\frac{\Delta l}{l} - R\frac{\Delta A}{A} + R\frac{\Delta \rho}{\rho}\end{aligned}$$

or,

$$\frac{\Delta R}{R} = \frac{\Delta l}{l} - \frac{\Delta A}{A} + \frac{\Delta \rho}{\rho} \qquad (6)$$

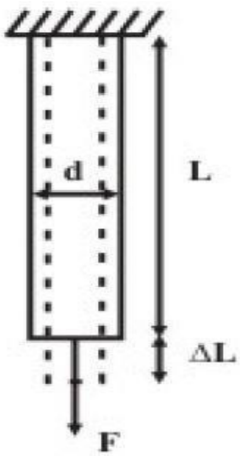


Fig:Change of Resistance with strain

Now, for a circular cross section, $A = \frac{\pi d^2}{4}$; from which, $\Delta A = \frac{\pi d}{2}\Delta d$. Alternatively,

$$\frac{\Delta A}{A} = 2\frac{\Delta d}{d}$$

Hence,

$$\frac{\Delta R}{R} = \frac{\Delta l}{l} - 2\frac{\Delta d}{d} + \frac{\Delta \rho}{\rho} \tag{7}$$

Now, the *Poisson's Ratio* is defined as:

$$\upsilon = -\frac{lateral \quad strain}{longitudinal \quad strain} = -\frac{\Delta d/d}{\Delta l/l}$$

The Poisson's Ratio is the property of the material, and does not depend on the dimension. So, (6) can be rewritten as:

$$\frac{\Delta R}{R} = (1+2\upsilon)\frac{\Delta l}{l} + \frac{\Delta \rho}{\rho}$$

Hence,

$$\frac{\Delta R/R}{\Delta l/l} = 1 + 2\upsilon + \frac{\Delta \rho/\rho}{\Delta l/l}$$

The last term in the right hand side of the above expression, represents the change in resistivity of the material due to applied strain that occurs due to the *piezo-resistance property* of the material. In fact, all the elements in the right hand side of the above equation are independent of the geometry of the wire, subjected to strain, but rather depend on the material property of the wire. Due to this reason, a term *Gage Factor* is used to characterize the performance of a strain gage. The Gage Factor is defined as:

$$G := \frac{\Delta R/R}{\Delta l/l} = 1 + 2\upsilon + \frac{\Delta \rho/\rho}{\Delta l/l} \tag{8}$$

For normal metals the Poisson's ratio υ varies in the range:

$$0.3 \le \upsilon \le 0.6,$$

while the piezo-resistance coefficient varies in the range:

$$0.2 \le \frac{\Delta \rho/\rho}{\Delta l/l} \le 0.6.$$

Thus, the Gage Factor of metallic strain gages varies in the range 1.8 to 2.6. However, the semiconductor type strain gages have a very large Gage Factor, in the range of

100-150. This is attained due to dominant piezo-resistance property of semiconductors. The commercially available strain gages have certain fixed resistance values, such as, 120Ω, 350 Ω, 1000 Ω, etc. The manufacturer also specifies the Gage Factor and the maximum gage current to avoid self-heating (normally in the range 15 mA to 100 mA).
The choice of material for a metallic strain gage should depend on several factors. The material should have low temperature coefficient of resistance. It should also have low coefficient for thermal expansion. Judging from all these factors, only few alloys qualify for a commercial metallic strain gage. They are:
Advance (55% Cu, 45% Ni): Gage Factor between 2.0 to 2.2
Nichrome (80% Ni, 20% Co): Gage Factor between 2.2 to 2.5

Apart from these two, *Isoelastic* -another trademarked alloy with Gage Factor around 3.5 is also in use. Semiconductor type strain gages, though having large Gage Factor, find limited use, because of their high sensitivity and nonlinear characteristics.

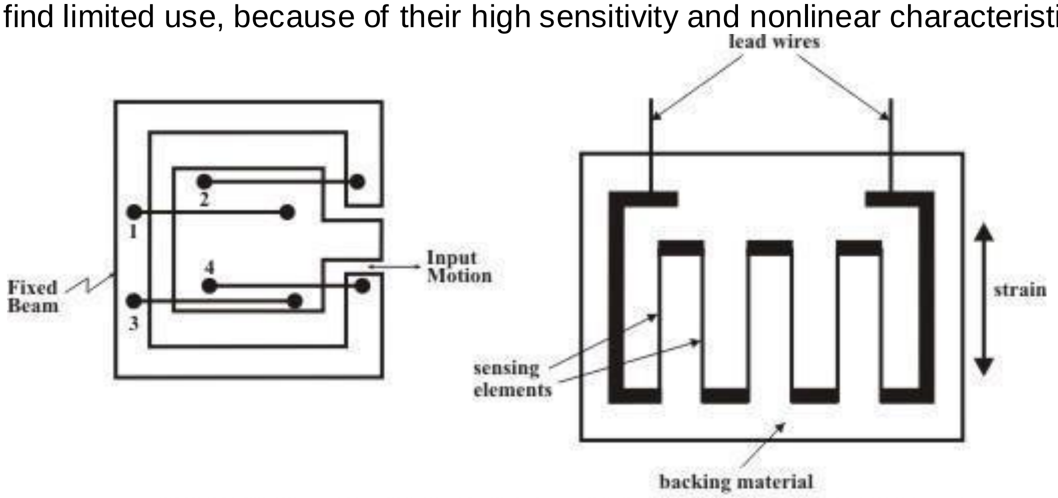


Fig. 6 (a) Unbonded metallic strain gage, (b) bonded metal foil type strain gage

Metallic Strain Gage

Most of the strain gages are metallic type. They can be of two types: *unbonded* and *bonded*. The unbonded strain gage is normally used for measuring strain (or displacement) between a fixed and a moving structure by fixing four metallic wires in such a way, so that two are in compression and two are in tension, as shown in fig. 6 (a). On the other hand, in the bonded strain gage, the element is fixed on a backing material, which is permanently fixed over a structure, whose strain has to be measured, with adhesive. Most commonly used bonded strain gages are *metal foil type*. The construction of such a strain gage is shown in fig. 6(b). The metal foil type strain gage is manufactured by photo-etching technique. Here the thin strips of the foil are the active elements of the strain gage, while the thick ones are for providing electrical connections. Because of large area of the thick portion, their resistance is small and they do not contribute to any change in resistance due to strain, but increase the heat dissipation area. Also it is easier to connect the lead wires with the strain gage. The strain gage in fig. 6(b) can measure strain in one direction only. But if we want to measure the strain in two or more directions at the same point, strain gage *rosette*, which is manufactured by stacking multiple strain gages in different directions, is used. Fig. 7 shows a three-element strain gage rosette stacked at 45^0.

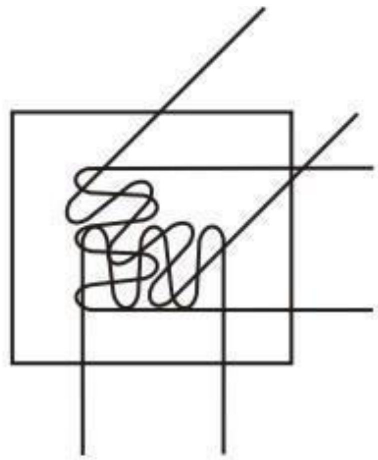


Fig. 7 Three-element strain gage rosette- 45° stacked.

The *backing material*, over which the strain gage is fabricated and which is fixed with the strain measuring structure has to satisfy several important properties. Firstly, it should have high mechanical strength; it should also have high dielectric strength. But the most important it should have is that it should be non-hygroscopic, otherwise, absorption of moisture will cause bulging and generate local strain. The backing materials normally used are impregnated paper, fibre glass, etc. The *bonding material* used for fixing the strain gage permanently to the structure should also be non-hygroscopic. Epoxy and Cellulose are the bonding materials normally used.

Semiconductor type Strain Gage

Semiconductor type strain gage is made of a thin wire of silicon, typically 0.005 inch to 0.0005 inch, and length 0.05 inch to 0.5 inch. They can be of two types: *p*-type and *n*-type. In the former the resistance increases with positive strain, while, in the later the resistance decreases with temperature. The construction and the typical characteristics of a semiconductor strain gage are shown in fig.8.
MEMS pressure sensors is now a days becoming increasingly popular for measurement of pressure. It is made of a small silicon diagram with four piezo-resistive strain gages mounted on it. It has an in- built signal conditioning circuits and delivers measurable output voltage corresponding to the pressure applied. Low weight and small size of the sensor make it suitable for measurement of pressure in specific applications.

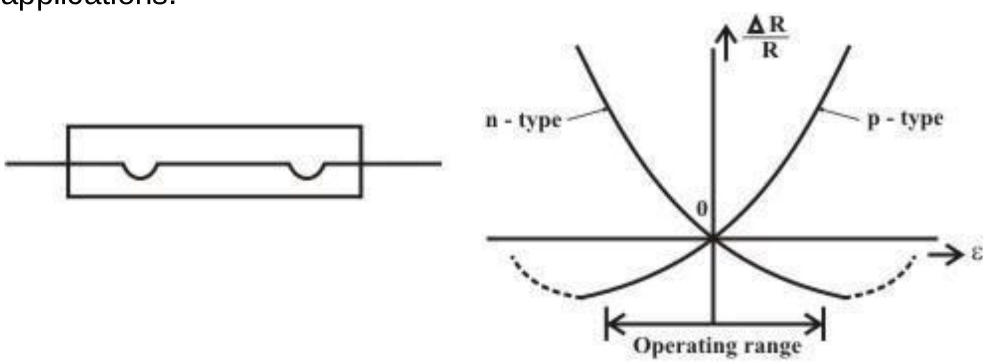
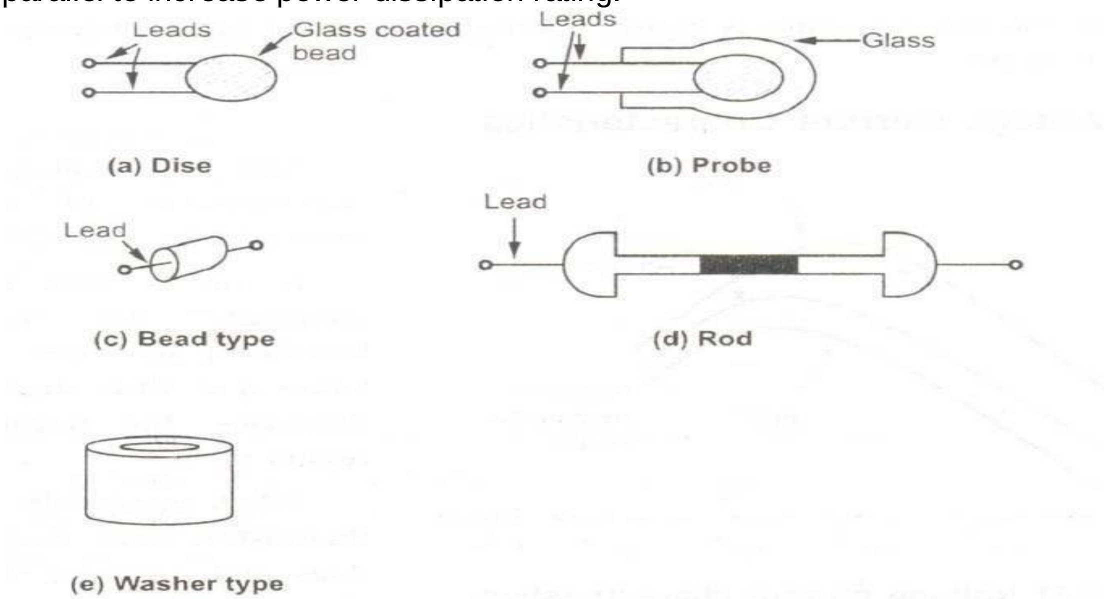


Fig. 8 (a) construction and (b) characteristics of a semiconductor strain gage

Thermistors:
Basically thermistor is a contraction of a word 'thermal resistors', The resistors depending on temperature are thermal resistors. Thus resistance thermometers are

also thermistors having positive -temperature coefficients. But generally the resistors having negative temperature coefficients (NTC) are called thermistors. The resistance of a thermistor decreases as temperature increases. The NTC of thermistors can be as large as few percent per degree celcius change in temperature. Thus the thermistors are very sensitive and can detect very small changes in temperature too.

Construction of thermistor:
Thermistors are composed of a sintered mixture of metallic oxides, such as manganese, nickel, cobalt, copper, iron, and uranium. Their resistances at ambient temperature may range from 100 n to 100 ill. Thermistors are available in a wide variety of shapes and sizes as shown in the Fig. Smallest in size arethe beads with a diameter of 0.15 mm to 1.25 mm. Beads may be sealed in the tips of solid glass rods toform probes. Disks and washers are made by pressing thermistor materia~ under high pressure into Hatcylindrical shapes. Washers can be placed in series or in parallel to increase power dissipation rating.



Thermistors are well suited for precision temperature measurement, temperature control, and temperaturecompensation, because of their. very large change in resistance with temperature. They are widely
used for measurements in the temperature range -1000 C to +2000 C. The measurement of the change inresistance with temperature is carried out with a Wheatstone bridge.

Inductive Transducer

Inductive transducers work on the principle of inductance change due to any appreciable change in the quantity to be measured i.e. measured. For example, LVDT, a kind of **inductive transducers**, measures displacement in terms of voltage difference between its two secondary voltages. Secondary voltages are nothing but the result of induction due to the flux change in the secondary coil with the displacement of the iron bar. Anyway LVDT is discussed here briefly to explain the principle of **inductive transducer**. LVDT will be explained in other article in more detail. For the time being let"s focus on basic introduction of **inductive transducers**. Now first our motive is to find how the **inductive transducers** can be made to work. This can be done by changing the flux with the help of measured and this changing flux obviously changes the inductance and this inductance change can be calibrated in terms of measured. Hence **inductive transducers** use one of the following principles for its working.

1. Change of self inductance
2. Change of mutual inductance
3. Production of eddy current

Change of Self Inductance of Inductive Transducer

We know very well that self inductance of a coil is given by $L = \frac{N^2}{R}$ Where, N = number of turns. R

$$R = \frac{l}{\mu A}$$

= reluctance of the magnetic circuit. Also we know that reluctance R is given by

$$L = \frac{N^2 \mu A}{l}$$

where μ = effective permeability of the medium in and around the coil.

$$L = N^2 \mu G$$

Where, G = A/l and called geometric form factor. A = area of cross-section of coil. l =length of the coil. So, we can vary self inductance by

- Change in number of turns, N,
- Changing geometric configuration, G,
- Changing permeability

For the sake of understanding we can say that if the displacement is to be measured by the **inductive transducers**, it should change any of the above parameter for causing in the change in self inductance.

Change of Mutual Inductance of Inductive Transducer

Here transducers, which work on change of mutual inductance principle, use multiple coils. We use here two coils for the sake of understanding. Both coils have their self inductance as well. So let"s denote their self inductance by L_1 and L_2. Mutual inductance between these two coils is given by

$$M = K\sqrt{L_1 L_2}$$

Thus mutual inductance can be changed by varying self inductance or by varying coefficient of coupling, K. The methods of changing self inductance we already discussed. Now coefficient of coupling depends on the distance and orientation between two coils. Thus for the measurement of displacement we can fix one coil and make other movable which moves with the source whose displacement is to be measured. With the change in distance in displacement coefficient of coupling changes and it causes the change in mutual inductance. This change in mutual inductance can becalibrated with the displacement and measurement can be done.

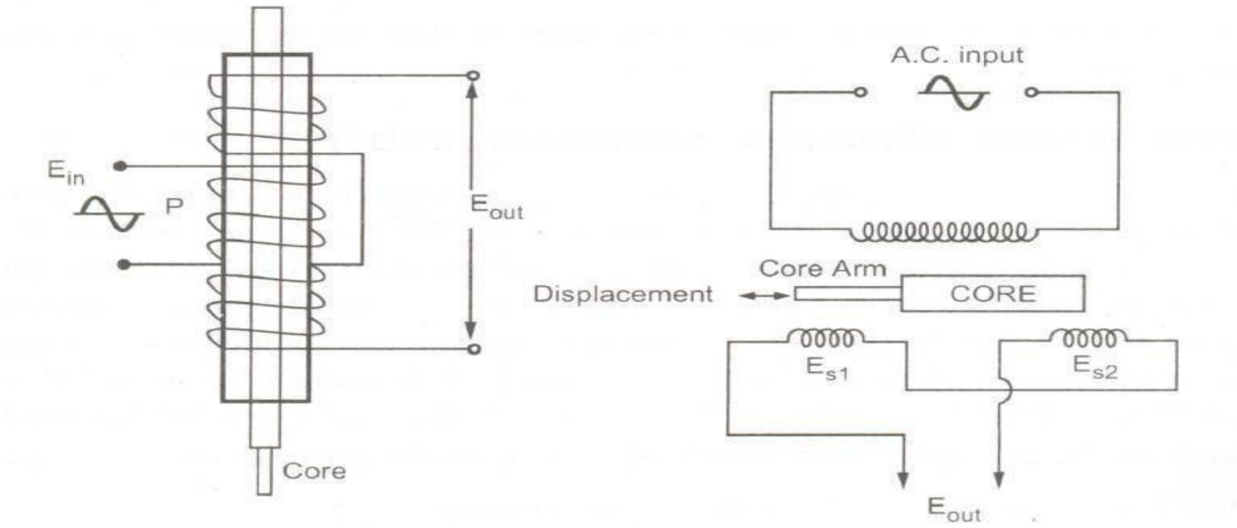
Production of Eddy Current of Inductive Transducer

We know that when a conducting plate is placed near a coil carrying alternating current, a circulating current is induced in the plate called “EDDY CURRENT”. This principle is used in such type of **inductive transducers**. Actually what happens? When a coil is placed near to coil carrying alternating current, a circulating current is induced in it which in turn produces its own flux which try to reduce the flux of the coil carrying the current and hence inductance of the coil changes. Nearer the plate is to the coil, higher will be eddy current and higher is the reduction in inductance and vice versa. Thus inductanceof coil varied with the variation of distance between coil and plate. Thus the movement of the plate can becalibrated in terms of inductance change to measure the quantity like displacement.

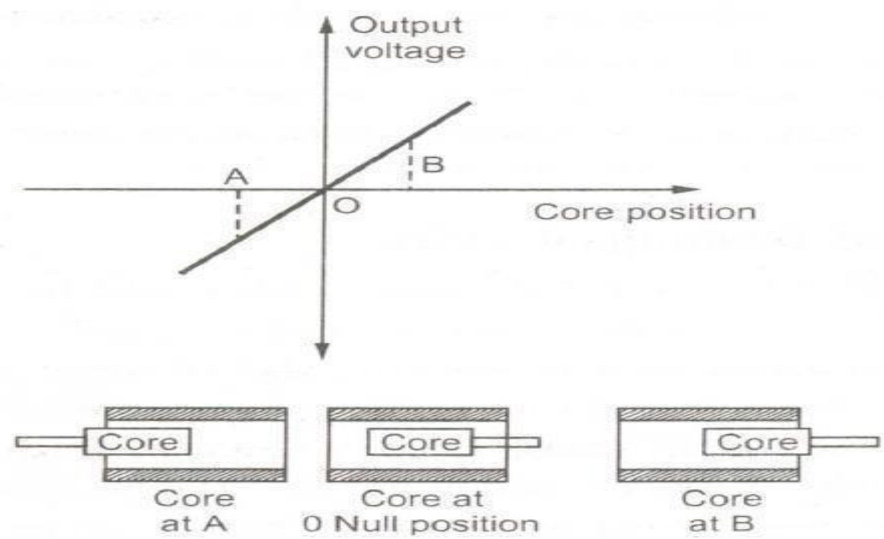
Real Life Application of Inductive Transducer

Inductive transducers find application in proximity sensors which are used for position measurement, dynamic motion measurement, touch pads etc. Particularly **inductive transducer** is used for the detectionof type of metal, finding missing parts or counting the number of objects.

Linear variable differential transformer (LVDT)



When an externally applied force moves the core to the left-hand position, more magnetic flux links theleft-hand coil than the righthand coil. The emf induced in the left-hand coil, ES], is therefore larger thanthe induced emf of the right-hand [oil, Es2' The magnitude of the output voltage is then equal to the difference between the two secondary voltages and it is in phase with the voltage of the left-hand coil.



Output voltage of LVDT at different core positions

Construction of LVDT

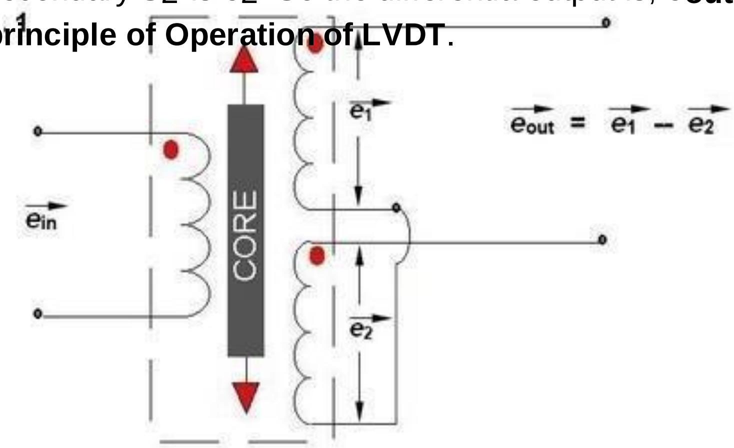
Main Features of Construction are as,

- The transformer consists of a primary winding P and two secondary winding S1 and S2 wound on acylindrical former(which is hollow in nature and will contain core).

- Both the secondary windings have equal number of turns and are identically placed on the either side of primary winding
- The primary winding is connected to an AC source which produces a flux in the air gap and voltages are induced in secondary windings.
- A movable soft iron core is placed inside the former and displacement to be measured is connected to the iron core.
- The iron core is generally of high permeability which helps in reducing harmonics and high sensitivity of LVDT.
- The LVDT is placed inside a stainless steel housing because it will provide electrostatic and electromagnetic shielding.
- The both the secondary windings are connected in such a way that resulted output is the difference of the voltages of two windings.

Principle of Operation and Working

As the primary is connected to an AC source so alternating current and voltages are produced in the secondary of the LVDT. The output in secondary S_1 is e_1 and in the secondary S_2 is e_2. So the differential output is, **$e_{out} = e_1 - e_2$** This equation explains the **principle of Operation of LVDT**.

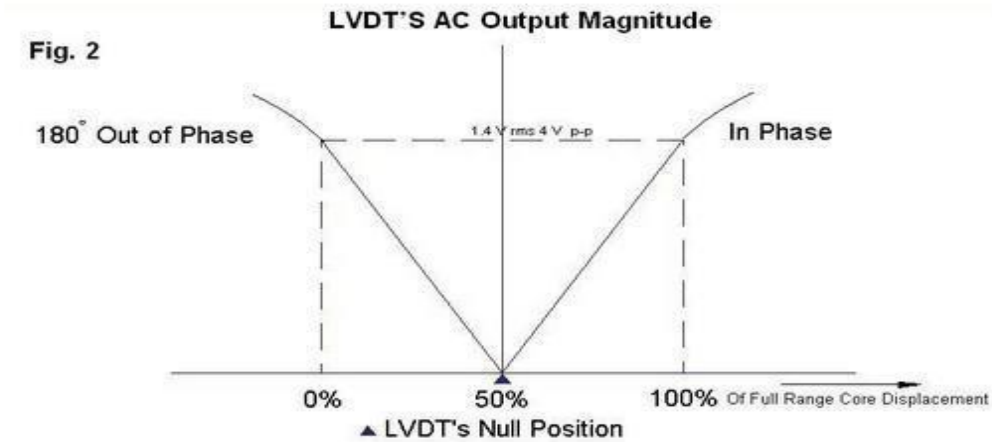


Now three cases arise according to the locations of core which explains the working of LVDT are discussed below as,

- **CASE I** When the core is at null position (for no displacement) When the core is at null position then the flux linking with both the secondary windings is equal so the induced emf is equal in both the windings. So for no displacement the value of output e_{out} is zero as e_1 and e_2 both are equal. So it shows that no displacement took place.
- **CASE II** When the core is moved to upward of null position (For displacement to the upward of reference point) In the this case the flux linking with secondary winding S_1 is more as compared to flux linking with S_2. Due to this e_1 will be more as that of e_2. Due to this output voltage e_{out} is positive.

- **CASE III** When the core is moved to downward of Null position (for displacement to the downward of reference point) In this case magnitude of e_2 will be more as that of e_1. Due to this output e_{out} will be negative and shows the output to downward of reference point.

Output VS Core Displacement A linear curve shows that output voltage varies linearly with displacement of core.



AC Output of Conventional LVDT Versus Core Displacement

Some important points about magnitude and sign of voltage induced in LVDT

- The amount of change in voltage either negative or positive is proportional to the amount of movement of core and indicates amount of linear motion.
- By noting the output voltage increasing or decreasing the direction of motion can be determined
- The output voltage of an LVDT is linear function of core displacement .

Advantages of LVDT

- **High Range -** The LVDTs have a very high range for measurement of displacement.they can usedfor measurement of displacements ranging from 1.25mm to 250mm
- **No Frictional Losses -** As the core moves inside a hollow former so there is no loss of displacementinput as frictional loss so it makes LVDT as very accurate device.
- **High Input and High Sensitivity -** The output of LVDT is so high that it doesn"t need anyamplification.the transducer posseses a high sensitivity which is typically about 40V/mm.
- **Low Hysteresis -** LVDTs show a low hysteresis and hence repeatability is excellent under allconditions

- **Low Power Consumption -** The power is about 1W which is very as compared to other transducers.
- **Direct Conversion to Electrical Signals -** They convert the linear displacement to electrical voltagewhich are easy to process

Disadvantages of LVDT

- LVDT is sensitive to stray magnetic fields so they always require a setup to protect them from straymagnetic fields.
- They are affected by vibrations and temperature.

It is concluded that they are advantageous as compared than any other inductive transducers.

Applications of LVDT

1. They are used in applications where displacements ranging from fraction of mm to few cm are to bemeasured. The LVDT acting as a primary Transducer converts the displacement to electrical signal directly.
2. They can also acts as the secondary transducers. E.g. the Bourbon tube which acts as a primary transducer and covert pressure into linear displacement.then LVDT coverts this displacement intoelectrical signal which after calibration gives the ideas of the pressure of fluid.

Capacitive Transducers

A capacitor consists of two conductors (plates) that are electrically isolated from one another by a nonconductor (dielectric). When the two conductors are at different potentials (voltages), the system is capable of storing an electric charge. The storage capability of a capacitor is measured in farads.The principle of operation of capacitive transducers is based upon the equation for capacitance of a parallel plate capacitor as

$$Capacitance \qquad C = \frac{\varepsilon A}{D}$$

shown in Fig.

Where, A = Overlapping area of plates;

m^2,d = Distance between two

plates; m,

ε = Permittivity (dielectric constant); F/m.

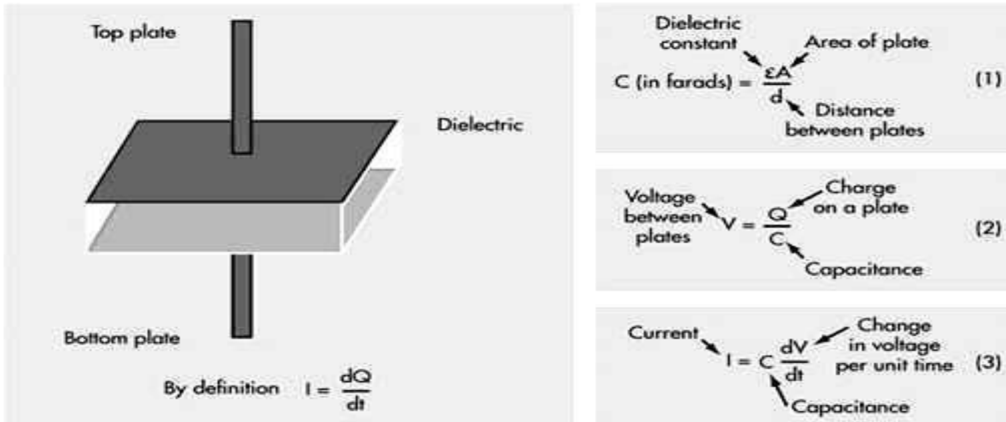


Fig. Parallel plate capacitor

The capacitance is measured with a bridge circuits. The output impedance Z of a capacitivetransducer is:

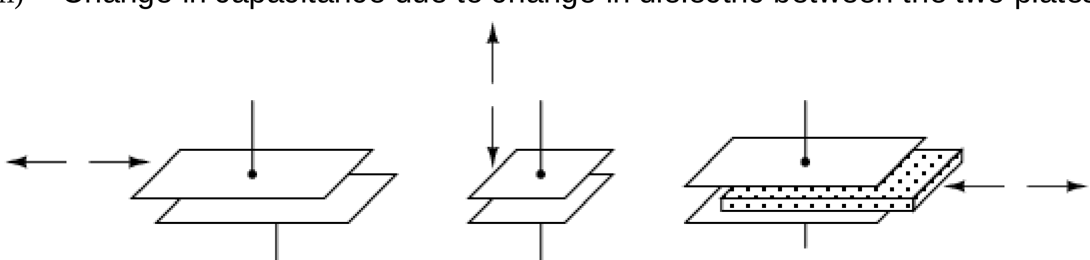
$$Z = 1/2\pi fC$$

Where: Z = Impedance
f = frequency, 50 Hz. C = capacitance

In general, the output impedance of a capacitive transducer is high. This fact calls for a careful design of the output circuitry. The capacitive transducers work on the principle of change in capacitance of the capacitor. This change in capacitance could be caused by change in overlapping area A of the plates, change in the distance d between the plates and change in dielectric constant ε.

In most of the cases the above changes are caused by the physical variables, such as, displacement, force or pressure. Variation in capacitance is also there when the dielectric medium between the plates changes, as in the case of measurement of liquid or gas levels. Therefore, the capacitive transducers are commonly used for measurement of linear displacement, by employing the following effects as shown in Fig a and fig b.

i) Change in capacitance due to change in overlapping area of plates.
ii) Change in capacitance due to change in distance between the two plates.
iii) Change in capacitance due to change in dielectric between the two plates



(a) (b) (c)

Fig.a Variable capacitive transducer varies; (a) area of overlap, (b) distance betweenplates, (c) amount of dielectric between plates

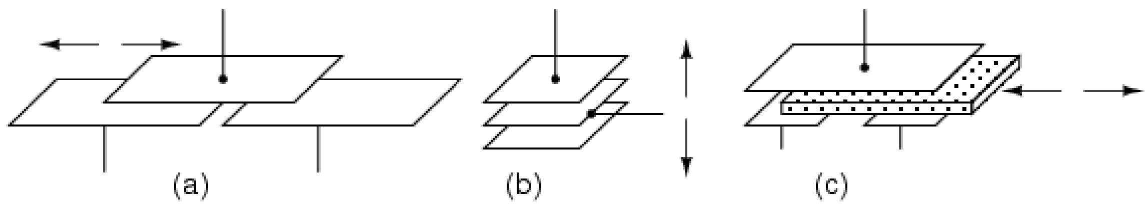


Fig.b Differential capacitive transducer varies capacitance ratio by changing: (a) area ofoverlap, (b) distance between plates, (c) dielectric between plates

As may be seen in Fig b, all of the differential devices have three wire connections rather than two: one wire for each of the end plates and one for the common plate. As the capacitance between one of the endplates and the common plate changes, the capacitance between the other end plate and the common plate also changes in the opposite direction.

a) Transducers Using Change in Area of Plates

Examining the equation for capacitance, it is found that the capacitance is directly proportional to the area, A of the plates. Thus, the capacitance changes linearly with change in area of plates. Hence this type of capacitive transducer is useful for measurement of moderate to large displacements say from 1 mm to several cm. The area changes linearly with displacement and also the capacitance.

For a parallel plate capacitor, the capacitance is:

$$C = \frac{\varepsilon A}{d} = \frac{\varepsilon l w}{d} F$$

Where, l = length of overlapping part of plates; m,
andw = width of overlapping part of plates;
m.

$$Sensitivity \qquad S = \frac{\partial C}{\partial l} = \varepsilon \frac{w}{d} F/m$$

The sensitivity is constant and therefore there is linear relationship between capacitance anddisplacement.

This type of a capacitive transducer is suitable for measurement of linear displacement rangingfrom 1 to 10 cm. The accuracy is as high as 0.005%.

b) Transducers Using Change in Distance between Plates

Fig. 17.2(b) shows the basic form of a capacitive transducer employing change in distance between the two plates to cause the change in capacitance. One plate is fixed and the displacement to be measured is applied to the other plate which is movable. Since, the capacitance, C, varies inversely as the distance d, between the plates the response of this transducer is not linear. Thus this transducer is useful only for measurement of extremely small displacements.

$$Sensitivity \qquad S = \frac{\partial C}{\partial l} = -\frac{\varepsilon A}{d^2}$$

Thus the sensitivity of this type of transducer is not constant but varies over the range of the transducer. The relationship between variations of capacitance with variation of distance between plates is hyperbolic and is only approximately linear over a small range of displacement. The linearity can be closely approximated by use of a piece of dielectric material like mica having a high dielectric constant, such as, a thin piece of mica.

c) Transducers Using Change in dielectric constant between Plates

If the area (A) of and the distance (d) between the plates of a capacitor remain constant, capacitance will vary only as a function of the dielectric constant (□) of the substance filling the gap between the plates. If the space between the plates of a capacitor is filled with an insulator, the capacitance of the capacitor will change compared to the situation in which there is vacuum between the plates. The change in the capacitance is caused by a change in the electric field between the plates.

The value of dielectric constant is initially set by design in the choice of dielectric material used to make the capacitor. Many factors will cause the to change, and this change in will vary for different materials. The major factors that will cause a change in are moisture, voltage, frequency, and temperature. The dielectric constant of a process material can change due to variations in temperature, moisture, humidity, material bulk density, and particle size etc. The in the basic formula is the effective dielectric constant of the total space between the electrodes. This space may consist of the dielectric material, air, and even moisture, if present. The figure shows that how in a capacitor the position of the dielectric is varied to vary the capacitance. Physical variables, such as, displacement, force or pressure can cause the movement of dielectric material in the capacitor plates, resulting in changes in the effective dielectric constant, which in turn will change the capacitance.

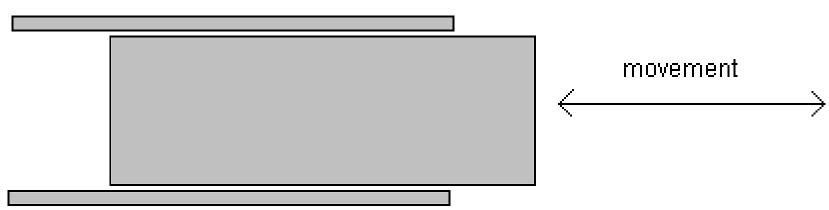


Fig. Change in capacitance due to movement of dielectric between plates

The major advantages of capacitive transducers are that they require extremely small forces to operate them and hence are very useful for use in small systems. They are extremely sensitive and require small power to operate them. Owing to their good frequency response they are very useful for dynamic studies.

The disadvantages of capacitive transducers include their non-linear behaviour on account of edge effects and the effects of stray capacitances especially when the transducers have a low value of capacitance. Therefore guard rings must be used to eliminate this effect. The metallic parts of the capacitive transducers must be insulated from each other. In order to reduce the effects of stray capacitances, the frames must be earthed.

Capacitive transducers can be used for measurement of both linear and angular displacements. The capacitive transducers are highly sensitive and can be used for measurement of extremely small displacements down to the order of molecular dimensions, i.e., 0.1×10^{-6} mm. On the other hand, they can be used for measurement of large displacements up to about 30 m as in aeroplane altimeters. The change in area method is used for measurement of displacements ranging from 10 to 100 mm. Capacitive transducers can be used for the measurement of force and pressure. The force and pressure to be measured are first converted to displacement which causes a change of capacitance. Capacitive transducers can also be used directly as pressure transducers in all those cases where the dielectric constant of a medium changes with pressure. They can be used for measurement of humidity in gases and moisture content in soil / food products etc.

Flow measurement

Flow may be measured by measuring the velocity of fluid over a known area. For very large flows, tracer methods may be used to deduce the flow rate from the change in concentration of a dye or radioisotope.

Turbine flowmeter

The turbine flowmeter (better described as an axial turbine) translates the mechanical action of the turbine rotating in the liquid flow around an axis into a user-readable rate of flow (gpm, lpm, etc.). The turbine tends to have all the flow traveling around it.

The turbine wheel is set in the path of a fluid stream. The flowing fluid impinges on the turbine blades, imparting a force to the blade surface and setting the rotor in motion. When a steady rotation speed has been reached, the speed is proportional to fluid velocity.

Turbine flowmeters are used for the measurement of natural gas and liquid flow.[4] Turbine meters are less accurate than displacement and jet meters at low flow rates, but the measuring element does not occupy or severely restrict the entire path of flow. The flow direction is generally straight through the meter, allowing for higher flow rates and less pressure loss than displacement-type meters. They are the meter of choice for large commercial users, fire protection, and as master meters for the water distribution system. Strainers are generally required to be installed in front of the meter to protect the measuring element from gravel or other debris that could enter the water distribution system. Turbine meters are generally available for 4 to 30 cm (1+½–12 in) or higher pipe sizes. Turbine meter bodies are commonly made of bronze, cast Iron, or ductile iron. Internal turbine elements can be plastic or non-corrosive metal alloys. They are accurate in normal working conditions but are greatly affected by the flow profile and fluid conditions.

Fire meters are a specialized type of turbine meter with approvals for the high flow rates required in fire protection systems. They are often approved by Underwriters Laboratories (UL) or Factory Mutual (FM) or similar authorities for use in fire protection. Portable turbine meters may be temporarily installed to measure water used from a fire hydrant. The meters are normally made of aluminum to be lightweight, and are usually 7.5 cm (3 in) capacity. Water utilities often require them for measurement of water used in construction, pool filling, or where a permanent meter is not yet installed.

Magnetic flowmeters

Magnetic flowmeters, often called "mag meter"s or "electromag"s, use a magnetic field applied to the metering tube, which results in a potential difference proportional to the flow velocity perpendicular to the flux lines. The potential difference is sensed by electrodes aligned perpendicular to the flow and the applied magnetic field. The physical principle at work is Faraday's law of electromagnetic induction. The magnetic flowmeter requires a conducting fluid and a nonconducting pipe liner. The electrodes must not corrode in contact with the process fluid; some magnetic flowmeters have auxiliary transducers installed to clean the electrodes in place. The applied magnetic field is pulsed, which allows the flowmeter to cancel out the effect of stray voltage in the piping system.

DRAWBACK - *ONLY MAGNETIC LIQUID FLOW CAN BE MEASURED BY THIS METOD*

Ultrasonic flowmeters

There are two main types of ultrasonic flowmeters: Doppler and transit time. While they both utilize ultrasound to make measurements and can be non-invasive (measure flow from outside the tube, pipe or vessel), they measure flow by very different methods.

Ultrasonic **Doppler** flowmeters measure the Doppler shift resulting from reflecting an ultrasonic beam off the particulates in flowing fluid. The frequency of the transmitted beam is affected by the movement of the particles; this frequency shift can be used to calculate the fluid velocity. For the Doppler principle to work, there must be a high enough density of sonically reflective materials such as solid particles or air bubbles suspended in the fluid. This is in direct contrast to an ultrasonic transit time flowmeter, where bubbles and solid particles reduce the accuracy of the measurement. Due to the dependency on these particles, there are limited applications for Doppler flowmeters. This technology is also known as acoustic Doppler velocimetry.

One advantage of ultrasonic flowmeters is that they can effectively measure the flow rates for a wide variety of fluids, as long as the speed of sound through that fluid is known. For example, ultrasonic flowmeters are used for the measurement of such diverse fluids as liquid natural gas (LNG) and blood.[23] One can also calculate the expected speed of sound for a given fluid; this can be compared to the speed of sound empirically measured by an ultrasonic flowmeter for the purposes of monitoring the quality of the flowmeter's measurements. A drop in quality (change in the measured speed of sound) is an indication that the meter needs servicing.

Anemometers

Anemometers The use of hot wire resistance transducer is to measure the flow rate of fluids by means of measuring velocity of non conducting liquids. In open channels and closed pipes it can be conveniently measured by suitably locating the hot wire filament. Hot wire filament is usually made by wire of platinum or tungsten material. It is suitably mounted in the flow channel by means of a support . The diameter of the wire may vary from 5 µm to 300 µm and the length is about half the diameter of the pipe. A smaller diameter gives a larger resistance per unit length and less inertia but is also less capable of sustaining large fluid pressure. The capability of its withstanding large pressure shocks is decided after actual testing. The hot wire element is centrally located inside the pipe such that the axis of the wire is normal to the direction of fluid flow.

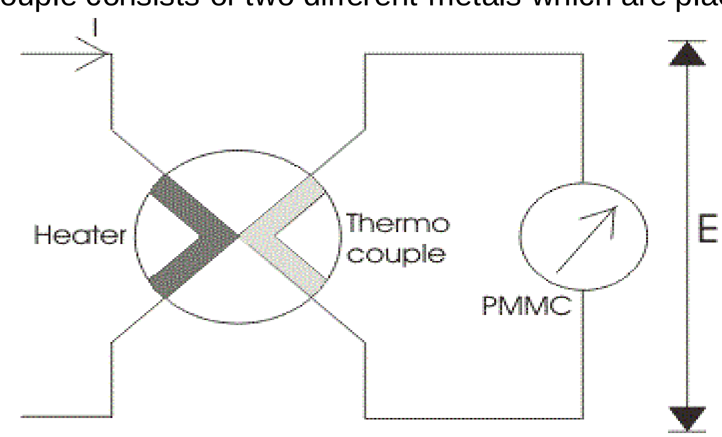
Principle In hot wire anemometers, the filament is heated initially by means of passing an electric current. This heated resistive filament mounted on a tube is exposed to air flow or wind, which is cooled because of fanning effect. Depending on the velocity of air flow, the amount of cooling varies. The resistance of the probe when it is hot is different from that when it is cooled. This difference in resistance or this variation in resistance is converted into a voltage variation and thereby the flow velocity is converted into a voltage variation.

Types of Hot Wire Anemometers Hot wire anemometers are commonly available in two forms : (i) Constant current type and (ii) Constant temperature type. Both

The measuring circuit of the constant-current anemometer can be used for the measurement of steady velocities as well as rapidly fluctuating components such as turbulent components superimposed on an average velocity.

Thermocouples

Basically thermocouple consists of two different metals which are placed in contact with each other as



shown in the diagram.

First part is called the heater element because when the current will flow through this, a heat is produced and thus the temperature will increased at the junction. At this junction an emf is produced which is approximately proportional to the temperature difference of hot and cold junctions.

The emf produced is a DC voltage which is directly proportional to root mean square value of electric current. A permanent magnet moving coil instrument is connected with the second part to read the current passing through the heater. One question must be arise in our mind that why we have used only a permanent magnet coil instrument? Answer to this question is very easy it is because PMMC instrument has greater accuracy and sensitivity towards the measurement of DC value. The **thermocouple type instruments** employ thermocouple in their construction. Thermocouple type instruments can be used for both ac and DC applications. Also thermocouple type of instruments has greater accuracy in measuring the current and voltages at very high frequency accurately.

Now we will look how the temperature difference is mathematically related to generated emf at the junction in thermocouple type of instruments. Let us consider temperature of the heater element be T_a and the temperature of cold metal be T_b. Now it is found that the generated emf at the junction is related to temperature difference as:

$$e = a(T_a - T_b) + b(T_a - T_b)^2$$

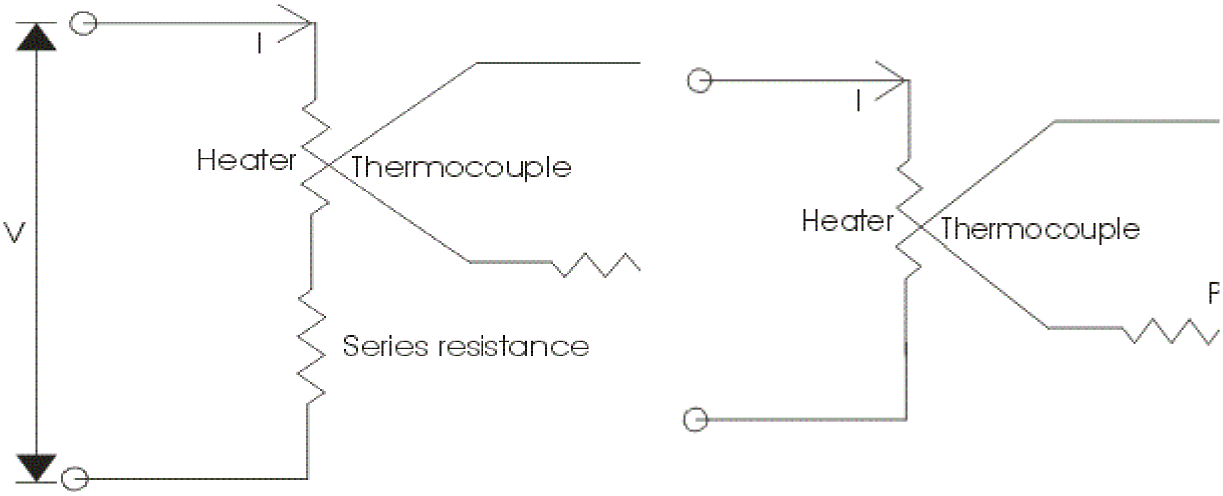
Where a and b are constant whose values completely depends upon the type of metal we are using. The above equation represents parabolic function. The approximated value of a is from 40 to 50 micro volts or more per degree Celsius rise in temperature and value of constant b is very small and can be neglected if the air gap field of permanent magnet moving coil is uniform. Thus we can approximate the above temperature emf relation as $e = a(T_a - T_b)$, here we have assume $b = 0$. The current flowing through the heater coil produces heat as I^2R where I is the root mean square value of current, if we assume the temperature of cold junction is maintained at room temperature then the rise in the temperature of the hot junction will be equal to temperature rise at the junction. Hence we can write $(T_a - T_b)$ is directly proportional to I^2R or we can say $(T_a - T_b) = kI^2R$. Now the deflection angle x in moving coil instrument is equal to; $x = Ke$ or $x = K[a(T_a - T_b)]$ hence we can write $k.K.a.I^2R = k_1I^2$, where k_1 is some constant.

From the above equation we see that the instrument shows the square law response.

Construction of Thermocouple Type Instrument

Now let us look at the construction of Thermocouple type Instruments. Broadly speaking the thermocouple type of instruments consists of two major parts which are written below: (a) Thermo electric elements: The thermocouple type of instruments consists of thermo electric elements which canbe of four types:

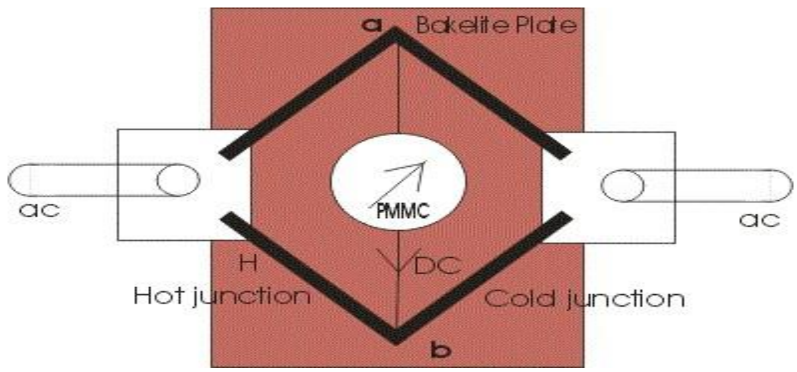
1. **Contact Type:** It has a separate heater which is shown in the diagram.



The action of thermocouple type instruments can be explained briefly as,

- At the junction the electrical energy is being converted to thermal energy in the heater element. A portion of the heat is transferred to the hot junction while most of the heat energy is dissipated away.
- The heat energy which is transferred to hot junction is again converted to electrical due to Seebeckeffect. Only a portion of electrical energy is converted into mechanical energy which is used to produce a deflecting torque. The overall efficiency of the system is low thus the instrument consumes high power. So there is a requirement of highly accurate and sensitive DC instrument.

2. **Non Contact Type:** In non contact type there is insulation between the heating element and the thermocouple i.e. there no direct contact between two. Due to this these instruments are not muchsensitive as compared contact type.
3. **Vacuum Thermo-elements:** These types of instruments are mostly employed for the measurementof electric current at very high frequency of the order of 100 Mega hertz or more as these instruments retain their accuracy even at such high frequency.
4. **Bridge Type:** These bridges are manufactured on the ac ratings usually from 100 mili amperes to 1 amperes. In this two thermocouple are connected to form a bridge which is shown in the figure given



below:

5. There is no requirement of heating element, the electric current which directly passing through thethermocouple raises the temperature which is directly proportional to the I^2R losses. The bridge works on balanced condition at which there will be no current in the arm ab. The connected meterwill show the potential difference between the junctions a and b.

Advantages of Thermocouple Type Instruments

Following are advantages of Thermocouple type of instruments,

1. The thermocouple type of instruments accurately indicates the root mean square value of current andvoltages irrespective of the waveform. There is a wide varieties of range of thermocouple instruments are available in the market.
2. Thermocouple type of instruments give very accurate reading even at high frequency, thus thesetypes of instruments are completely free from frequency errors.
3. The measurement of quantity under these instruments is not affected by stray magnetic fields.

4. These instruments are known for their high sensitivity.
5. Usually for measuring the low value of current bridge type of arrangement is used i.e. ranging from

   0.5 Amperes to 20 Amperes while for measuring the higher value of current heater element isrequired to retain accuracy.

Disadvantages of Thermocouple Type Instruments

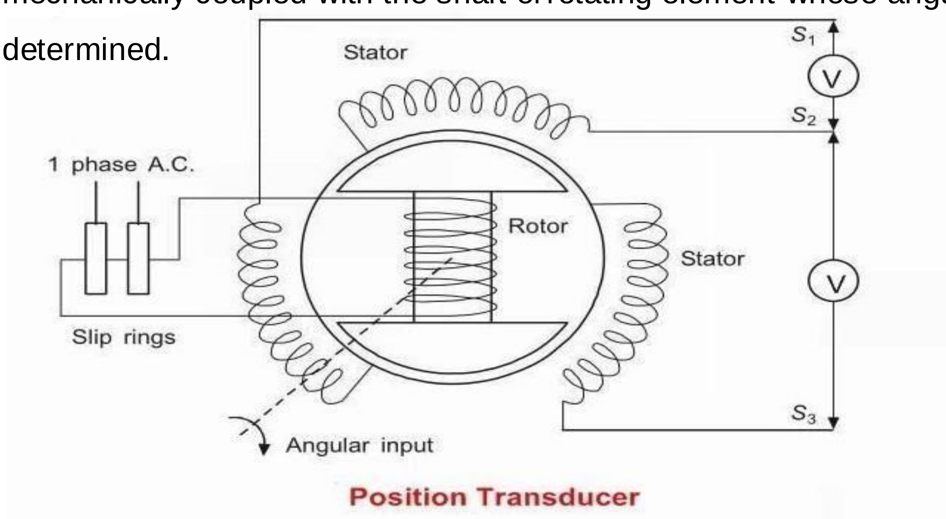
Instead of many advantages these type of instruments posses one disadvantage, The over load capacity of thermocouple type of instrument is small, even fuse is not able to the heater wire because heater wire maybburn out before the fuse blows out.

Synchro Position Transducer Working Principle

We know that Syncro is an inductive device which works on the principle of rotating transformer. Here the term *rotating transformer* means the primary to secondary coupling can be changed by physically changing the relative orientation of the winding. So based on this working principle of syncro we can useit as position transducer.

Construction Of Position Transducer:

Position transducer is one of the basic application of the Synchro. It uses dumb-bell shaped rotor. Singlephase ac supply is given to the rotor of the Synchro. This rotor is mechanically coupled with the shaft ofrotating element whose angular position is to be determined.



Position Transducer

Position Transducer Working Principle:

We know that the stator of the synchro has three windings. These three winding of the stator are connected in star connection. Remaining ends of each winding are taken out to connect them with the voltmeter as shown in the figure. When the angle of the rotor changes the output voltage i.e. the statorvoltages of each winding is given by,

E1 = Eom cosθ sin wt = instantaneous voltage for stator windings S1.

E1 = Eom cos(θ+120) sin wt = instantaneous voltage for stator windings S2.

E1 = Eom cos(θ+240) sin wt = instantaneous voltage for stator

windings S3.Where

- θ= angular position of the rotor
- Eom = peak value of voltage of each winding
- w= 2πf
- f= frequency of the rotor
- t = time in seconds.

All instantaneous voltages are sinusoidal in nature. But they give different values of voltages at differentposition of rotor.

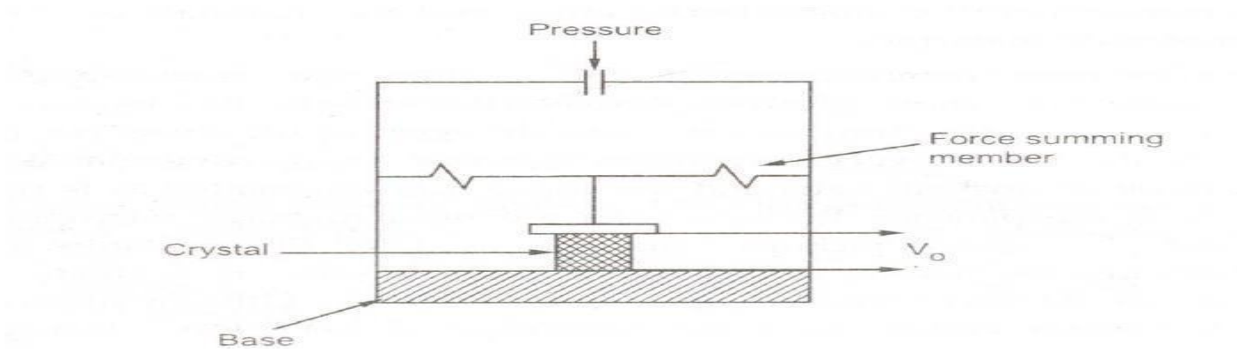
Thus using these three values of stator voltages we can easily measure the position of the rotor. HenceSynchro can be used as a position transducer.

Applications Of Position Transducer:

1) For measuring the angle of the rotating machine like antenna platform.

2) Position transducer can be used as ratary position sensor for aircraft control surfaces

Piezoelectric transducer:

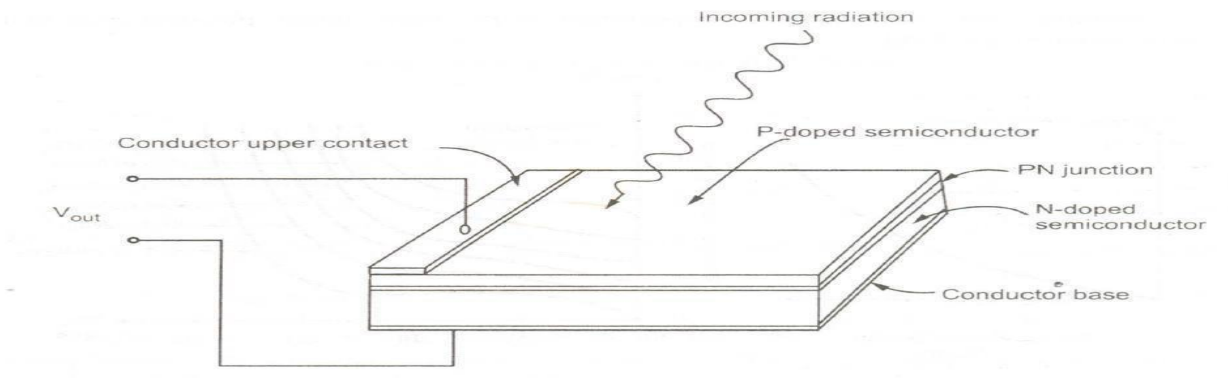


A piezoelectric quartz crystal is hexagonal prism shaped crystal, which has pyramids Jt both ends. This isshown in the Fig. (a). The marking of co-ordinate axes are fixed for such crystals. The axis passing through the end points of pyramids is called optic axis or z axis. The axis passing through corners is called electrical axis or x axis while the aXIs passing through midpoints of opposite sides is called mechanical axis or y axis.

The axes are shown in the fig.

Photovoltaic cell:

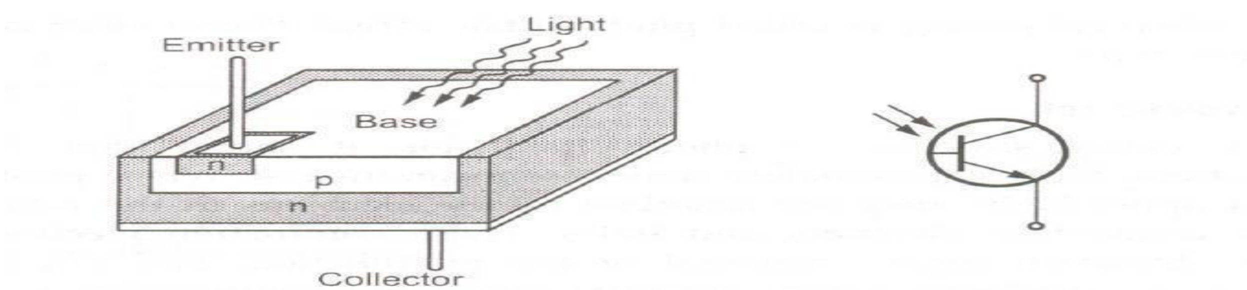
Fig shows structure of photovoltaic cell. It shows that cell is actually a PN-junction diode with appropriately doped semiconductors. When photons strike on the thin p-doped upper layer, they are absorbed by the electrons in the n-layer; which causes formation of conduction electrons and holes. Theseconduction electrons and holes are separated by depletion region potential of the pn junction. When il load is connected across the cell, the depletion region potential causes the photocurrent to flow through the load N



Phototransistor:

The photo transistor has a light sensitive collector to base junction. A lens is used in a transistor packageto expose base to an incident light. When no light is incident, a small leakage current flows from collector to emitter called IeEO, due to small thermal generation. This is very small current, of the orderof nA. This is called a dark current. When the base is exposed to the light, the base current is produced which is proportional to the light intensity. Such photoinduced base current is denoted as I)...The resulting collector current is given by, The structure of a phototransistor is shown in the Fig. (a) while the symbolis shown in the Fig.

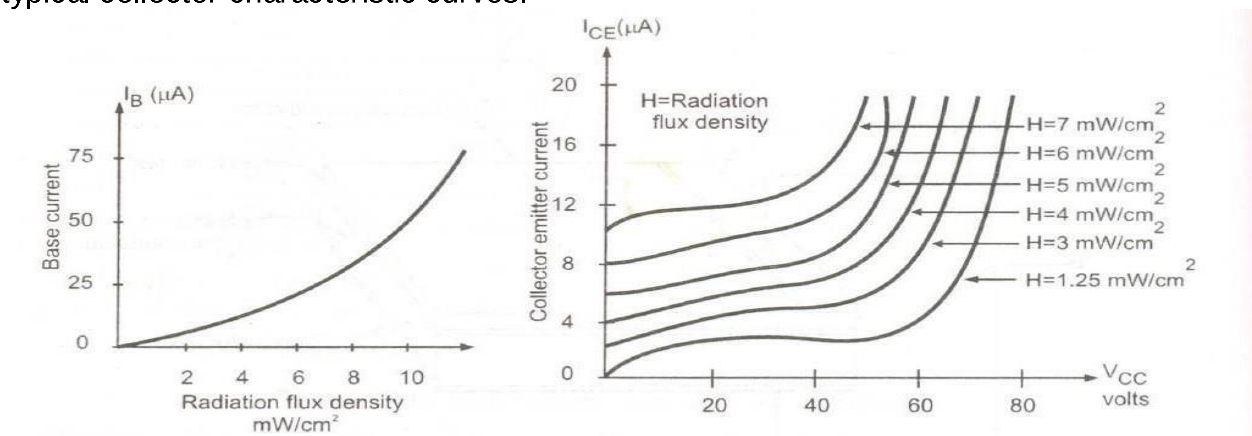
$$\cdot\ I_C \approx h_{fe} I_\lambda$$

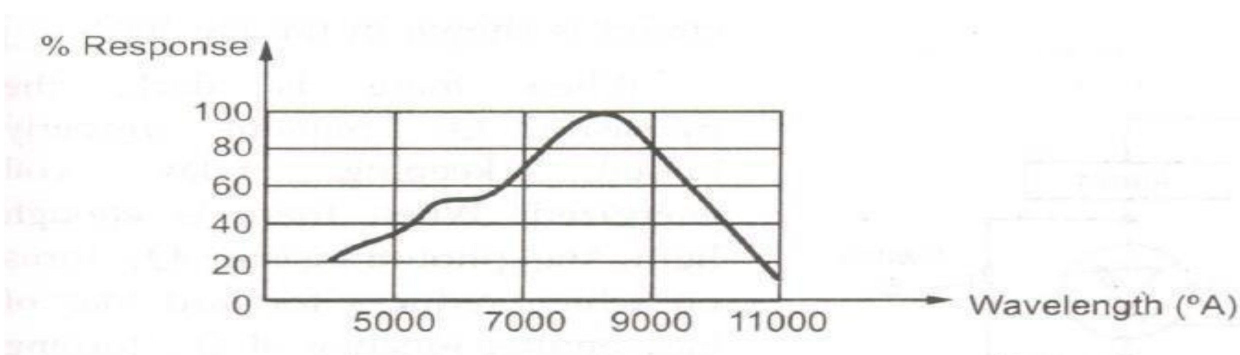


(a) Construction (b) Symbol

To generate more base current proportional to the light, larger physical area of the base is exposed to thelight. The fig .shows the graph of base current against· the radiation flux density measured in mW/ cm2.The Fig. (b) shows the collector characteristics of a phototransistor. As light intensity increases, the base current increases exponentially. Similarly the collector current also increases corresponding to the increase in the light intensity. A phototransistor can be either a two lead or a three

lead device. In a
three lead device, the base lead is brought out so that it can be used as a conventional BJT with or withoutthe light sensitivity feature. In a two lead device, the base is not electrically available and the device use is totally light dependent. The use of phototransistor as a two lead device is shown in the Fig. (a) while theFig. (b) shows the typical collector characteristic curves.





Spectral response

Each curve on the characteristic graph is related to specific light intensity. The collector current level increases corresponding to increase in the light intensity. In most of the applications the phototransistor isused as a two lead device. The phototransistor is not sensitive to all the light but sensitive to light within acertain range. The graph of response against wavelength is called spectral response. A typical spectral response is shown in the Fig.

Unit 6

Temperature measurement

Temperature is a physical quantity that expresses hot and cold. It is the manifestation of thermal energy, present in all matter, which is the source of the occurrence of heat, a flow of energy, when a body is in contact with another that is colder or hotter.

Temperature is measured with a thermometer.

The most common scales are the Celsius scale (formerly called *centigrade*, denoted as °C), the Fahrenheit scale (denoted as °F), and the Kelvin scale (denoted as K),

Kelvin scale is predominantly used for scientific purposes by conventions of the International System of Units (SI).

The lowest theoretical temperature is absolute zero, at which no more thermal energy can be extracted from a body.

The following table shows the temperature conversion formulas for conversions to and from the Celsius scale.

Temperature conversions

	from Celsius	to Celsius
Fahrenheit	[°F] = [°C] × 9⁄5 + 32	[°C] = ([°F] − 32) × 5⁄9
Kelvin	[K] = [°C] + 273.15	[°C] = [K] − 273.15
Rankine	[°R] = ([°C] + 273.15) × 9⁄5	[°C] = ([°R] − 491.67) × 5⁄9
Delisle	[°De] = (100 − [°C]) × 3⁄2	[°C] = 100 − [°De] × 2⁄3
Newton	[°N] = [°C] × 33⁄100	[°C] = [°N] × 100⁄33
Réaumur	[°Ré] = [°C] × 4⁄5	[°C] = [°Ré] × 5⁄4
Rømer	[°Rø] = [°C] × 21⁄40 + 7.5	[°C] = ([°Rø] − 7.5) × 40⁄21

Different methods of temperature measurement

1.Mechanical methods:

Mercury in glass thermometers

This consists of a glass tube of very fine bore joined to a reservoir at the bottom and sealed at the top. A measured quantity of mercury is the enclosed. When the thermometer is heated the mercury expands much more than the glass and is therefore forced to rise up in the tubing A scale is fixed at the side.

Bimetallic Thermometer

Bimetallic strip thermometer Two dissimilar metals are bonded together into what is called a bimetallic strip as shown below. Suppose metal A has a smaller coefficient of thermal expansion than does metal B. As temperature increases, metal B expands more than does metal A, causing the bimetallic strip to curl upwards as shown below. One common application of bimetallic strips is in air-conditioning thermostats, where a bimetallic strip is used as the arm of a switch between electrical contacts. As the room temperature changes, the bimetallic strip bends as discussed above. When the bimetalic strip bends far enough, it makes contact with electrical leads which turn the heat or air conditioning on or off.

Two metals whose coefficient of linear expansion is different are welded and rolled together to the desire thickness. The actual movement of a bimetal is its flexibility with one end fixed, a straight bimetal strip deflects in proportion to its temperature, to the square of its length and inversely with its thickens

2.Electrical method of temperature measurement

Thermocouples

A thermocouple is made up of two dissimilar metals, joined together at one end, that produce a voltage (expressed in millivolts) with a change in temperature. The junction of the two metals, called the sensing junction, is connected to extension wires. Any two dissimilar metals may be used to make a thermocouple.

Principle of Operation

1. When two dissimilar metals are connected together, a small voltage called a thermojunction voltage is generated at the junction. This is called the Peltier effect.
2. If the temperature of the junction changes, it causes voltage to change too, which can be measured by the input circuits of an electronic controller. The output is a voltage proportional to the temperature difference between the junction and the free ends. This is called the Thompson effect.
3. Both of these effects can be combined to measure temperature. By holding one junction at a known temperature (reference junction) and measuring the voltage, the temperature at the sensing junction can be deduced. The voltage generated is directly proportional to the temperature difference. The combined effect is known as the thermo-junction effect or the Seebeck effect.

It is a simple device consisting of a dissimilar metal wires joined at their ends. When an of each wire is connected to a measuring instrument thermocouples becomes an accurate and sensitive temperature measuring device

.

Resistance – Temperature Detectors (RTD)

RTD's are generally used for precise temperature measurement. It consists of a five wire wrapped around an insulator and enclosed in a metal. The most sheath of a resistance thermometer resembles that of bimetallic thermometer
bulb.

PRINCIPLE: "Resistance increases as temperature increase"
Rt. = Ro $(1 + \alpha t)$
Rt. = Resistance of Temperature to measured.
Ro. = Resistance of zero temperature.
α = Co. off of thermal (expansion).
t = Temperature to be measured.

These metals have a positive temperature co-efficient of expansion. Therefore resistance increases as the temperature increases.
Types of material used: (1) Platinum (2) Nickel
These metals have a positive temperature co-efficient of expansion. Therefore resistance increases as the temperature increases.

Thermistors

Thermistors are temperature sensitive semiconductors that exhibit a large change in resistance over a relatively small range of temperature. There are two main types of thermistors

: positive temperature coefficient (PTC) and

negative temperature coefficient (NTC).

NTC thermistors exhibit the characteristic of resistance falling with increasing temperature. These are most commonly used for temperature measurement.

A thermistor is similar to an RTD, but a semiconductor material is used instead of a metal. A thermistor is a solid state device and has larger sensitivity than does an RTD. Unlike RTD's, the temperature-resistance characteristic of a thermistor is non-linear, and cannot be characterized by a single coefficient. Furthermore, unlike RTDs, the resistance of a thermistor decreases with increasing temperature. Thermistors cannot be used to measure high temperatures compared to RTDs. In fact, the maximum temperature of operation is sometimes only 100 or 200°C.

3. RADIATIVE TEMPERATURE MEASURING DEVICES

Two types of radiative measuring devices are:

1. Infrared pyrometers, and

2. Optical pyrometers.

Infrared Pyrometer

Infrared temperature sensors also known as pyrometers or non-contact temperature sensors are used to measure the temperature of an object without contact. This is different from most temperature measurement devices, which require direct contact with the measured media. Noncontact methods of temperature measurement are advantageous when contact methods are impossible or impractical, such as when the target is inaccessible or so hot that contact devices will not survive.

Principle of Operation

Infrared temperature sensors use the principle that any object emits an amount of energy that is a function of its temperature. This function dictates that as the temperature of an object rises, so does the amount of energy it emits. “An infrared temperature sensor determines temperature by measuring the intensity of energy given off by an object.”

Optical Pyrometer

An optical pyrometer is useful for measuring very high temperatures (even flames). The optical pyrometer uses an infrared radiation-sensitive sensor, e.g. a photodiode or a photoresistor, to compare the radiation from the unknown with that of the radiation from an internal incandescent source. The accuracy of the optical pyrometer is very much a function of the emissivity of the device that is radiating the heat. The obvious advantage in using an optical pyrometer at very high temperatures is that the measurement is non-contacting. This approach is very expensive,

Summarizing

The two most common type of temperature sensors are Thermocouples and RTD’s. Although these sensors have overlapping temperature ranges, each has certain application-dependent advantages. These are summarized below. An RTD is the sensor of choice when sensitivity and application flexibility are the most important criteria. When it comes to component cost, an RTD is more expensive than a thermocouple.

Medical Electronics

BY- Sh. Yogender singh(lect.)

Unit-1

Medical Electronics Engineering that advances knowledge in Engineering and Medicine is one of the most extensively growing fields of this era where people prefer the best and precise health service. Medical Science examines treats and cures almost all the disease meeting the need of accuracy by the people by using sophisticated and precise equipments. A medical electronics engineer can be called as the doctor of the equipments without which a doctor, a surgeon or a multi specialty hospital is helpless. He/She is the person, who knows the mechanism, operation and maintenance of these equipments.

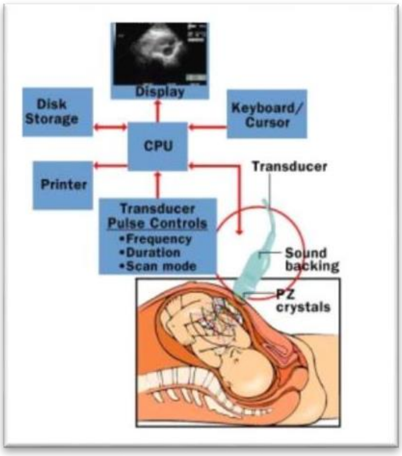
Medical Electronics engineers are the one who design devices and measures that solve medical and health-related problems by combining their knowledge of biology and medicine with engineering principles and practices. Many do research, along with life scientists, chemists, and medical scientists, to develop and evaluate systems and products such as artificial organs, prostheses (artificial devices that replace missing body parts), instrumentation, medical information systems, health management and care delivery systems.

Different type of medical electronics equipment are

- Medical electronic devices include pacemakers, defibrillators, drug-releasing pumps, hearing aids**, and diagnostic equipment for measuring, monitoring, and recording body functions such as heartbeat and brain waves.**

Ultrasonography

- Transducer probe - probe that sends and receives the sound waves
- Central processing unit (CPU) - computer that does all of the calculations and contains the electrical power supplies for itself and the transducer probe
- Transducer pulse controls - changes the amplitude, frequency and duration of the pulses emitted from the transducer probe
- Display - displays the image from the ultrasound data processed by the CPU
- Keyboard/cursor - inputs data and takes measurements from the display
- Disk storage device (hard, floppy, CD) - stores the acquired images
- Printer - prints the image from the displayed data



The transducer probe is the main part of the ultrasound machine. The transducer probe makes the sound waves and receives the echoes. It is, so to speak, the mouth and ears of the ultrasound machine. The transducer probe generates and receives sound waves using a principle called the piezoelectric (pressure electricity) effect, which was discovered by Pierre and Jacques Curie in 1880. In the probe, there are one or more quartz crystals called piezoelectric crystals. When an electric current is applied to these crystals, they change shape rapidly. The rapid shape changes, or vibrations, of the crystals produce sound waves that travel outward. Conversely, when sound or pressure waves hit the crystals, they emit electrical currents. Therefore, the same crystals can be used to send and receive sound waves. The probe also has a sound absorbing substance to eliminate back reflections from the probe itself, and an acoustic lens to help focus the emitted sound waves.

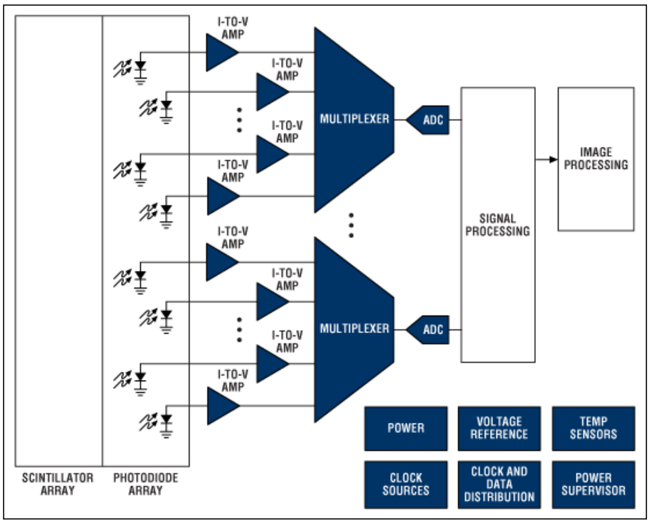
Transducer probes come in many shapes and sizes, as shown in the photo above. The shape of the probe determines its field of view, and the frequency of emitted sound waves determines how deep the sound waves penetrate and the resolution of the image. Transducer probes may contain one or more crystal elements; in multiple-element probes, each crystal has its own circuit. Multiple-element probes have the advantage that the ultrasound beam can be "steered" by changing the timing in which each element gets pulsed; steering the beam is especially important for cardiac ultrasound (see Basic Principles of Ultrasound for details on transducers). In addition to probes that can be moved across the surface of the body, some probes are designed to be inserted through various openings of the body (vagina, rectum, oesophagus) so that they can get closer to the organ being examined (uterus, prostate gland, stomach); getting closer to the organ can allow for more detailed views.

The CPU is the brain of the ultrasound machine. The CPU is basically a computer that contains the microprocessor, memory, amplifiers and power supplies for the microprocessor and transducer probe. The CPU sends electrical currents to the transducer probe to emit sound waves, and also receives the electrical pulses from the probes that were created from the returning echoes. The CPU does all of the calculations involved in processing the data. Once the raw data are processed, the CPU forms the image on the monitor. The CPU can also store the processed data and/or image on disk.

The transducer pulse controls allow the operator, called the ultra-sonographer, to set and change the frequency and duration of the ultrasound pulses, as well as the scan mode of the machine. The commands from the operator are translated into changing electric currents that are applied to the piezoelectric crystals in the transducer probe.

Computed tomography (CT)

medical-imaging systems generate three-dimensional (3-D) images of internal body



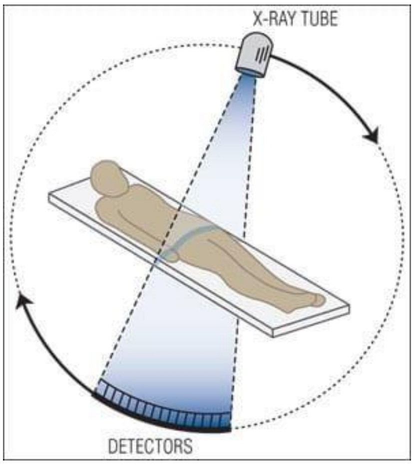
structures using complex x-ray and computer-aided tomographic imaging techniques.

The x-ray images used to generate the tomographic images are generated first by exposing the patient to a fan-shaped x-ray beam and then detecting the projected image on a thin semi-circular, digital x-ray detector. The patient is placed between the source and detector, and the detector is configured with its geometric centre located at the x-ray source. Each image is an x-ray projection of a very thin transverse slice of the body. To collect the multitude of x-ray projections necessary to generate a tomographic CT image, both the x-ray source and detector are rotated about a patient within a supporting gantry. While the source and detector rotate, images are collected and stored. As in a traditional x-ray, the signal levels in the image slice represent the relative radio density of the patient along a line from the x-ray source to the corresponding pixel location.

To improve image-capture times and resolution, manufacturers utilize multislice CT imaging techniques. Instead of a single 2D detector array which provides only a single image slice, multislice imaging uses a 3-D array. The added imaging dimension allows the system to generate multiple slices in parallel. Photodetector arrays used in CT imaging have as many as 1000 detectors in the long dimension along the semi-circular detector arch; 16 or more detectors are positioned in the shorter dimension tangential to the arch. The number of detectors in the short dimension determines the number of available image slices.

The patient is exposed to a fan-shaped x-ray beam and the projected image is detected on a thin, semi-circular digital x-ray detector.

Modern CT imaging systems can also generate images in any plane within the body by using a technique called spiral CT. In a spiral-CT system the patient is slowly moved into the centre of the gantry while the x-ray source and detector rotate about the patient. Very-high-speed computers are necessary to process the images collected in this manner. Sophisticated tomographic imaging techniques are used to produce the required image.



<u>X-Ray Detection</u>

Early CT imaging systems accomplished x-ray detection using both scintillation crystals and photo-multiplier tubes. The scintillation crystals converted x-rays to light and the photomultiplier tubes converted these light signals to a usable electrical signal. Modern CT systems now employ more sophisticated scintillation crystal materials and solid-state photodetector diodes for this purpose.

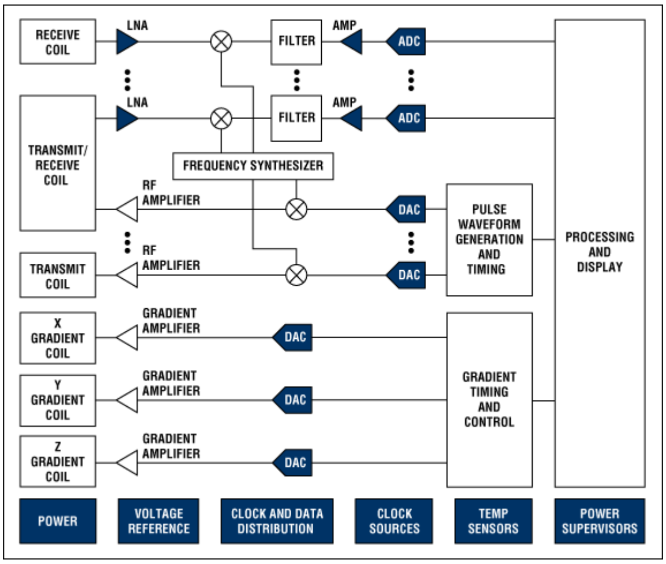
The output from each photodiode is a current proportional to the light striking the diode. These currents can be directly converted to a voltage by a low-noise trans impedance amplifier (TIA), or integrated over time using a capacitor or active integrator op-amp circuit to produce a voltage output. Integration of the current from each diode can be accomplished in multiple ways. Capacitance in the photodiode detector array itself can be used for this purpose. The signals from these capacitors are multiplexed using FET switches in the diode-array detector. The signals are then routed to the digital acquisition system (DAS) which amplifies and converts the signals to a digital format using high-resolution analog-to-digital converters (ADCs). An alternative method routes the signals from every photodiode to an integrator in the DAS. In these implementations, the integrated current signals are converted to a voltage, sampled at the same time, and multiplexed into the input of an ADC.

Tomographic Imaging

The resulting x-ray image data set is converted to an image by the image processor. The image processor is typically a very-high-speed computer which performs the massive calculations required for the tomographic image reconstruction. The resulting image will commonly have a very large dynamic range (i.e., 16-bit grayscale images). Further image processing is necessary to map this large dynamic range most effectively into the limited visible display range.

MRI scan

Magnetic resonance imaging (MRI) systems provide highly detailed images of tissue in the body. The systems detect and process the signals generated when hydrogen atoms, which are abundant in tissue, are placed in a strong magnetic field and excited by a resonant



magnetic excitation pulse.

Magnetic resonance imaging (MRI) system.

Proper stimulation by a resonant magnetic or RF field at the resonant frequency of the hydrogen nuclei can force the magnetic moments of the nuclei to partially, or completely, tip into a plane perpendicular to the applied field. When the applied RF-excitation field is removed, the magnetic moments of the nuclei process in the static field as they realign. This realignment generates an RF signal at a resonant frequency determined by the magnitude of the applied field. This signal is detected by the MRI imaging system and used to generate an image.

Static Magnetic Field

MRI imaging requires the patient to be placed in a strong magnetic field in order to align the hydrogen nuclei. There are typically three methods to generate this field: fixed magnets, resistive magnets (current passing through a traditional coil of wire), and super-conducting magnets. Fixed magnets and resistive magnets are generally restricted to field strengths below 0.4T and cannot generate the higher field strengths typically necessary for high-resolution imaging. As a result, most high-resolution imaging systems use super-conducting magnets. The super-conducting magnets are large and complex; they need the coils to be soaked in liquid Helium to reduce their temperature to a value close to absolute zero.

The magnetic fields generated by these methods must not only be strong, but also highly uniform in space and stable in time. A typical system must have less than 10ppm variation over the imaging area. To achieve this accuracy, most systems generate weaker static magnetic fields using specialized shim coils to "shim" or "tweak" the static field from the super conductor and thereby correct for field inaccuracies.

Gradient Coils

To produce an image, the MRI system must first stimulate hydrogen nuclei in a specific 2D image plane in the body, and then determine the location of those nuclei within that plane as they process back to their static state. These two tasks are accomplished using gradient coils which cause the magnetic field within a localized area to vary linearly as a function of spatial location.

RF Receiver

An RF receiver is used to process the signals from the receiver coils. Most modern MRI systems have six or more receivers to process the signals from multiple coils. The signals range from approximately 1MHz to 300MHz, with the frequency range highly dependent on applied-static magnetic field strength. The bandwidth of the received signal is small, typically less than 20kHz, and dependent on the magnitude of the gradient field.

A traditional MRI receiver configuration has a low-noise amplifier (LNA) followed by a mixer. The mixer mixes the signal of interest to a low-frequency IF frequency for conversion by a high-resolution, low-speed, 12-bit to 16-bit analog-to-digital converter (ADC). In this receive architecture, the ADCs used have relatively low sample rates below 1MHz. Because of the low-bandwidth requirements, ADCs with higher 1MHz to 5MHz sample rates can be used to convert multiple channels by time-multiplexing the receive channels through an analog multiplexer into a single ADC.

With the advent of higher-performance ADCs, newer receiver architectures are now possible. High-input-bandwidth, high-resolution, 12-bit to 16-bit ADCs with samples rates up to 100MHz can also be used to directly sample the signals, thereby eliminating the need for analog mixers in the receive chain.

Transmitter

The MRI transmitter generates the RF pulses necessary to resonate the hydrogen nuclei. The range of frequencies in the transmit excitation pulse and the magnitude of the gradient field determine the width of the image slice. A typical transmit pulse will produce an output signal with a relatively narrow ±1kHz bandwidth. The time-domain waveform required to produce this narrow frequency band typically resembles a traditional sync function. This waveform is usually generated digitally at baseband and then up converted by a mixer to the appropriate centre frequency. Traditional transmit implementations require relatively low-speed digital-to-analog converters (DACs) to generate the baseband waveform, as the bandwidth of this signal is relatively small.

Image Signal Processing

Both frequency and phase data are collected in what is commonly referred to as the k-space. A two-dimensional Fourier transform of this k-space is computed by a display processor/computer to produce a grey-scale image.

UNIT -2
(MEDICAL ELECTRONICS)

A **biosignal** is any signal in living beings that can be continually measured and monitored. The term biosignal is often used to refer to bioelectrical signals, but it may refer to both electrical and non-electrical signals

The electrical, chemical, and mechanical activity that occurs during this **biological** event often produces **signals** that can be measured and analyzed.

Bioelectrical signals are **generated** from the complex self-regulatory system and can be measured through changes in electrical potential across a cell or an organ

Bioelectric potentials are generated at cellular level and the **source** of these **potentials** is ionic in **nature**. The prominent ions are K^+, Na^+, and Cl^-. Electronic **potential** is seen in commonly used cells for example the Galvanic cel.

Bioelectrodes:-

Bioelectric events have to be picked up from the surface of the body before they can be put to the amplifier for subsequent record or display. This is done by use of electrodes. The potentials produced at different points are measured by placing electrodes at various points on the body.

They carry the currents produced due to potential differences to instrumentation amplifiers, where the signals are amplified and further processed by signal processing systems.

When a measurement is made outside the body by placing surface electrodes it is called **in vitro** measurement.

When measurement is made by inserting a needle electrode inside the tissue, it is called **in vivo** measurement.

Bioelectrodes can be classified as:

- **Surface electrodes:** These electrodes pick up potentials from the surface of the tissue.
- **Deep seated electrodes:** These electrodes are inserted inside a live tissue or cell.

Why electrolyte or a jelly is applied to the area where the electrodes make contact:-

Due to the movement of electrodes, noise signals are generated. They are referred to as **artifacts.** To avoid artifacts and establish a low impedance path, an electrolyte or a jelly is applied to the area where the electrodes make contact. The skin is first cleaned by rubbing alcohol; all hair at that portion is removed. A jelly is then applied and the electrode is placed.

Properties of Bioelectrodes

Bioelectrodes should possess the following properties:

- They should be good conductors
- They should have low impedance

- They should not polarize when a current flows through them
- They should establish a good contact with the body and not cause motion
- Potentials generated at the metal electrolyte (jelly) surface should be low.
- They should not cause itching, swelling or discomfort to the patient for example the metal should not be toxic
- They should be mechanically rugged
- They should be chemically inert
- They should be easy clean.

The Skin Contact Impedance

The outer layer of the skin has large impedance which is much greater than the electrical impedance of body tissue beneath the skin. The outer layer skin is responsible for the bulk of the skin contact impedance and therefore, a careful skin preparation is essential in order to obtain best results.

Types of Electrodes used in Biomedical Measurements

The three basic types of biopotential electrodes used in biomedical measurements are:

1. Microelectrodes
2. Skin surface electrodes
3. Needle electrodes

Microelectrodes

These electrodes are designed to measure bioelectric potentials near or within the cell. These electrodes are much smaller in cross-sectional area as compared to the size of the cell in which they are to be inserted so that their penetration should not damage the cell.

A cell is rarely larger than 500 microns, therefore these microelectrodes should have tip dimension of about 5 microns and the tip should be strong enough to penetrate the cell without damage.

- Microelectrodes can be of two types :-

Metal

Micropipette

Body Surface Electrodes

These electrodes are designed to measure ECG, EEG, and EMG potentials from the surface of the skin; hence these electrodes are least traumatic. Body Surface electrodes are designed and used to measure bioelectric potentials from the surface of the body. They are available in many forms and sizes. The larger electrodes are usually used for sensing of ECG

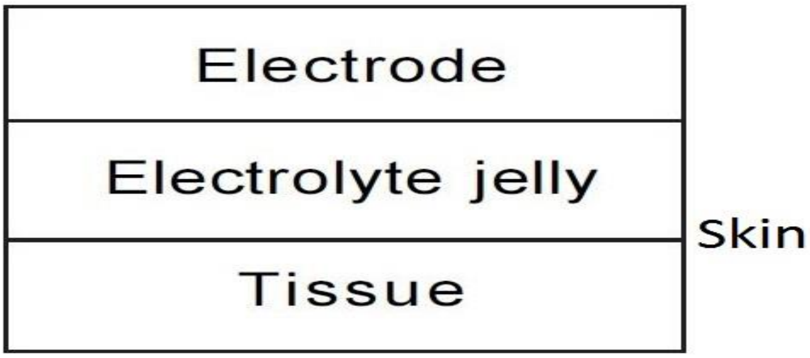
Needle Electrodes

They are generally made of stainless steel. These electrodes are designed to penetrate the skin surface of the body to some depth to record EEG potentials of a region of the brain or EMG potentials of a muscle.

The figure below shows how the electrodes make contact with the skin surface:

OR

ELECTRODE TISSUE INTERFACE



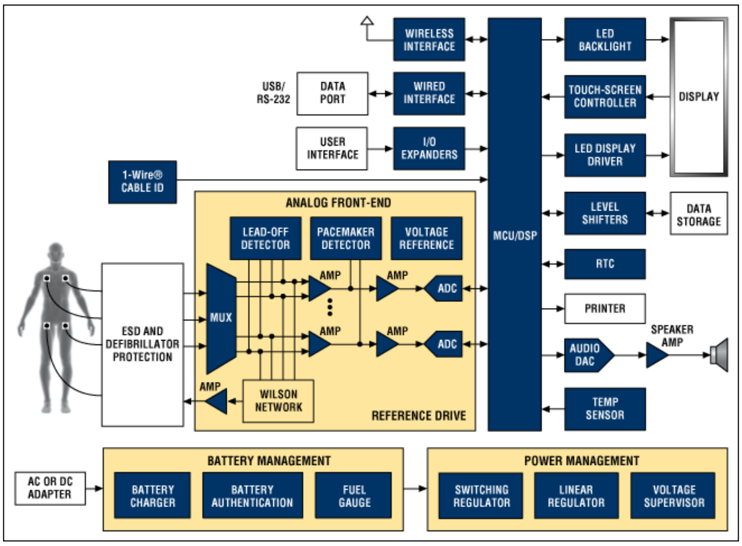
Unit-3,4 & 5

ECG

ll ECGs pick up heart signals through electrodes connected externally to specific locations on the body. The heart signals are generated by the body and have amplitudes of a few millivolts. The specific locations of the electrodes allow the heart's electrical activity to be viewed from different angles, each of which is displayed as a channel on the ECG printout. Each channel represents the differential voltage between two of the electrodes, or the differential voltage between one electrode and the average voltage from several electrodes. The different combinations of electrodes allow more channels to be displayed than there are electrodes. The channels are commonly referred to as "leads," so a 12-lead ECG device has 12 separate channels displayed graphically. The number of leads varies from 1 to 12 depending on the application. Unfortunately, the wires running to the electrodes are occasionally referred to as leads as well. This can create confusion, as a 12-lead (12-channel) ECG device only requires 10 electrodes (10 wires), so be careful of the context in which "lead" is used. In addition to the biological signals, most ECGs also detect two manmade signals. The most important of these signals comes from implanted pacemakers and is referred to simply as "pace." The pace signal is relatively short, tens of microseconds to a couple of milliseconds, with an amplitude ranging from a few millivolts to nearly a volt. Often, the ECG must detect the presence of a pace signal while simultaneously preventing it from distorting the signals from the heart.

The second manmade signal is for detecting "lead-off," which is when an electrode is making poor electrical contact. Many ECG devices must provide an alert when this poor contact occurs. Therefore, the ECG device generates a signal to measure the impedance between the electrode and the body for detecting a lead-off occurrence. The measurement may be AC, DC, or both. In some ECG devices, respiration rate is also detected by analysing the impedance from the lead-off measurement. Lead-off detection is continuous and should not interfere with accurate measurement of the heart signals.

Modern heart rate monitors usually comprise two elements: a chest strap transmitter



[needs update] and a wrist receiver (which usually is a smartwatch). In early plastic straps, water or liquid was required to get good performance. Early units have used conductive smart fabric with built-in microprocessors that analyse the electrical activity to determine the heart rate similar to an EKG. More recent devices use optics to measure heart rate by which measures changes in blood flow by shining a light from an LED through the skin and measuring how it scatters off blood vessels.

Pulse rate monitor

Strapless heart rate monitors (often referred to as "wearables") now allow the user to just touch two sensors on a smartwatch display for a few seconds to view heart rate data. These are popular for comfort and ease of use, though they don't give as much detail as monitors that use a chest strap. Some models of these variations of heart rate monitors use either infrared light or red visible light to measure the heart rate, as opposed to two or more electrodes. In addition to measuring the heart rate, devices using this technology are able to measure blood oxygen saturation (SpO2)

More advanced models offer measurements of heart rate variability, activity, and breathing rate to assess parameters relating to a subject's fitness. Sensor fusion algorithms allow these monitors to detect core temperature and dehydration.

Another style of heart rate monitor replaces the plastic around-the-chest strap with fabric sensors - the most common of these is a sports bra that includes sensors in the fabric.

In old versions, when a heartbeat is detected a radio signal is transmitted, which the receiver uses to determine the current heart rate. This signal can be a simple radio pulse or a unique coded signal from the chest strap (such as Bluetooth, ANT, or other low-power radio link); the latter prevents one user's receiver from using signals from other nearby transmitters (known as cross-talk interference).

Newer versions include a microprocessor, which simultaneously monitors heart rate, SpO2, and other parameters. These may include sensors such as accelerometers, gyroscopes, and GPS to detect speed, location and distance eliminating the need for ankle worn devices.

There are a wide number of receiver designs, with various features. These include average heart rate over exercise period, time in a specific heart rate zone, calories burned, breathing rate, built-in speed and distance, and detailed logging that can be downloaded to a computer. The receiver can be built into a smartwatch or smartphone. Bracelets with integrated sensors work optically, and have poor accuracy.

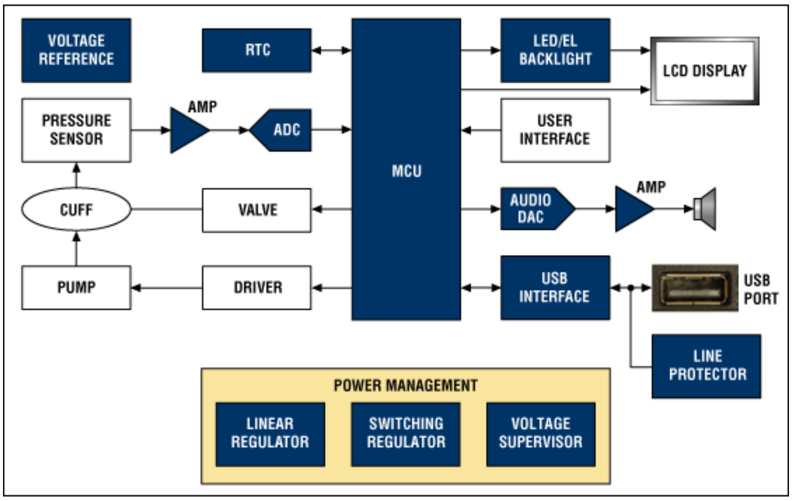
Blood pressure monitor

A blood pressure monitor, or sphygmomanometer, uses an inflatable air-bladder cuff and a listening device or pressure sensor to measure blood pressure in an artery. This monitoring can be performed by using either of two methods: a manually inflated cuff with a stethoscope for listening to arterial wall sounds (the auscultatory method), or a blood pressure monitor that contains a pressure sensor for sensing arterial wall vibrations (the oscillometric method).

Automatic Monitor Types

The two main types of automatic blood pressure monitors are upper-arm and wrist models. The upper-arm model has a cuff that is placed on the upper arm; the cuff is connected by a tube to the monitor that rests on a surface near the arm. The wrist model is smaller and the entire unit wraps around the wrist—this is a much more space-critical design. Some upper-arm models require manual inflation of the cuff, but most upper-arm and all wrist models are fully automatic.

Measurement Techniques



An automatic blood pressure monitor inflates a cuff surrounding an arm with sufficient pressure to prevent blood flow in the local main artery. This pressure is gradually released until the moment that the blood begins to flow through the artery, the measurement of which determines the systolic pressure. Pulse rate is also sensed at this time. The measurement taken when the blood flow is no longer restricted determines the diastolic pressure. This complete measurement cycle is performed automatically with a pump, cuff, valve, and pressure sensor.

The signal from the pressure sensor is conditioned with an op-amp circuit or by an instrumentation amplifier before data conversion by an analog-to-digital converter (ADC). The systolic pressure, diastolic pressure, and pulse rate are then calculated in the digital domain using a method appropriate for the type of monitor and sensor utilized. The resulting systolic, diastolic, and pulse-rate measurements are displayed on a liquid-crystal display (LCD), time/date-stamped, and stored in non-volatile memory.

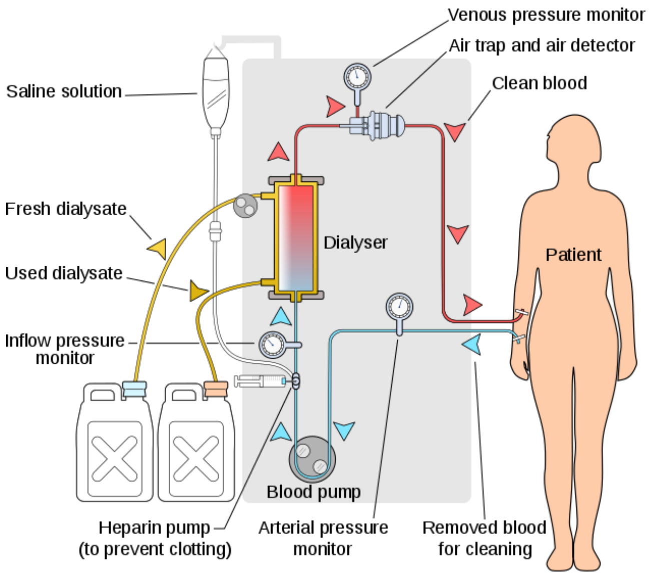
Haemodialysis

Haemodialysis, commonly called kidney dialysis or simply dialysis, is a process of purifying the blood of a person whose kidneys are not working normally. This type of dialysis achieves the extracorporeal removal of waste products such as creatinine and urea and free water from the blood when the kidneys are in a state of kidney failure.

The principle of haemodialysis is the same as other methods of dialysis; it involves diffusion of solutes across a semipermeable membrane. Haemodialysis utilizes counter current flow, where the dialysate is flowing in the opposite direction to blood flow in the extracorporeal circuit. Counter-current flow maintains the concentration gradient across the membrane at a maximum and increases the efficiency of the dialysis.

Fluid removal (ultrafiltration) is achieved by altering the hydrostatic pressure of the dialysate compartment, causing free water and some dissolved solutes to move across the membrane along a created pressure gradient.

The dialysis solution that is used may be a sterilized solution of mineral ions. Urea and other waste products, potassium, and phosphate diffuse into the dialysis solution. However, concentrations of sodium and chloride are similar to those of normal plasma to prevent loss. Sodium bicarbonate is added in a higher concentration than plasma to correct blood acidity. A small amount of glucose is also commonly used.



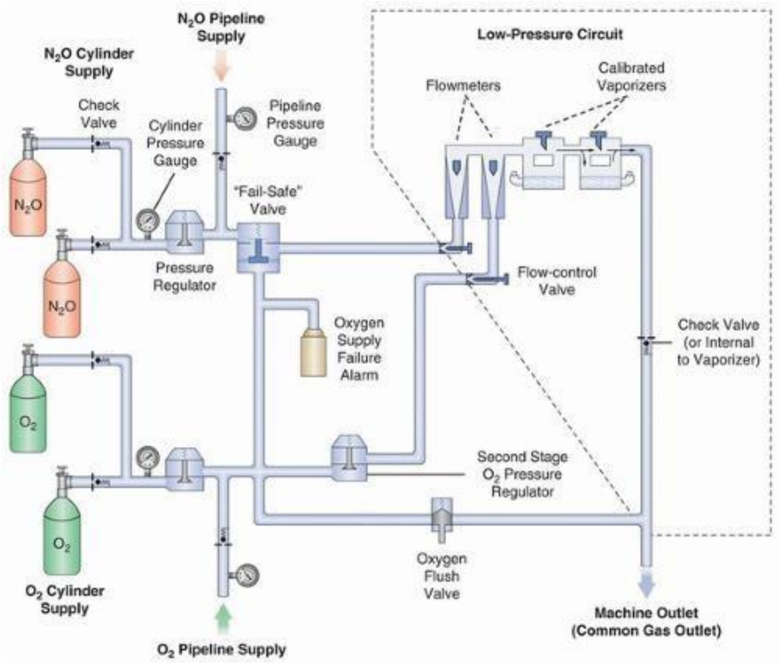
Dialyzer

The dialyzer is the piece of equipment that actually filters the blood. Almost all dialyzers in use today are of the hollow-fibre variety. A cylindrical bundle of hollow fibres, whose walls are composed of semi-permeable membrane, is anchored at each end into potting compound (a sort of glue). This assembly is then put into a clear plastic cylindrical shell with four openings. One opening or blood port at each end of the cylinder communicates with each end of the bundle of hollow fibres. This forms the "blood compartment" of the dialyzer. Two other ports are cut into the side of the cylinder. These communicate with the space around the hollow fibres, the "dialysate compartment." Blood is pumped via the blood ports through this bundle of very thin capillary-like tubes, and the dialysate is pumped through the space surrounding the fibres. Pressure gradients are applied when necessary to move fluid from the blood to the dialysate compartment.

Anaesthesia Machine

Anaesthesia machine (US English) or Boyle's machine is used independently by physician anaesthesiologists and nurse anaesthetists. Anaesthesiologist assistants also use anaesthesia machines under the direct supervision of physician anaesthesiologists.

Anaesthesia machines are used to support the administration of anaesthesia. The most common type of anaesthetic machine in use in the developed world is the continuous-flow anaesthetic machine, which is designed to provide an accurate and continuous supply of medical gases (such as oxygen and nitrous oxide), mixed with an accurate concentration of anaesthetic vapour (such as isoflurane), and deliver this to the patient at a safe pressure and flow. Modern machines incorporate a ventilator, suction unit, and patient monitoring devices.



<u>A modern anaesthesia machine includes the following components</u>:

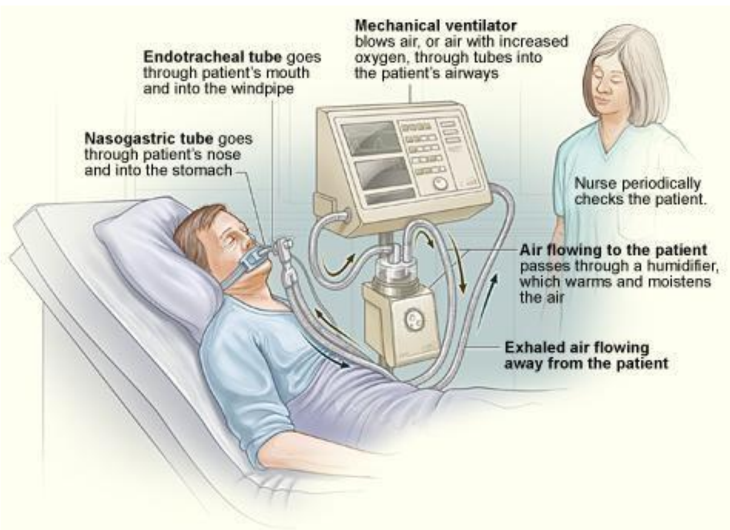
- Connections to piped hospital oxygen, medical air, and nitrous oxide.
- Reserve gas cylinders of oxygen, air, and nitrous oxide attached via a specific yoke with a Bodok seal.
- A high-flow oxygen flush which provides pure oxygen at 30-75 litres/minute
- Pressure gauges, regulators and 'pop-off' valves, to protect the machine components and patient from high-pressure gases
- Flow meters (rotameters) for oxygen, air, and nitrous oxide, low Flow meters' oxygen nitrous oxide
- Updated vaporizers to provide accurate dosage control when using volatile anaesthetics such as isoflurane and sevoflurane
- An integrated ventilator to properly ventilate the patient during administration of anaesthesia
- A manual ventilation bag in combination with an Adjustable Pressure Limiting (APL) valve
- Systems for monitoring the gases being administered to, and exhaled by the patient

- Systems for monitoring the patient's heart rate, ECG, blood pressure and oxygen saturation, in some cases with additional options for monitoring end-tidal carbon dioxide and temperature
- Breathing circuits, circle attachment, or a Bain's breathing system

Ventilator

A medical ventilator (or simply ventilator in context) is a mechanical ventilator, a machine designed to move breathable air into and out of the lungs, to provide breathing for a patient who is physically unable to breathe, or breathing insufficiently. While modern ventilators are computerized machines, patients can be ventilated with a simple, hand-operated bag valve mask. Ventilators are chiefly used in intensive care medicine, home care, and emergency medicine (as standalone units) and in anaesthesia (as a component of an anaesthesia machine).

Function: - In its simplest form, a modern positive pressure ventilator consists of a compressible air reservoir or turbine, air and oxygen supplies, a set of valves and tubes, and a disposable or reusable "patient circuit". The air reservoir is pneumatically compressed several times a minute to deliver room-air, or in most cases, an air/oxygen mixture to the patient. If a turbine is used, the turbine pushes air through the ventilator, with a flow valve adjusting pressure to meet patient-specific parameters. When over pressure is released, the patient will exhale passively due to the lungs' elasticity, the exhaled air being released usually through a one-way valve within the patient circuit called the patient manifold. Ventilators may also be equipped with monitoring and alarm systems for patient-related parameters (e.g. pressure, volume, and flow) and ventilator function (e.g. air leakage, power failure, mechanical failure), backup batteries, oxygen tanks, and remote control. The pneumatic system is nowadays often replaced by a computer-controlled turbo pump. Modern ventilators are electronically controlled by a small embedded system to allow exact



adaptation of pressure and flow characteristics to an individual patient's needs. Fine-tuned ventilator settings also serve to make ventilation more tolerable and comfortable for the patient. In Canada and the United States, respiratory therapists are responsible for tuning these settings, while biomedical technologists are responsible for the maintenance. The patient circuit usually consists of a set of three durables, yet lightweight plastic tubes, separated by function (e.g. inhaled air, patient pressure, exhaled air). Determined by the type of ventilation needed, the patient-end of the circuit may be either non-invasive or invasive. Non-invasive methods, which are adequate for patients who require a ventilator only while sleeping and resting, mainly employ a nasal mask. Invasive methods require intubation, which for long-term ventilator dependence will normally be a tracheotomy cannula, as this is much more comfortable and practical for long-term care than is larynx or nasal intubation.

Life-critical system: -Because failure may result in death, mechanical ventilation systems are classified as a life-critical system, and precautions must be taken to ensure that they are highly reliable, including their power-supply. Mechanical ventilators are therefore carefully designed so that no single point of failure can endanger the patient. They may have manual backup mechanisms to enable hand-driven respiration in the absence of power (such as the mechanical ventilator integrated into an anaesthetic machine). They may also have safety valves, which open to atmosphere in the absence of power to act as an anti-suffocation valve for spontaneous breathing of the patient. Some systems are also equipped with compressed-gas tanks, air compressors, and/or backup batteries to provide ventilation in case of power failure or defective gas supplies, and methods to operate or call for help if their mechanisms or software fail.

Heart Lung Machine

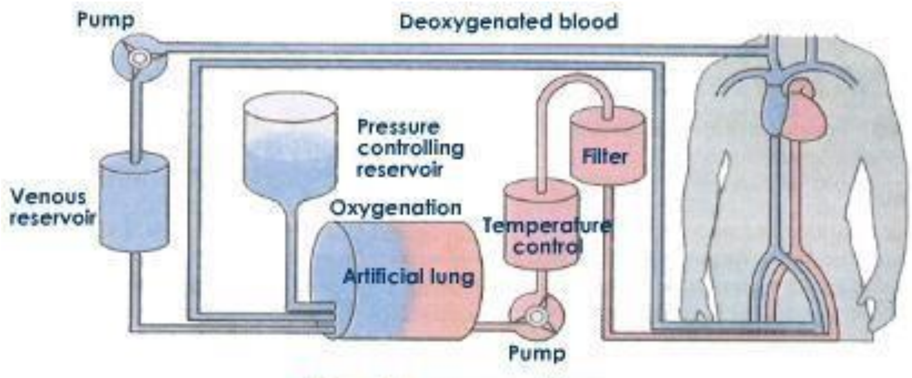
The heart–lung machine is a system which takes over the function of the heart and the lungs with sufficient safety to maintain life while the heart is stopped or opened to allow surgery on the coronary arteries or the heart valves, or to allow repair of congenital abnormalities.

While in theory it is only necessary to bypass the function of the heart, it soon became apparent that in practice it is simpler to bypass the function of both the heart and the lungs. The main components of a heart–lung machine are a pump (to provide the driving force to the blood in the arterial system), an oxygenator (for exchange of oxygen and carbon dioxide), and a heat exchanger (to allow control of temperature of the body). The connecting tubing and filter are other components of the heart–lung bypass circuit.

Venous blood is siphoned from the body via a tube in the right atrium of the heart, or via two tubes in the major veins which converge on the heart. It is pumped through the oxygenator and heat exchanger, and returned via a plastic tube into the arterial system of the body — usually at the upper portion of the ascending aorta.

The design of pump which is in most common use today is the roller pump — a simple rotating arm carrying rollers which compress a loop of polymeric tubing against a solid surface. Speed of rotation of the roller-bearing arm is controlled to allow a pumping rate similar to that of the normal heart at rest (about 2.4 litres/min/m2 body surface — or typically about 5 litres/min in an adult).

There are two main types of oxygenator in use at present. 'Bubble oxygenators' expose the passing blood to a stream of gaseous bubbles composed of 95% oxygen and 5% carbon dioxide. Gas exchange with the blood occurs on the surface of the bubbles and results in reasonably normal levels of oxygenation of the blood and maintains carbon dioxide in the



Heart-lung machine

normal physiological range. The bubble oxygenator has a sponge-like filter and reservoir to enable gaseous bubbles to be removed from the oxygenated blood before it is pumped back to the body.

Membrane oxygenators consist of a series of fine tubes which allow diffusion of oxygen and carbon dioxide between the blood flowing through them and the ventilating gas surrounding them (or vice versa).

The oxygenator also combines with a heat exchanger — a system of tubes through which the blood passes, surrounded by circulating water at controlled temperature. This allows the blood temperature to be maintained (counteracting the heat loss during the passage of blood through the heart–lung machine). It also allows deliberate cooling and subsequent rewarming of the blood, giving the surgeon the option of reducing, or even stopping, the circulation of the blood around the body for a period of time with safety, because the oxygen requirement of the body is reduced by hypothermia.

The connecting tubes, the oxygenator, and the pump tubing are all filled with a physiologically compatible fluid (priming fluid) prior to final connection with the circulation of the body. Avoidance of air bubbles in the heart–lung circuit is of vital importance. Exposure of blood to the foreign surfaces of the heart–lung machine initiates the natural clotting mechanisms of the body, and this must be inhibited by giving the drug heparin to the patient before allowing the circulation to be taken over by the heart–lung machine. Normal blood clotting is restored after the operation by the administration of protamine, which neutralizes the heparin.

Muscle Simulator

Electrical muscle stimulation (EMS), also known as neuromuscular electrical stimulation (NMES) or electromyostimulation, is the elicitation of muscle contraction using electric impulses. EMS has received an increasing amount of attention in the last few years for many reasons: it can be utilized as a strength training tool for healthy subjects and athletes; it could be used as a rehabilitation and preventive tool for partially or totally immobilized patients; it could be utilized as a testing tool for evaluating the neural and/or muscular function in vivo; it could be used as a post-exercise recovery tool for athletes.[1] The impulses are generated by a device and are delivered through electrodes on the skin near to the muscles being stimulated. The electrodes are generally pads that adhere to the skin. The impulses mimic the action potential that comes from the central nervous system, causing the muscles to contract.

Electronic muscle stimulation (EMS) may help you to strengthen weak muscles. There are several theories on how EMS may assist muscle strengthening.

One potential reason is that when you maximally contract a muscle, at best, only 30% of all your muscle fibres are in a state of contraction. The remaining 70% are dormant and awaiting recruitment when the contracting fibres fatigue. With EMS you can potentially electrically stimulate these resting muscle fibres to improve their strength. Clinically, EMS appears to be more effective when the muscles are very weak and you have difficulty performing normal anti-gravity exercises.

Another reason that EMS potentially works is via an improvement in the recruitment of nerve conduction rates. Explained simply, it takes approximately 10000 repetitions for your brain to learn how to quickly send a message to your muscles via the quickest nerve pathways. This contraction pattern becomes your "memory engram". The more frequent your muscle is recruited the better your body becomes at finding the quickest way to recruit that muscle. EMS can potentially provide you with repeated contractions to accelerate this learning process.

To achieve your best outcome, we recommend that you seek professional advice on how to best utilise your EMS machine from your local physiotherapist who specialises in EMS muscle retraining.

Hearing Aids

A hearing aid is a device designed to improve hearing by making sound audible to a person with hearing loss. Hearing aids are classified as medical devices in most countries, and regulated by the respective regulations. Small audio amplifiers such as PSAPs or other plain sound reinforcing systems cannot be sold as "hearing aids".

Early devices, such as ear trumpets or ear horns, were passive amplification cones designed to gather sound energy and direct it into the ear canal. Modern devices are computerised electroacoustic systems that transform environmental sound to make it audible, according to audio metrical and cognitive rules. Modern devices also utilize sophisticated digital signal processing to try and improve speech intelligibility and comfort for the user. Such signal processing includes feedback management, wide dynamic range compression, directionality, frequency lowering, and noise reduction.

Modern hearing aids require configuration to match the hearing loss, physical features, and lifestyle of the wearer. This process is called "fitting" and is performed by audiologists. The amount of benefit a hearing aid delivers depends in large part on the quality of its fitting. Almost all hearing aids in use in the US are digital hearing aids. Devices similar to hearing aids include the Osseo integrated auditory prosthesis (formerly called the bone anchored hearing aid) and cochlear implant.

Hearing aids are used for a variety of pathologies including sensorineural hearing loss, conductive hearing loss, and single-sided deafness. Hearing aid candidacy is typically determined by an audiologist, who will also fit the device based on the nature and degree of the hearing loss being treated.

The amount of benefit experienced by the user of the hearing aid is multi-factorial, depending on the type, severity, and aetiology of the hearing loss, the technology and fitting of the device, and on the motivation, personality, lifestyle, and overall health of the user.

Nebulizer

In medicine, a nebulizer is a drug delivery device used to administer medication in the form of a mist inhaled into the lungs. Nebulizers are commonly used for the treatment of cystic fibrosis, asthma, COPD and other respiratory diseases or disorders.

Analytical nebulizers are another form of nebulizer and are used primarily in laboratory settings for elemental analysis.

Nebulizers use oxygen, compressed air or ultrasonic power to break up solutions and suspensions into small aerosol droplets that can be directly inhaled from the mouthpiece of the device. An aerosol is a mixture of gas and solid or liquid particles.

Soft mist inhaler The medical company Boehringer Ingelheim also invented a new device named Respimat Soft Mist Inhaler in 1997. This new technology provides a metered dose to the user, as the liquid bottom of the inhaler is rotated clockwise 180 degrees by hand, adding a build-up tension into a spring around the flexible liquid container. When the user activates the bottom of the inhaler, the energy from the spring is released and imposes pressure on the flexible liquid container, causing liquid to spray out of 2 nozzles, thus forming a soft mist to be inhaled.

Jet nebulizer The most commonly used nebulizers are jet nebulizers, which are also called "atomizers". Jet nebulizers are connected by tubing to a compressor, that causes compressed air or oxygen to flow at high velocity through a liquid medicine to turn it into an aerosol, which is then inhaled by the patient. Currently there seems to be a tendency among physicians to prefer prescription of a pressurized Metered Dose Inhaler (pMDI) for their patients, instead of a jet nebulizer that generates a lot more noise (often 60 dB during use) and is less portable due to a heavier weight. However, jet nebulizers are commonly used for patients in hospitals who have difficulty using inhalers, such as in serious cases of respiratory disease, or severe asthma attacks. The main advantage of the jet nebulizer is related to its low operational cost.

Ultrasonic wave nebulizer Ultrasonic wave nebulizers were invented in 1964[citation needed] as a new type of portable nebulizer. The technology inside an ultrasonic wave nebulizer is to have an electronic oscillator generate a high frequency ultrasonic wave, which causes the mechanical vibration of a piezoelectric element. This vibrating element is in contact with a liquid reservoir and its high frequency vibration is sufficient to produce a vapour mist. As they create aerosols from ultrasonic vibration instead of using a heavy air compressor, they only have a weight around 170 grams (6.0 oz.). Another advantage is that the ultrasonic vibration is almost silent.

Vibrating mesh technology, a new significant innovation was made in the nebulizer market around 2005, with creation of the ultrasonic Vibrating Mesh Technology (VMT). With this technology a mesh/membrane with 1000-7000 laser drilled holes vibrates at the top of the liquid reservoir, and thereby pressures out a mist of very fine droplets through the holes. This technology is more efficient than having a vibrating piezoelectric element at the bottom of the liquid reservoir, and thereby shorter treatment times are also achieved. The old problems found with the ultrasonic wave nebulizer, having too much liquid waste and undesired heating of the medical liquid, have also been solved by the new vibrating mesh nebulizers.

Biosignals

A bio signal is any signal in living beings that can be continually measured and monitored. The term bio signal is often used to refer to bioelectrical signals, but it may refer to both electrical and non-electrical signals. The usual understanding is to refer only to time-varying signals, although spatial parameter variations (e.g. the nucleotide sequence determining the genetic code) are sometimes subsumed as well.

Electrical biosignals[edit]

Electrical biosignals, or bioelectrical time signals, usually refers to the change in electric current produced by the sum of an electrical potential difference across a specialized tissue, organ or cell system like the nervous system. Thus, among the best-known bioelectrical signals are:

- Electroencephalogram (EEG)
- Electrocardiogram (ECG)
- Electromyogram (EMG)
- Mechanomyogram (MMG)
- Electrooculography (EOG)
- Galvanic skin response (GSR)
- Magneto encephalogram (MEG)

EEG, ECG, EOG and EMG are measured with a differential amplifier which registers the difference between two electrodes attached to the skin. However, the galvanic skin response measures electrical resistance and the MEG measures the magnetic field induced by electrical currents (electroencephalogram) of the brain.

With the development of methods for remote measurement of electric fields using new sensor technology, electric biosignals such as EEG and ECG can be measured without electric contact with the skin. This can be applied for example for remote monitoring of brain waves and heartbeat of patients who must not be touched, in particular patients with serious burns.

Electrical currents and changes in electrical resistances across tissues can also be measured from plants.

Biosignals may also refer to any non-electrical signal that is capable of being monitored from biological beings, such as mechanical signals (e.g. the mechanomyogram or MMG), acoustic signals (e.g. phonetic and non-phonetic utterances, breathing), chemical signals (e.g. pH, oxygenation) and optical signals (e.g. movements).

Bio electrodes

Bio electrodes function as an interface between biological structures and electronic systems. Electrical activity within the biological structure is either sensed or stimulated. The electrical systems are either passively sensing (measuring) or actively stimulating (inducing) electrical potentials within the biological structure or unit.

Electrical currents are generated by many biological structures. Currents give rise to potential differences that can be measured using electrodes and can be interpreted to gain insight in the functioning of the source structure. Conversely, current can be applied to the biological structure through electrodes to affect the target.

The same electrode may function either passively or actively, depending on the purpose and the electronic system controls. An example seen on TV is the large defibrillation paddles used by paramedics to resuscitate people in cardiac distress. When the paddles are applied to a patient, the electrical system is programmed to first passively sense the electrical activity (or lack of) within the heart. Then the electrical system uses algorithms to determine if a stimulation (shock) is required, and finally to provide the appropriate electrical stimulation.

The size of bio electrodes ranges from microscopic intra-cellular research electrodes to large (3 x 5-inch) defibrillation paddles.

Most bio electrodes are made of metal, but the microscopic intra-cellular research electrodes are glass capillary tubes filled with a conductive saline solution.

Contact impedance

Electrical Impedance Tomography (EIT) applies current and measures the resulting voltage on the surface of a target. In biomedical applications, this current is applied, and voltage is measured, through electrodes attached to the body. Models are used to represent these electrode connections in the reconstruction of the conductivity image, tying circuit models to Finite Element Method (FEM) simulations. Changes in the contact impedance or boundary shape relative to the electrode's surface area can introduce artifacts in the reconstructed image. The quantity and quality of these artifacts is dependent upon the electrode model and the properties assigned to that model. The electrode models were originally formulated in the context of mathematical proofs of solution existence and uniqueness for EIT (Calder´on 2006, Nachman 1996). The Complete Electrode Model (CEM) allows a complex impedance for each electrode that models the metal electrode, conductive gel and chemical interaction at the skin electrode interface (Cheng et al. 1989, Somersalo et al. 1992). The FEM is used in the numerical solution of EIT images. The simplest electrode model to implement in the FEM is the Point Electrode Model (PEM) which applies current and measures voltage at single nodes on the boundary and requires no further equations to implement. The PEM does not consider the geometry or contact impedance of an electrode. To reconstruct accurate images from in vivo data, an accurate electrode model is frequently required, and thus, the CEM is generally preferred (Cheng et al. 1989). (Figure 1)

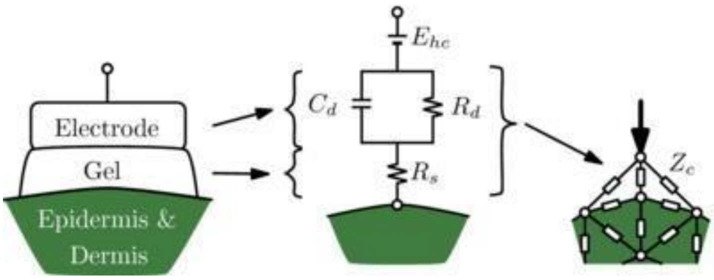


Figure 1 Generalized electrode model

Electrodes in Biomedical Instrumentation

Electrodes are devices that convert ionic potentials into electronic potentials. The type of electrode used for the measurements depends on the anatomical location of the bioelectric event to be measured. In order to process the signal in electronic circuits, it will be better to convert ionic conduction into electronic conduction. So simply bio-electrodes are a class of sensors that transduces ionic conduction into electronic conduction. The purpose of bio-electrodes is to acquire bioelectrical signals such as ECG, EMG, EEG etc.

Electrodes are mainly classified into two. They are perfectly polarized electrodes and perfectly non-polarized electrodes. There are a wide variety of electrodes which can be used to measure bioelectric events. The three main classes of electrodes are Microelectrodes, Body Surface electrodes and Needle electrodes.

A. **Microelectrodes** are electrodes with tips having tips sufficiently small enough to penetrate a single cell in order to obtain readings from within the cell. The tips must be small enough to permit penetration without damaging the minute cell. The main functions of microelectrodes are potential recording and current injection. Microelectrodes are having high impedances in mega ohm range because of their smaller size. Microelectrodes are generally of two types. With the use of a microelectrode or an array of microelectrodes, researchers can gather all sort of information regarding living organism.

- Metal type
- Micropipette type

a. **Metal microelectrode:** Metal microelectrodes are formed by electrolytic ally etching the tip of fine tungsten to the desired size and dimension. Then the wire is coated almost to the tip with any type of insulating material. The metal-ion interface takes place where the metal tip contacts the electrolyte. The main features of metal microelectrodes are

1. Very good S/N ratio
2. Strong enough to penetrate
3. High biocompatibility

b. **Micropipette**: The micropipette type of microelectrode is a glass micropipette with its tip drawn out to the desired size. The micropipette is filled with an electrolyte which should be compatible with the cellular fluids. A micropipette is a small and extremely fine pointed pipette used in making microinjections. A commercial type of micropipette is shown in figure below.

B. Body Surface Electrodes:

Surface electrodes are those which are placed in contact with the skin of the subject in order to obtain bioelectric potentials from the surface. Body surface electrodes are of many sizes and types. In spite of the type, any surface electrode can be used to sense ECG, EEG, EMG etc. The various types of body surface electrodes are discussed below. Major body surface electrodes are

1. **Immersion electrodes:** They are one of the first type of bioelectric measuring electrodes. Immersion electrodes were simply buckets of saline solution in which the subject placed his hands and feet. So it was not a comfortable type of measurement and hence it was replaced with plate electrodes.

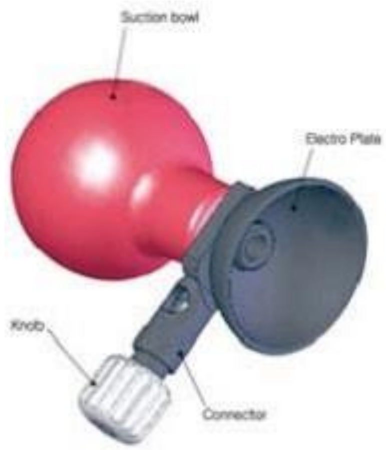
2. **Plate electrodes**: These electrodes were separated from subject's skin by cotton pads socked in a strong saline solution. The plate electrodes have generally smaller contact area and they do not totally seal on the patient. The electrode slippage and displacement of plates were the major difficulties faced by these type of electrodes because they have a tendency to lose their adhesive ability as a result of contact with fluids on or near the patient. Since these types of electrodes were very sensitive, it led to measurement errors.

3. **Floating electrodes:** These types of electrodes can eliminate the movement errors (called

artifacts) which is a main problem with plate electrodes. This is done by avoiding any direct contact of the metal with the skin. So the main advantage of floating electrodes is mechanical reliability. Here the conductive path between the metal and the skin is the electrolyte paste or jelly.

4. **Disposable electrodes**: Normally plate electrodes, floating electrodes etc. can be used more than one time. This requires the cleaning and cares after each use. We can use disposable electrodes which can be used only once and be disposed after the use. These types of electrodes are now widely used.

5. **Suction electrodes**: These type of electrodes are well suited for the attachment to flat surfaces of body and to regions where the underlying tissue is soft, due to the presence of contact surface. An advantage of these type of electrodes is that it has a small surface area. These types of electrodes are mainly used for the measurement of ECG. Suction electrodes used a plastic syringe barrel to house suction tubing and input cables to an AC amplifier.



6. **Ear clip & Scalp electrodes**: These type of electrodes is widely used in the measurement of EEG exclusively. Scalp electrodes can provide EEG easily by placing it over bare head. A typical ear clip electrode is shown in figure below. The most common method for EEG measurement is 10 – 20 electrode placement system and here we use scalp electrode usually. They can avoid measurement errors and movement errors. During labour internal monitoring may be needed and is usually in the form of an electrode placed under the baby's scalp. It is called fetal scalp electrode which is used to monitor baby's heartbeat while still in uterus.

C. **<u>Needle Electrodes</u>**:

To reduce the interface and noise (artifact) caused due to electrode movement, during the measurement of EEG, EMG etc. we can use small sub-dermal needle electrodes which penetrate the scalp. Actually the needle electrodes are not inserted into the brain. They nearly penetrate the skin. Generally they are simply inserted through a small section of the skin just beneath the skin parallel to it.

The needle electrodes for EMG measurement consist of fine insulated wires placed in such a way that their tips are in contact with the muscle, nerve or other tissues from which the measurement is made. The needle creates the hole necessary for insertion and the wires forming the electrodes are carried inside it. A typical EEG needle electrode is shown in figure.

One of the main advantage of needle electrodes is that they are less susceptible to movement errors than surface electrodes. Also the needle electrodes have lower impedances when compared to surface electrodes as it makes direct contact with the sub-dermal tissues or intracellular fluid.

Telemetry Bio potential (ECG, EMG, EEG) Data

Bio potential telemetry measures the electrical potential between two electrodes/leads. Depending on where the leads are placed, bio potential telemetry can be used to measure long-term electrocardiogram (ECG), electroencephalogram (EEG) and electromyogram

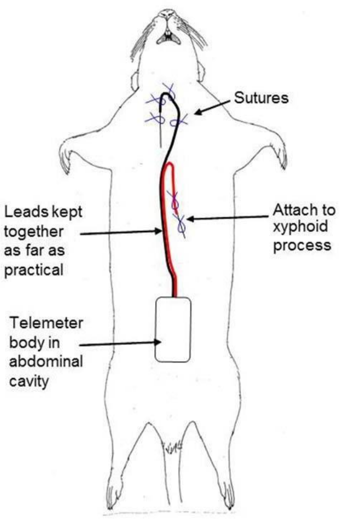
(EMG) in rats and mice. The Rat Telemetry System offers solutions for the long term recording of a single bio potential signal using our TR50B or simultaneous recording of two bio potential signals (e.g. EEG+EMG) with our TR50BB telemeters. For mice, Mouse Telemetry offers the MT10B telemeter for the long term recording of a single bio potential signal.

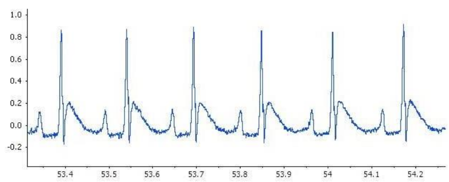
Electrocardiography – ECG or EKG

ECG records the net change in electrical activity across the heart with each heartbeat over time. Under normal conditions, the mechanical activity is temporally linked with the electrical activity. Changes in the electrical potential from the depolarization of the myocardium result in muscular contraction in the region of depolarization while repolarization results in the relaxation. The wave of depolarization in each heart beat originates from the sinoatrial node, travels through the atria to the atrial-ventricular node, then travels down the ventricular septum and up walls of the ventricles.

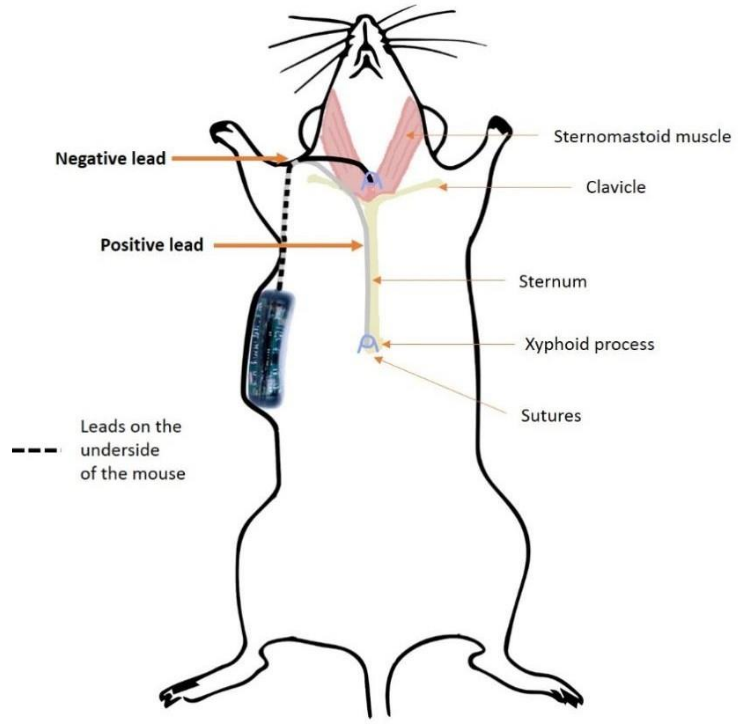
The P, QRS and T waves recorded in an ECG represent the change in electrical potential during the depolarization of the atria, the depolarization of the ventricles (and the simultaneous repolarization of the atria), and the repolarization of the ventricles respectively.

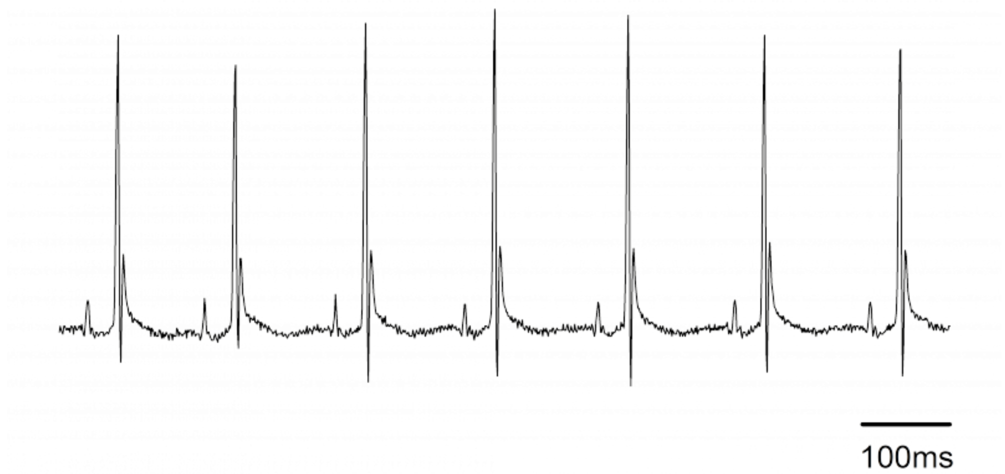
Clinically, ECG is recorded using 12 lead positions to generate information about how the electrical activity propagates in 3 dimensions. In rats and mice, telemetry ECG is often recorded using two electrodes recording a single lead configuration. Depending on the position of the electrodes and lead configuration, different ECG profiles are seen. The modified lead II position of an ECG in a rat is shown in the figure below. The lead wire positions have been adjusted slightly from the standard lead II position, which is most in line with the ventricles, to improve signal quality when the rat or mouse is moving around.





Example placement of electrode leads for recording of high quality ECG in the modified lead II configuration (left) and signal recorded (above).





Adjusting the position of the electrodes to record from different ECG lead configurations changes the relative size of the P, QRS and T waves, which allows detailed focus on one event. Good muscle contact will be required for large clean telemetry ECG signals as electrical conductance through connective tissue is poor. It is also

important to keep the lead wires together as far as practical and avoid curling the leads into a loop to minimize interference from ambient electrical noise.

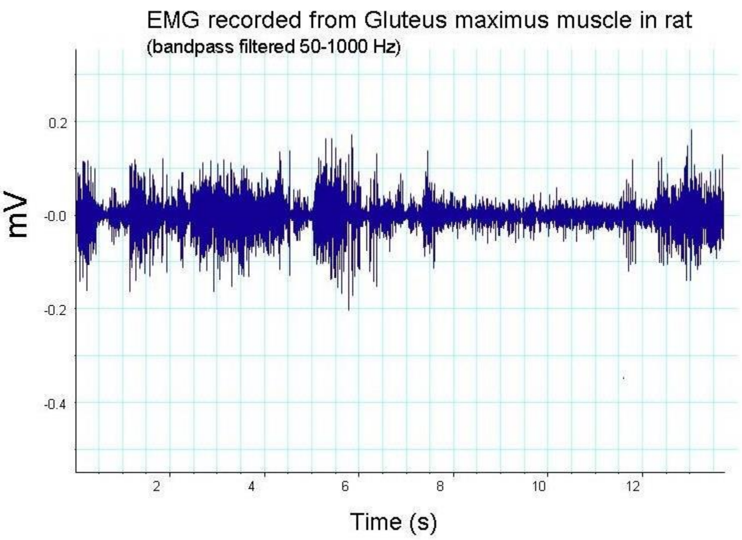
The most common use for ECG is the measurement of heart rate, which can be calculated by the number of R waves in a given period of time. In addition to the detection of the heart rate, the ECG can be used to record:

- heart rate variability (R-R intervals)
- duration of depolarization of the atria (P wave duration) or ventricles (QRS duration)
- arrhythmias and counts of P, QRS, T waves
- electrical conductance across the heart and waveform abnormalities
- relative amplitudes of the waveforms can also be calculated
- These measurements can be useful in basic cardiovascular physiology, behavioural or pharmacological/toxicology studies.

Electromyography - EMG

EMG records the electrical potential and the change in potential (electrical activity) across the skeletal muscle. The electrical potential changes during contraction and relaxation of the skeletal muscle as the muscle depolarizes and then repolarizes. EMG can be used to assess the muscle contraction, mechanics or inferring movement rates across the muscle. For example, attaching the lead wires to the diaphragm allows for the measurement of respiration rate, while attachment of the lead wires to the gluteus Maximus muscle can be used to assess walking or movement. Other potential applications include measurement of gastrointestinal peristaltic movements and movement of specific muscles such as nuchal EMG.

Below is an example of the EMG signal recorded from the gluteus Maximus muscle in a walking rat.



EMG can be recorded from a muscle or muscle groups, provided that the size of the muscle is large enough for the attachment of the two electrodes with a gap to prevent contact between the electrodes. It is important that the electrodes do not contact as this results in electrical shortening, and a signal will not be detected. It is also important that the electrodes are tied securely to the muscle that is being recorded from to minimize movement artifacts.

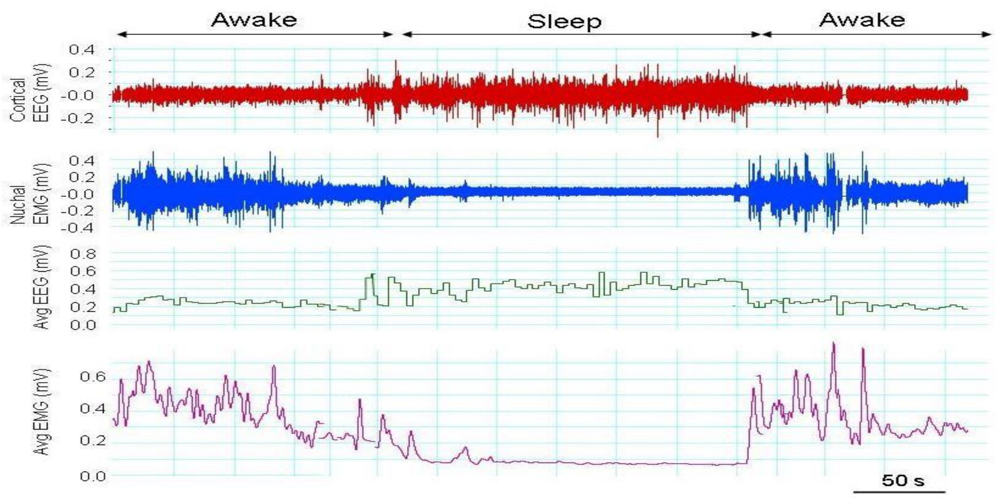
Telemetry is advantageous over tethered systems for measuring EMG as the rats and mice are allowed to move freely in their home cage eliminating behavioural adaptations to tethering. The battery of the rat telemeters also allows recording of EMG away from the home cage and tethers, opening up opportunities such as EMG recordings during behavioural tests. Rat bio potential telemeters provide up to 4 hours of continuous data transmission powered by a rechargeable battery when the rat telemeter is not actively being powered by the Smart Pad.

Electroencephalography - EEG

EEG records the electrical activity over an area of brain tissue. Electrical activity in the brain is generated by depolarization of neurons. As the electrical activity provided by a single neuron is too small to be detected, recorded telemetry EEG represents the summation of electrical activity from a large number of neurons in an area.

Electrodes can be placed on the surface of the brain for cortical surface EEG, or deeper into the tissue to measure from a particular area of interest. EEG can be applied to the study of neurological disorders, such as seizure detection or epilepsy studies and pharmacological studies into neurostimulating drugs. EEG can also be used in behavioural research to study brain activity patterns in response to stimuli or in studies of sleep-wake cycles. Both the Rat and Mouse Telemetry Systems allow recording of EEG from animals living in their home cages free from the stress and restriction of using tethers.

Below is an example of EEG recorded in a rat using the TR50BB telemeter with simultaneous nuchal EMG recordings in a study of the sleep-**wake cycle.**



In addition to EMG, it is possible to record EEG in conjunction with brain oxygen (TR50B + TR57Y) or intracranial pressure (TRM54PB) using the Rat Telemetry System.

Unit 6
Temperature measurement

Temperature is a physical quantity that expresses hot and cold. It is the manifestation of thermal energy, present in all matter, which is the source of the occurrence of heat, a flow of energy, when a body is in contact with another that is colder or hotter.

Temperature is measured with a thermometer.

The most common scales are the Celsius scale (formerly called *centigrade*, denoted as °C), the Fahrenheit scale (denoted as °F), and the Kelvin scale (denoted as K),

Kelvin scale is predominantly used for scientific purposes by conventions of the International System of Units (SI).

The lowest theoretical temperature is absolute zero, at which no more thermal energy can be extracted from a body.

The following table shows the temperature conversion formulas for conversions to and from the Celsius scale.

Temperature conversions

	from Celsius	to Celsius
Fahrenheit	[°F] = [°C] × $\frac{9}{5}$ + 32	[°C] = ([°F] − 32) × $\frac{5}{9}$
Kelvin	[K] = [°C] + 273.15	[°C] = [K] − 273.15
Rankine	[°R] = ([°C] + 273.15) × $\frac{9}{5}$	[°C] = ([°R] − 491.67) × $\frac{5}{9}$
Delisle	[°De] = (100 − [°C]) × $\frac{3}{2}$	[°C] = 100 − [°De] × $\frac{2}{3}$
Newton	[°N] = [°C] × $\frac{33}{100}$	[°C] = [°N] × $\frac{100}{33}$
Réaumur	[°Ré] = [°C] × $\frac{4}{5}$	[°C] = [°Ré] × $\frac{5}{4}$
Rømer	[°Rø] = [°C] × $\frac{21}{40}$ + 7.5	[°C] = ([°Rø] − 7.5) × $\frac{40}{21}$

Different methods of temperature measurement

1.Mechanical methods:

Mercury in glass thermometers

This consists of a glass tube of very fine bore joined to a reservoir at the bottom and sealed at the top. A measured quantity of mercury is the enclosed. When the thermometer is heated the mercury expands much more than the glass and is therefore forced to rise up in the tubing A scale is fixed at the side.

Bimetallic Thermometer

Bimetallic strip thermometer Two dissimilar metals are bonded together into what is called a bimetallic strip as shown below. Suppose metal A has a smaller coefficient of thermal expansion than does metal B. As temperature increases, metal B expands more than does metal A, causing the bimetallic strip to curl upwards as shown below. One common application of bimetallic strips is in air-conditioning thermostats, where a bimetallic strip is used as the arm of a switch between electrical contacts. As the room temperature changes, the bimetallic strip bends as discussed above. When the bimetalic strip bends far enough, it makes contact with electrical leads which turn the heat or air conditioning on or off.

Two metals whose coefficient of linear expansion is different are welded and rolled together to the desire thickness. The actual movement of a bimetal is its flexibility with one end fixed, a straight bimetal strip deflects in proportion to its temperature, to the square of its length and inversely with its thickens

2.Electrical method of temperature measurement

Thermocouples

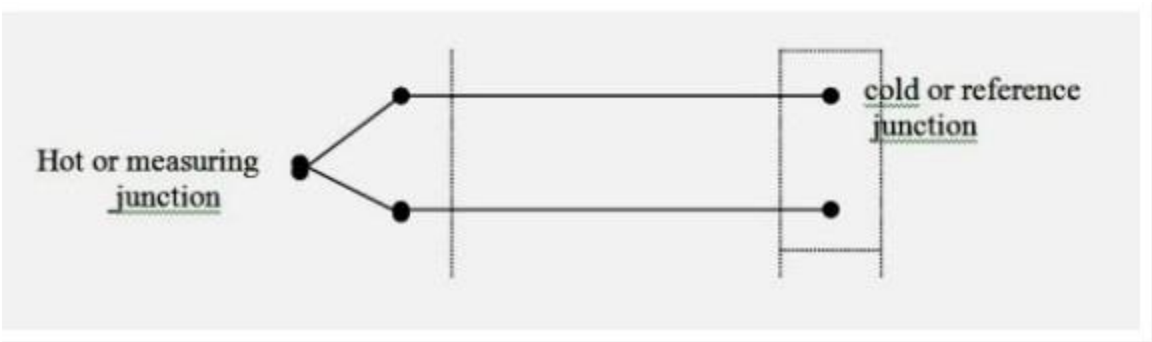
A thermocouple is made up of two dissimilar metals, joined together at one end, that produce a voltage (expressed in millivolts) with a change in temperature. The junction of the two metals, called the sensing junction, is connected to extension wires. Any two dissimilar metals may be used to make a thermocouple.

Principle of Operation

1. When two dissimilar metals are connected together, a small voltage called a thermojunction voltage is generated at the junction. This is called the Peltier effect.
2. If the temperature of the junction changes, it causes voltage to change too, which can be measured by the input circuits of an electronic controller. The output is a voltage proportional to the temperature difference between the junction and the free ends. This is called the Thompson effect.
3. Both of these effects can be combined to measure temperature. By holding one junction at a known temperature (reference junction) and measuring the voltage, the temperature at the sensing junction can be deduced. The voltage generated is directly proportional to the temperature difference. The combined effect is known as the thermo-junction effect or the Seebeck effect.

It is a simple device consisting of a dissimilar metal wires joined at their ends. When an of each wire is connected to a measuring instrument thermocouples becomes an accurate and sensitive temperature measuring device

.



Resistance – Temperature Detectors (RTD)

RTD’s are generally used for precise temperature measurement. It consists of a five wire wrapped around an insulator and enclosed in a metal. The most sheath of a resistance thermometer resembles that of bimetallic thermometer
bulb.

PRINCIPLE: “Resistance increases as temperature increase”
Rt. = Ro (1 + α t)
Rt. = Resistance of Temperature to measured.
Ro. = Resistance of zero temperature.
α = Co. off of thermal (expansion).
t = Temperature to be measured.

These metals have a positive temperature co-efficient of expansion. Therefore resistance increases as the temperature increases.
Types of material used: (1) Platinum (2) Nickel
These metals have a positive temperature co-efficient of expansion. Therefore resistance increases as the temperature increases.

Thermistors

Thermistors are temperature sensitive semiconductors that exhibit a large change in resistance over a relatively small range of temperature. There are two main types of thermistors

: positive temperature coefficient (PTC) and

negative temperature coefficient (NTC).

NTC thermistors exhibit the characteristic of resistance falling with increasing temperature. These are most commonly used for temperature measurement.

A thermistor is similar to an RTD, but a semiconductor material is used instead of a metal. A thermistor is a solid state device and has larger sensitivity than does an RTD. Unlike RTD's, the temperature-resistance characteristic of a thermistor is non-linear, and cannot be characterized by a single coefficient. Furthermore, unlike RTDs, the resistance of a thermistor decreases with increasing temperature. Thermistors cannot be used to measure high temperatures compared to RTDs. In fact, the maximum temperature of operation is sometimes only 100 or 200°C.

3. RADIATIVE TEMPERATURE MEASURING DEVICES

Two types of radiative measuring devices are:

1. Infrared pyrometers, and

2. Optical pyrometers.

Infrared Pyrometer

Infrared temperature sensors also known as pyrometers or non-contact temperature sensors are used to measure the temperature of an object without contact. This is different from most temperature measurement devices, which require direct contact with the measured media. Noncontact methods of temperature measurement are advantageous when contact methods are impossible or impractical, such as when the target is inaccessible or so hot that contact devices will not survive.

Principle of Operation

Infrared temperature sensors use the principle that any object emits an amount of energy that is a function of its temperature. This function dictates that as the temperature of an object rises, so does the amount of energy it emits. “An infrared temperature sensor determines temperature by measuring the intensity of energy given off by an object.”

Optical Pyrometer

An optical pyrometer is useful for measuring very high temperatures (even flames). The optical pyrometer uses an infrared radiation-sensitive sensor, e.g. a photodiode or a photoresistor, to compare the radiation from the unknown with that of the radiation from an internal incandescent source. The accuracy of the optical pyrometer is very much a function of the emissivity of the device that is radiating the heat. The obvious advantage in using an optical pyrometer at very high temperatures is that the measurement is non-contacting. This approach is very expensive,

Summarizing

The two most common type of temperature sensors are Thermocouples and RTD’s. Although these sensors have overlapping temperature ranges, each has certain application-dependent advantages. These are summarized below. An RTD is the sensor of choice when sensitivity and application flexibility are the most important criteria. When it comes to component cost, an RTD is more expensive than a thermocouple.

UNIT - 1 : INTRODUCTION TO MICROPROCESSOR

1. Introduction and evolution of microprocessors
2. Architecture of 8085 processor
3. Pin configuration of 8085
4. Bus organization
5. Basic instruction sets.

1. INTRODUCTION TO MICROPROCESSOR AND MICROCOMPUTER ARCHITECTURE:

A microprocessor is a programmable electronics chip that has computing and decision making capabilities similar to central processing unit of a computer. Any microprocessor-based systems having limited number of resources are called microcomputers. Nowadays, microprocessor can be seen in almost all types of electronics devices like mobile phones, printers, washing machines etc.

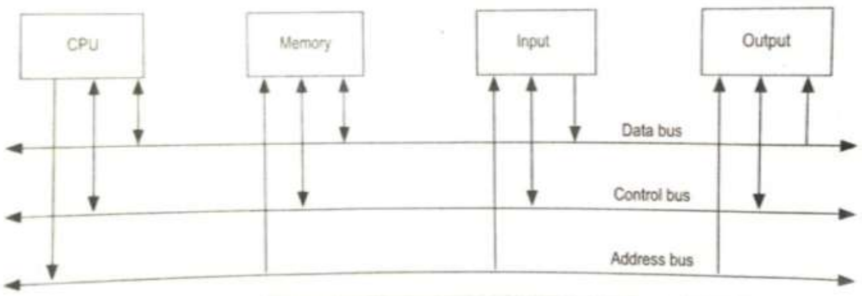


Fig.1 Microprocessor-based system

- **Address Bus**: It carries the address, which is a unique binary pattern used to identify a memory location or an I/O port. For example, an eight bit address bus has eight lines and thus it can address $2^8 = 256$ different locations. The locations in hexadecimal format can be written as 00H – FFH.
- **Data Bus**: The data bus is used to transfer data between memory and processor or between I/O device and processor. For example, an 8-bit processor will generally have an 8-bit data bus and a 16-bit processor will have 16-bit data bus.
- **Control Bus**: The control bus carry control signals, which consists of signals for selection of memory or I/O device from the given address, direction of data transfer and synchronization of data transfer in case of slow devices.

A typical microprocessor consists of arithmetic and logic unit (ALU) in association with control unit to process the instruction execution. Almost all the microprocessors are based on the principle of store-program concept. In store-program concept, programs or instructions are sequentially stored in the memory locations that are to be executed. To do any task using a microprocessor, it is to be programmed by the user. So the programmer must have idea about its internal resources, features and supported instructions. Each microprocessor has a set of instructions, a list which is provided by the microprocessor manufacturer. The instruction set of a microprocessor is provided in two forms: binary machine code and mnemonics.

Microprocessor communicates and operates in binary numbers 0 and 1. The set of instructions in the form of binary patterns is called a machine language and it is difficult for us to understand. Therefore, the binary patterns are given abbreviated names, called mnemonics, which forms the assembly language. The conversion of assembly-level language into binary machine-level language is done by using an application called assembler.

Evolution of Microprocessors

4-bit Microprocessors

The first microprocessor was introduced in 1971 by Intel Corp. It was named Intel 4004 as it was a 4 bit processor. It was a processor on a single chip. It could perform simple arithmetic and logic operations such as addition, subtraction, boolean AND and boolean OR. It had a control unit capable of performing control functions like fetching an instruction from memory, decoding it, and generating control pulses to execute it. It was able to operate on 4 bits of data at a time. This first microprocessor was quite a success in industry. Soon other microprocessors were also introduced. Intel introduced the enhanced version of 4004, the 4040.

8-bit Microprocessors

The first 8 bit microprocessor which could perform arithmetic and logic operations on 8 bit words was introduced in 1973 again by Intel. This was Intel 8008 and was later followed by an improved version, Intel 8088. Some other 8 bit processors are Zilog-80 and Motorola M6800.

16-bit Microprocessors

The 8-bit processors were followed by 16 bit processors. They are Intel 8086 and 80286.

32-bit Microprocessors

The 32 bit microprocessors were introduced by several companies but the most popular one is Intel 80386.

Pentium Series

Instead of 80586, Intel came out with a new processor namely Pentium processor. Its performance is closer to RISC performance. Pentium was followed by Pentium Pro CPU. Pentium Pro allows multiple CPUs in a single system in order to achieve multiprocessing. The MMX extension was added to Pentium Pro and the result was Pentiuum II.

The Pentium III provided high performance floating point operations for certain types of computations by using the SIMD extensions to the instruction set. These new instructions makes the Pentium III faster than high-end RISC CPUs.

NAME	YEAR	TRANSISTORS	DATA WIDTH	CLOCK SPEED
8080	1974	6,000	8 bits	2 MHz
8085	1976	6,500	8 bits	5 MHz
8086	1978	29,000	16 bits	5 MHz
8088	1979	29,000	8 bits	5 MHz
80286	1982	134,000	16 bits	6 MHz
80386	1985	275,000	32 bits	16 MHz
80486	1989	1,200,000	32 bits	25 MHz
PENTIUM	1993	3,100,000	32/64 bits	60 MHz
PENTIUM II	1997	7,500,000	64 bits	233 MHz
PENTIUM III	1999	9,500,000	64 bits	450 MHz
PENTIUM IV	2000	42,000,000	64 bits	1.5 GHz

UNIT - 2 : ARCHITECTURE OF MICROPROCESSOR

The 8085 microprocessor is an 8-bit processor available as a 40-pin IC package and uses +5 V for power. It can run at a maximum frequency of 3 MHz. Its data bus width is 8-bit and address bus width is 16-bit, thus it can address 2^{16} = 64 KB of memory. The internal architecture of 8085 is shown is Fig. 2.

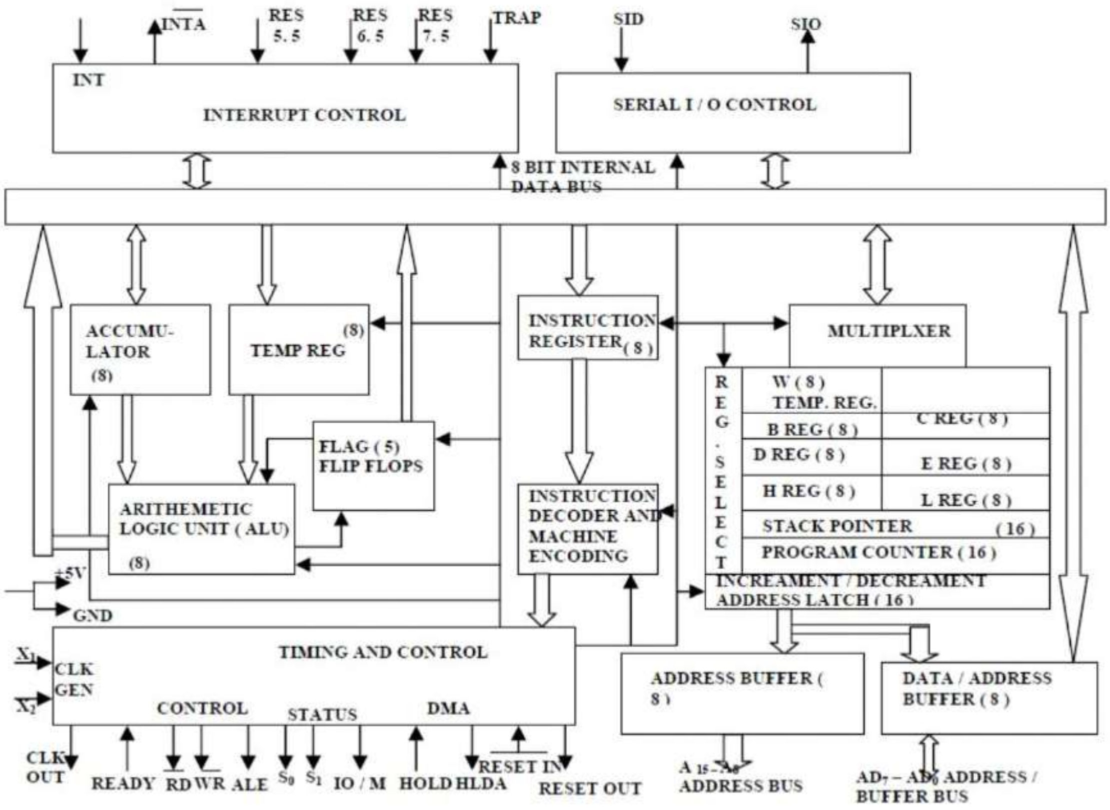


Fig. 2 Internal Architecture of 8085

Based on architecture:

- Reduced Instruction Set Computer (RISC) processors
- Complex Instruction Set Computer (CISC) processors

2. 8085 MICROPROCESSOR ARCHITECTURE

The 8085 microprocessor is an 8-bit processor available as a 40-pin IC package and uses +5 V for power. It can run at a maximum frequency of 3 MHz. Its data bus width is 8-bit and address bus width is 16-bit, thus it can address 2^{16} = 64 KB of memory. The internal architecture of 8085 is shown is Fig. 2.

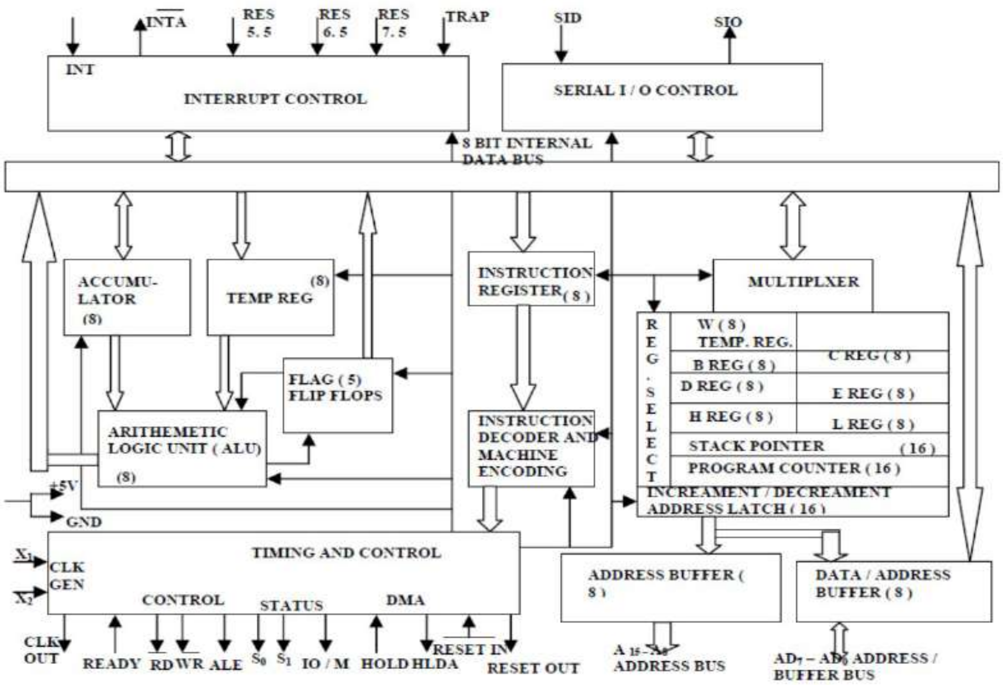


Fig. 2 Internal Architecture of 8085

Arithmetic and Logic Unit

The ALU performs the actual numerical and logical operations such as Addition (ADD), Subtraction (SUB), AND, OR etc. It uses data from memory and from Accumulator to perform operations. The results of the arithmetic and logical operations are stored in the accumulator.

Registers

The 8085 includes six registers, one accumulator and one flag register, as shown in Fig. 3. In addition, it has two 16-bit registers: stack pointer and program counter. They are briefly described as follows.

The 8085 has six general-purpose registers to store 8-bit data; these are identified as B, C, D, E, H and L. they can be combined as register pairs - BC, DE and HL to perform some

16-bit operations. The programmer can use these registers to store or copy data into the register by using data copy instructions.

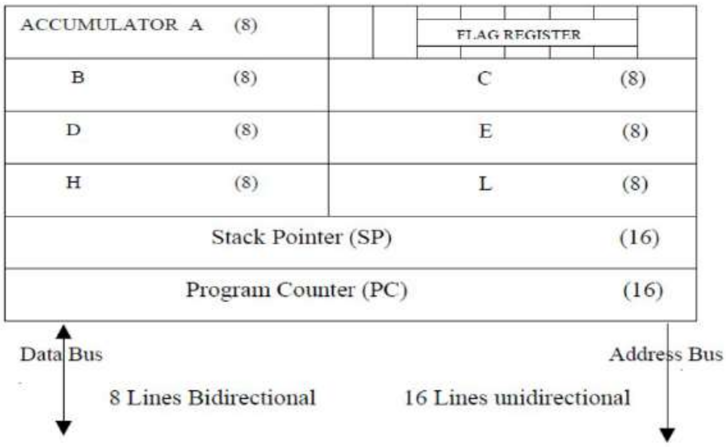


Fig. 3 Register organisation

Accumulator

The accumulator is an 8-bit register that is a part of ALU. This register is used to store 8-bit data and to perform arithmetic and logical operations. The result of an operation is stored in the accumulator. The accumulator is also identified as register A.

Flag register

The ALU includes five flip-flops, which are set or reset after an operation according to data condition of the result in the accumulator and other registers. They are called Zero (Z), Carry (CY), Sign (S), Parity (P) and Auxiliary Carry (AC) flags. Their bit positions in the flag register are shown in Fig. 4. The microprocessor uses these flags to test data conditions.

D_7	D_6	D_5	D_4	D_3	D_2	D_1	D_0
S	**Z**		**AC**		**P**		**CY**

Fig. 4 Flag register

For example, after an addition of two numbers, if the result in the accumulator is larger than 8-bit, the flip-flop uses to indicate a carry by setting CY flag to 1. When an arithmetic operation results in zero, Z flag is set to 1. The S flag is just a copy of the bit D7 of the accumulator. A negative number has a 1 in bit D7 and a positive number has a 0 in 2's complement representation. The AC flag is set to 1, when a carry result from bit D3 and passes to bit D4. The P flag is set to 1, when the result in accumulator contains even number of 1s.

Program Counter (PC)

This 16-bit register deals with sequencing the execution of instructions. This register is a memory pointer. The microprocessor uses this register to sequence the execution of the instructions. The function of the program counter is to point to the memory address from which the next byte is to be fetched. When a byte is being fetched, the program counter is automatically incremented by one to point to the next memory location.

Stack Pointer (SP)

The stack pointer is also a 16-bit register, used as a memory pointer. It points to a memory location in R/W memory, called stack. The beginning of the stack is defined by loading 16-bit address in the stack pointer.

Instruction Register/Decoder

It is an 8-bit register that temporarily stores the current instruction of a program. Latest instruction sent here from memory prior to execution. Decoder then takes instruction and decodes or interprets the instruction. Decoded instruction then passed to next stage.

Control Unit

Generates signals on data bus, address bus and control bus within microprocessor to carry out the instruction, which has been decoded. Typical buses and their timing are described as follows:

- *Data Bus*: Data bus carries data in binary form between microprocessor and other external units such as memory. It is used to transmit data i.e. information, results of arithmetic etc between memory and the microprocessor. Data bus is bidirectional in nature. The data bus width of 8085 microprocessor is 8-bit i.e. 2^8 combination of binary digits and are typically identified as D0 – D7. Thus size of the data bus determines what arithmetic can be done. If only 8-bit wide then largest number is 11111111 (255 in decimal). Therefore, larger numbers have to be broken down into chunks of 255. This slows microprocessor.
- *Address Bus*: The address bus carries addresses and is one way bus from microprocessor to the memory or other devices. 8085 microprocessor contain 16-bit address bus and are generally identified as A0 - A15. The higher order address lines (A8 – A15) are unidirectional and the lower order lines (A0 – A7) are multiplexed (time-shared) with the eight data bits (D0 – D7) and hence, they are bidirectional.
- *Control Bus*: Control bus are various lines which have specific functions for coordinating and controlling microprocessor operations. The control bus carries control signals partly unidirectional and partly bidirectional. The following control and status signals are used by 8085 processor:
    I. ALE (output): Address Latch Enable is a pulse that is provided when an address appears on the AD0 – AD7 lines, after which it becomes 0.

II. $\overline{RD}$ (active low output): The Read signal indicates that data are being read from the selected I/O or memory device and that they are available on the data bus.

III. $\overline{WR}$ (active low output): The Write signal indicates that data on the data bus are to be written into a selected memory or I/O location.

IV. $IO/\overline{M}$ (output): It is a signal that distinguished between a memory operation and an I/O operation. When $IO/\overline{M} = 0$ it is a memory operation and $IO/\overline{M} = 1$ it is an I/O operation.

V. S1 and S0 (output): These are status signals used to specify the type of operation being performed; they are listed in Table 1.

Table 1 Status signals and associated operations

S1	S0	States
0	0	Halt
0	1	Write
1	0	Read
1	1	Fetch

The schematic representation of the 8085 bus structure is as shown in Fig. 5. The microprocessor performs primarily four operations:

I. Memory Read: Reads data (or instruction) from memory.
II. Memory Write: Writes data (or instruction) into memory.
III. I/O Read: Accepts data from input device.
IV. I/O Write: Sends data to output device.

The 8085 processor performs these functions using address bus, data bus and control bus as shown in Fig. 5.

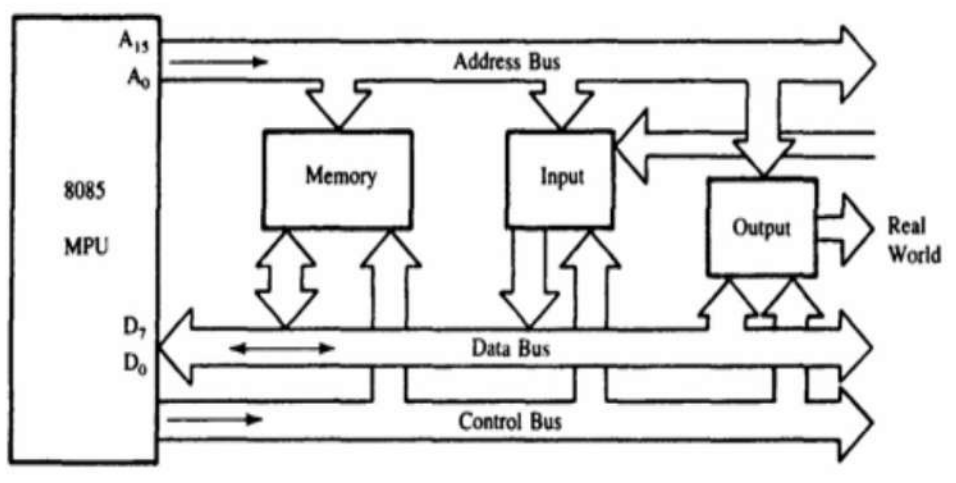


Fig. 5 The 8085 bus structure

3. 8085 PIN DESCRIPTION

Properties:

- It is a 8-bit microprocessor
- Manufactured with N-MOS technology
- 40 pin IC package
- It has 16-bit address bus and thus has 2^{16} = 64 KB addressing capability.
- Operate with 3 MHz single-phase clock
- +5 V single power supply

The logic pin layout and signal groups of the 8085nmicroprocessor are shown in Fig. 6. All the signals are classified into six groups:

- Address bus
- Data bus
- Control & status signals
- Power supply and frequency signals
- Externally initiated signals
- Serial I/O signals

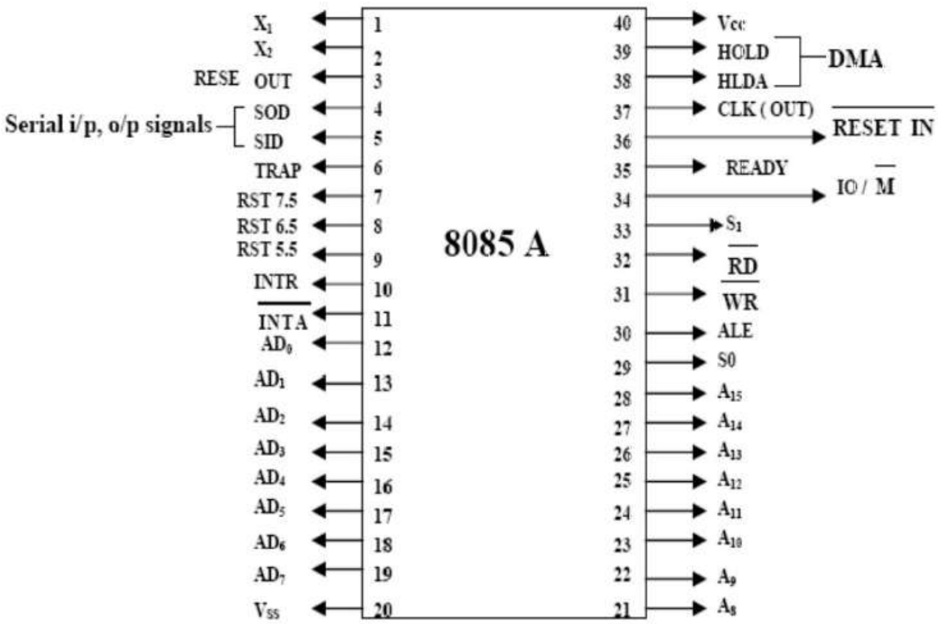


Fig. 6 8085 microprocessor pin layout and signal groups

Address and Data Buses:

- A8 – A15 (output, 3-state): Most significant eight bits of memory addresses and the eight bits of the I/O addresses. These lines enter into tri-state high impedance state during HOLD and HALT modes.
- AD0 – AD7 (input/output, 3-state): Lower significant bits of memory addresses and the eight bits of the I/O addresses during first clock cycle. Behaves as data bus

during third and fourth clock cycle. These lines enter into tri-state high impedance state during HOLD and HALT modes.

Control & Status Signals:

- ALE: Address latch enable
- $\overline{RD}$: Read control signal.
- $\overline{WR}$: Write control signal.
- $IO/\overline{M}$, S1 and S0 : Status signals.

Power Supply & Clock Frequency:

- Vcc: +5 V power supply
- Vss: Ground reference
- X1, X2: A crystal having frequency of 6 MHz is connected at these two pins
- CLK: Clock output

Externally Initiated and Interrupt Signals:

- $\overline{RESET\ \ IN}$: When the signal on this pin is low, the PC is set to 0, the buses are tri-stated and the processor is reset.
- RESET OUT: This signal indicates that the processor is being reset. The signal can be used to reset other devices.
- READY: When this signal is low, the processor waits for an integral number of clock cycles until it goes high.
- HOLD: This signal indicates that a peripheral like DMA (direct memory access) controller is requesting the use of address and data bus.
- HLDA: This signal acknowledges the HOLD request.
- INTR: Interrupt request is a general-purpose interrupt.
- $\overline{INTA}$: This is used to acknowledge an interrupt.
- RST 7.5, RST 6.5, RST 5,5 – restart interrupt: These are vectored interrupts and have highest priority than INTR interrupt.
- TRAP: This is a non-maskable interrupt and has the highest priority.

Serial I/O Signals:

- SID: Serial input signal. Bit on this line is loaded to D7 bit of register A using RIM instruction.
- SOD: Serial output signal. Output SOD is set or reset by using SIM instruction.
-

UNIT - 3 : PROGRAMMING

4. INSTRUCTION SET AND EXECUTION IN 8085

Based on the design of the ALU and decoding unit, the microprocessor manufacturer provides instruction set for every microprocessor. The instruction set consists of both machine code and mnemonics.

An instruction is a binary pattern designed inside a microprocessor to perform a specific function. The entire group of instructions that a microprocessor supports is called instruction set. Microprocessor instructions can be classified based on the parameters such functionality, length and operand addressing.

Classification based on functionality:

I. Data transfer operations: This group of instructions copies data from source to destination. The content of the source is not altered.
II. Arithmetic operations: Instructions of this group perform operations like addition, subtraction, increment & decrement. One of the data used in arithmetic operation is stored in accumulator and the result is also stored in accumulator.
III. Logical operations: Logical operations include AND, OR, EXOR, NOT. The operations like AND, OR and EXOR uses two operands, one is stored in accumulator and other can be any register or memory location. The result is stored in accumulator. NOT operation requires single operand, which is stored in accumulator.
IV. Branching operations: Instructions in this group can be used to transfer program sequence from one memory location to another either conditionally or unconditionally.
V. Machine control operations: Instruction in this group control execution of other instructions and control operations like interrupt, halt etc.

Classification based on length:

I. One-byte instructions: Instruction having one byte in machine code. Examples are depicted in Table 2.
I. Two-byte instructions: Instruction having two byte in machine code. Examples are depicted in Table 3
II. Three-byte instructions: Instruction having three byte in machine code. Examples are depicted in Table 4.

Table 2 Examples of one byte instructions

Opcode	Operand	Machine code/Hex code
MOV	A, B	78
ADD	M	86

Table 3 Examples of two byte instructions

Opcode	Operand	Machine code/Hex code	Byte description
MVI	A, 7FH	3E	First byte
		7F	Second byte
ADI	0FH	C6	First byte
		0F	Second byte

Table 4 Examples of three byte instructions

Opcode	Operand	Machine code/Hex code	Byte description
JMP	9050H	C3	First byte
		50	Second byte
		90	Third byte
LDA	8850H	3A	First byte
		50	Second byte
		88	Third byte

Addressing Modes in Instructions:

The process of specifying the data to be operated on by the instruction is called addressing. The various formats for specifying operands are called addressing modes. The 8085 has the following five types of addressing:

I. Immediate addressing
II. Memory direct addressing
III. Register direct addressing
IV. Indirect addressing
V. Implicit addressing

Immediate Addressing:

In this mode, the operand given in the instruction - a byte or word – transfers to the destination register or memory location.

Ex: MVI A, 9AH

- The operand is a part of the instruction.
- The operand is stored in the register mentioned in the instruction.

Memory Direct Addressing:

Memory direct addressing moves a byte or word between a memory location and register. The memory location address is given in the instruction.

Ex: LDA 850FH

This instruction is used to load the content of memory address 850FH in the accumulator.

Register Direct Addressing:

Register direct addressing transfer a copy of a byte or word from source register to destination register.

Ex: MOV B, C

It copies the content of register C to register B.

Indirect Addressing:

Indirect addressing transfers a byte or word between a register and a memory location.

Ex: MOV A, M

Here the data is in the memory location pointed to by the contents of HL pair. The data is moved to the accumulator.

Implicit Addressing

In this addressing mode the data itself specifies the data to be operated upon.

Ex: CMA

The instruction complements the content of the accumulator. No specific data or operand is mentioned in the instruction.

5. INSTRUCTION SET OF 8085

Data Transfer Instructions:

Opcode	Operand	Description
Copy from source to destination		
MOV	Rd, Rs M, Rs Rd, M	This instruction copies the contents of the source register into the destination register; the contents of the source register are not altered. If one of the operands is a memory location, its location is specified by the contents of the HL registers. Example: MOV B, C or MOV B, M
Move immediate 8-bit		
MVI	Rd, data M, data	The 8-bit data is stored in the destination register or memory. If the operand is a memory location, its location is specified by the contents of the HL registers. Example: MVI B, 57 or MVI M, 57
Load accumulator		
LDA	16-bit address	The contents of a memory location, specified by a 16-bit address in the operand, are copied to the accumulator. The contents of the source are not altered. Example: LDA 2034 or LDA XYZ
Load accumulator indirect		
LDAX	B/D Reg. pair	The contents of the designated register pair point to a memory location. This instruction copies the contents of that memory location into the accumulator. The contents of either the register pair or the memory location are not altered. Example: LDAX B
Load register pair immediate		
LXI	Reg. pair, 16-bit data	The instruction loads 16-bit data in the register pair designated in the operand. Example: LXI H, 2034
Load H and L registers direct		
LHLD	16-bit address	The instruction copies the contents of the memory location pointed out by the 16-bit address into register L and copies the contents of the next memory location into register H. The contents of source memory locations are not altered. Example: LHLD 2040

Store accumulator direct		
STA	16-bit address	The contents of the accumulator are copied into the memory location specified by the operand. This is a 3-byte instruction, the second byte specifies the low-order address and the third byte specifies the high-order address. Example: STA 4350 or STA XYZ
Store accumulator indirect		
STAX	Reg. pair	The contents of the accumulator are copied into the memory location specified by the contents of the operand (register pair). The contents of the accumulator are not altered. Example: STAX B
Store H and L registers direct		
SHLD	16-bit address	The contents of register L are stored into the memory location specified by the 16-bit address in the operand and the contents of H register are stored into the next memory location by incrementing the operand. The contents of registers HL are not altered. This is a 3-byte instruction, the second byte specifies the low-order address and the third byte specifies the high-order address. Example: SHLD 2470
Exchange H and L with D and E		
XCHG	none	The contents of register H are exchanged with the contents of register D, and the contents of register L are exchanged with the contents of register E. Example: XCHG
Copy H and L registers to the stack pointer		
SPHL	none	The instruction loads the contents of the H and L registers into the stack pointer register, the contents of the H register provide the high-order address and the contents of the L register provide the low-order address. The contents of the H and L registers are not altered. Example: SPHL
Exchange H and L with top of stack		
XTHL	none	The contents of the L register are exchanged with the stack location pointed out by the contents of the stack pointer register. The contents of the H register are exchanged with the next stack location (SP+1); however, the contents of the stack pointer register are not altered. Example: XTHL

Push register pair onto stack		
PUSH	Reg. pair	The contents of the register pair designated in the operand are copied onto the stack in the following sequence. The stack pointer register is decremented and the contents of the high-order register (B, D, H, A) are copied into that location. The stack pointer register is decremented again and the contents of the low-order register (C, E, L, flags) are copied to that location. Example: PUSH B or PUSH A
Pop off stack to register pair		
POP	Reg. pair	The contents of the memory location pointed out by the stack pointer register are copied to the low-order register (C, E, L, status flags) of the operand. The stack pointer is incremented by 1 and the contents of that memory location are copied to the high-order register (B, D, H, A) of the operand. The stack pointer register is again incremented by 1. Example: POP H or POP A
Output data from accumulator to a port with 8-bit address		
OUT	8-bit port address	The contents of the accumulator are copied into the I/O port specified by the operand. Example: OUT 87
Input data to accumulator from a port with 8-bit address		
IN	8-bit port address	The contents of the input port designated in the operand are read and loaded into the accumulator. Example: IN 82

Arithmetic Instructions:

Opcode	Operand	Description
Add register or memory to accumulator		
ADD	R M	The contents of the operand (register or memory) are added to the contents of the accumulator and the result is stored in the accumulator. If the operand is a memory location, its location is specified by the contents of the HL registers. All flags are modified to reflect the result of the addition. Example: ADD B or ADD M
Add register to accumulator with carry		
ADC	R M	The contents of the operand (register or memory) and the Carry flag are added to the contents of the accumulator and the result is stored in the accumulator. If the operand is a memory location, its location is specified by the contents of the HL registers. All flags are modified to reflect the result of the addition. Example: ADC B or ADC M
Add immediate to accumulator		
ADI	8-bit data	The 8-bit data (operand) is added to the contents of the accumulator and the result is stored in the accumulator. All flags are modified to reflect the result of the addition. Example: ADI 45
Add immediate to accumulator with carry		
ACI	8-bit data	The 8-bit data (operand) and the Carry flag are added to the contents of the accumulator and the result is stored in the accumulator. All flags are modified to reflect the result of the addition. Example: ACI 45
Add register pair to H and L registers		
DAD	Reg. pair	The 16-bit contents of the specified register pair are added to the contents of the HL register and the sum is stored in the HL register. The contents of the source register pair are not altered. If the result is larger than 16 bits, the CY flag is set. No other flags are affected. Example: DAD H

Subtract register or memory from accumulator

SUB	R M	The contents of the operand (register or memory) are subtracted from the contents of the accumulator, and the result is stored in the accumulator. If the operand is a memory location, its location is specified by the contents of the HL registers. All flags are modified to reflect the result of the subtraction. Example: SUB B or SUB M

Subtract source and borrow from accumulator

SBB	R M	The contents of the operand (register or memory) and the Borrow flag are subtracted from the contents of the accumulator and the result is placed in the accumulator. If the operand is a memory location, its location is specified by the contents of the HL registers. All flags are modified to reflect the result of the subtraction. Example: SBB B or SBB M

Subtract immediate from accumulator

SUI	8-bit data	The 8-bit data (operand) is subtracted from the contents of the accumulator and the result is stored in the accumulator. All flags are modified to reflect the result of the subtraction. Example: SUI 45

Subtract immediate from accumulator with borrow

SBI	8-bit data	The 8-bit data (operand) and the Borrow flag are subtracted from the contents of the accumulator and the result is stored in the accumulator. All flags are modified to reflect the result of the subtracion. Example: SBI 45

Increment register or memory by 1

INR	R M	The contents of the designated register or memory) are incremented by 1 and the result is stored in the same place. If the operand is a memory location, its location is specified by the contents of the HL registers. Example: INR B or INR M

Increment register pair by 1

INX	R	The contents of the designated register pair are incremented by 1 and the result is stored in the same place. Example: INX H

Decrement register or memory by 1		
DCR	R M	The contents of the designated register or memory are decremented by 1 and the result is stored in the same place. If the operand is a memory location, its location is specified by the contents of the HL registers. Example: DCR B or DCR M
Decrement register pair by 1		
DCX	R	The contents of the designated register pair are decremented by 1 and the result is stored in the same place. Example: DCX H
Decimal adjust accumulator		
DAA	none	The contents of the accumulator are changed from a binary value to two 4-bit binary coded decimal (BCD) digits. This is the only instruction that uses the auxiliary flag to perform the binary to BCD conversion, and the conversion procedure is described below. S, Z, AC, P, CY flags are altered to reflect the results of the operation. If the value of the low-order 4-bits in the accumulator is greater than 9 or if AC flag is set, the instruction adds 6 to the low-order four bits. If the value of the high-order 4-bits in the accumulator is greater than 9 or if the Carry flag is set, the instruction adds 6 to the high-order four bits. Example: DAA

BRANCHING INSTRUCTIONS

Opcode	Operand	Description
Jump unconditionally		
JMP	16-bit address	The program sequence is transferred to the memory location specified by the 16-bit address given in the operand. Example: JMP 2034 or JMP XYZ

Jump conditionally

Operand: 16-bit address

The program sequence is transferred to the memory location specified by the 16-bit address given in the operand based on the specified flag of the PSW as described below.
Example: JZ 2034 or JZ XYZ

Opcode	Description	Flag Status
JC	Jump on Carry	CY = 1
JNC	Jump on no Carry	CY = 0
JP	Jump on positive	S = 0
JM	Jump on minus	S = 1
JZ	Jump on zero	Z = 1
JNZ	Jump on no zero	Z = 0
JPE	Jump on parity even	P = 1
JPO	Jump on parity odd	P = 0

Unconditional subroutine call

CALL 16-bit address

The program sequence is transferred to the memory location specified by the 16-bit address given in the operand. Before the transfer, the address of the next instruction after CALL (the contents of the program counter) is pushed onto the stack.
Example: CALL 2034 or CALL XYZ

Call conditionally

Operand: 16-bit address

The program sequence is transferred to the memory location specified by the 16-bit address given in the operand based on the specified flag of the PSW as described below. Before the transfer, the address of the next instruction after the call (the contents of the program counter) is pushed onto the stack.
Example: CZ 2034 or CZ XYZ

Opcode	Description	Flag Status
CC	Call on Carry	CY = 1
CNC	Call on no Carry	CY = 0
CP	Call on positive	S = 0
CM	Call on minus	S = 1
CZ	Call on zero	Z = 1
CNZ	Call on no zero	Z = 0
CPE	Call on parity even	P = 1
CPO	Call on parity odd	P = 0

Return from subroutine unconditionally

RET none

The program sequence is transferred from the subroutine to the calling program. The two bytes from the top of the stack are copied into the program counter, and program execution begins at the new address.
Example: RET

Return from subroutine conditionally

Operand: none

The program sequence is transferred from the subroutine to the calling program based on the specified flag of the PSW as described below. The two bytes from the top of the stack are copied into the program counter, and program execution begins at the new address.
Example: RZ

Opcode	Description	Flag Status
RC	Return on Carry	CY = 1
RNC	Return on no Carry	CY = 0
RP	Return on positive	S = 0
RM	Return on minus	S = 1
RZ	Return on zero	Z = 1
RNZ	Return on no zero	Z = 0
RPE	Return on parity even	P = 1
RPO	Return on parity odd	P = 0

Load program counter with HL contents

PCHL	none	The contents of registers H and L are copied into the program counter. The contents of H are placed as the high-order byte and the contents of L as the low-order byte. Example: PCHL

Restart

RST	0-7	The RST instruction is equivalent to a 1-byte call instruction to one of eight memory locations depending upon the number. The instructions are generally used in conjunction with interrupts and inserted using external hardware. However these can be used as software instructions in a program to transfer program execution to one of the eight locations. The addresses are:

Instruction	Restart Address
RST 0	0000H
RST 1	0008H
RST 2	0010H
RST 3	0018H
RST 4	0020H
RST 5	0028H
RST 6	0030H
RST 7	0038H

The 8085 has four additional interrupts and these interrupts generate RST instructions internally and thus do not require any external hardware. These instructions and their Restart addresses are:

Interrupt	Restart Address
TRAP	0024H
RST 5.5	002CH
RST 6.5	0034H
RST 7.5	003CH

LOGICAL INSTRUCTIONS

Opcode	Operand	Description

Compare register or memory with accumulator

CMP	R
	M
		The contents of the operand (register or memory) are compared with the contents of the accumulator. Both contents are preserved . The result of the comparison is shown by setting the flags of the PSW as follows:
		if (A) < (reg/mem): carry flag is set, s=1
		if (A) = (reg/mem): zero flag is set, s=0
		if (A) > (reg/mem): carry and zero flags are reset, s=0
		Example: CMP B or CMP M

Compare immediate with accumulator

CPI	8-bit data	The second byte (8-bit data) is compared with the contents of the accumulator. The values being compared remain unchanged. The result of the comparison is shown by setting the flags of the PSW as follows:
		if (A) < data: carry flag is set, s=1
		if (A) = data: zero flag is set, s=0
		if (A) > data: carry and zero flags are reset, s=0
		Example: CPI 89

Logical AND register or memory with accumulator

ANA	R M	The contents of the accumulator are logically ANDed with the contents of the operand (register or memory), and the result is placed in the accumulator. If the operand is a memory location, its address is specified by the contents of HL registers. S, Z, P are modified to reflect the result of the operation. CY is reset. AC is set. Example: ANA B or ANA M

Logical AND immediate with accumulator

ANI	8-bit data	The contents of the accumulator are logically ANDed with the 8-bit data (operand) and the result is placed in the accumulator. S, Z, P are modified to reflect the result of the operation. CY is reset. AC is set. Example: ANI 86

Exclusive OR register or memory with accumulator

XRA	R M	The contents of the accumulator are Exclusive ORed with the contents of the operand (register or memory), and the result is placed in the accumulator. If the operand is a memory location, its address is specified by the contents of HL registers. S, Z, P are modified to reflect the result of the operation. CY and AC are reset. Example: XRA B or XRA M

Exclusive OR immediate with accumulator

XRI	8-bit data	The contents of the accumulator are Exclusive ORed with the 8-bit data (operand) and the result is placed in the accumulator. S, Z, P are modified to reflect the result of the operation. CY and AC are reset. Example: XRI 86

Logical OR register or memory with accumulaotr

ORA	R M	The contents of the accumulator are logically ORed with the contents of the operand (register or memory), and the result is placed in the accumulator. If the operand is a memory location, its address is specified by the contents of HL registers. S, Z, P are modified to reflect the result of the operation. CY and AC are reset. Example: ORA B or ORA M

Logical OR immediate with accumulator

ORI	8-bit data	The contents of the accumulator are logically ORed with the 8-bit data (operand) and the result is placed in the accumulator. S, Z, P are modified to reflect the result of the operation. CY and AC are reset. Example: ORI 86

Rotate accumulator left

RLC	none	Each binary bit of the accumulator is rotated left by one position. Bit D_7 is placed in the position of D_0 as well as in the Carry flag. CY is modified according to bit D_7. S, Z, P, AC are not affected. Example: RLC

Rotate accumulator right

RRC	none	Each binary bit of the accumulator is rotated right by one position. Bit D_0 is placed in the position of D_7 as well as in the Carry flag. CY is modified according to bit D_0. S, Z, P, AC are not affected. Example: RRC

Rotate accumulator left through carry

RAL none

Each binary bit of the accumulator is rotated left by one position through the Carry flag. Bit D_7 is placed in the Carry flag, and the Carry flag is placed in the least significant position D_0. CY is modified according to bit D_7. S, Z, P, AC are not affected.
Example: RAL

Rotate accumulator right through carry

RAR none

Each binary bit of the accumulator is rotated right by one position through the Carry flag. Bit D_0 is placed in the Carry flag, and the Carry flag is placed in the most significant position D_7. CY is modified according to bit D_0. S, Z, P, AC are not affected.
Example: RAR

Complement accumulator

CMA none

The contents of the accumulator are complemented. No flags are affected.
Example: CMA

Complement carry

CMC none

The Carry flag is complemented. No other flags are affected.
Example: CMC

Set Carry

STC none

The Carry flag is set to 1. No other flags are affected.
Example: STC

CONTROL INSTRUCTIONS

Opcode Operand Description

No operation

NOP none

No operation is performed. The instruction is fetched and decoded. However no operation is executed.
Example: NOP

Halt and enter wait state

HLT none

The CPU finishes executing the current instruction and halts any further execution. An interrupt or reset is necessary to exit from the halt state.
Example: HLT

Disable interrupts

DI none

The interrupt enable flip-flop is reset and all the interrupts except the TRAP are disabled. No flags are affected.
Example: DI

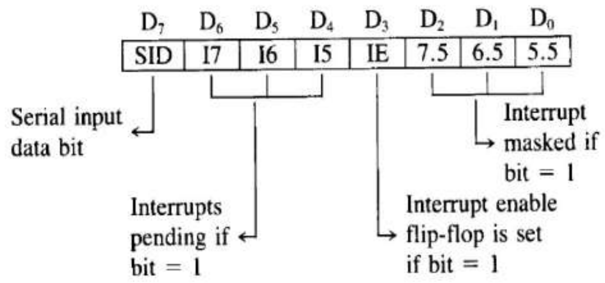
Enable interrupts

EI none

The interrupt enable flip-flop is set and all interrupts are enabled. No flags are affected. After a system reset or the acknowledgement of an interrupt, the interrupt enable flip-flop is reset, thus disabling the interrupts. This instruction is necessary to reenable the interrupts (except TRAP).
Example: EI

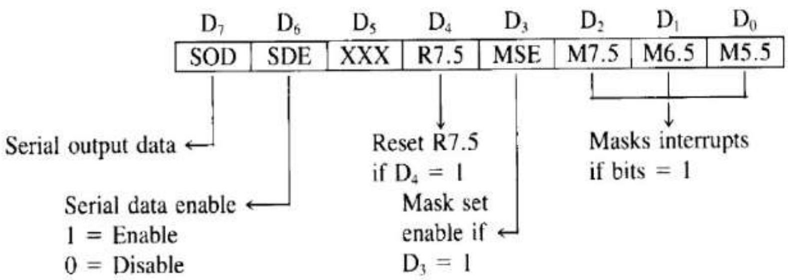
Read interrupt mask
RIM none

This is a multipurpose instruction used to read the status of interrupts 7.5, 6.5, 5.5 and read serial data input bit. The instruction loads eight bits in the accumulator with the following interpretations.
Example: RIM



Set interrupt mask
SIM none

This is a multipurpose instruction and used to implement the 8085 interrupts 7.5, 6.5, 5.5, and serial data output. The instruction interprets the accumulator contents as follows.
Example: SIM



- SOD — Serial Output Data: Bit D_7 of the accumulator is latched into the SOD output line and made available to a serial peripheral if bit $D_6 = 1$.
- SDE — Serial Data Enable: If this bit = 1, it enables the serial output. To implement serial output, this bit needs to be enabled.
- XXX — Don't Care
- R7.5 — Reset RST 7.5: If this bit = 1, RST 7.5 flip-flop is reset. This is an additional control to reset RST 7.5.
- MSE — Mask Set Enable: If this bit is high, it enables the functions of bits D_2, D_1, D_0. This is a master control over all the interrupt masking bits. If this bit is low, bits D_2, D_1, and D_0 do not have any effect on the masks.
- M7.5 — D_2 = 0, RST 7.5 is enabled.
  = 1, RST 7.5 is masked or disabled.
- M6.5 — D_1 = 0, RST 6.5 is enabled.
  = 1, RST 6.5 is masked or disabled.
- M5.5 — D_0 = 0, RST 5.5 is enabled.
  = 1, RST 5.5 is masked or disabled.

UNIT - 4 : INSTRUMENTATION TIMING AND CYCLE

6. INSTRUCTION EXECUTION AND TIMING DIAGRAM:

Each instruction in 8085 microprocessor consists of two part- operation code (opcode) and operand. The opcode is a command such as ADD and the operand is an object to be operated on, such as a byte or the content of a register.

Instruction Cycle: The time taken by the processor to complete the execution of an instruction. An instruction cycle consists of one to six machine cycles.

Machine Cycle: The time required to complete one operation; accessing either the memory or I/O device. A machine cycle consists of three to six T-states.
T-State: Time corresponding to one clock period. It is the basic unit to calculate execution of instructions or programs in a processor.

To execute a program, 8085 performs various operations as:

- Opcode fetch
- Operand fetch
- Memory read/write
- I/O read/write

External communication functions are:

- Memory read/write
- I/O read/write
- Interrupt request acknowledge

Opcode Fetch Machine Cycle:

It is the first step in the execution of any instruction. The timing diagram of this cycle is given in Fig. 7.

The following points explain the various operations that take place and the signals that are changed during the execution of opcode fetch machine cycle:

T1 clock cycle

i. The content of PC is placed in the address bus; AD0 - AD7 lines contains lower bit address and A8 – A15 contains higher bit address.
ii. $IO/\overline{M}$ signal is low indicating that a memory location is being accessed. S1 and S0 also changed to the levels as indicated in Table 1.
iii. ALE is high, indicates that multiplexed AD0 – AD7 act as lower order bus.

T2 clock cycle

i. Multiplexed address bus is now changed to data bus.
ii. The $\overline{RD}$ signal is made low by the processor. This signal makes the memory device load the data bus with the contents of the location addressed by the processor.

T3 clock cycle

i. The opcode available on the data bus is read by the processor and moved to the instruction register.
ii. The signal is deactivated by making it logic 1.

T4 clock cycle

i. The processor decode the instruction in the instruction register and generate the necessary control signals to execute the instruction. Based on the instruction further operations such as fetching, writing into memory etc takes place.

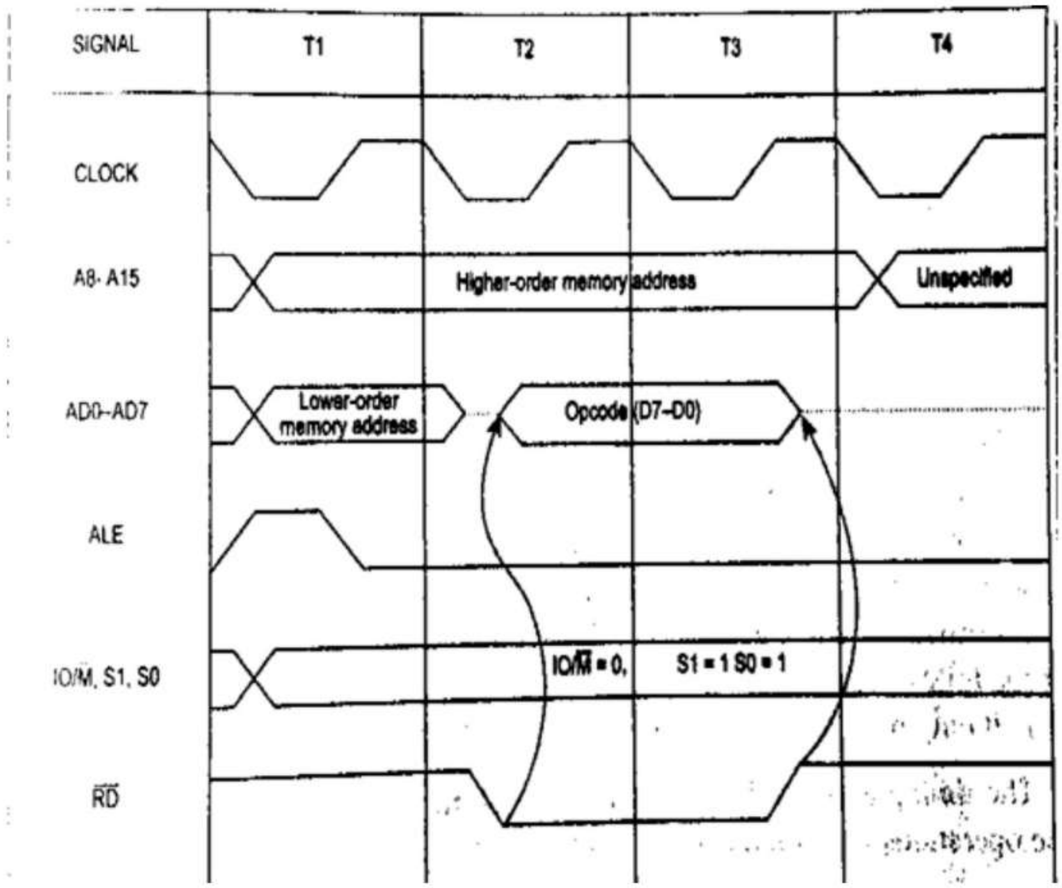


Fig. 7 Timing diagram for opcode fetch cycle

Memory Read Machine Cycle:

The memory read cycle is executed by the processor to read a data byte from memory. The machine cycle is exactly same to opcode fetch except: a) It has three T-states b) The S0 signal is set to 0. The timing diagram of this cycle is given in Fig. 8.

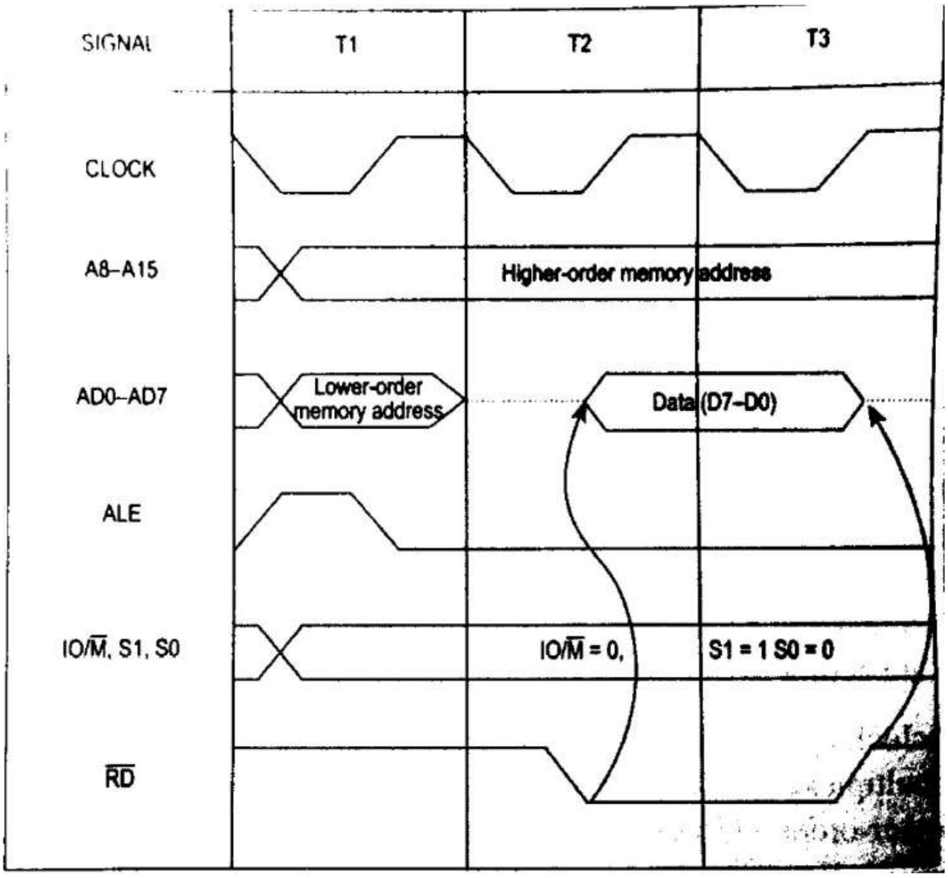


Fig. 8 Timing diagram for memory read machine cycle

Memory Write Machine Cycle:

The memory write cycle is executed by the processor to write a data byte in a memory location. The processor takes three T-states and $\overline{WR}$ signal is made low. The timing diagram of this cycle is given in Fig. 9.

I/O Read Cycle:

The I/O read cycle is executed by the processor to read a data byte from I/O port or from peripheral, which is I/O mapped in the system. The 8-bit port address is placed both in the lower and higher order address bus. The processor takes three T-states to execute this machine cycle. The timing diagram of this cycle is given in Fig. 10.

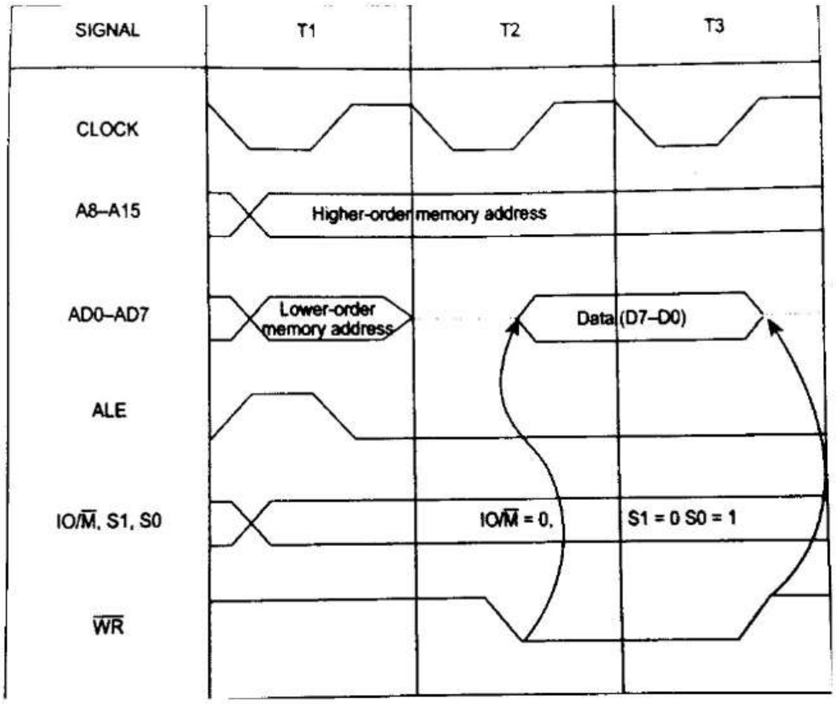


Fig. 9 Timing diagram for memory write machine cycle

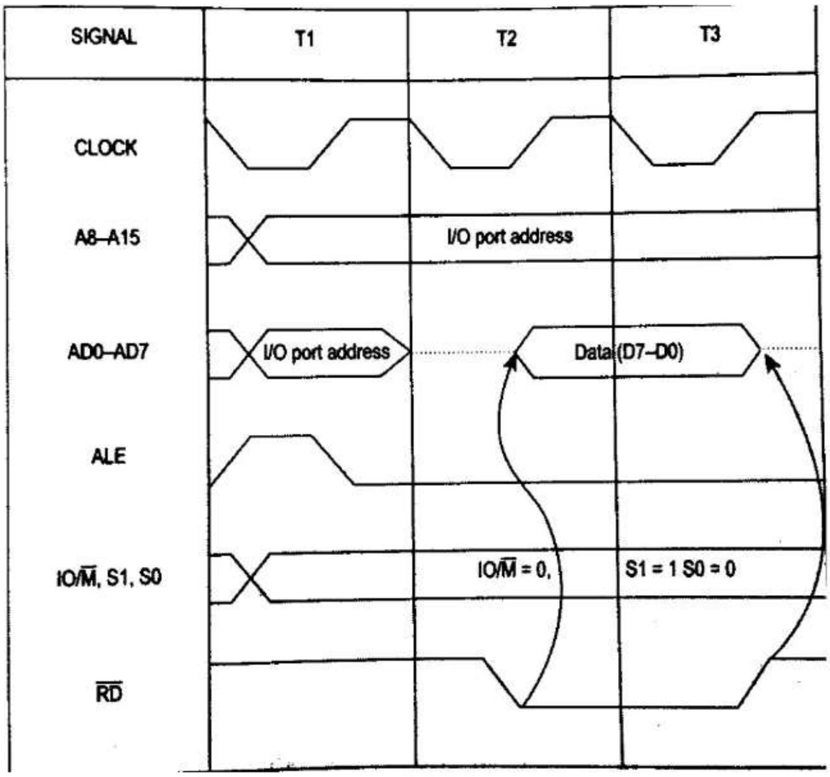


Fig. 10 Timing diagram I/O read machine cycle

I/O Write Cycle:

The I/O write cycle is executed by the processor to write a data byte to I/O port or to a peripheral, which is I/O mapped in the system. The processor takes three T-states to execute this machine cycle. The timing diagram of this cycle is given in Fig. 11.

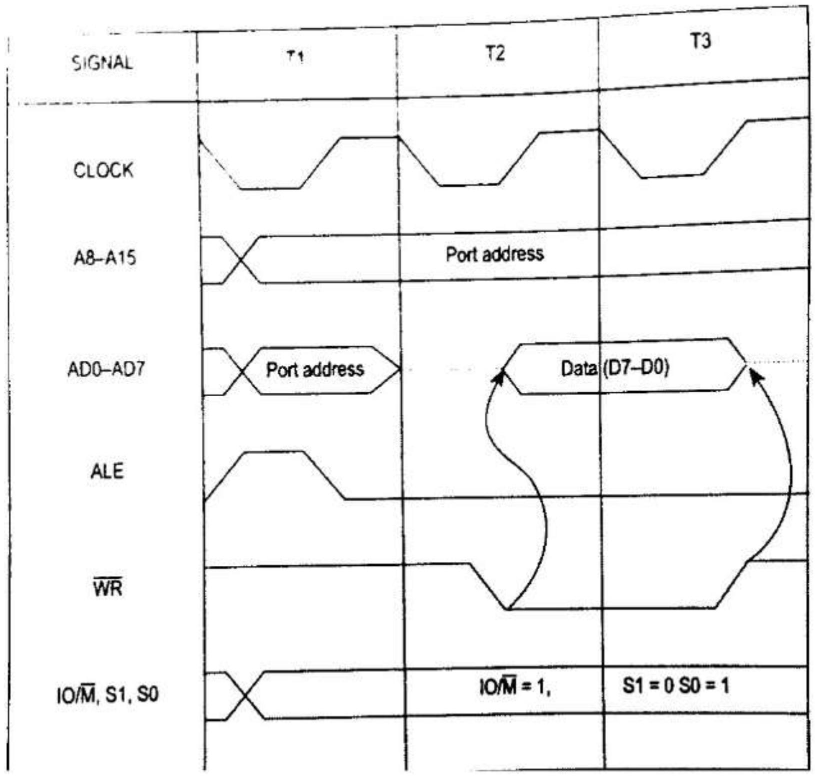


Fig. 11 Timing diagram I/O write machine cycle

Ex: Timing diagram for IN 80H.

The instruction and the corresponding codes and memory locations are given in Table 5.

Table 5 IN instruction

Address	Mnemonics	Opcode
800F	IN 80H	DB
8010		80

i. During the first machine cycle, the opcode DB is fetched from the memory, placed in the instruction register and decoded.
ii. During second machine cycle, the port address 80H is read from the next memory location.
iii. During the third machine cycle, the address 80H is placed in the address bus and the data read from that port address is placed in the accumulator.

The timing diagram is shown in Fig. 12.

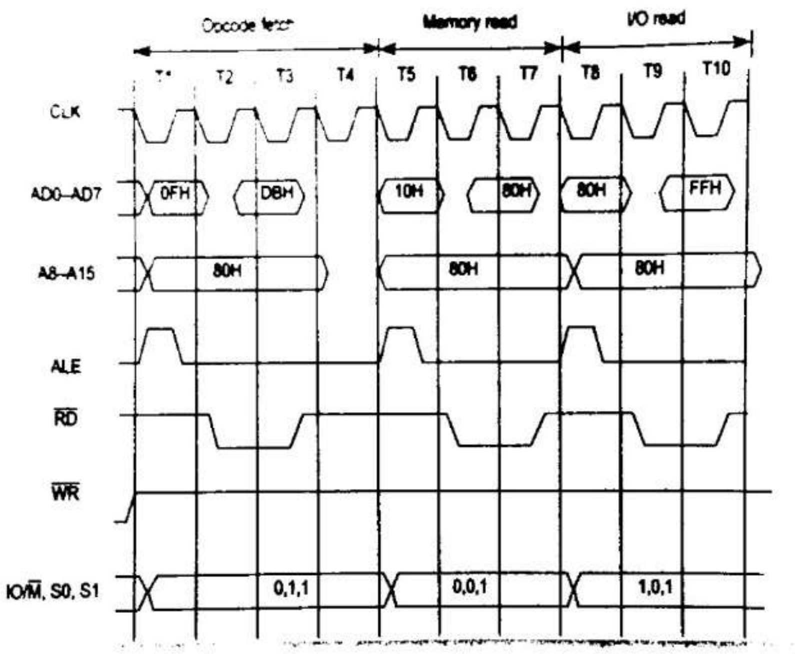


Fig. 12 Timing diagram for the IN instruction

UNIT - 5 : INTERRUPT

Interrupt Structure:

Interrupt is the mechanism by which the processor is made to transfer control from its current program execution to another program having higher priority. The interrupt signal may be given to the processor by any external peripheral device.

The program or the routine that is executed upon interrupt is called interrupt service routine (ISR). After execution of ISR, the processor must return to the interrupted program. Key features in the interrupt structure of any microprocessor are as follows:

i. Number and types of interrupt signals available.
ii. The address of the memory where the ISR is located for a particular interrupt signal. This address is called interrupt vector address (IVA).
iii. Masking and unmasking feature of the interrupt signals.
iv. Priority among the interrupts.
v. Timing of the interrupt signals.
vi. Handling and storing of information about the interrupt program (status information).

Types of Interrupts:

Interrupts are classified based on their maskability, IVA and source. They are classified as:

i. Vectored and Non-Vectored Interrupts
   - Vectored interrupts require the IVA to be supplied by the external device that gives the interrupt signal. This technique is vectoring, is implemented in number of ways.
   - Non-vectored interrupts have fixed IVA for ISRs of different interrupt signals.

ii. Maskable and Non-Maskable Interrupts
   - Maskable interrupts are interrupts that can be blocked. Masking can be done by software or hardware means.
   - Non-maskable interrupts are interrupts that are always recognized; the corresponding ISRs are executed.

iii. Software and Hardware Interrupts
   - Software interrupts are special instructions, after execution transfer the control to predefined ISR.
   - Hardware interrupts are signals given to the processor, for recognition as an interrupt and execution of the corresponding ISR.

Interrupt Handling Procedure:

The following sequence of operations takes place when an interrupt signal is recognized:

i. Save the PC content and information about current state (flags, registers etc) in the stack.
ii. Load PC with the beginning address of an ISR and start to execute it.
iii. Finish ISR when the return instruction is executed.
iv. Return to the point in the interrupted program where execution was interrupted.

Interrupt Sources and Vector Addresses in 8085:

Software Interrupts:

8085 instruction set includes eight software interrupt instructions called Restart (RST) instructions. These are one byte instructions that make the processor execute a subroutine at predefined locations. Instructions and their vector addresses are given in Table 6.

Table 6 Software interrupts and their vector addresses

Instruction	Machine hex code	Interrupt Vector Address
RST 0	C7	0000H
RST 1	CF	0008H
RST 2	D7	0010H
RST 3	DF	0018H
RST 4	E7	0020H
RST 5	EF	0028H
RST 6	F7	0030H
RST 7	FF	0032H

The software interrupts can be treated as CALL instructions with default call locations. The concept of priority does not apply to software interrupts as they are inserted into the program as instructions by the programmer and executed by the processor when the respective program lines are read.

Hardware Interrupts and Priorities:

8085 have five hardware interrupts – INTR, RST 5.5, RST 6.5, RST 7.5 and TRAP. Their IVA and priorities are given in Table 7.

Table 7 Hardware interrupts of 8085

Interrupt	Interrupt vector address	Maskable or non-maskable	Edge or level triggered	priority
TRAP	0024H	Non-makable	Level	1
RST 7.5	003CH	Maskable	Rising edge	2
RST 6.5	0034H	Maskable	Level	3
RST 5.5	002CH	Maskable	Level	4
INTR	Decided by hardware	Maskable	Level	5

Masking of Interrupts:

Masking can be done for four hardware interrupts INTR, RST 5.5, RST 6.5, and RST 7.5. The masking of 8085 interrupts is done at different levels. Fig. 13 shows the organization of hardware interrupts in the 8085.

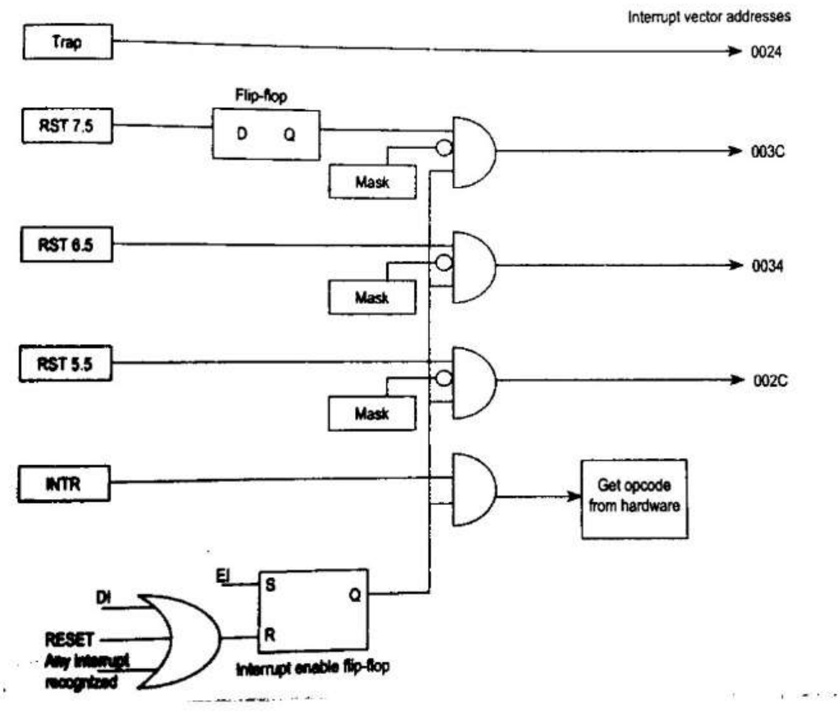


Fig. 13 Interrupt structure of 8085

The Fig. 13 is explained by the following five points:

i. The maskable interrupts are by default masked by the Reset signal. So no interrupt is recognized by the hardware reset.
ii. The interrupts can be enabled by the EI instruction.
iii. The three RST interrupts can be selectively masked by loading the appropriate word in the accumulator and executing SIM instruction. This is called software masking.
iv. All maskable interrupts are disabled whenever an interrupt is recognized.
v. All maskable interrupts can be disabled by executing the DI instruction.

RST 7.5 alone has a flip-flop to recognize edge transition. The DI instruction reset interrupt enable flip-flop in the processor and the interrupts are disabled. To enable interrupts, EI instruction has to be executed.

SIM Instruction:

The SIM instruction is used to mask or unmask RST hardware interrupts. When executed, the SIM instruction reads the content of accumulator and accordingly mask or unmask the interrupts. The format of control word to be stored in the accumulator before executing SIM instruction is as shown in Fig. 14.

Bit position	**D7**	**D6**	**D5**	**D4**	**D3**	**D2**	**D1**	**D0**
Name	SOD	SDE	X	R7.5	MSE	M7.5	M6.5	M5.5
Explanation	Serial data to be sent	Serial data enable— set to 1 for sending	Not used	Reset RST 7.5 flip-flop	Mask set enable— Set to 1 to mask interrupts	Set to 1 to mask RST 7.5	Set to 1 to mask RST 6.5	Set to 1 to mask RST 5.5

Fig. 14 Accumulator bit pattern for SIM instruction

In addition to masking interrupts, SIM instruction can be used to send serial data on the SOD line of the processor. The data to be send is placed in the MSB bit of the accumulator and the serial data output is enabled by making D6 bit to 1.

RIM Instruction:

RIM instruction is used to read the status of the interrupt mask bits. When RIM instruction is executed, the accumulator is loaded with the current status of the interrupt masks and the pending interrupts. The format and the meaning of the data stored in the accumulator after execution of RIM instruction is shown in Fig. 15.

In addition RIM instruction is also used to read the serial data on the SID pin of the processor. The data on the SID pin is stored in the MSB of the accumulator after the execution of the RIM instruction.

Bit position	D7	D6	D5	D4	D3	D2	D1	D0
Name	SID	I7.5	I6.5	I5.5	IE	M7.5	M6.5	M5.5
Explanation	Serial input data in the SID pin	Set to 1 if RST 7.5 is pending	Set to 1 if RST 6.5 is pending	Set to 1 if RST 5.5 is pending	Set to 1 if interrupts are enabled	Set to 1 if RST 7.5 is masked	Set to 1 if RST 6.5 is masked	Set to 1 if RST 5.5 is masked

Fig. 15 Accumulator bit pattern after execution of RIM instruction

Ex: Write an assembly language program to enables all the interrupts in 8085 after reset.

EI : Enable interrupts

MVI A, 08H : Unmask the interrupts

SIM : Set the mask and unmask using SIM instruction

Timing of Interrupts:

The interrupts are sensed by the processor one cycle before the end of execution of each instruction. An interrupts signal must be applied long enough for it to be recognized. The longest instruction of the 8085 takes 18 clock periods. So, the interrupt signal must be applied for at least 17.5 clock periods. This decides the minimum pulse width for the interrupt signal.

The maximum pulse width for the interrupt signal is decided by the condition that the interrupt signal must not be recognized once again. This is under the control of the programmer.

QUESTIONS:

1. What is the function of a microprocessor in a system?
2. Why is the data bus in 8085 bidirectional?
3. How does microprocessor differentiate between data and instruction?
4. How long would the processor take to execute the instruction LDA 1753H if the T-state duration is 2μs?
5. Draw the timing diagram of the instruction LDAX B.
6. Sketch and explain the various pins of the 8085.
7. Explain direct addressing mode of 8085 with an example?
8. Draw and explain the timing diagram of the instruction IN 82H.
9. What is meant by 'priority of the interrupts'? Explain the operation of the interrupts structure of the 8085, with the help of a circuit diagram.
10. Explain the bit pattern for SIM instruction. Write the assembly language program lines to enable all the interrupts in the 8085 after reset.
11. Write the logical instructions which affect and which does not affect flags in 8085.
12. Write an ALP in 8085 MPU to reject all the negative readings and add all the positive reading from a set of ten reading stored in memory locations starting at XX60H. When the sum exceeds eight bits produce output FFH to PORT1 to indicate overload otherwise display the sum.
13. Write an ALP in 8085 to eliminate the blanks (bytes with zero value) from a string of eight data bytes. Use two memory pointers: one to get a byte and the other to store the byte.
14. Design an up-down counter to count from 0 to 9 and 9 to 0 continuously with a 1.5 second delay between each count, and display the count at one of the output ports.

UNIT - 6 : MEMORY AND I/O INTERFACING

1. INTERFACING MEMORY AND I/O DEVICES WITH 8085

The programs and data that are executed by the microprocessor have to be stored in ROM/EPROM and RAM, which are basically semiconductor memory chips. The programs and data that are stored in ROM/EPROM are not erased even when power supply to the chip is removed. Hence, they are called non-volatile memory. They can be used to store permanent programs.

In a RAM, stored programs and data are erased when the power supply to the chip is removed. Hence, RAM is called volatile memory. RAM can be used to store programs and data that include, programs written during software development for a microprocessor based system, program written when one is learning assembly language programming and data enter while testing these programs.

Input and output devices, which are interfaced with 8085, are essential in any microprocessor based system. They can be interfaced using two schemes: I/O mapped I/O and memory-mapped I/O. In the I/O mapped I/O scheme, the I/O devices are treated differently from memory. In the memory-mapped I/O scheme, each I/O device is assumed to be a memory location.

2. INTERFACING MEMORY CHIPS WITH 8085

8085 has 16 address lines (A0 - A15), hence a maximum of 64 KB (= 2^{16} bytes) of memory locations can be interfaced with it. The memory address space of the 8085 takes values from 0000H to FFFFH.

The 8085 initiates set of signals such as $IO/\overline{M}$, $\overline{RD}$ and $\overline{WR}$ when it wants to read from and write into memory. Similarly, each memory chip has signals such as $\overline{CE}$ or $\overline{CS}$ (chip enable or chip select), $\overline{OE}$ or $\overline{RD}$ (output enable or read) and $\overline{WE}$ or $\overline{WR}$ (write enable or write) associated with it.

Generation of Control Signals for Memory:

When the 8085 wants to read from and write into memory, it activates $IO/\overline{M}$, $\overline{RD}$ and $\overline{WR}$ signals as shown in Table 8.

Table 8 Status of $IO/\overline{M}$, $\overline{RD}$ and $\overline{WR}$ signals during memory read and write operations

$IO/\overline{M}$	$\overline{RD}$	$\overline{WR}$	Operation
0	0	1	8085 reads data from memory
0	1	0	8085 writes data into memory

Using $IO/\overline{M}$, $\overline{RD}$ and $\overline{WR}$ signals, two control signals $\overline{MEMR}$ (memory read) and $\overline{MEMW}$ (memory write) are generated. Fig. 16 shows the circuit used to generate these signals.

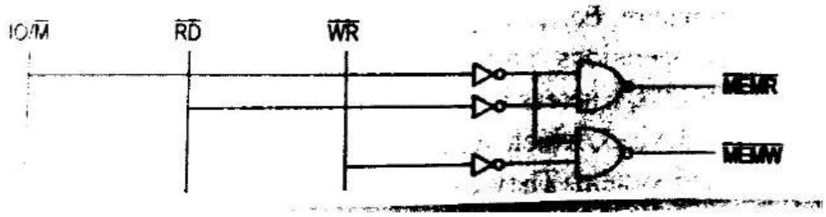


Fig. 16 Circuit used to generate and signals

When is high, both memory control signals are deactivated irrespective of the status of and signals.

Ex: Interface an IC 2764 with 8085 using NAND gate address decoder such that the address range allocated to the chip is 0000H – 1FFFH.

Specification of IC 2764:

- 8 KB (8×2^{10} byte) EPROM chip
- 13 address lines (2^{13} bytes = 8 KB)

Interfacing:

- 13 address lines of IC are connected to the corresponding address lines of 8085.
- Remaining address lines of 8085 are connected to address decoder formed using logic gates, the output of which is connected to the pin of IC.
- Address range allocated to the chip is shown in Table 9.
- Chip is enabled whenever the 8085 places an address allocated to EPROM chip in the address bus. This is shown in Fig. 17.

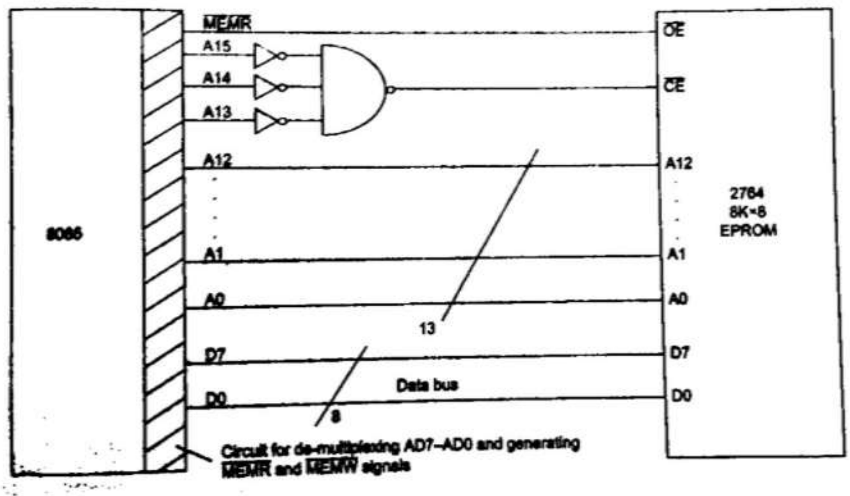


Fig. 17 Interfacing IC 2764 with the 8085

Table 9 Address allocated to IC 2764

A15	A14	A13	A12	A11	A10	A9	A8	A7	A6	A5	A4	A3	A2	A1	A0	Address
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0000H
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0001H
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1FFEH
0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1FFFH

Ex: Interface a 6264 IC (8K x 8 RAM) with the 8085 using NAND gate decoder such that the starting address assigned to the chip is 4000H.

Specification of IC 6264:

- 8K x 8 RAM
- 8 KB = 2^{13} bytes
- 13 address lines

The ending address of the chip is 5FFFH (since 4000H + 1FFFH = 5FFFH). When the address 4000H to 5FFFH are written in binary form, the values in the lines A15, A14, A13 are 0, 1 and 0 respectively. The NAND gate is designed such that when the lines A15 and A13 carry 0 and A14 carries 1, the output of the NAND gate is 0. The NAND gate output is in turn connected to the $\overline{CE1}$ pin of the RAM chip. A NAND output of 0 selects the RAM chip for read or write operation, since CE2 is already 1 because of its connection to +5V. Fig. 18 shows the interfacing of IC 6264 with the 8085.

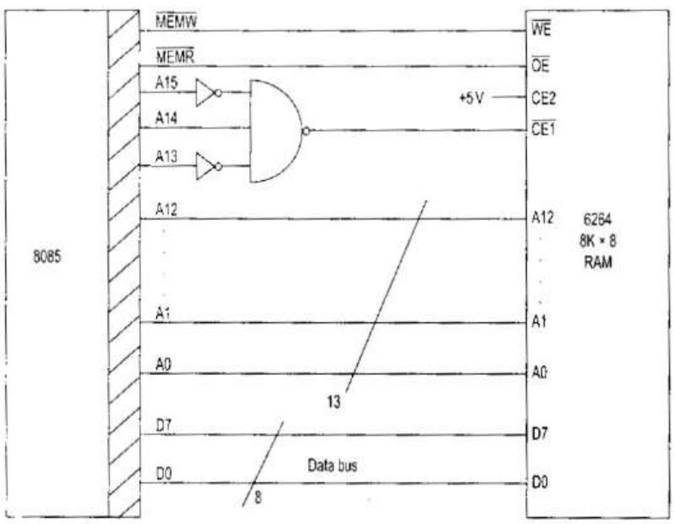


Fig. 18 Interfacing 6264 IC with the 8085

Ex: Interface two 6116 ICs with the 8085 using 74LS138 decoder such that the starting addresses assigned to them are 8000H and 9000H, respectively.

Specification of IC 6116:

- 2 K x 8 RAM
- 2 KB = 2^{11} bytes
- 11 address lines

6116 has 11 address lines and since 2 KB, therefore ending addresses of 6116 chip 1 is and chip 2 are 87FFH and 97FFH, respectively. Table 10 shows the address range of the two chips.

Table 10 Address range for IC 6116

A15	A14	A13	A12	A11	A10	A9	A8	A7	A6	A5	A4	A3	A2	A1	A0	Address
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8000H
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	87FFH (RAM chip 1)
1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	9000H
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	97FFH (RAM chip 2)

Interfacing:

- Fig. 19 shows the interfacing.
- A0 – A10 lines of 8085 are connected to 11 address lines of the RAM chips.
- Three address lines of 8085 having specific value for a particular RAM are connected to the three select inputs (C, B and A) of 74LS138 decoder.
- Table 10 shows that A13=A12=A11=0 for the address assigned to RAM 1 and A13=0, A12=1 and A11=0 for the address assigned to RAM 2.
- Remaining lines of 8085 which are constant for the address range assigned to the two RAM are connected to the enable inputs of decoder.
- When 8085 places any address between 8000H and 87FFH in the address bus, the select inputs C, B and A of the decoder are all 0. The Y0 output of the decoder is also 0, selecting RAM 1.
- When 8085 places any address between 9000H and 97FFH in the address bus, the select inputs C, B and A of the decoder are 0, 1 and 0. The Y2 output of the decoder is also 0, selecting RAM 2.

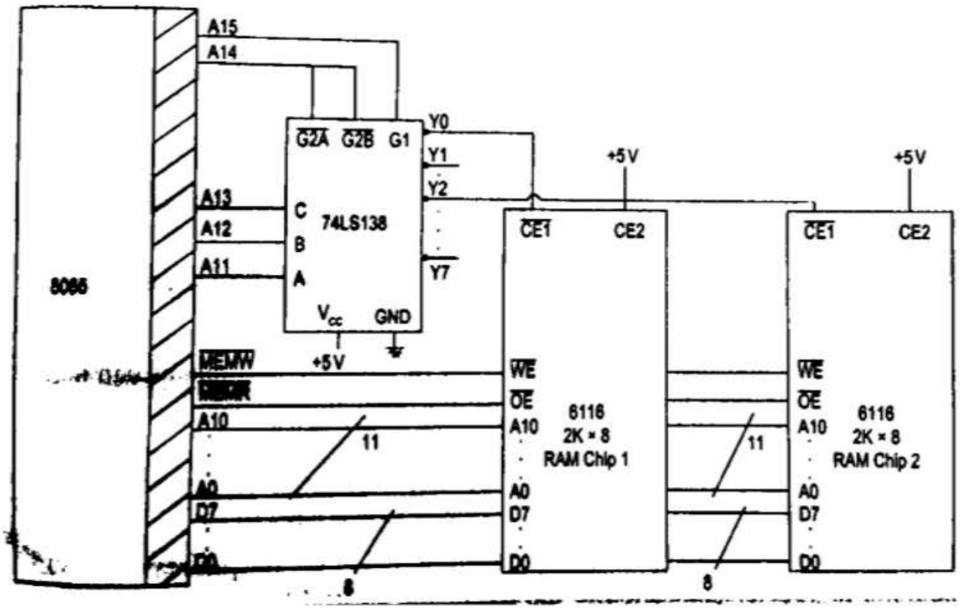


Fig. 19 Interfacing two 6116 RAM chips using 74LS138 decoder

3. PERIPHERAL MAPPED I/O INTERFACING

In this method, the I/O devices are treated differently from memory chips. The control signals I/O read ($\overline{IOR}$) and I/O write ($\overline{IOW}$), which are derived from the , and signals of the 8085, are used to activate input and output devices, respectively. Generation of these control signals is shown in Fig. 20. Table 11 shows the status of , and signals during I/O read and I/O write operation.

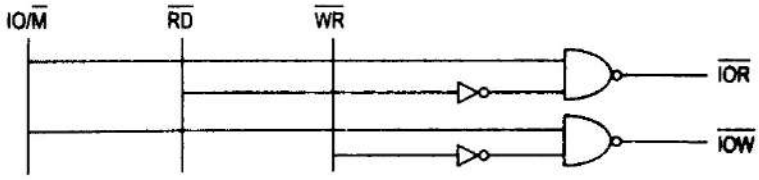


Fig. 20 Generation of $\overline{IOR}$ and $\overline{IOW}$ signals

IN instruction is used to access input device and OUT instruction is used to access output device. Each I/O device is identified by a unique 8-bit address assigned to it. Since the control signals used to access input and output devices are different, and all I/O device use 8-bit address, a maximum of 256 (2^8) input devices and 256 output devices can be interfaced with 8085.

Table 11 Status of $\overline{IOR}$ and $\overline{IOW}$ signals in 8085.

			$\overline{IOR}$	$\overline{IOW}$	Operation
1	0	1	0	1	I/O read operation
1	1	0	1	0	I/O write operation
0	X	X	1	1	Memory read or write operation

Ex: Interface an 8-bit DIP switch with the 8085 such that the address assigned to the DIP switch if F0H.

IN instruction is used to get data from DIP switch and store it in accumulator. Steps involved in the execution of this instruction are:

i. Address F0H is placed in the lines A0 – A7 and a copy of it in lines A8 – A15.
ii. The $\overline{\text{IOR}}$ signal is activated ($\overline{\text{IOR}}$ = 0), which makes the selected input device to place its data in the data bus.
iii. The data in the data bus is read and store in the accumulator.

Fig. 21 shows the interfacing of DIP switch.

A7	A6	A5	A4	A3	A2	A1	A0	
1	1	1	1	0	0	0	0	= F0H

A0 – A7 lines are connected to a NAND gate decoder such that the output of NAND gate is 0. The output of NAND gate is ORed with the $\overline{\text{IOR}}$ signal and the output of OR gate is connected to $\overline{\text{1G}}$ and $\overline{\text{2G}}$ of the 74LS244. When 74LS244 is enabled, data from the DIP switch is placed on the data bus of the 8085. The 8085 read data and store in the accumulator. Thus data from DIP switch is transferred to the accumulator.

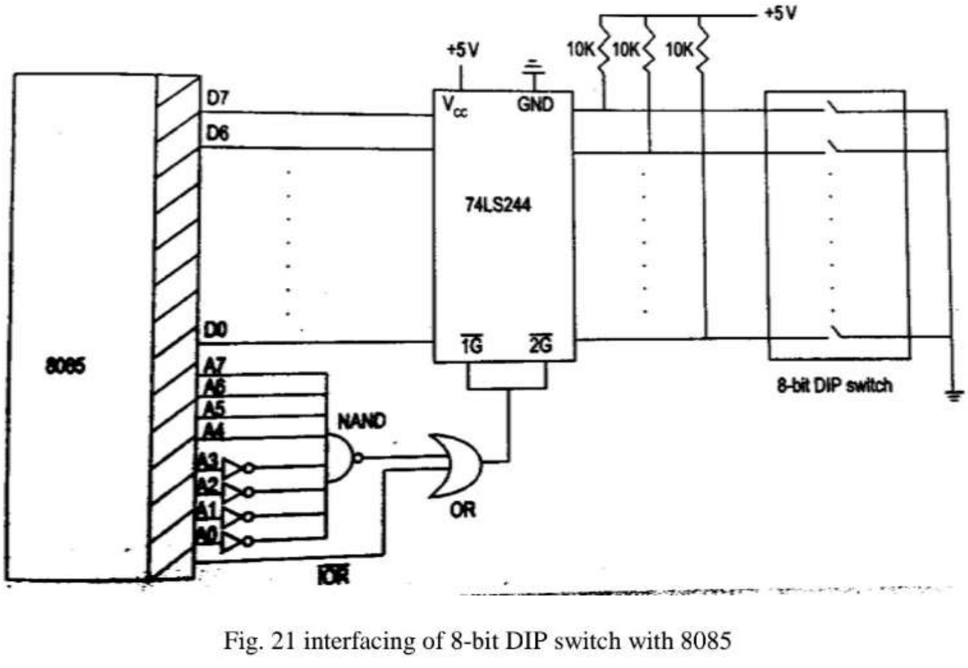


Fig. 21 interfacing of 8-bit DIP switch with 8085

4. **MEMORY MAPPED I/O INTERFACING**

In memory-mapped I/O, each input or output device is treated as if it is a memory location. The and control signals are used to activate the devices. Each input or output device is identified by unique 16-bit address, similar to 16-bit address assigned to memory location. All memory related instruction like LDA 2000H, LDAX B, MOV A, M can be used.

Since the I/O devices use some of the memory address space of 8085, the maximum memory capacity is lesser than 64 KB in this method.

Ex: Interface an 8-bit DIP switch with the 8085 using logic gates such that the address assigned to it is F0F0H.

Since a 16-bit address has to be assigned to a DIP switch, the memory-mapped I/O technique must be used. Using LDA F0F0H instruction, the data from the 8-bit DIP switch can be transferred to the accumulator. The steps involved are:

i. The address F0F0H is placed in the address bus A0 – A15.
ii. The signal is made low for some time.
iii. The data in the data bus is read and stored in the accumulator.

Fig. 22 shows the interfacing diagram.

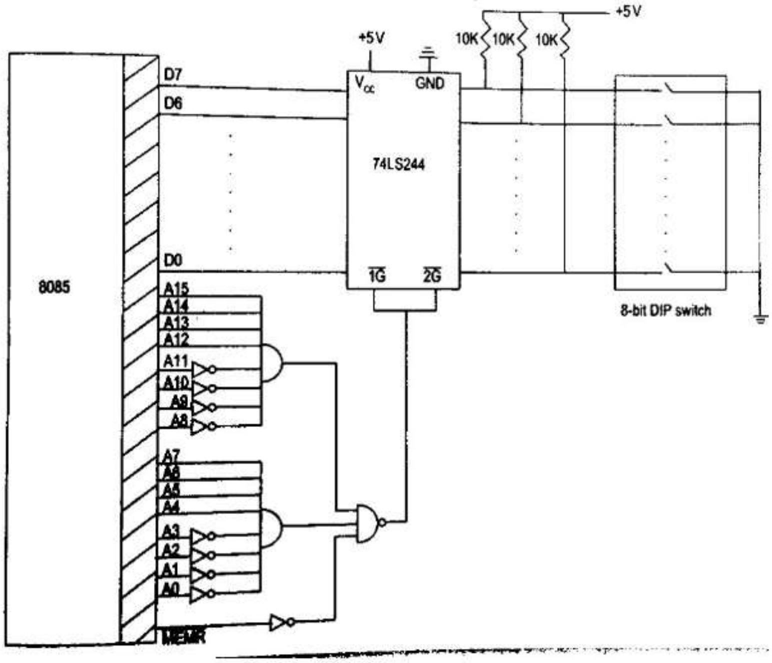


Fig. 22 Interfacing 8-bit DIP switch with 8085

When 8085 executes the instruction LDA F0F0H, it places the address F0F0H in the address lines A0 – A15 as:

A15	A14	A13	A12	A11	A10	A9	A8	A7	A6	A5	A4	A3	A2	A1	A0	
1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	= F0F0H

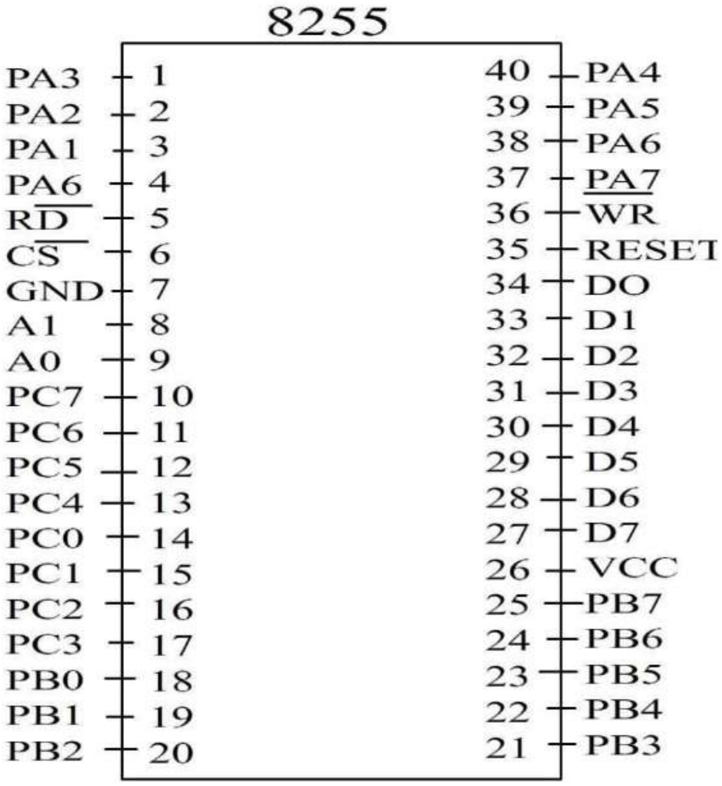
The address lines are connected to AND gates. The output of these gates along with signal are connected to a NAND gate, so that when the address F0F0H is placed in the address bus and = 0 its output becomes 0, thereby enabling the buffer 74LS244. The data from the DIP switch is placed in the 8085 data bus. The 8085 reads the data from the data bus and stores it in the accumulator.

UNIT - 7 : PERIPHERAL DEVICE

INTEL 8255: (Programmable Peripheral Interface)

The 8255A is a general purpose programmable I/O device designed for use with Intel microprocessors. It consists of three 8-bit bidirectional I/O ports (24I/O lines) that can be configured to meet different system I/O needs. The three ports are PORT A, PORT B & PORT C. Port A contains one 8-bit output latch/buffer and one 8-bit input buffer. Port B is same as PORT A or PORT B. However, PORT C can be split into two parts PORT C lower (PC_0-PC_3) and PORT C upper (PC_7-PC_4) by the control word. The three ports are divided in two groups Group A (PORT A and upper PORT C) Group B (PORT B and lower PORT C). The two groups can be programmed in three different modes. In the first mode (mode 0), each group may be programmed in either input mode or output mode (PORT A, PORT B, PORT C lower, PORT C upper). In mode 1, the second's mode, each group may be programmed to have 8-lines of input or output (PORT A or PORT B) of the remaining 4-lines (PORT C lower or PORT C upper) 3-lines are used for hand shaking and interrupt control signals. The third mode of operation (mode 2) is a bidirectional bus mode which uses 8-line (PORT A only for a bidirectional bus and five lines (PORT C upper 4 lines and borrowing one from other group) for handshaking.

The 8255 is contained in a 40-pin package, whose pin out is shown below:



PIN Names

RESET – Reset input

$\overline{CS}$ - Chip selected

$\overline{RD}$ - Read input

$\overline{WR}$ - Write input

A_0 A_1 – Port Address

$PA_7 – PA_0$ – PORT A

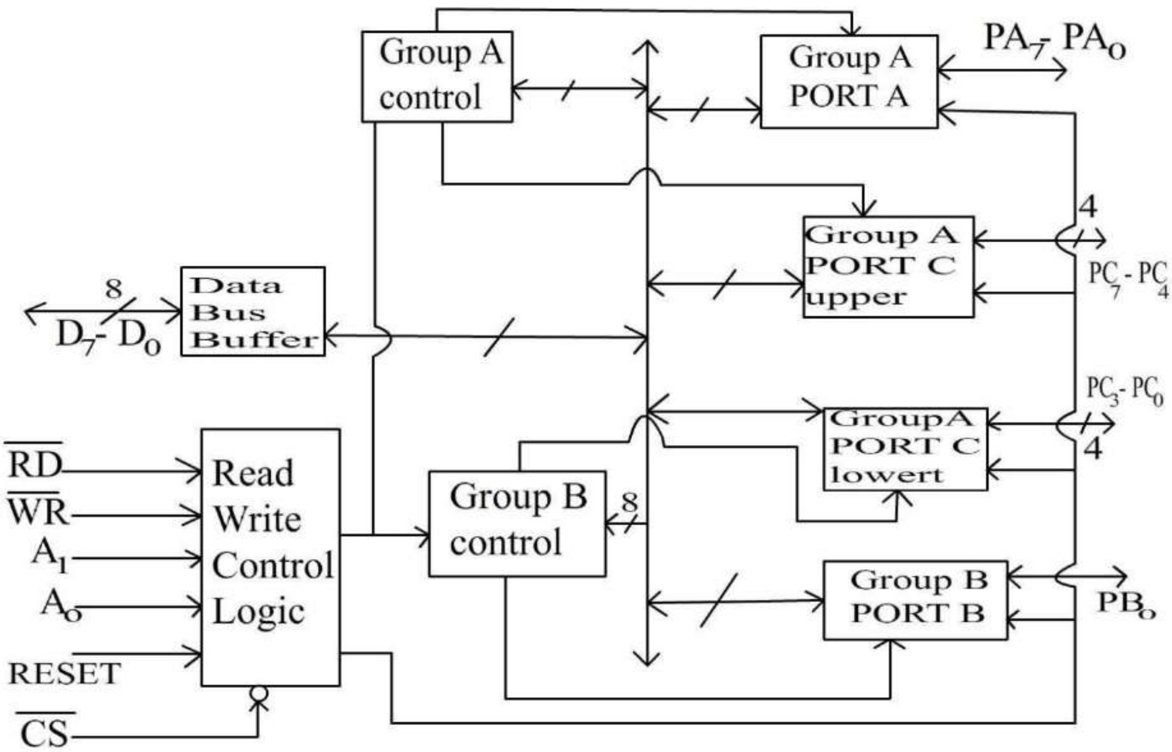
$PB_7 – PB_0$ – PORT B

$PC_7 – PC_0$ – PORT C

VCC - +5v

GND - Ground

The block diagram is shown below:



Functional Description:

This support chip is a general purpose I/O component to interface peripheral equipment to the microcomputer system bus. It is programmed by the system software so that normally no external logic is necessary to interface peripheral devices or structures.

Data Bus Buffer:

It is a tri-state 8-bit buffer used to interface the chip to the system data bus. Data is transmitted or received by the buffer upon execution of input or output instructions by the CPU. Control words and status information are also transferred through the data bus buffer. The data lines are connected to BDB of μp.

Read/Write and logic control:

The function of this block is to control the internal operation of the device and to control the transfer of data and control or status words. It accepts inputs from the CPU address and control buses and in turn issues command to both the control groups.

$\overline{CS}$ Chip Select:

A low on this input selects the chip and enables the communication between the 8255 A & the CPU. It is connected to the output of address decode circuitry to select the device when it $\overline{RD}$ (Read). A low on this input enables the 8255 to send the data or status information to the CPU on the data bus.

$\overline{WR}$ (Write):

A low on this input pin enables the CPU to write data or control words into the 8255 A.

A_1, A_0 port select:

These input signals, in conjunction with the $\overline{RD}$ and $\overline{WR}$ inputs, control the selection of one of the three ports or the control word registers. They are normally connected to the least significant bits of the address bus (A_0 and A_1).

Following Table gives the basic operation,

A_1	A_0	$\overline{RD}$	$\overline{WR}$	$\overline{CS}$	Input operation
0	0	0	1	0	PORT A $\longrightarrow$ Data bus
0	1	0	1	0	PORT B $\longrightarrow$ Data bus
1	0	0	1	0	PORT C $\longrightarrow$ Data bus
0	0	1	0	0	Output operation Data bus $\longrightarrow$ PORT A
0	1	1	0	0	Data bus $\longrightarrow$ PORT B
1	0	1	0	0	Data bus $\longrightarrow$ PORT C
1	1	1	0	0	Data bus $\longrightarrow$ control

All other states put data bus into tri-state/illegal condition.

RESET:

A high on this input pin clears the control register and all ports (A, B & C) are initialized to input mode. This is connected to RESET OUT of 8255. This is done to prevent destruction of circuitry connected to port lines. If port lines are initialized as output after a power up or reset, the port might try to output into the output of a device connected to same inputs might destroy one or both of them.

PORTs A, B and C:

The 8255A contains three 8-bit ports (A, B and C). All can be configured in a variety of functional characteristic by the system software.

PORTA:

One 8-bit data output latch/buffer and one 8-bit data input latch.

PORT B:

One 8-bit data output latch/buffer and one 8-bit data input buffer.

PORT C:

One 8-bit data output latch/buffer and one 8-bit data input buffer (no latch for input). This port can be divided into two 4-bit ports under the mode control. Each 4-bit port contains a 4-bit latch and it can be used for the control signal outputs and status signals inputs in conjunction with ports A and B.

Group A & Group B control:

The functional configuration of each port is programmed by the system software. The control words outputted by the CPU configure the associated ports of the each of the two groups. Each control block accepts command from Read/Write content logic receives control words from the internal data bus and issues proper commands to its associated ports.

Control Group A – Port A & Port C upper

Control Group B – Port B & Port C lower

The control word register can only be written into No read operation if the control word register is allowed.

Operation Description:

Mode selection:

There are three basic modes of operation that can be selected by the system software.

Mode 0: Basic Input/output

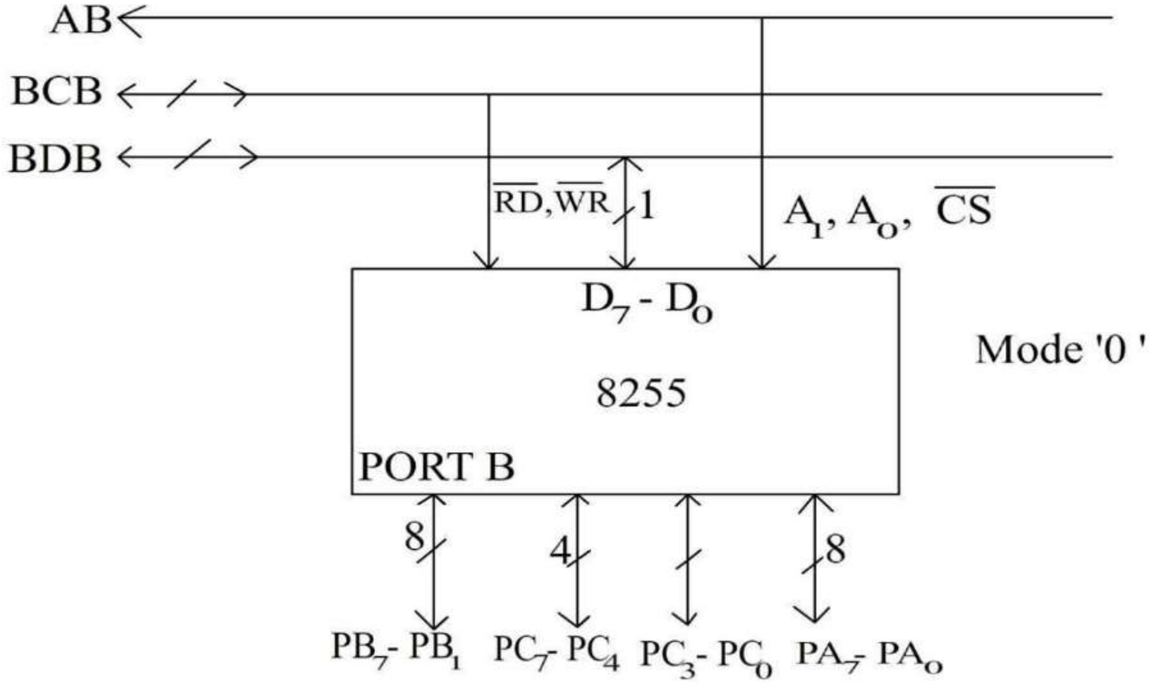
Mode 1: Strobes Input/output

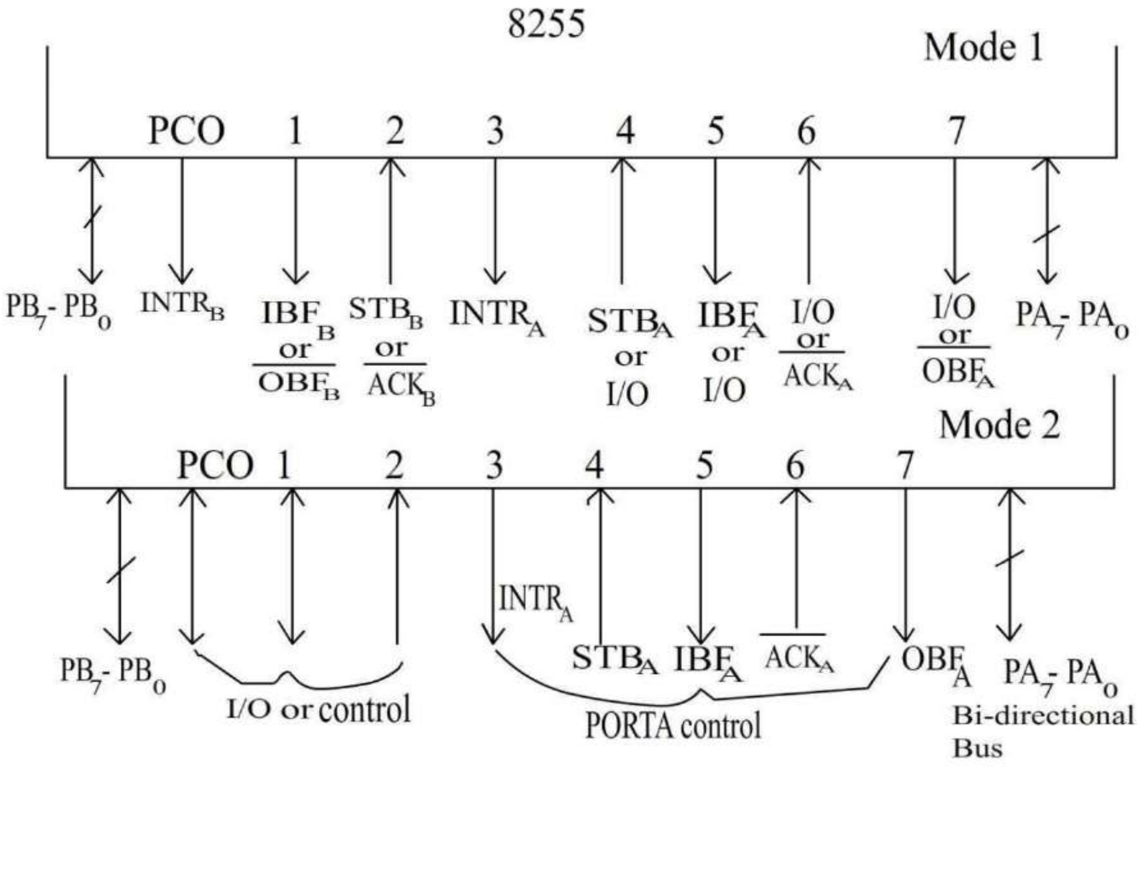
Mode 2: Bi-direction bus.

When the reset input goes HIGH all poets are set to mode'0' as input which means all 24 lines are in high impedance state and can be used as normal input. After the reset is removed the 8255A remains in the input mode with no additional initialization. During the execution of the program any of the other modes may be selected using a single output instruction.

The modes for PORT A & PORT B can be separately defined, while PORT C is divided into two portions as required by the PORT A and PORT B definitions. The ports are thus divided into two groups Group A & Group B. All the output register, including the status flip-flop will be reset whenever the mode is changed. Modes of the two group may be combined for any desired I/O operation e.g. Group A in mode '1' and group B in mode '0'.

The basic mode definitions with bus interface and the mode definition format are given in fig (a) & (b),


AB
BCB
BDB
$\overline{RD}$, $\overline{WR}$
1
A_1, A_0, $\overline{CS}$
D_7 - D_0
8255
Mode '0 '
PORT B
8
4
8
PB_7 - PB_1
PC_7 - PC_4
PC_3 - PC_0
PA_7 - PA_0

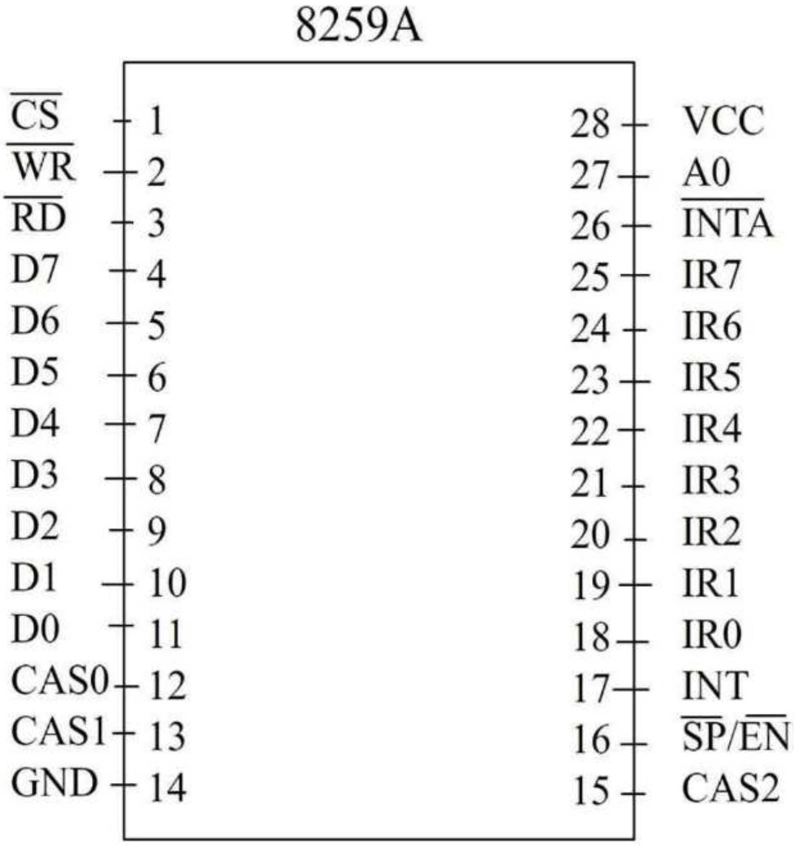

8255
Mode 1
PCO 1 2 3 4 5 6 7
PB_7 - PB_0
$INTR_B$
IBF_B or $\overline{OBF_B}$
STB_B or $\overline{ACK_B}$
$INTR_A$
STB_A or I/O
IBF_A or I/O
I/O or $\overline{ACK_A}$
I/O or $\overline{OBF_A}$
PA_7 - PA_0
Mode 2
PCO 1 2 3 4 5 6 7
$INTR_A$
PB_7 - PB_0
STB_A
IBF_A
$\overline{ACK_A}$
OBF_A
PA_7 - PA_0
Bi-directional
Bus
I/O or control
PORTA control

INTEL 8259A Programmable Interrupt Controller

The 8259A is a programmable interrupt controller designed to work with Intel microprocessor 8080 A, 8085, 8086, 8088. The 8259 A interrupt controller can

1) Handle eight interrupt inputs. This is equivalent to providing eight interrupt pins on the processor in place of one INTR/INT pin.
2) Vector an interrupt request anywhere in the memory map. However, all the eight interrupt are spaced at the interval of either four or eight location. This eliminates the major drawback, 8085 interrupt, in which all interrupts are vectored to memory location on page 00_H.
3) Resolve eight levels of interrupt priorities in a variety of modes.
4) Mask each interrupt request individually.
5) Read the status of pending interrupts, in service interrupts, and masked interrupts.
6) Be set up to accept either the level triggered or edge triggered interrupt request.
7) Mine 8259 as can be cascade in a master slave configuration to handle 64 interrupt inputs.

The 8259 A is contained in a 28-element in line package that requires only a compatible with 8259. The main difference between the two is that the 8259 A can be used with Intel 8086/8088 processor. It also induces additional features such as level triggered mode, buffered mode and automatic end of interrupt mode. The pin diagram and interval block diagram is shown below:



The pins are defined as follows:

<u>$\overline{CS}$: Chip select</u>

To access this chip, $\overline{CS}$ is made low. A LOW on this pin enables $\overline{RD}$ & $\overline{WR}$ communication between the CPU and the 8259A. This pin is connected to address bus through the decoder logic circuits. INTA functions are independent of $\overline{CS}$

<u>$\overline{WR}$:</u>

A low on this pin. When $\overline{CS}$ is low enables the 8259 A to accept command words from CPU.

$\overline{RD}$:

A low on this pin when $\overline{CS}$ is low enables these 8259 A to release status on to the data bus for the CPU. The status in dudes the contents of IMR, ISR or TRR register or a priority level.

D_7-D_0:

Bidirectional data bus control status and interrupt in a this bus. This bus is connected to BDB of 8085.

CAS_0-CAS_2:

Cascade lines: The CAS lines form a private 8259A bus to control a multiple 8259A structure ie to identify a particular slave device. These pins are outputs of a master 8259A and inputs for a slave 8259A.

$\overline{SP}/\overline{EN}$: Salve program/enable buffer:

This is a dual function pin. It is used as an input to determine whether the 8259A is to a master ($\overline{SP}/\overline{EN}$ = 1) or as a slave ($\overline{SP}/\overline{EN}$ = 0). It is also used as an output to disable the data bus transceivers when data are being transferred from the 8259A to the CPU. When in buffered mode, it can be used as an output and when not in the buffered mode it is used as an input.

INT:

This pin goes high whenever a valid interrupt request is asserted. It is used to interrupt the CPU, thus it is connected to the CPU's interrupt pin (INTR).

$\overline{INTA}$:

Interrupt: Acknowledge. This pin is used to enable 8259A interrupt vector data on the data bus by a sequence of interrupt request pulses issued by the CPU.

IR_0-IR_7:

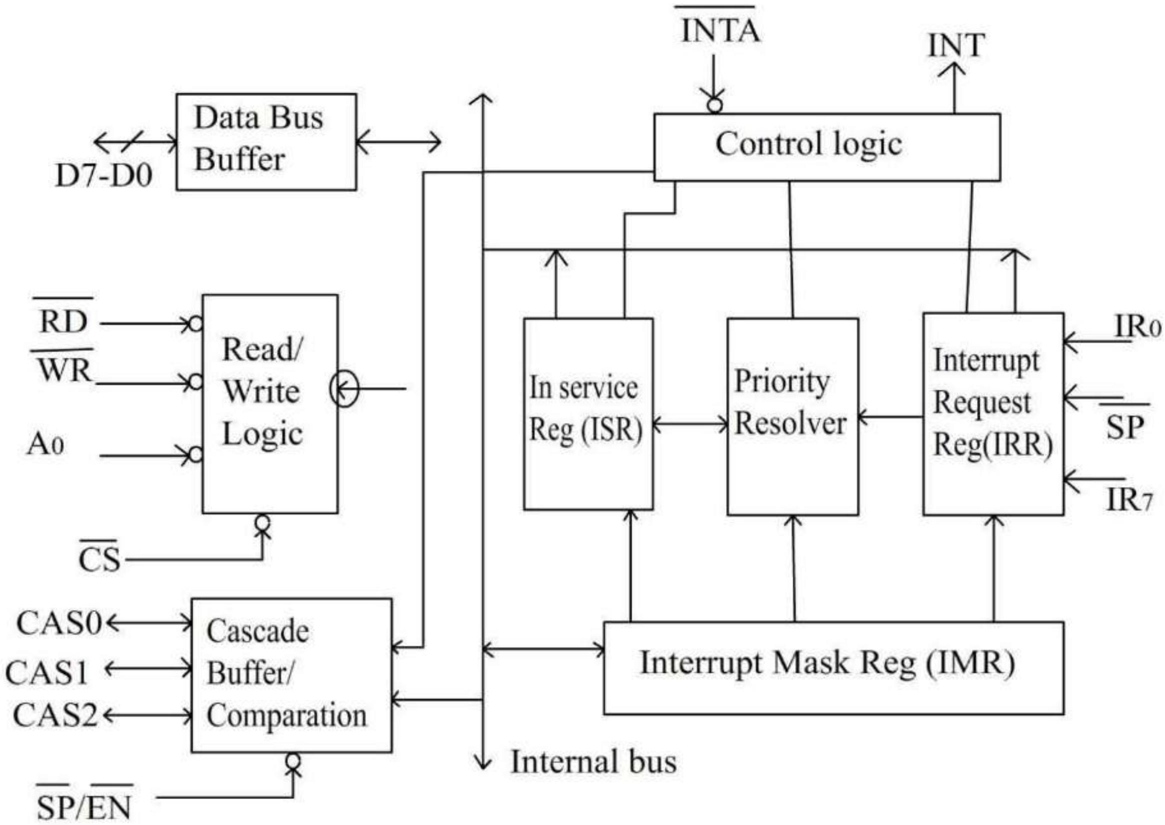
Interrupt Requests: Asynchronous interrupt inputs. An interrupt request is executed by raising an IR input (low to high), and holding it high until it is acknowledged. (Edge triggered mode).or just by a high level on an IR input (levels triggered mode).

A_0:

A_0 address line: This pin acts in conjunction with the $\overline{RD}$, $\overline{WR}$ & $\overline{CS}$ pins. It is used by the 8259A to send various command words from the CPU and to read the status. If is connected to the CPU A_0 address line. Two addresses must be reserved in the I/O address space for each 8259 in the system.

Functional Description:

The 8259 A has eight interrupt request inputs, TR2 IR0. The 8259 A uses its INT output to interrupt the 8085A via INTR pin. The 8259A receives interrupt acknowledge pulses from the μp at its $\overline{INTA}$ input. Vector address used by the 8085 A to transfer control to the service subroutine of the interrupting device, is provided by the 8259 A on the data bus. The 8259A is a programmable device that must be initialized by command words sent by the. After initialization the 8259 A mode of operation can be changed by operation command words from the.



The descriptions of various blocks are,

Data bus buffer:

This 3- state, bidirectional 8-bit buffer is used to interface the 8259A to the system data bus. Control words and status information are transferred through the data bus buffer.

Read/Write & control logic:

The function of this block is to accept OUTPUT commands from the CPU. It contains the initialization command word (ICW) register and operation command word (OCW) register which store the various control formats for device operation. This function block also allows the status of 8159A to be transferred to the data bus.

Interrupt request register (IRR):

IRR stores all the interrupt inputs that are requesting service. Basically, it keeps track of which interrupt inputs are asking for service. If an interrupt input is unmasked, and has an interrupt signal on it, then the corresponding bit in the IRR will be set.

Interrupt mask register (IMR):

The IMR is used to disable (Mask) or enable (Unmask) individual interrupt inputs. Each bit in this register corresponds to the interrupt input with the same number. The IMR operation on the IRR. Masking of higher priority input will not affect the interrupt request lines of lower priority. To unmask any interrupt the corresponding bit is set '0'.

In service register (ISR):

The in service registers keeps tracks of which interrupt inputs are currently being serviced. For each input that is currently being serviced the corresponding bit will be set in the in service register. Each of these 3-reg can be read as status reg.

Priority Resolver:

This logic block determines the priorities of the set in the IRR. The highest priority is selected and strobed into the corresponding bit of the ISR during $\overline{INTA}$ pulse.

Cascade buffer/comparator:

This function blocks stores and compare the IDS of all 8259A's in the reg. The associated 3-I/O pins (CAS0-CAS2) are outputs when

8259A is used a master. Master and are inputs when 8259A is used as a slave. As a master, the 8259A sends the ID of the interrupting slave device onto the cas2-cas0. The slave thus selected will send its pre-programmed subroutine address on to the data bus during the next one or two successive $\overline{INTA}$ pulses.

8257: Direct Memory Access Controller

The *Direct Memory Access* or DMA mode of data transfer is the fastest amongst all the modes of data transfer. In this mode, the device may transfer data directly to/from memory without any interference from the CPU. The device requests the CPU (through a DMA controller) to hold its data, address and control bus, so that the device may transfer data directly to/from memory.

The DMA data transfer is initiated only after receiving HLDA signal from the CPU. Intel's 8257 is a four channel DMA controller designed to be interfaced with their family of microprocessors. The 8257, on behalf of the devices, requests the CPU for bus access using local bus request input i.e. HOLD in minimum mode. In maximum mode of the microprocessor RQ/GT pin is used as bus request input.

On receiving the HLDA signal (in minimum mode) or RQ/GT signal (in maximum mode) from the CPU, the requesting devices gets the access of the bus, and it completes the required number of DMA cycles for the data transfer and then hands over the control of the bus back to the CPU.

Internal Architecture of 8257

The internal architecture of 8257 is shown in figure. The chip support four DMA channels, i.e. four peripheral devices can independently request for DMA data transfer through these channels at a time. The DMA controller has 8-bit internal data buffer, a read/write unit, a control unit, a priority resolving unit along with a set of registers.

Register Organization of 8257

Table *8257 Register Selection*

Register	*Byte*	*Address Inputs*				*F/L*	*BI-Directional Data Bus*							
		A_3	A_2	A_1	A_0		D_7	D_6	D_5	D_4	D_3	D_2	D_1	D_0
CH-0 DMA Address	LSB	0	0	0	0	0	A_7	A_6	A_5	A_4	A_3	A_2	A_1	A_0
	MSB	0	0	0	0	1	A_{15}	A_{14}	A_{13}	A_{12}	A_{11}	A_{10}	A_9	A_8
CH-0 Terminal Count	LSB	0	0	0	1	0	C_7	C_6	C_5	C_4	C_3	C_2	C_1	C_0
	MSB	0	0	0	1	1	Rd	Wr	C_{13}	C_{12}	C_{11}	C_{10}	C_9	C_8
CH-1 DMA Address	LSB	0	0	1	0	0	A_7	A_6	A_5	A_4	A_3	A_2	A_1	A_0
	MSB	0	0	1	0	1	A_{15}	A_{14}	A_{13}	A_{12}	A_{11}	A_{10}	A_9	A_8
CH-1 Terminal Count	LSB	0	0	1	1	0	C_7	C_6	C_5	C_4	C_3	C_2	C_1	C_0
	MSB	0	0	1	1	1	Rd	Wr	C_{13}	C_{12}	C_{11}	C_{10}	C_9	C_8
CH-2 DMA Address	LSB	0	1	0	0	0	A_7	A_6	A_5	A_4	A_3	A_2	A_1	A_0
	MSB	0	1	0	0	1	A_{15}	A_{14}	A_{13}	A_{12}	A_{11}	A_{10}	A_9	A_8
CH-2 Terminal Count	LSB	0	1	0	1	0	C_7	C_6	C_5	C_4	C_3	C_2	C_1	C_0
	MSB	0	1	0	1	1	Rd	Wr	C_{13}	C_{12}	C_{11}	C_{10}	C_9	C_8
CH-3 DMA Address	LSB	0	1	1	0	0	A_7	A_6	A_5	A_4	A_3	A_2	A_1	A_0
	MSB	0	1	1	0	1	A_{15}	A_{14}	A_{13}	A_{12}	A_{11}	A_{10}	A_9	A_8
CH-3 Terminal Count	LSB	0	1	1	1	0	C_7	C_6	C_5	C_4	C_3	C_2	C_1	C_0
	MSB	0	1	1	1	1	Rd	Wr	C_{13}	C_{12}	C_{11}	C_{10}	C_9	C_8
MODE SET (Programme only)	—	1	0	0	0	0	AL	TCS	EW	RP	EN3	EN2	EN1	EN0
STATUS (Read only)	—	1	0	0	0	0	0	0	0	UP	TC3	TC2	TC1	TC0

The 8257 performs the DMA operation over four independent DMA channels. Each of four channels of 8257 has a pair of two 16-bit registers, viz. DMA *address* *register* and *terminal count register*.

There are two common registers for all the channels, namely, *mode set register* and *status register*. Thus there are a total of ten registers. The CPU selects one of these ten registers using address lines A_0-A_3. Table shows how the A_0-A_3 bits may be used for selecting one of these registers.

DMA Address Register

Each DMA channel has one DMA address register. The function of this register is to store the address of the starting memory location, which will be accessed by the DMA channel. Thus the starting address of the memory block which will be accessed by the device is first loaded in the DMA address register of the channel.
The device that wants to transfer data over a DMA channel, will access the block of the memory with the starting address stored in the DMA Address Register.

Terminal Count Register

Each of the four DMA channels of 8257 has one terminal count register (TC). This 16-bit register isused for ascertaining that the data transfer through a DMA channel ceases or stops after the required number of DMA cycles. The low order 14-bits of the terminal count register are initialised with the binary equivalent of the number of required DMA cycles minus one.
After each DMA cycle, the terminal count register content will be decremented by one and finally it becomes zero after the required number of DMA cycles are over. The bits 14 and 15 of this register indicate the type of the DMA operation (transfer). If the device wants to write data into the memory, the DMA operation is called DMA write operation. Bit 14 of the register in this case will be set to one and bit 15 will be set to zero.
Table gives detail of DMA operation selection and corresponding bit configuration of bits 14 and 15 of the TC register.

Table *DMA Operation Selection Using A_{15}/RD and A_{14}/WR*

Bit 15	*Bit 14*	*Type of DMA Operation*
0	0	Verify DMA Cycle
0	1	Write DMA Cycle
1	0	Read DMA Cycle
1	1	(Illegal)

Mode Set Register

The mode set register is used for programming the 8257 as per the requirements of the system. The function of the mode set register is to enable the DMA channels individually and also to set the various modes of operation.

The DMA channel should not be enabled till the DMA address register and the terminal count register contain valid information, otherwise, an unwanted DMA request may initiate a DMA cycle, probably destroying the valid memory data. The bits D_0-D_3 enable one of the four DMA channels of 8257. for example, if D_0 is '1', channel 0 is enabled. If

bit 4 is set, rotating priority is enabled, otherwise, the normal, i.e. fixed priority is enabled.

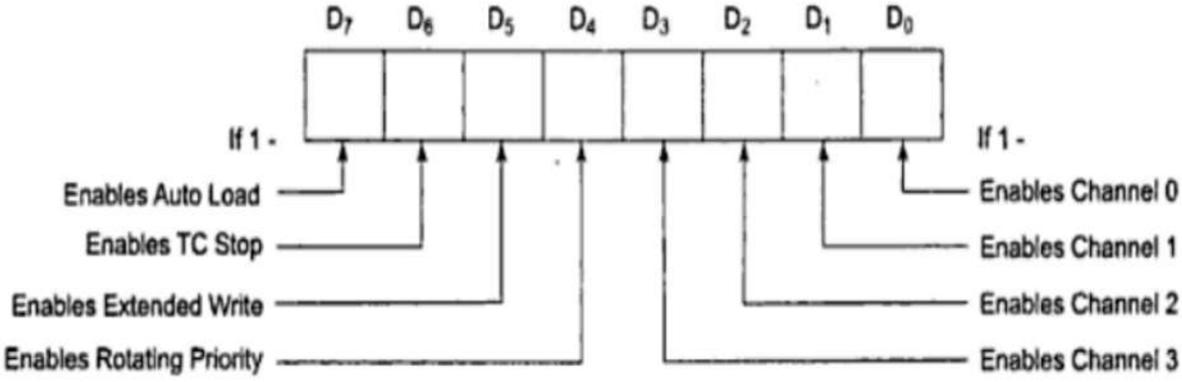


Fig. *Bit Definitions of the Mode Set Register*

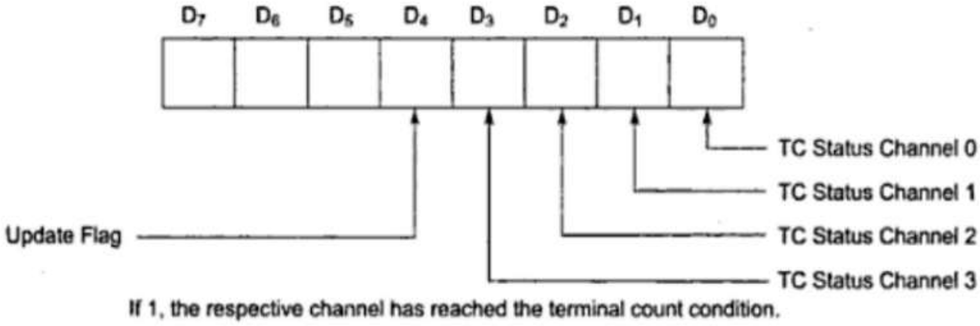
If the TC STOP bit is set, the selected channel is disabled after the *terminal count* condition is reached, and it further prevents any DMA cycle on the channel. To enable the channel again, this bit must be reprogrammed. If the TC STOP bit is programmed to be zero, the channel is not disabled, even after the count reaches zero and further request are allowed on the same channel.

The auto load bit, if set, enables channel 2 for the repeat block chaining operations, without immediate software intervention between the two successive blocks. The channel 2 registers are used as usual, while the channel 3 registers are used to store the block reinitialisation parameters, i.e. the DMA starting address and terminal count. After the first block is transferred using DMA, the channel 2 registers are reloaded with the corresponding channel 3 registers for the next block transfer, if the *update* flag is set. The extended write bit, if set to '1', extends the duration of MEMW and IOW signals by activating them earlier, this is useful in interfacing the peripherals with different access times.

If the peripheral is not accessed within the stipulated time, it is expected to give the 'NOT READY' indication to 8257, to request it to add one or more wait states in the DMA CYCLE. The mode set register can only be written into.

Status Register

The status register of 8257 is shown in figure. The lower order 4-bits of this register contain the terminal count status for the four individual channels. If any of these bits is set, it indicates that the specific channel has reached the terminal count condition.



If 1, the respective channel has reached the terminal count condition.

These bits remain set till either the status is read by the CPU or the 8257 is reset. The update flag is not affected by the read operation. This flag can only be cleared by resetting 8257 or by resetting the auto load bit of the mode set register. If the update flag is set, the contents of the channel 3 registers are reloaded to the corresponding registers of channel 2 whenever the channel 2 reaches a terminal count condition, after transferring one block and the next block is to be transferred using the autoload feature of 8257.

The update flag is set every time, the channel 2 registers are loaded with contents of the channel 3 registers. It is cleared by the completion of the first DMA cycle of the new block. This register can only read.

Data Bus Buffer, Read/Write Logic, Control Unit and Priority Resolver

The 8-bit. Tristate, bidirectional buffer interfaces the internal bus of 8257 with the external system bus under the control of various control signals.

In the slave mode, the read/write logic accepts the I/O Read or I/O Write signals, decodes the Ao-A3 lines and either writes the contents of the data bus to the addressed internal register or reads the contents of the selected register depending upon whether IOW or IOR signal is activated.

In master mode, the read/write logic generates the IOR and IOW signals to control the data flow to or from the selected peripheral. The control logic controls the sequences of operations and generates the required control signals like AEN, ADSTB, MEMR, MEMW, TC and MARK along with the address lines A4-A7, in master mode. The priority resolver resolves the priority of the four DMA channels depending upon whether normal priority or rotating priority is programmed.

Signal Description of 8257

DRQo-DRQ3 :

These are the four individual channel DMA request inputs, used by the peripheral devices for requesting the DMA services. The DRQo has the highest priority while DRQ**3** has the lowest one, if the fixed priority mode is selected.

DACKo-DACK3 :

These are the active-low DMA acknowledge output lines which inform the requesting peripheral that the request has been honoured and the bus is relinquished by the CPU. These lines may act as strobe lines for the requesting devices.

Do-D7:

These are bidirectional, data lines used to interface the system bus with the internal data bus of 8257. These lines carry command words to 8257 and status word from 8257, in slave mode, i.e. under the control of CPU.

The data over these lines may be transferred in both the directions. When the 8257 is the bus master (master mode, i.e. not under CPU control), it uses Do-D7 lines to send higher byte of the generated address to the latch. This address is further latched using ADSTB signal. the address is transferred over Do-D7 during the first clock cycle of the DMA cycle. During the rest of the period, data is available on the data bus.

IOR:

This is an active-low bidirectional tristate input line that acts as an input in the slave mode. In slave mode, this input signal is used by the CPU to read internal registers of 8257.this line acts output in master mode. In master mode, this signal is used to read data from a peripheral during a memory write cycle.

IOW :

This is an active low bidirection tristate line that acts as input in slave mode to load the contents of the data bus to the 8-bit mode register or upper/lower byte of a 16-bit DMA address register or terminal count register. In the master mode, it is a control output that loads the data to a peripheral during DMA memory read cycle (write to peripheral).

CLK:

This is a clock frequency input required to derive basic system timings for the internal operation of 8257.

RESET :

This active-high asynchronous input disables all the DMA channels by clearing the mode register and tristates all the control lines.

Ao-A3:

These are the four least significant address lines. In slave mode, they act as input which select one of the registers to be read or written. In the master mode, they are the four least significant memory address output lines generated by 8257.

CS:

This is an active-low chip select line that enables the read/write operations from/to 8257, in slave mode. In the master mode, it is automatically disabled to prevent the chip from getting selected (by CPU) while performing the DMA operation.

A4-A7 :

This is the higher nibble of the lower byte address generated by 8257 during the master mode of DMA operation.

READY:

This is an active-high asynchronous input used to stretch memory read and write cycles of 8257 by inserting wait states. This is used while interfacing slower peripherals..

HRQ:

The hold request output requests the access of the system bus. In the non-cascaded 8257 systems, this is connected with HOLD pin of CPU. In the cascade mode, this pin of a slave is connected with a DRQ input line of the master 8257, while that of the master is connected with HOLD input of the CPU.

HLDA :

The CPU drives this input to the DMA controller high, while granting the bus to the device. This pin is connected to the HLDA output of the CPU. This input, if high, indicates to the DMA controller that the bus has been granted to the requesting peripheral by the CPU.

MEMR:

This active –low memory read output is used to read data from the addressed memory locations during DMA read cycles.

MEMW :

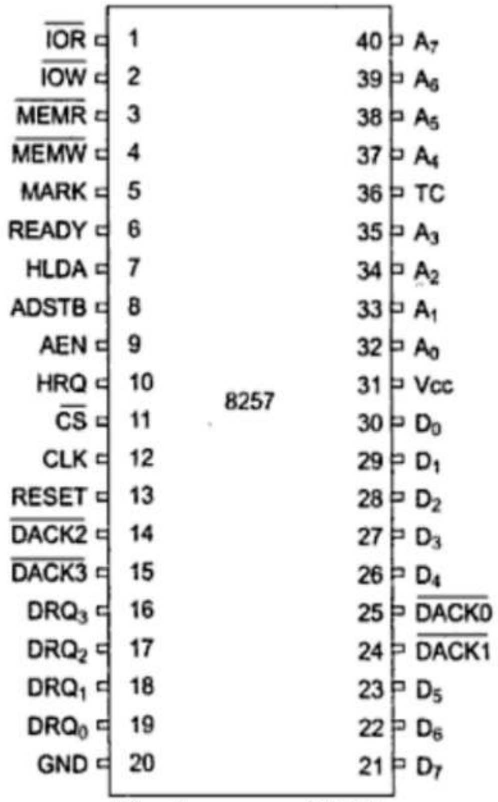
This active-low three state output is used to write data to the addressed memory location during DMA write operation.

ADST :

This output from 8257 strobes the higher byte of the memory address generated by the DMA controller into the latches.

AEN:

This output is used to disable the system data bus and the control the bus driven by the CPU, this may be used to disable the system address and data bus by using the enable input of the bus drivers to inhibit the non-DMA devices from responding during DMA operations. If the 8257 is I/O mapped, this should be used to disable the other I/O devices, when the DMA controller addresses is on the address bus.



Pin diagram of 8257

TC:

Terminal count output indicates to the currently selected peripherals that the present DMA cycle is the last for the previously programmed data block. If the TC STOP bit in the mode set register is set, the selected channel will be disabled at the end of the DMA cycle.

The TC pin is activated when the 14-bit content of the terminal count register of the selected channel becomes equal to zero. The lower order 14 bits of the terminal count register are to be programmed with a 14-bit equivalent of (n-1), if n is the desired number of DMA cycles.

MARK :

The modulo 128 mark output indicates to the selected peripheral that the current DMA cycle is the 128^{th} cycle since the previous MARK output. The mark will be activated after each 128 cycles or integral multiples of it from the beginning if the data block (the first DMA cycle), if the total number of the required DMA cycles (n) is completely divisible by 128.

Vcc :

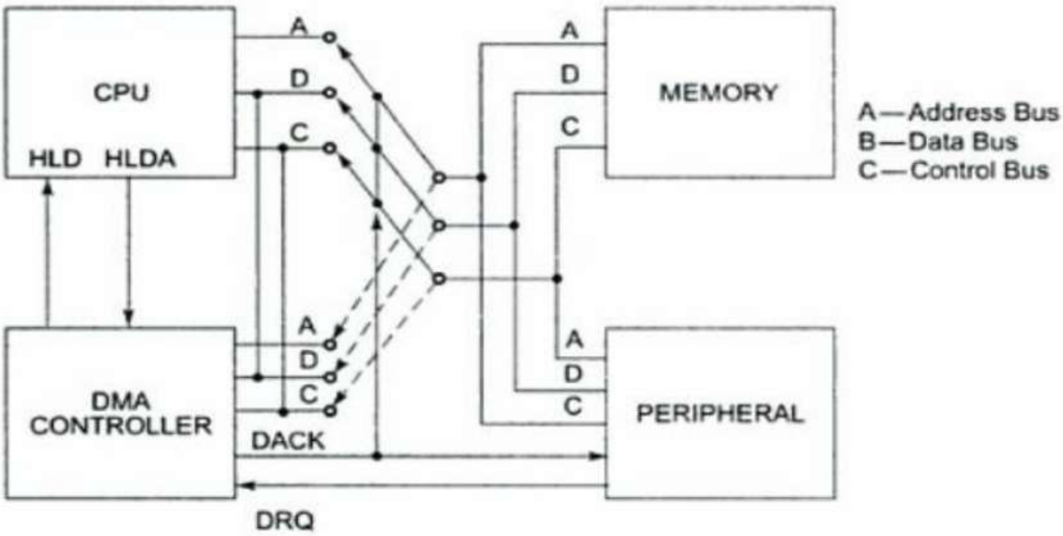
This is a +5v supply pin required for operation of the circuit.

GND :

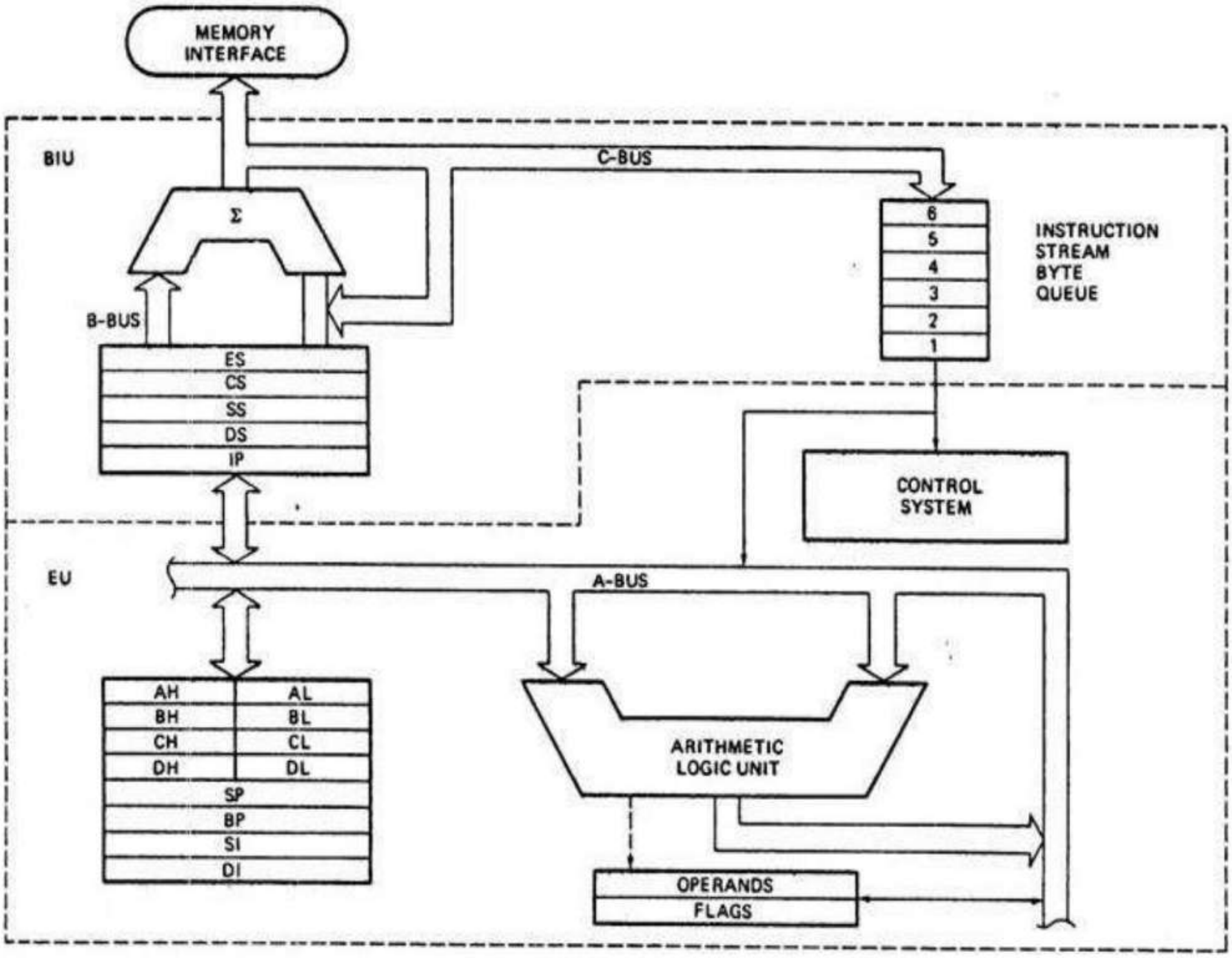
This is a return line for the supply (ground pin of the IC).

Interfacing 8257 with 8086

Once a DMA controller is initialised by a CPU property, it is ready to take control of the system bus on a DMA request, either from a peripheral or itself (in case of memory-to-memory transfer). The DMA controller sends a HOLD request to the CPU and waits for the CPU to assert the HLDA signal. The CPU relinquishes the control of the bus before asserting the HLDA signal.



A conceptual implementation of the system is shown in Figure



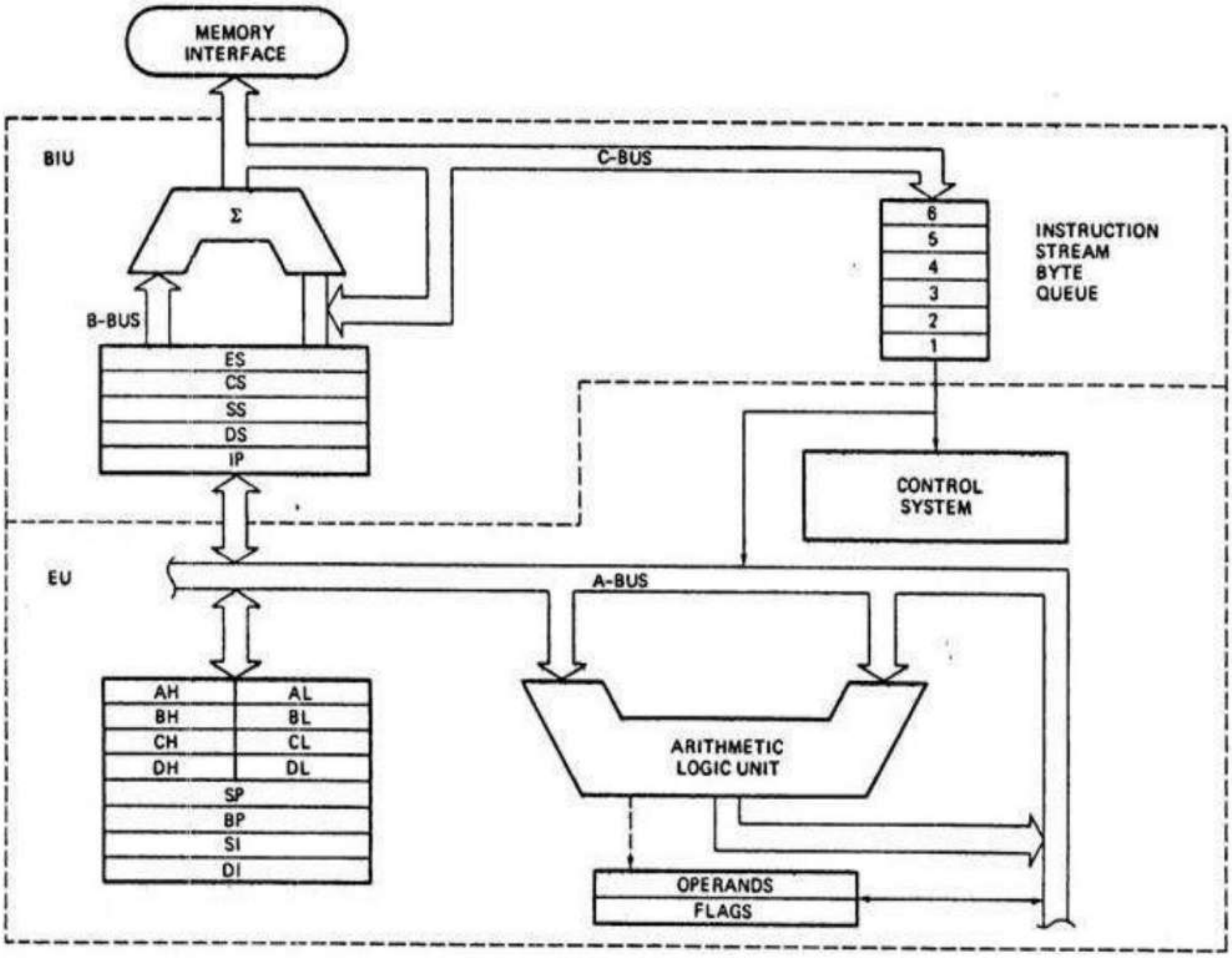
The Execution Unit (EU):

- The execution unit of the 8086 tells the BIU where to fetch instructions or data from, decodes instructions, and executes instructions.
- The EU contains **control circuitry,** which directs internal operations.
- A decoder in the EU translates instructions fetched from memory into a series of actions, which the EU carries out.
- The EU has a 16-bit **arithmetic logic unit** (ALU) which can add, subtract, AND, OR, XOR, increment, decrement, complement or shift binary numbers.
- The main functions of EU are:
  - Decoding of Instructions
  - Execution of instructions
    - ✓ Steps
      - EU extracts instructions from top of queue in BIU
      - Decode the instructions
      - Generates operands if necessary
      - Passes operands to BIU & requests it to perform read or write bus cycles to memory or I/O
      - Perform the operation specified by the instruction on operands

Bus Interface Unit (BIU):

- The BIU sends out addresses, fetches instructions from memory, reads data from ports and memory, and writes data to ports and memory.

UNIT – 8 : ARCHITECTURE OF MICROPROCESSOR



The Execution Unit (EU):

- The execution unit of the 8086 tells the BIU where to fetch instructions or data from, decodes instructions, and executes instructions.
- The EU contains **control circuitry,** which directs internal operations.
- A decoder in the EU translates instructions fetched from memory into a series of actions, which the EU carries out.
- The EU has a 16-bit **arithmetic logic unit** (ALU) which can add, subtract, AND, OR, XOR, increment, decrement, complement or shift binary numbers.
- The main functions of EU are:
  - Decoding of Instructions
  - Execution of instructions
    - ✓ Steps
      - EU extracts instructions from top of queue in BIU
      - Decode the instructions
      - Generates operands if necessary
      - Passes operands to BIU & requests it to perform read or write bus cycles to memory or I/O
      - Perform the operation specified by the instruction on operands

Bus Interface Unit (BIU):

- The BIU sends out addresses, fetches instructions from memory, reads data from ports and memory, and writes data to ports and memory.

- In simple words, the BIU handles all transfers of data and addresses on the buses for the execution unit.

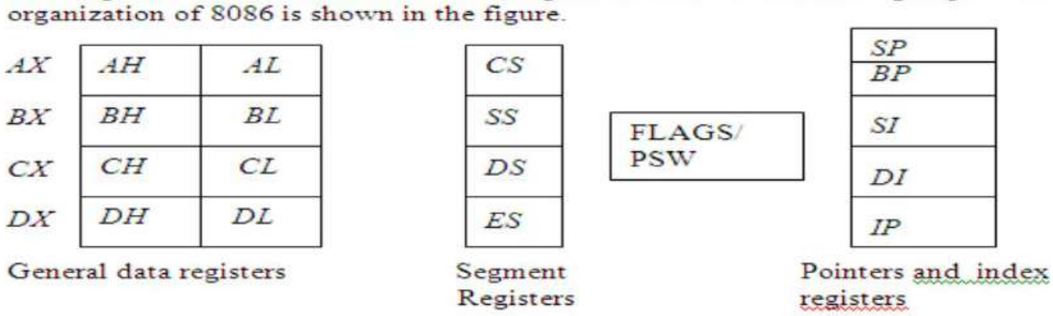
8086 HAS PIPELINING ARCHITECTURE:

- While the EU is decoding an instruction or executing an instruction, which does not require use of the buses, the BIU fetches up to six instruction bytes for the following instructions.
- The BIU stores these pre-fetched bytes in a first-in-first-out register set called a *queue.*
- When the EU is ready for its next instruction from the queue in the BIU. This is much faster than sending out an address to the system memory and waiting for memory to send back the next instruction byte or bytes.
- Except in the case of JMP and CALL instructions, where the queue must be dumped and then reloaded starting from a new address, this pre-fetch and queue scheme greatly speeds up processing.
- Fetching the next instruction while the current instruction executes is called **pipelining**.

Register organization:

- 8086 has a powerful set of registers known as *general purpose registers* and *special purpose registers.*
- All of them are 16-bit registers.
- *General purpose registers:*
  - These registers can be used as either 8-bit registers or 16-bit registers.
  - They may be either used for holding data, variables and intermediate results temporarily or for other purposes like a counter or for storing offset address for some particular addressing modes etc.
- *Special purpose registers:*
  - These registers are used as segment registers, pointers, index registers or as offset storage registers for particular addressing modes.
- The 8086 registers are classified into the following types:
  - General Data Registers
  - Segment Registers
  - Pointers and Index Registers
  - Flag Register

The register set of 8086 can be categorized into 4 different groups. The register organization of 8086 is shown in the figure.



Register organization of 8086

General Data Registers:

- The registers *AX, BX, CX* and *DX* are the general purpose 16-bit registers.
- *AX* is used as 16-bit accumulator. The lower 8-bit is designated as *AL* and higher 8-bit is designated as *AH*. *AL* can be used as an 8-bit accumulator for 8-bit operation.
- All data register can be used as either 16 bit or 8 bit. *BX* is a 16 bit register, but *BL* indicates the lower 8-bit of *BX* and *BH* indicates the higher 8-bit of *BX*.
- The register *BX* is used as offset storage for forming physical address in case of certain addressing modes.
- The register *CX* is used default counter in case of string and loop instructions.
- *DX* register is a general purpose register which may be used as an implicit operand or destination in case of a few instructions.

Segment Registers:

- There are 4 segment registers. They are:
  - Code Segment Register(CS)
  - Data Segment Register(DS)
  - Extra Segment Register(ES)
  - Stack Segment Register(SS)
- The 8086 architecture uses the concept of **segmented memory**. 8086 able to address a memory capacity of 1 megabyte and it is byte organized. This 1 megabyte memory is divided into 16 logical segments. Each segment contains 64 kbytes of memory.
- <u>Code segment register (CS):</u> is used for addressing memory location in the code segment of the memory, where the executable program is stored.
- <u>Data segment register (DS):</u> points to the data segment of the memory where the data is stored.
- <u>Extra Segment Register (ES)</u> : also refers to a segment in the memory which is another data segment in the memory.
- <u>Stack Segment Register (SS):</u> is used for addressing stack segment of the memory. The stack segment is that segment of memory which is used to store stack data.
- While addressing any location in the memory bank, the **physical address** is calculated from two parts:

  Physical address= segment address + offset address
- The first is segment address, the segment registers contain 16-bit segment base addresses, related to different segment.
- The second part is the offset value in that segment.

Pointers and Index Registers:

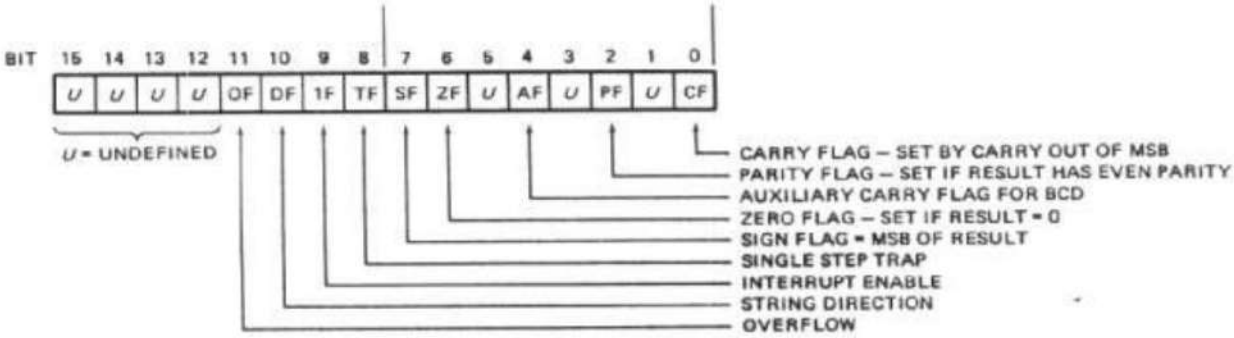
- The index and pointer registers are given below:
  - IP—Instruction pointer-store memory location of next instruction to be executed
  - BP—Base pointer
  - SP—Stack pointer
  - SI—Source index
  - DI—Destination index
- The pointers registers contain offset within the particular segments.
  - The pointer register *IP* contains offset within the code segment.
  - The pointer register *BP* contains offset within the data segment.
  - Thee pointer register *SP* contains offset within the stack segment.
- The index registers are used as general purpose registers as well as for offset storage in case of indexed, base indexed and relative base indexed addressing modes.
- The register *SI* is used to store the offset of source data in data segment.
- The register *DI* is used to store the offset of destination in data or extra segment.
- The index registers are particularly useful for string manipulation.

8086 flag register and its functions:

- The 8086 flag register contents indicate the results of computation in the *ALU*. It also contains some flag bits to control the *CPU* operations.
- A 16 bit flag register is used in 8086. It is divided into two parts .
  - Condition code or status flags
  - Machine control flags
- The **condition code flag register** is the lower byte of the 16-bit flag register. The condition code flag register is identical to 8085 flag register, with an additional overflow flag.

- The **control flag register** is the higher byte of the flag register. It contains three flags namely direction flag (*D*), interrupt flag (*I*) and trap flag (*T*).

Flag register configuration



The description of each flag bit is as follows:

***SF*- Sign Flag:** This flag is set, when the result of any computation is negative. For signed computations the sign flag equals the MSB of the result.

***ZF*- Zero Flag:** This flag is set, if the result of the computation or comparison performed by the previous instruction is zero.

***PF*- Parity Flag:** This flag is set to 1, if the lower byte of the result contains even number of 1's.

***CF*- Carry Flag:** This flag is set, when there is a carry out of MSB in case of addition or a borrow in case of subtraction.

***AF*-Auxilary Carry Flag:** This is set, if there is a carry from the lowest nibble, i.e, bit three during addition, or borrow for the lowest nibble, i.e, bit three, during subtraction.

***OF*- Over flow Flag:** This flag is set, if an overflow occurs, i.e, if the result of a signed operation is large enough to accommodate in a destination register. The result is of more than 7-bits in size in case of 8-bit signed operation and more than 15-bits in size in case of 16-bit sign operations, and then the overflow will be set.

***TF*- Tarp Flag:** If this flag is set, the processor enters the single step execution mode. The processor executes the current instruction and the control is transferred to the Trap interrupt service routine.

***IF*- Interrupt Flag:** If this flag is set, the mask able interrupts are recognized by the CPU, otherwise they are ignored.

***D*- Direction Flag:** This is used by string manipulation instructions. If this flag bit is '0', the string is processed beginning from the lowest address to the highest address, i.e., auto incrementing mode. Otherwise, the string is processed from the highest address towards the lowest address, i.e., auto decrementing mode.

Memory Segmentation:

- The memory in an 8086 based system is organized as segmented memory.
- The CPU 8086 is able to access 1MB of physical memory. The complete 1MB of memory can be divided into 16 segments, each of 64KB size and is addressed by one of the segment register.
- The 16-bit contents of the segment register actually point to the starting location of a particular segment. The address of the segments may be assigned as 0000H to F000h respectively.
- To address a specific memory location within a segment, we need an offset address. The offset address values are from 0000H to FFFFH so that the physical addresses range from 00000H to FFFFFH.

Physical address is calculated as below:

Ex: Segment address -------→ 1005H

Offset address ----------→ 5555H

Segment address -------→ 1005H ----- 0001 0000 0000 0101

Shifted left by 4 Positions------ 0001 0000 0000 0101 0000

+

Offset address --- 5555H ------ 0101 0101 0101 0101

Physical address -------155A5H 0001 0101 0101 1010 0101

Physical address = Segment address * 10H + Offset address.

The main advantages of the segmented memory scheme are as follows:

1. Allows the memory capacity to be 1MB although the actual addresses to be handled are of 16-bit size.
2. Allows the placing of code, data and stack portions of the same program in different parts (segments) of memory, for data and code protection.
3. Permits a program and/or its data to be put into different areas of memory each time the program is executed, i.e., provision for relocation is done.

Overlapping and Non-overlapping Memory segments:

➢ In the overlapping area locations physical address = CS1+IP1 = CS2+IP2. Where '+' indicates the procedure of physical address formation.

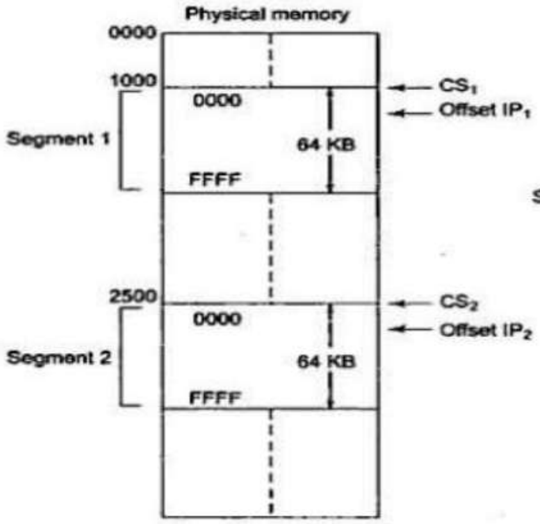


Fig. 1.3(a) *Non-overlapping Segments*

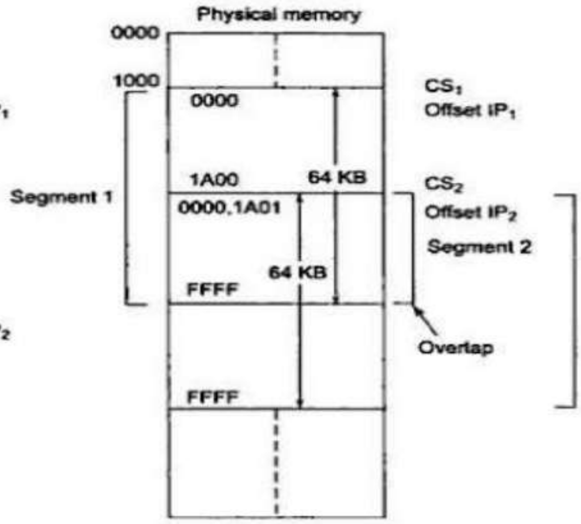


Fig. 1.3(b) *Overlapping Segments*

Addressing modes of 8086:

- Addressing mode indicates a way of locating data or operands.
- The addressing modes describe the types of operands and the way they are accessed for executing an instruction.
- According to the flow of instruction execution, the instructions may be categorized as
  - i) Sequential control flow instructions and
  - ii) Control transfer instructions

Sequential control flow instructions are the instructions, which after execution, transfer control to the next instruction appearing immediately after it (in the sequence) in the program. For example, the arithmetic, logic, data transfer and processor control instructions are sequential control flow instructions.

The **control transfer instructions**, on the other hand, transfer control to some predefined address or the address somehow specified in the instruction, after their execution. For example, INT, CALL, RET and JUMP instructions fall under this category.

The addressing modes for sequential control transfer instructions are:

1. **Immediate:** In this type of addressing, immediate data is a part of instruction and appears in the form of successive byte or bytes.

Ex: MOV AX, 0005H

In the above example, 0005H is the immediate data. The immediate data may be 8-bit or 16-bit in size.

2. **Direct:** In the direct addressing mode a 16-bit memory address (offset) is directly specified in the instruction as a part of it.

Ex: MOV AX, [5000H]

Here, data resides in a memory location in the data segment, whose effective address may be completed using 5000H as the offset address and content of DS as segment address. The effective address here, is 10H * DS + 5000H.

3. **Register:** In register addressing mode, the data is stored in a register and is referred using the particular register. All the registers, except IP, may be used in this mode.

Ex: MOV BX, AX

4. **Register Indirect:** Sometimes, the address of the memory location, which contains data or operand, is determined in an indirect way, using the offset register. This mode of addressing is known as register indirect mode. In this addressing mode, the offset address of data is in either BX or SI or DI register. The default segment is either DS or ES. The data is supposed to be available at the address pointed to by the content of any of the above registers in the default data segment.

Ex: MOV AX, [BX]

Here, data is present in a memory location in DS whose offset address is in BX. The effective address of the data is given as 10H * DS+[BX].

5. **Indexed:** In this addressing mode, offset of the operand is stored in one of the index registers. DS and ES are the default segments for index registers, SI and DI respectively. This is a special case of register indirect addressing mode.

Ex: MOV AX, [SI]

Here, data is available at an offset address stored in SI in DS. The effective address, in this case, is computed as 10*DS+[SI].

6. **Register Relative:** In this addressing mode, the data is available at an effective address formed by adding an 8-bit or 16-bit displacement with the content of any one of the registers BX, BP, SI and DI in the default (either DS or ES) segment.

Ex: MOV AX, 50H[BX]

Here, the effective address is given as 10H *DS+50H+[BX]

7. **Based Indexed:** The effective address of data is formed, in this addressing mode, by adding content of a base register (any one of BX or BP) to the content of an index register (any one of SI or DI). The default segment register may be ES or DS.

Ex: MOV AX, [BX][SI]

Here, BX is the base register and SI is the index register the effective address is computed as 10H * DS + [BX] + [SI].

8. **Relative Based Indexed:** The effective address is formed by adding an 8 or 16-bit displacement with the sum of the contents of any one of the base register (BX or BP) and any one of the index register, in a default segment.

Ex: MOV AX, 50H [BX] [SI]

Here, 50H is an immediate displacement, BX is base register and SI is an index register the effective address of data is computed as

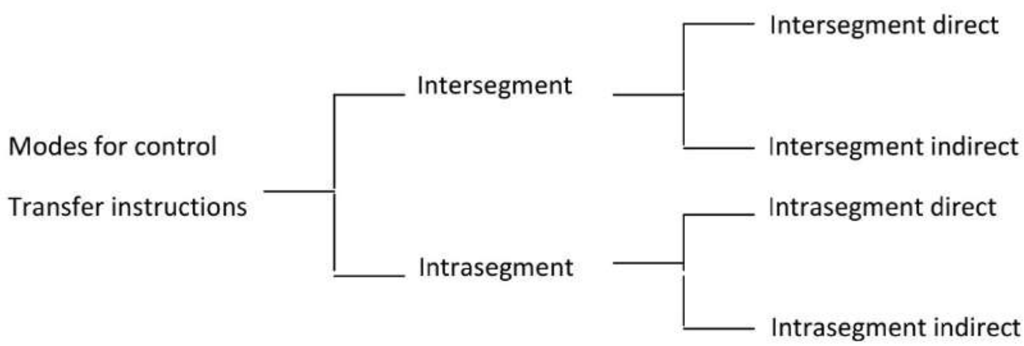
10H * DS + [BX] + [SI] + 50H

For control transfer instructions, the addressing modes depend upon whether the destination is within the same segment or different one. It also depends upon the method of passing the destination address to the processor.

Basically, there are two addressing modes for the control transfer instructions, **intersegment** addressing and **intrasegment** addressing modes.

If the location to which the control is to be transferred lies in a different segment other than the current one, the mode is called intersegment mode.

If the destination location lies in the same segment, the mode is called intrasegment mode.



Addressing modes for Control Transfer Instructions

9. **Intrasegment Direct Mode:** In this mode, the address to which the control is to be transferred lies in the same segment in which the control transfer instruction lies and appears directly in the instruction as an immediate displacement value. In this addressing mode, the displacement is computed relative to the content of the instruction pointer IP.

The effective address to which the control will be transferred is given by the sum of 8 or 16-bit displacement and current content of IP. In the case of jump instruction, if the signed displacement (d) is of 8-bits (i.e $-128<d<+128$) we term it as short jump and if it is of 16-bits (i.e-32, $768<d<+32,768$) it is termed as long jump.

10. **Intrasegment Indirect Mode:** In this mode, the displacement to which the control is to be transferred, is in the same segment in which the control transfer instruction lies, but it is passed to the instruction indirectly. Here, the branch address is found as the content of a register or a memory location. This addressing mode may be used in unconditional branch instructions.

11. **Intersegment Direct:** In this mode, the address to which the control is to be transferred is in a different segment. This addressing mode provides a means of branching from one code segment to another code segment. Here, the CS and IP of the destination address are specified directly in the instruction.

12. **Intersegment Indirect:** In this mode, the address to which the control is to be transferred lies in a different segment and it is passed to the instruction indirectly, i.e contents of a memory block containing four bytes, i.e IP (LSB), IP(MSB), CS(LSB) and CS (MSB) sequentially. The starting address of the memory block may be referred using any of the addressing modes, except immediate mode.

DS : [BX + SI]

10H*DS ⇔ 10000

[BX] ⇔ +2000

[SI] ⇔ +3000

____

15000H – Effective address

____

v. **Relative based index:**

MOV AX, 5000[BX][SI]

DS : [BX+SI+5000]

10H*DS ⇔ 10000

[BX] ⇔ +2000

[SI] ⇔ +3000

+5000

____

1A000H – Effective address

____

Pin Diagram of 8086:

Signal description of 8086:

- The 8086 is a 16-bit microprocessor. This microprocessor operates in single processor or multiprocessor configurations to achieve high performance.
- The pin configuration of 8086 is shown in the figure. Some of the pins serve a particular function in minimum mode (single processor mode) and others function in maximum mode (multiprocessor mode).

The 8086 signals are categorized into 3 types:

1. Common signals for both minimum mode and maximum mode.
2. Special signals which are meant only for minimum mode
3. Special signals which are meant only for maximum mode

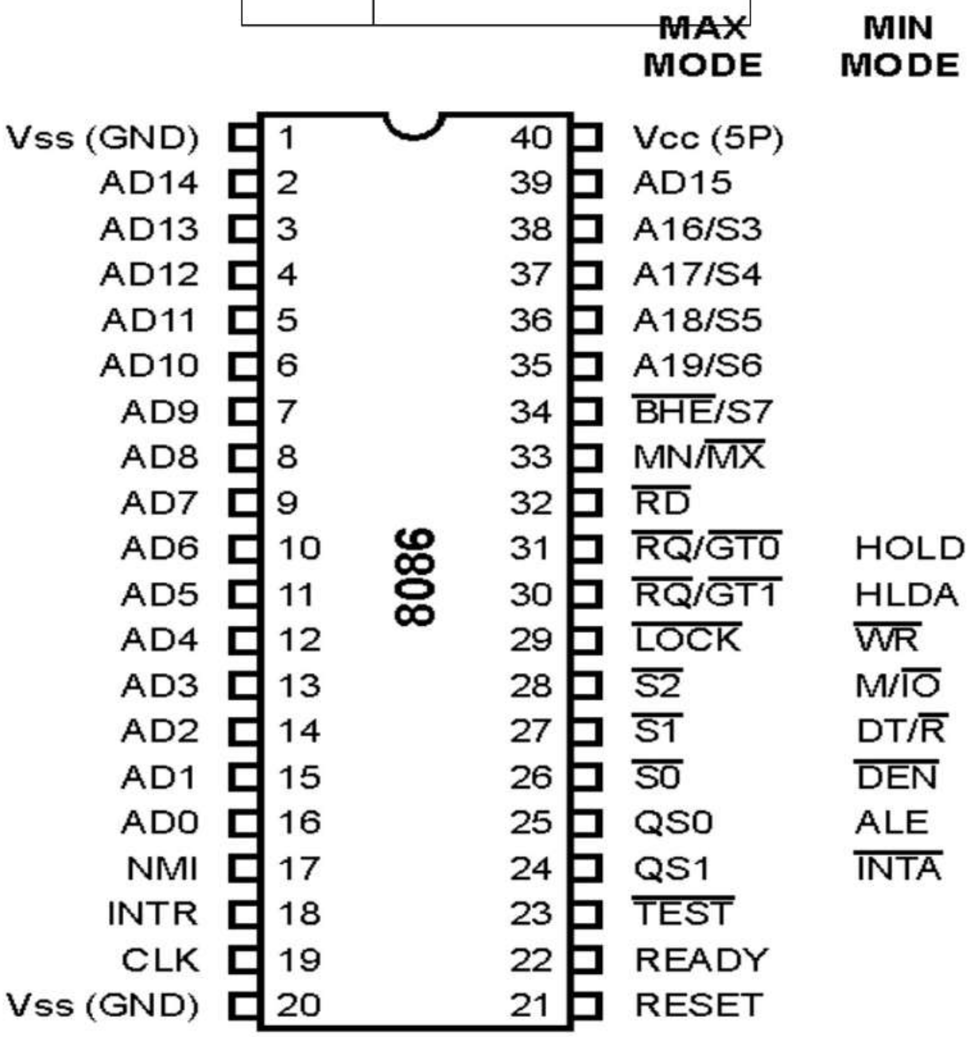
Common Signals for both Minimum mode and Maximum mode:

$AD_7 - AD_0$: The address/ data bus lines are the multiplexed address data bus and contain the right most eight bit of memory address or data. The address and data bits are separated by using *ALE* signal.

$AD_{15} - AD_8$: The address/data bus lines compose the upper multiplexed address/data bus. This lines contain address bit $A_{15} - A_8$ or data bus $D_{15} - D_8$. The address and data bits are separated by using *ALE* signal.

$A_{19} / S_6 - A_{18} / S_3$ The address/status bus bits are multiplexed to provide address signals $A_{19} - A_{16}$ and also status bits $S_6 - S_3$. The address bits are separated from the status bits using the ALE signals. The status bit S_6 is always a logic 0, bit S_5 indicates the condition of the interrupt flag bit. The S_4 and S_3 indicate which segment register is presently being used for memory access.

S_4	S_3	Type of segment register used
0	0	Extra segment
0	1	Stack segment
1	0	Code or no segment
1	1	Data Segment



$\overline{BHE}/S_7$ The bus high enable (BHE) signal is used to indicate the transfer of data over the higher order $(D_{15}-D_8)$ data bus. It goes low for the data transfer over $D_{15}-D_8$ and is used to derive chip select of odd address memory bank or peripherals.

$\overline{BHE}$	A_0	Indication
0	0	Whole word
0	1	Upper byte from or to odd address
1	0	Lower byte from or to even address
1	1	None

$\overline{RD}$: Read: whenever the read signal is at logic

0, the data bus receives the data from the memory or I/O devices connected to the system

READY: This is the acknowledgement from the slow devices or memory that they have completed the data transfer operation. This signal is active high.

INTR: Interrupt Request: Interrupt request is used to request a hardware interrupt of *INTR* is held high when interrupt enable flag is set, the 8086 enters an interrupt acknowledgement cycle after the current instruction has completed its execution.

$\overline{TEST}$: This input is tested by "*WAIT*" instruction. If the *TEST* input goes low; execution will continue. Else the processor remains in an idle state.

NMI- Non-maskable Interrupt: The non-maskable interrupt input is similar to *INTR* except that the *NMI* interrupt does not check for interrupt enable flag is at logic 1, i.e, *NMI* is not maskable internally by software. If *NMI* is activated, the interrupt input uses interrupt vector 2.

RESET: The reset input causes the microprocessor to reset itself. When 8086 reset, it restarts the execution from memory location *FFFF0H*. The reset signal is active high and must be active for at least four clock cycles.

CLK: Clock input: The clock input signal provides the basic timing input signal for processor and bus control operation. It is asymmetric square wave with 33% duty cycle.

V_{CC} (+5V): Power supply for the operation of the internal circuit

GND: Ground for the internal circuit

$MN/\overline{MX}$: The minimum/maximum mode signal to select the mode of operation either in minimum or maximum mode configuration. Logic 1 indicates minimum mode.

Minimum mode Signals: The following signals are for minimum mode operation of 8086.

$M/\overline{IO}$ - Memory/*IO* $M/\overline{IO}$ signal selects either memory operation or I/O operation. This line indicates that the microprocessor address bus contains either a memory address or an I/O port address. Signal high at this pin indicates a memory operation. This line is logically equivalent to $\overline{S_2}$ in maximum mode.

$\overline{INTA}$ - Interrupt acknowledge: The interrupt acknowledge signal is a response to the INTR input signal. The $\overline{INTA}$ signal is normally used to gate the interrupt vector number onto the data bus in response to an interrupt request.

ALE- Address Latch Enable: This output signal indicates the availability of valid address on the address/data bus, and is connected to latch enable input of latches.

$DT/\overline{R}$: Data transmit/Receive: This output signal is used to decide the direction of date flow through the bi-directional buffer. $DT/\overline{R}=1$ Indicates transmitting and $DT/\overline{R}=0$ indicates receiving the data.

$\overline{DEN}$ Data Enable: Data bus enable signal indicates the availability of valid data over the address/data lines.

$\overline{WR}$ Write: whenever the write signal is at logic 0, the data bus transmits the data to the memory or I/O devices connected to the system.

HOLD: The hold input request a direct memory access (DMA). If the hold signal is at logic 1, the micro process stops its normal execution and places its address, data and control bus at the high impedance state.

HLDA: Hold acknowledgement indicates that 8086 has entered into the hold state.

Maximum mode signal: The following signals are for maximum mode operation of 8086.

$\overline{S}_2, \overline{S}_1, \overline{S}_0$ - Status lines: These are the status lines that reflect the type of operation being carried out by the processor.

These status lines are encoded as follows

$\overline{S}_2$	$\overline{S}_1$	$\overline{S}_0$	Function
0	0	0	Interrupt Acknowledge
0	0	1	Read I/O port
0	1	0	Write I/O port
0	1	1	Halt
1	0	0	Code Access
1	0	1	Read memory
1	1	0	Write memory
1	1	1	Passive (In active)

$\overline{LOCK}$: The lock output is used to lock peripherals off the system, i.e, the other system bus masters will be prevented from gaining the system bus.

QS_1 and QS_0 - Queue status: The queue status bits shows the status of the internal instruction queue. The encoding of these signals is as follows

QS_1	QS_0	Function
0	0	No operation, queue is idle
0	1	First byte of opcode
1	0	Queue is empty
1	1	Subsequent byte of opcode

$\overline{RQ}/\overline{GT1}$ and $\overline{RQ}/\overline{GT0}$ - request/Grant: The request/grant pins are used by other local bus masters to force the processor to release the local bus at the end of the processors current bus cycle. These lines are bi-directional and are used to both request and grant a *DMA* operation. $\overline{RQ}/\overline{GT0}$ is having higher priority than $\overline{RQ}/\overline{GT1}$

8086 Minimum mode system operation with timing diagrams:

- In a minimum mode 8086 system, the microprocessor 8086 is operated in minimum mode by strapping its MN/MX$^{|}$ pin to logic1.

UNIT – 9 : DATA TRANSFER TECHNIQUES

Why is Data Transfer needed?

We can connect several I/O devices and memory peripherals to a microprocessor. However, since different technologies are involved, there will be differences in the speed of operation and of data transfer.

Usually, when memory is connected with the microprocessor, there isn't a stark difference in the processing speed since semiconductor memories are generally easily compatible with microprocessors. However, this might not always be the case.

But the problem often arises when external peripherals are connected as I/O devices. A slow I/O device won't be able to transfer data at a satisfactory rate whenever the microprocessor requests for data transfer. And usually, the peripheral devices *do* have slower transfer rates than the processor. Maybe the processor might send two units of information per second, but the external device might only accept 1 unit of information per second. Or conversely, perhaps the external device expects quicker transfers, but our microprocessor might be sending information a bit slower.

This might lead to severe data losses, or the devices might get damaged, or there might even be chances for the system to slow down all over, thus affecting the overall efficiency of the system.

To avoid this problem, a number of ***data transfer techniques*** have been devised.

Since constant communication is required between the external device and the processor, these data transfer techniques play a crucial role in the efficient functioning of the system with the externally connected devices.

Classification of Data Transfer Schemes

We can broadly classify the data transfer schemes into two modes – Serial Data Transfer and Parallel Data Transfer.

Data Transfer Techniques

Parallel Data Transfer

1. Programmed Data Transfer
   (a) Synchronous Data Transfer
   (b) Asynchronous Data Transfer
2. Interrupt Driven Data Transfer
3. Device or DMA Data Transfer
   (a) Burst or Block Transfer DMA
   (b) Cycle steal or single byte transfer DMA
   (c) Transparent or hidden DMA

Serial Data Transfer

1. Synchronous Data Transfer
2. Asynchronous Data Transfer

Data Transfer

As discussed earlier, we can broadly classify data transfer schemes into parallel data transfer techniques and serial data transfer techniques. Let us understand the differences between them.

SERIAL DATA TRANSFER TECHNIQUES	PARALLEL DATA TRANSFER TECHNIQUES
Under this scheme, the data is transferred one bit at a time.	Under this scheme, the data is transferred several bits the same time.
It is a slower mode of data transfer.	Data is transferred much quicker.
Serial data transfer is preferred when data is to be sent over a long distance and the cost of cables would be too expensive.	Parallel data transfer is the preferred technique for sho distance communication.
For this mode, the transmitter first performs parallel – to – serial conversion and the serial – to – parallel conversion at the receiver.	No such conversions are required at both the transmission and reception endpoints.
This mode requires a single line to transfer information.	This mode requires multiple lines for data transfer.
Noise and errors are much lesser.	As multiple bits are transmitted at the same time, there scope for more error and noise.
Cables used for serial communication are much thinner, longer, and very economical.	Here, the cables are much shorter, and thicker compar to the Serial communication cables.
This is a very reliable, inexpensive, and straightforward	It is considered a little unreliable, expensive, and

process.	complicated process.

Parallel Data Transfer Techniques

We know that under the parallel data transfer scheme, multiple data bits can be transmitted at the same time. Thus, for the Intel 8085, 8 bits of data are sent all together using eight parallel lines. Let us go through the different types of parallel data transfer schemes. We have:

- Programmed I/O Data Transfer
- Interrupt Driven I/O Data Transfer
- Device or Direct Memory Access (DMA) Data Transfer

Let us study each of these transfer schemes in detail.

Programmed I/O Data Transfer

- This is a straightforward scheme under parallel data transfer mechanisms. This mode is generally preferred for simple, small microprocessor systems where speed is critical.
- This method can work under the synchronous and asynchronous mode, depending on the speed and architecture of the I/O devices.
- We also prefer this method when there is a small amount of information to be exchanged between the microprocessor and other devices that are placed near to the microprocessor. Example: Computer, printer, etc.

- Using the IN andand OUT instructions, data transfer is carried out between the microprocessor and I/O devices.
- The processor reads the data from an input port or input device using the IN command. The processor sends data out from the CPU to the output port or the output device using the OUT instruction.
- Thus, as the speeds of the processor and external device match, the data transferring process is carried out using the IN and OUT instructions.

Synchronous Data Transfer Method

- The word '**Synchronous**' means 'taking place at the same time.'
- Thus, to establish communication between our processor and the device, we need to set a common clock pulse. This common pulse ***synchronizes*** the peripheral device with the 8085 microprocessor.

- This method is used when the speed of the microprocessor, Intel 8085, in this case, and the external peripheral device match with each other.
- If the device is ready to send data, it can indicate via the **READY** pin of 8085. Once the speeds match, the data transfer immediately begins, once a signal is issued by the microprocessor to begin transferring. The microprocessor need not wait for an extended period because of the matching speeds.
- This technique of data transfer is seldom used to communicate with I/O devices though. Because I/O devices compatible with the microprocessor's speed are usually not found.
- Hence, this method of data transfer is most commonly employed for communicating with compatible memory devices.

Asynchronous Data Transfer Method

- When the speed of the I/O device is slower than that of the microprocessor, we prefer the Asynchronous Data Transfer Method. As the speeds of both the devices differ, the I/O device's internal timing is entirely independent of the microprocessor.
- Thus, they are termed to be '***asynchronous***' from each other. The term asynchronous means 'at irregular intervals.'