SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF BIO AND CHEMICAL ENGINEERING
DEPARTMENT OF BIOMEDICAL ENGINEERING

UNIT – I - MEDICAL ELECTRONICS AND E-HEALTH– SBMA7004

I-CHEMICAL AND OPTICAL TRANSDUCERS

PH, PO2, PCO2, HCO3 electrodes, Ion sensor, Anion and Cation sensor, Liquid and solid ion exchange membrane electrode, Enzyme electrode, Principle of fiber optic cable, fiber optic sensors, Photo acoustic sensors, PPG sensors.

1.1 PH ELECTRODES:

A glass electrode is a type of ion-selective electrode made of a doped glass membrane that is sensitive to a specific ion. The most common application of ion-selective glass electrodes is for the measurement of pH. The pH electrode is an example of a glass electrode that is sensitive to hydrogen ions. Glass electrodes play an important part in the instrumentation for chemical analysis and physico-chemical studies. The voltage of the glass electrode, relative to some reference value, is sensitive to changes in the activity of certain type of ions.

Construction

Scheme of typical pH glass electrode.

A typical modern pH probe is a combination electrode, which combines both the glass and reference electrodes into one body. The combination electrode consists of the following parts (see the drawing):

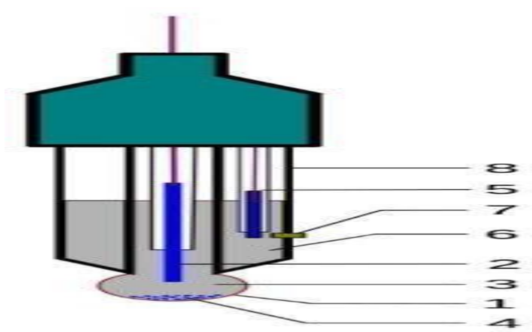

Fig:1.1-PH Glass electrode

1. a sensing part of electrode, a bulb made from a specific glass
2. internal electrode, usually silver chloride electrode or calomel electrode
3. internal solution, usually a pH=7 buffered solution of 0.1 mol/L KCl for pH electrodes or 0.1 mol/L MeCl for pMe electrodes
4. when using the silver chloride electrode, a small amount of AgCl can precipitate inside the glass electrode.

5. reference electrode, usually the same type as 2
6. reference internal solution, usually 0.1 mol/L KCl
7. Junction with studied solution, usually made from ceramics or capillary with asbestos or quartz fiber.
8. Body of electrode, made from non-conductive glass or plastics.

The bottom of a pH electrode balloons out into a round thin glass bulb. The pH electrode is best thought of as a tube within a tube. The inner tube contains an unchanging 1×10^{-7} mol/L HCl solution. Also inside the inner tube is the cathode terminus of the reference probe. The anodic terminus wraps itself around the outside of the inner tube and ends with the same sort of reference probe as was on the inside of the inner tube. It is filled with a reference solution of KCl and has contact with the solution on the outside of the pH probe by way of a porous plug that serves as a salt bridge.

Galvanic cell schematic representation

This section describes the functioning of two distinct types of electrodes as one unit which combines both the glass electrode and the reference electrode into one body. It deserves some explanation.

This device is essentially a galvanic cell that can be schematically represented as:
Glass electrode || Reference Solution || Test Solution || Glass electrode

Ag(*s*) | AgCl(*s*) | KCl(*aq*) || 1×10^{-7}M H^+ solution || glass membrane || Test Solution || junction || KCl(*aq*) | AgCl(*s*) | Ag(*s*)

In this schematic representation of the galvanic cell, one will note the symmetry between the left and the right members as seen from the center of the row occupied by the "Test Solution" (the solution whose pH must be measured). In other words, the glass membrane and the ceramic junction occupies both the same relative place in each respective electrode (indicative (sensing) electrode or reference electrode). The double "pipe symbol" (||) indicates a diffusive barrier that prevents (glass membrane), or slowing down (ceramic junction), the mixing of the different solutions. By using the same electrodes on the left and right, any potentials generated at the interfaces cancel each other (in principle), resulting in the system voltage being dependent only on the interaction of the glass membrane and the test solution.

The measuring part of the electrode, the glass bulb on the bottom, is coated both inside and out with a ~10 nm layer of a hydrated gel. These two layers are separated by a layer of dry glass. The silica glass structure (that is, the conformation of its atomic structure) is shaped in such a way that it allows Na^+ ions some mobility. The metal cations (Na^+) in the hydrated gel diffuse out of the glass and into solution while H^+ from solution can diffuse into the hydrated gel. It is the hydrated gel, which makes the pH electrode an ion-selective electrode.

H^+ does not cross through the glass membrane of the pH electrode, it is the Na^+ which crosses and allows for a change in free energy. When an ion diffuses from a region of activity to another region of activity, there is a free energy change and this is what the pH meter actually measures. The hydrated gel membrane is connected by Na^+ transport and thus the concentration of H^+ on the outside of the membrane is 'relayed' to the inside of the membrane by Na^+.

All glass pH electrodes have extremely high electric resistance from 50 to 500 MΩ. Therefore, the glass electrode can be used only with a high input-impedance measuring device like a pH meter, or, more generically, a high input-impedance voltmeter which is called an electrometer.

Storage

Between measurements any glass and membrane electrodes should be kept in a solution of its own ion It is necessary to prevent the glass membrane from drying out because the performance is dependent on the existence of a hydrated layer, which forms slowly.

1.2 PRINCIPLES OF PO2 MEASUREMENT WITH THE CLARK ELECTRODE

THE CLARK OXYGEN ELECTRODE

- A silver anode and platinum cathode are suspended in an electrolyte.
- Oxygen is dissolved in the electrolyte.
- A voltage of known magnitude (about 700 mV) is applied to the electrodes.
- Oxygen is reduced at the cathode and silver is oxidised at the anode.
- The resulting current increases as the voltage increases.
- The current reaches a plateau when the rate of reaction is determined by the diffusion of oxygen rather than the voltage.
- This plateau correlates to the oxygen tension in the electrolyte.

The major difference between this electrode and the earlier oxygen cathode is the addition of an oxygen-permeable membrane. Something resembling the original patent application diagram can be found here.

Its butchered representation can be found below.

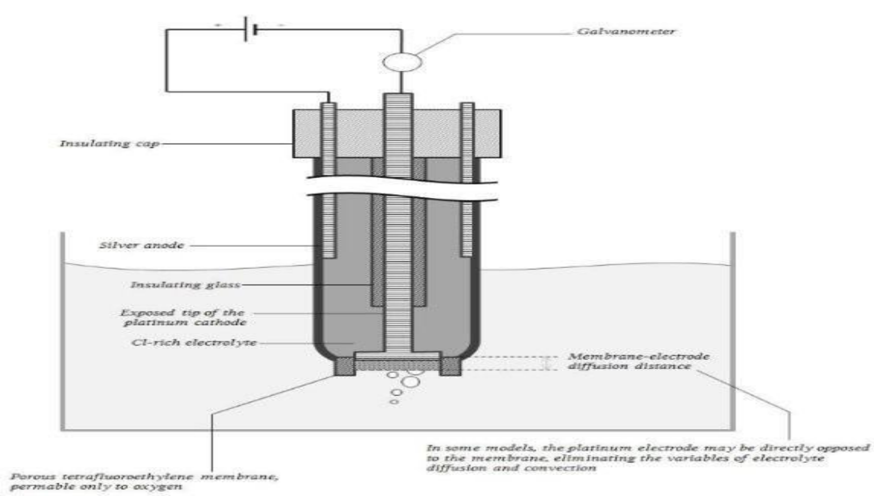

Fig: 1.2- The Clark Oxygen Electrode

A number of design flaws of the platinum oxygen cathode have been addressed by Clark's design;

The membrane is the major change. Its presence both protects the platinum from becoming encrusted in proteinaceous debris, and offers a predictable diffusion distance for the oxygen, without the chances of convection. This protects it from some sources of error (though it must be mentioned that the electrode can still occasionally give confusing results when it starts reducing halothane, for example).

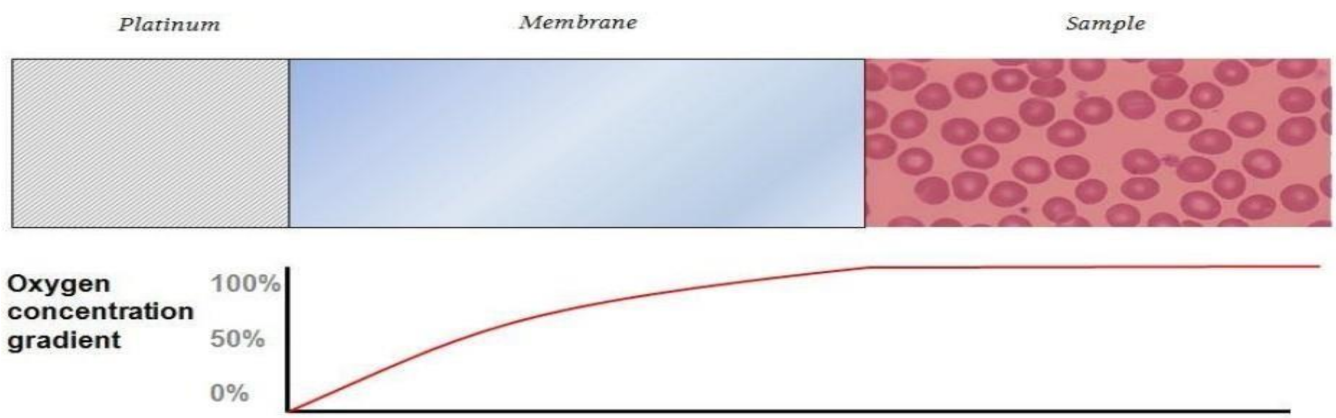

Fig.1.3: Oxygen concentration gradient

The rate of response of the electrode obviously depends on the membrane thickness. It takes time for those little molecules to make their way to the cathode. This diffusion is obviously going to take longer if the membrane is thicker, or if there is a post-membrane layer of electrolyte to negotiate (that is one of the reasons the electrodes these days are right up against the membrane). The response time of a 5µm Teflon membrane is about 1 second, and this can be increased to 0.4 seconds if the sample is heated to 80° C.

1.3 PCO2 MEASUREMENT

The Severinghaus Electrode

This thing is essentially a slightly modified glass electrode. Observe:

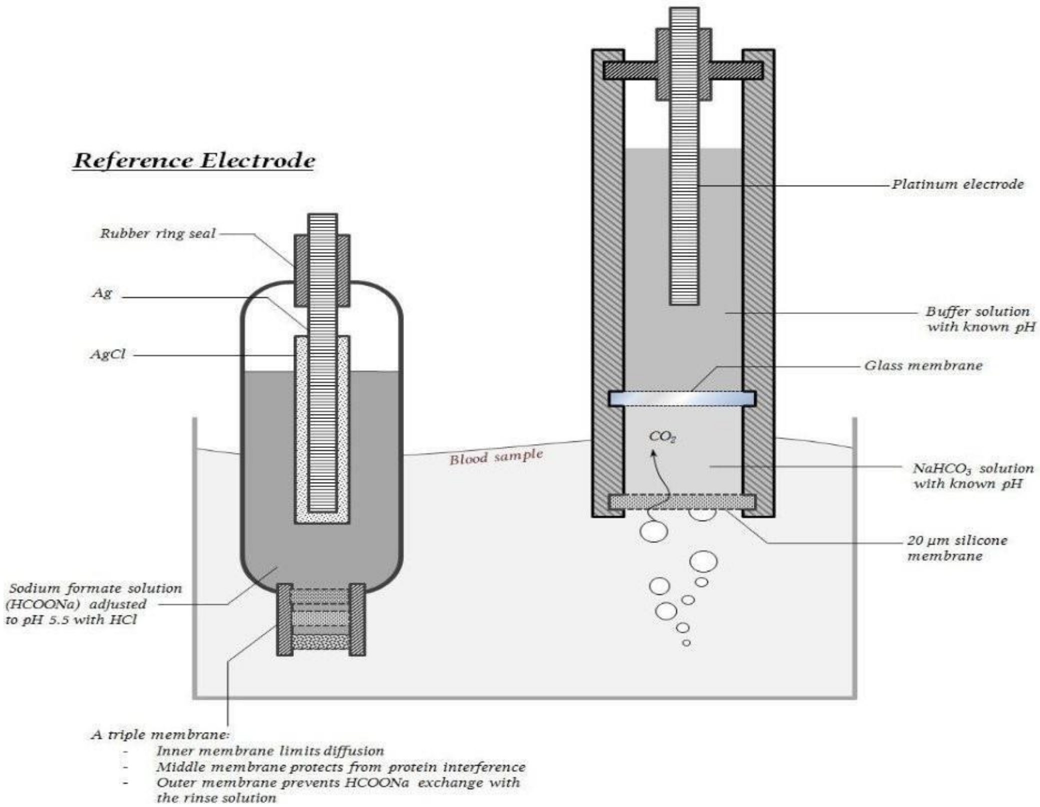

Fig: 1.4- Pco2 Measurement

The CO_2 dissolved in the sample diffuses into the middle compartment of the electrode via a thin membrane. After experimenting with several thin films of different polymers, Severinghaus settled on Teflon; however subsequent generations have changed over to silicon polymer.

Once inside, the CO_2 finds itself in an aqueous solution. For convenience, there may or may not be a bicarbonate solution added to this chamber. The reaction which takes place is an old familiar carbonic acid dissociation equilibrium:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

Thus, the pH of the solution in the middle chamber changes. The change in pH is completely dependent on the pCO_2, provided the temperature and pressure remain constant:

$$\boxed{pH = pK_a + \log \frac{cHCO_3^-}{pCO_2 \times a_{CO2}}}$$

The pKa is the dissociation equilibrium constant for the dissociation of carbonic acid in water

a_{CO2} is the solubility coefficient for CO_2 in water

Thus, from the change in pH, one can calculate the pCO_2

1.4 ION SENSOR

Chemical sensors are miniaturized analytical devices, which can deliver real-time and on-line information on the presence of specific compounds or ions in complex samples. Usually an analyte recognition process takes place followed by the conversion of chemical information into an electrical or optical signal. Among various classes of chemical sensorsion-selective electrodes (ISE) are one of the most frequently used potentiometric sensors during laboratory analysis as well as in industry, process control, physiological measurements, and environmental monitoring. The principle of ion-selective electrodes operation is quite well investigated and understood.

An ion-selective membrane is the key component of all potentiometric ion sensors. It establishes the preference with which the sensor responds to the analyte in the presence of various interfering ions from the sample. If ions can penetrate the boundary between two phases, then an electrochemical equilibrium will be reached, in which different potentials in the two phases are formed. If only one type of an ion can be exchanged between the two phases, then the potential difference formed between the phases is governed only by the activities of this target ion in these phases. When the membrane separates two solutions of different ionic activities (a_1 and a_2) and provided the membrane is only permeable to this single type of ion, the potential difference (E) across the membrane is described by the Nernst equation:

$$\mathbf{E = RT/zF \cdot \ln (a_2/a_1)}$$

If the activity of the target ion in phase 1 is kept constant, the unknown activity in phase 2 ($a_1 = a_x$) is related to (E) by:

$$\mathbf{E = RT/z_xF \cdot \ln (a_x/a_1) = const + S \cdot \log (a_x)}$$

where S=59.16/z [mV] at 298 K and z_x - the charge of the analyte. The potential difference can be measured between two identical reference electrodes placed in the two phases. In practice the potential difference i.e. the electromotive force is measured between an ion selective electrode and a reference electrode, placed in the sample solution. An examplary set-up for the measurement of electromotive force is presented in figure 1. It is important to note that this is a measurement at zero current i.e. under equilibrium conditions. Equilibrium means that the transfer of ions from the membrane into solution is equal to the transfer from the solution to the membrane. The measured signal is the sum of different potentials generated at all solid- solid, solid-liquid and liquid-liquid interfaces.

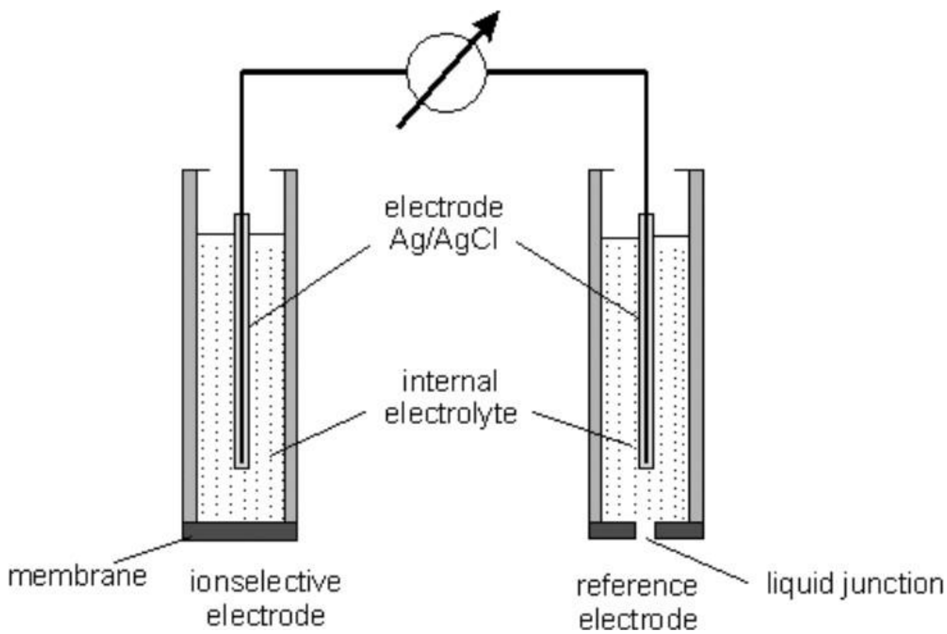

Fig.1.5 Measurement set-up of ion selective electrode (ISE)

Using a series of calibrating solutions the response curve or calibration curve of an ion-selective electrode can be measured and plotted as the signal (electromotive force) versus the activity of the analyte. Typical calibration curve of a potentiometric sensor determined in this way is shown in figure 2. The linear range of the calibration curve is usually applied to determine the activity of the target ion in any unknown solution. However, it should be pointed out that only at constant ionic strength a linear relationship between the signal measured and the concentration of the analyte is maintained (because of the clear cut relationship between ion activity and concentration, occurring in such condition).

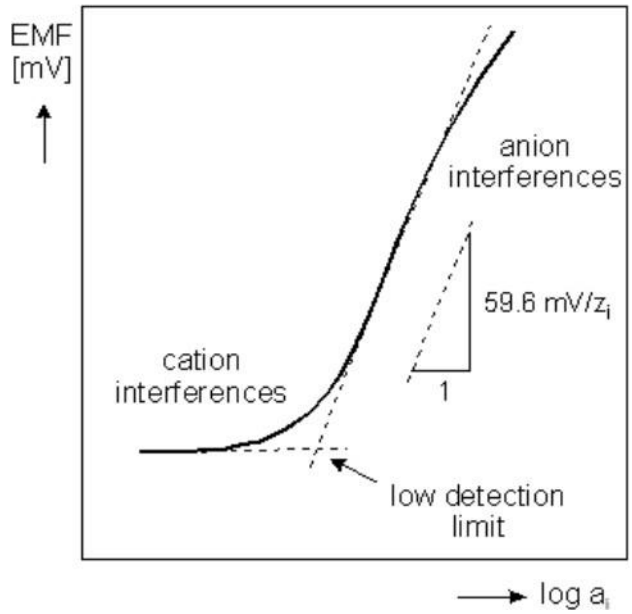

Fig.1.6 Typical calibration curve of an ion-selective electrode

Ions, present in the sample, for which the membrane is non-permeable (i.e. non-selective), will have no effect on the measured potential difference. However, a membrane truly selective for a single type of an ion and completely non-selective for other ions does not exist. For this reason the potential of such a membrane is governed mainly by the activity of the primary (target) ion and also by the activity of other secondary (interfering) ions. The influence of the presence of interfering species in a sample solution on the measured potential difference is taken into consideration in the Nikolski-Eisenman formalism:

$$\mathbf{E = const + S \cdot (log\ (a_x) + (z_x/z_y) \cdot log\ (K_{xy} \cdot a_y))}$$

where (a_y) is the activity of an interfering ion, (z_y) its charge and (K_{xy}) the selectivity coefficient (determined empirically).

Characterization of an ion-selective electrode

The properties of an ion-selective electrode are characterized by parameters like:

Selectivity. The selectivity is one of the most important characteristics of an electrode, as it often determines whether a reliable measurement in the sample is possible or not. The selectivity coefficient (K_{xy}) has been introduced in the Nikolski-Eisenman equation. Most often it is expressed as the logarithm of (K_{xy}). Negative values indicate a preference for the target ion relative to the interfering ion. Positive values of log Kxy indicate the preference of an electrode for the interfering ion. The experimental selectivity coefficients depend on the activity and a method of their determination. Different methods of the selectivity determination can be found in the literature. The IUPAC suggests two methods: separate solution method (SSM) and fixed interference method (FIM). There is also an alternative method of the selectivity determination called matched potential method (MPM). Each of them has got advantages and drawbacks, and there are not general rules pointing which method gives the true result. The methods proposed by IUPAC with several precautions will give meaningful data.

- **Slope** of the linear part of the measured calibration curve of the electrode. The theoretical value according to the Nernst equation is: 59.16 [mV/log(a_x)]at 298 K for a single charged ion or 59.16/2 = 29.58 [mV per decade] for a double charged ion. A useful slope can be regarded as 50-60 [mV per decade] (25-30 [mV per decade] for double charged ion respectively). However, in certain applications the value of the electrode slope is not critical and worse value does not exclude its usefulness.
- **Range of linear response.** At high and very low target ion activities there are deviations from linearity. Typically, the electrode calibration curve exhibits linear response range between 10^{-1}M and 10^{-5}M.
- **Detection limit.** According the IUPAC recommendation the detection limit is defined by the cross-section of the two extrapolated linear parts of the ion-selective calibration curve. In practice, detection limit on the order of 10^{-5}-10^{-6}M is measured for most of ion- selective electrodes. The observed detection limit is often governed by the presence of other interfering ions or impurities. If for example metal buffers are used to eliminate the effects which lead to the contamination of very dilute solutions it is possible to enhance the effects which lead to the detection limit down to 10^{-10}M.

- **Response time.** In earlier IUPAC recommendations, it was defined as the time between the instant at which the ion-selective electrode and a reference electrode are dipped in the sample solution (or the time at which the ion concentration in a solution is changed on contact with ISE and a reference electrode) and the first instant at which the potential of the cell becomes equal to its steady-state value within 1 [mV] or has reached 90% of the final value (in certain cases also 63% or 95%).This definition can be extended to consider the drift of the system. In this case, the second time instant is defined as the one at which the EMF/time slope becomes equal to a limiting value. However, it should be pointed out that a single time constant does not describe the form of the electrode response function. Moreover, in many investigations the response time of the overall measuring system is determined, which influences on the response time of the ISE.

Construction of the ion-selective electrodes

Typical for all kinds of electrodes used in potentiometric ion measurements is a ion-sensitive membrane. This membrane can be prepared as:

- **Solid membrane** (e.g. glass membrane or crystal membrane)
- **Liquid membrane** (based on e.g. classical ion-exchanger, neutral or charged carrier)
- **Membrane in a special electrode** (gas-sensing or enzyme electrode). Typically such a membrane contains an analyte-selective component which is responsible for the recognition process.

According to the nature of the binding sites, the membranes can be classified as: membranes containing fixed sites and membranes containing mobile ion-exchangers or ionospheres (carriers). The binding sites are incorporated in the membrane matrix, which determines the internal polarity, lipophilicity, transport and other mechanical properties of the membrane.

Glass-membrane electrodes

The best known example of this type of electrodes is the glass electrode (pH electrode), in which the anionic fixed sites are created by defects in the SiO_2 membrane and the cationic vacancies due to the non silicon constituents of glass. When the glass membrane is exposed to water a thick hydrated layer is formed (5-100 nm), which exhibits improved mobility of the ions. The concentration of anionic binding is estimated between 3 and 10M, which determine the wide linear range of the ISE calibration curve (typically 2-12 pH). The membrane is manufactured as a bulb of typical wall thickness of 0.05-0.2 mm (the optimum thickness is the result of compromise between the mechanical properties and electrical resistance). Two processes occur during the interaction of glass hydrated membrane and the sample solution (which both influence on the value of selectivity coefficient): ion-exchange and diffusion of all participating ions.

The example of an glass-membrane electrode is the sodium-selective electrode (pNa electrode). The construction of this electrode is essentially identical to that of pH electrode, with the exception of the glass used (silica with 10% Na_2O and Al_2O_3) and the fact that the inner reference solution has a fixed sodium ion activity. Interferences from hydrogen and silver cations are high, therefore the activity of such ions should be maintained some four order of magnitude less than the sodium ion activity to be measured. With proper precautions the range of measurable activities extends from 1 down to 10^{-8}M. Although the pNa electrode is insensitive to ions other than H^+ and Ag^+, glass membranes selective for other ions (e.g.K^+) have not been constructed.

Solid-state-membrane electrodes

Other types of membranes with fixed sites include single crystals of sparsely soluble salt and heterogeneous membranes in which the insoluble salt is incorporated in some suitable inert binder. In order to consider these layers at equilibrium it is necessary to use saturated solutions. In practice, these electrodes are applied in non-saturated solutions, so in this case the insoluble membrane slowly dissolves. Insoluble inorganic matererials as: Ag_2S, CuS, CdS, PbS, LaF_3, AgCl, AgBr, AgI and AgSCN have all been tested as cation exchange membranes, incorporated into an electrode body in the form of single crystal or compressed powder discs.

These materials are ionic conductors, though the conductivity is extremely small and mainly takes place through the migration of point defects in the lattice. The response time of this membranes can be increased by incorporating aliovalent ions into the lattice (e.g. the fluoride-selective membrane LaF_3 can be doped with Eu^{2+} ions). Sensors for the detection of: Ag^+, $Cu2^+$, Cd^{2+}, Pb^{2+}, S^{2-}, F^-, Br^-, I^-, SCN^- and CN^- ions can be constructed from such membranes. The sensitivity to ions of these electrodes arise from the desolation equilibrium at the membrane surface. The measurement ranges of such electrodes lies in the range of $1\text{-}10^{-6}$M, but interference effects are frequently encountered.

It was also proved that it is possible to fabricate sensors by directly contacting the membrane with a wire (so called coated-wire electrodes) to form an ohmic contact. Such systems are found to exhibit complex behavior (time and temperature dependences), requiring frequent re-calibration, but they are extremely simple to construct. An example is a silver-selective electrode, that has been designed by attaching a wire to the back of graphite/PTFE disc, on the front side of which the silver halide is rubbed in to the surface. Another exemplary coated-wire electrode was propagated by simply coating a wire with a membrane film of PVC containing an ion-exchanger.

1.5 Liquid-membrane electrodes

In addition to solid membranes, immiscible liquid (organic) phases with ion-exchange properties can be used, with such phases stabilized against the external solution phase within a polymer or ceramic membrane. The main component of electro active membrane is neutral or charged compound, which is able to complex ions reversibly and to transfer them through an organic membrane by carrier translocation. This compound is called as an ionophore or an ion carrier. There are two kinds of ionophores: charged one (usually termed liquid exchanger) and neutral carriers. They are mobile in both free and complexed forms, so the mobilities of all speciesare part of the selectivity coefficient together with ion-exchange equilibrium. The mobile binding sites are dissolved in a suitable solvent and usually trapped in a matrix of organic polymer (gel). Ion activity measurements are performed predominantly in aqueous media, so all membrane constituents are lipophilic. Therefore, the primary interaction between the ion in water and the lipophilic membrane containing the ionophore is the extraction process.

Typical polymeric membranes are based on plasticized poly(vinylchloride) (PVC) and contain approximately 66% of an plasticizer and 33% of PVC. Such a membrane is quite similar to liquid phase, because diffusion coefficients for dissolved low molecular weight ionophores are in the order of 10^{-7}-$10^{-8}cm^2/s$. An appropriate plasticizer is added to a membrane in order to ensure the mobility of the free and complxed ionophore. It determines the membrane polarity and provides suitable mechanical properties of the membrane. The ionophore is usually present in 1% amount (approximately $10^{-2}M$), which is relatively low as compared to the glass electrode. A cation selective membrane can contain a salt of lipophylic anion and hydrophilic cation (additive), which improves performances of a membrane. Although other polymers like: polisiloxane, polystyrene, PMMA, polyamide or polyimide can be used as a membrane matrix, PVC is the most widely used matrix due to simplicity of membrane preparation.

Among the ion carriers, electrically neutral ionospheres have found a wide field of applications as components in ion-selective liquid-membrane electrodes, e.g. in clinical chemistry, electrophysiology, as detectors in ion chromatography, in highly selective transport processes through artificial membranes (also biological membranes). As a results of the introduction of natural as well synthetic ionospheres in ion-selective membranes, ISEs for direct measurement of various cations and anions were designed.

Modified-membrane electrodes

Additional selectivity can be attained by using composite membranes, in which an enzyme present in the outer part of the membrane catalyses a specific chemical reaction to generate product ions. These ions can be detected by an internal ion-selective membrane. The well- known example is the selective detection of urea using unease as the enzyme catalyst. The ammonia generated, can then be detected by an ammonia or ammonium-selective electrode described above. Similarly, enzyme reactions generating protons can be followed with glass or other proton-selective membranes. There is a multiplicity of enzyme-electrodes that can be made in this way, with substrates including aliphatic alcohols, acetylcholine, amygdalin, asparganine, glucose, glutamin, penicillin and other.

Modified electrode can be designed also as potentiometric gas sensors. The original concept was made for sensing carbon dioxide (Severinghaus electrode), but the principle on which this

Electrode operates is general for other gas sensors for the detection of: NH_3, SO_2, NO_2, HCN etc. These electrodes are based on the measurement of local ion-cativity variation, caused by permeation of gas molecules (through a hydrophobic gas-permeable membrane) to the inner electrode compartment and their subsequent interaction with an internal solution. In the case of CO_2 electrode the mechanism can be described by series of equilibrium: the partitioning of the gas molecules between the sample and the electrode (solubility equilibrium) and their hydrolysis inside the internal solution (e.g. 0.1M $NaHCO_3$), which influences on the pH of this solution. The pH change is detected by an internal pH electrode (a bicarbonate-selective electrode may also be applied). It is important to note, that if the detected species is hydrogen ion, then all acid/base species will interfere. Improved selectivity is obtained by an appropriate choice of the internal electrode and by the differential gas permeability of the hydrophobic membrane.

Liquid Membrane Ion-Selective Electrode

Another class of ion-selective electrodes uses a hydrophobic membrane containing a liquid organic complexing agent that reacts selectively with the analyte. Three types of organic complexing agents have been used: cation exchangers, anion exchangers, and neutral ionophores.

One example of a liquid-based ion-selective electrode is that for Ca^{2+}, which uses a porous plastic membrane saturated with the cation exchanger di-(*n*-decyl) phosphate. As shown here, the membrane is placed at the end of a non-conducting cylindrical tube, and is in contact with two reservoirs. The outer reservoir contains di-(*n*-decyl) phosphate (shown in red) in di-*n*-octylphenylphosphonate, which soaks into the porous membrane. The inner reservoir contains a standard aqueous solution of Ca^{2+} and a Ag/AgCl reference electrode. Calcium ion-selective electrodes are also available in which the di-(*n*-decyl) phosphate is immobilized in a polyvinyl chloride (PVC) membrane, eliminating the need for the outer reservoir containing di-(*n*-decyl) phosphate.

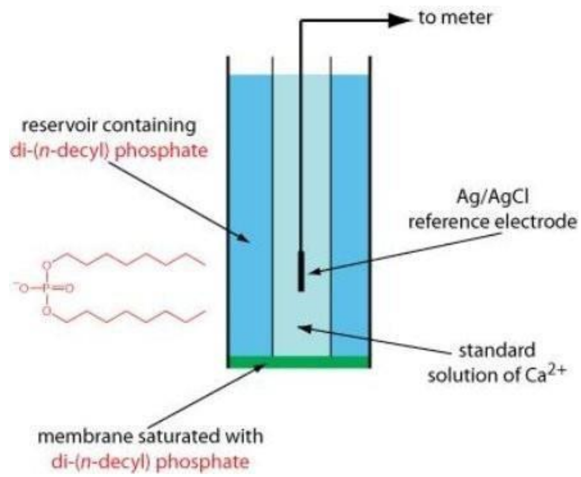

Fig: 1.7-Solid-State Ion-Selective Electrode

Schematic diagram of a solid-state electrode, the membrane for which is either a polycrystalline inorganic salt or a single crystal of an inorganic salt. We can fashion a polycrystalline solid-state ion-selective electrode by sealing a 1–2 mm thick pellet of Ag_2S—or a mixture of Ag_2S and a second silver salt or another metal sulfide—into the end of a no conducting plastic cylinder, filling the cylinder with an internal solution containing a fixed activity of analyte, and placing a reference electrode into the internal solution.

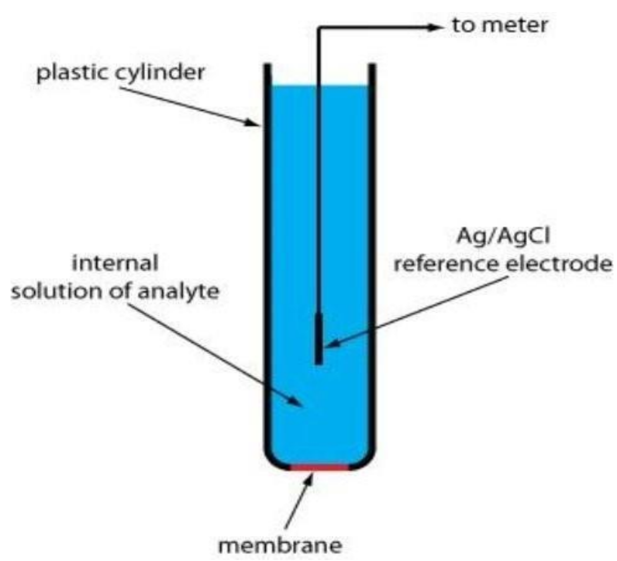

Fig: 1.8- Reference Electrode

1.6 ENZYME ELECTRODE

A gas-sensing electrode can be modified to create a potentiometric electrode that respond to a biochemically important species, such as enzymes. One example of an enzyme electrode is the urea electrode, which is based on the catalytic hydrolysis of urea by urease

$$CO(NH)_2(aq) + 2H_2O(l) \leftrightarrow 2NH_4^+(aq) + CO_3^{2-}(aq)$$

The schematic diagram here shows one version of the urea electrode, which modifies a gas-sensing NH_3 electrode by adding a dialysis membrane that traps a pH 7.0 buffered solution of urease between the dialysis membrane and the gas permeable membrane.

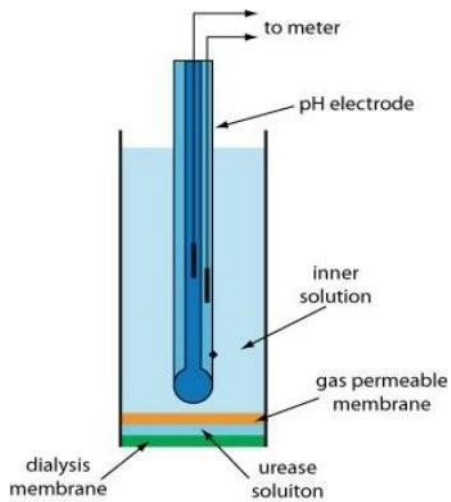

Fig: 1.9- Enzyme electrode

When immersed in the sample, urea diffuses through the dialysis membrane where it reacts with the enzyme urease to form the ammonium ion, NH_4^+, which, at a pH of 7, is largely present as NH_3. The NH_3, in turn, diffuses through the gas permeable membrane where a pH electrode measures the resulting change in pH. The response of the electrode to the concentration of urea is

$$E_{cell} = K - 0.05916\log(a_{urea})$$

Another version of the urea electrode immobilizes the enzyme urease in a polymer membrane formed directly on the tip of a glass pH electrode, as shown here.

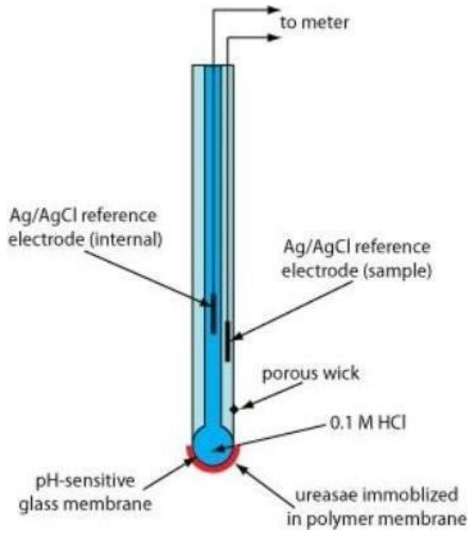

Fig: 1.10- Glass PH Electrode

1.7 PRINCIPLE OF FIBER OPTICS

Fiber-optics use light pulses to transmit information down fiber lines instead of using electronic pulses to transmit information down copper lines. Looking at the components in a fiber-optic chain will give a better understanding of how the system works in conjunction with wire based systems.

At one end of the system is a transmitter. This is the place of origin for information coming on to fiber-optic lines. The transmitter accepts coded electronic pulse information coming from copper wire. It then processes and translates that information into equivalently coded light pulses. A light-emitting diode (LED) or an injection-laser diode (ILD) can be used for generating the light pulses. Using a lens, the light pulses are funneled into the fiber-optic medium where they transmit themselves down the line.

Light pulses move easily down the fiber-optic line because of a principle known as total internal reflection. "This principle of total internal reflection states that when the angle of incidence exceeds a critical value, light cannot get out of the glass; instead, the light bounces back in. When this principle is applied to the construction of the fiber-optic strand, it is possible to transmit information down fiber lines in the form of light pulses.

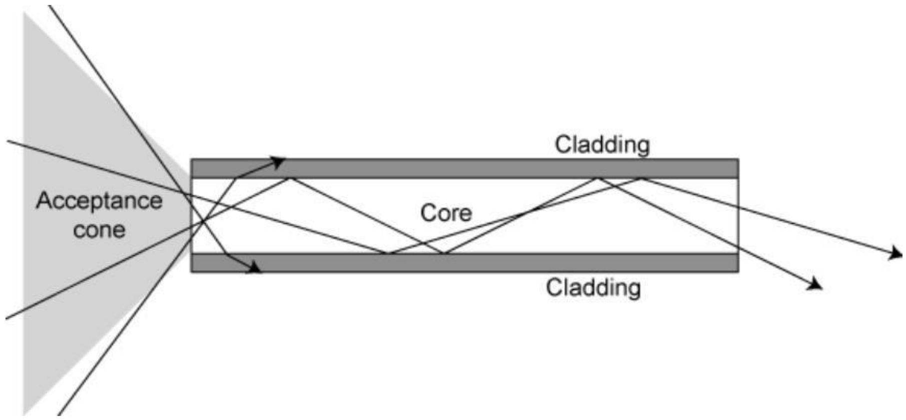

Fig: 1.11- How light is guided through an optical fiber

Incident rays which fall within the acceptance cone of the fiber are transmitted, whereas those which fall outside of the acceptance cone are lost in the cladding.

Types of Fiber Optics

There are three types of fiber optic cable commonly used: single mode, multimode and plastic optical fiber the optical fiber can be used as a medium for telecommunication and networking because it is flexible and can be bundled as cables. Although fibers can be made out of either transparent plastic (POF = plastic optical fibers) or glass, the fibers used in long-distance telecommunications applications are always glass, because of the lower optical absorption. The light transmitted through the fiber is confined due to total internal reflection within the material. This is an important property that eliminates signal crosstalk between fibers within the cable and allows the routing of the cable with twists and turns. In telecommunications applications, the light used is typically infrared light, at wavelengths near to the minimum absorption wavelength

Fibers are generally used in pairs, with one fiber of the pair carrying a signal in each direction, however bidirectional communications is possible over one strand by using two different wavelengths (colors) and appropriate coupling/splitting devices.

Single Mode Fiber

Single Mode cable is a single stand of glass fiber with a diameter of 8.3 to 10 microns that has one mode of transmission. Single Mode Fiber with a relatively narrow diameter, through which only one mode will propagate typically 1310 or 1550nm. Carries higher bandwidth than multimode fiber, but requires a light source with a narrow spectral width. Synonyms mono- mode optical fiber, single-mode fiber, single-mode optical waveguide, uni-mode fiber.

Single-mode fiber gives you a higher transmission rate and up to 50 times more distance than multimode, but it also costs more. Single-mode fiber has a much smaller core than multimode. The small core and single light-wave virtually eliminate any distortion that could result from overlapping light pulses, providing the least signal attenuation and the highest transmission speed of any fiber cable type.

Single-mode optical fiber is an optical fiber in which only the lowest order bound mode can propagate at the wavelength of interest typically 1300 to 1320nm.

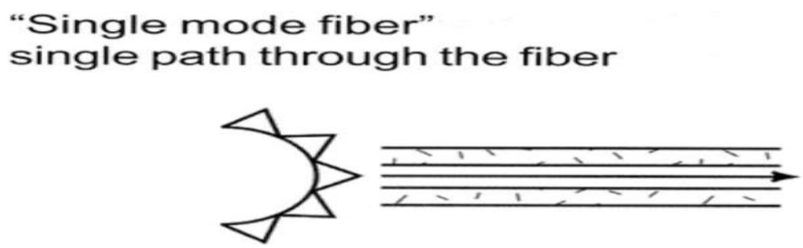

Fig: 1.12- Single mode Fiber

Multimode Fiber

Multimode cable is made of of glass fibers, with a common diameters in the 50-to-100 micron range for the light carry component (the most common size is 62.5). POF is a newer plastic- based cable which promises performance similar to glass cable on very short runs, but at a lowercost.

Multimode fiber gives you high bandwidth at high speeds over medium distances. Light waves are dispersed into numerous paths, or modes, as they travel through the cable's core typically 850 or 1300nm. Typical multimode fiber core diameters are 50, 62.5, and 100 micrometers.

However, in long cable runs (greater than 3000 feet [914.4 ml), multiple paths of light can cause signal distortion at the receiving end, resulting in an unclear and incomplete data transmission.

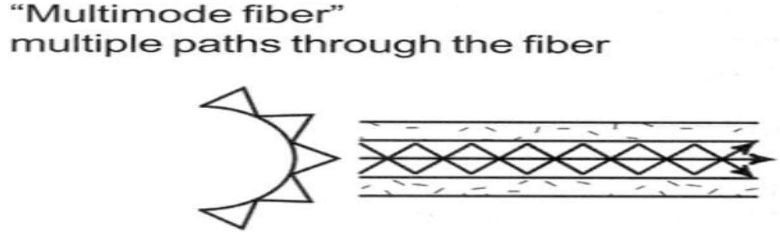

Fig: 1.13- Multi mode Fiber

Construction

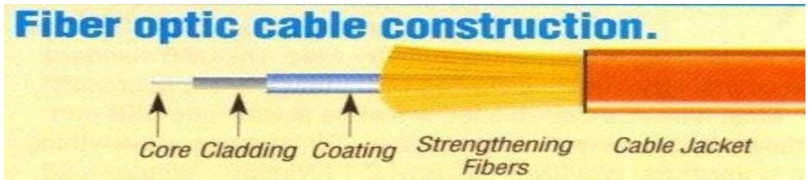

Fig: 1.14- Fiber optic cable

1.8 FIBER OPTIC SENSORS:

Fiber Optic Bio Sensors can be divided into two, sensors based on a bio catalyzed reaction and those based on a selective binding reaction.

For bio catalytic biosensors, an isolated enzyme is immobilized within the sensing region of an optical fiber. The selective bio catalytic reaction is catalyzed as the analyte approaches the immobilized enzyme and a product of this reaction is monitored through fibre optic probe.

FOBSs have been developed for the detection of ammonia, carbon dioxide, oxygen, hydrogen peroxide etc.

Research and development activities on FOBSs based on selective binding proteins are also employed on a large scale. Binding reactions between antibodies and antigens, lectins and carbohydrates, membrane receptors and their substrates have been used as chemical recognition components for sensors. Evanescent waves are also used for these types of sensors by replacing cladding and covalently attaching the antibody to the exposed fibre core.

Evanescent field excites the labelled molecules within the evanescent zone surrounding the exposed core. If the antibody is labelled, a large fluorescence signal is recorded in the absence of antigen. Binding of the antigen to the antibody either quenches or enhances the measured fluorescence intensity.

1.9 PHOTO ACOUSTIC SENSOR

The **photo acoustic effect** or **optoacoustic effect** is the formation of sound waves following light absorption in a material sample. In order to obtain this effect the light intensity must vary, either periodically (*modulated light*) or as a single flash (*pulsed light*). The photoacoustic effect is quantified by measuring the formed sound (pressure changes) with appropriate detectors, such as microphones or piezoelectric sensors. The time variation of the electric output (current or voltage) from these detectors is the photoacoustic signal. These measurements are useful to determine certain properties of the studied sample.

PPG SENSOR

Photoplethysmography (PPG) is a simple and low-cost optical technique that can be used to detect blood volume changes in the microvascular bed of tissue. It is often used non-invasively to make measurements at the skin surface. The PPG waveform comprises a pulsatile ('AC') physiological waveform attributed to cardiac synchronous changes in the blood volume with each heart beat, and is superimposed on a slowly varying ('DC') baseline with various lower frequency components attributed to respiration, sympathetic nervous system activity and thermoregulation. Although the origins of the components of the PPG signal are not fully understood, it is generally accepted that they can provide valuable information about the cardiovascular system.

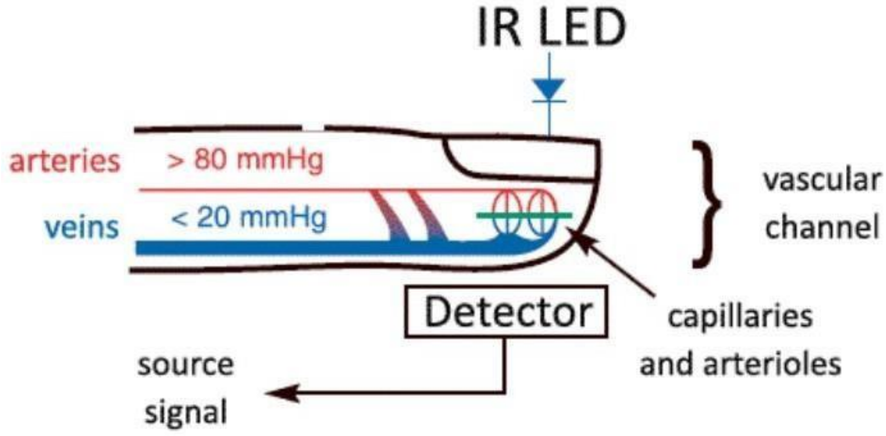

Fig: 1.15 PPG sensor

Infrared rays passing through the finger are recorded by a photo detector, which converts the light into a voltage (light/voltage converter) or frequency (light/frequency converter).

The signal is recorded either during the passage of photons from the light source to the photo detector through the tissue, or during the reflection, the light is reflected from the tissue back in direction of the photo detector. In the first case, the sensor is installed on the terminal phalanx of the finger, or earlobes, in the second it is installed with an adhesive layer on any part of the skin surface. Sensor operating in mode of light transmission has a better signal / noise ratio and it is more often used in pulse oximetry. Reflective sensor has two main advantages: no restriction on site and minimum compression of the tissue segment. The reflective sensor has significant advantages in cases when the signal is to be monitored during a long period of time (while intensive therapy or observation of a patient), as it is necessary to change the location of sensor, made in a form of “peg”, only after a several hours of test. However, the sensors operating in the mode of light transmission are optimal in case of short-term (few minutes) testing. Pulse waves recorded by the optical sensor mounted on the finger.

Fig: 1.16 Pulse waves

References:

1. Reinaldo Perez, Design of Medical Electronic Device, Elsevier Science, 1st Edition.
2. John G.Webster, Medical Instrumentation, Application and Design, 3rd Edition, John Willey and Sons,2005

Question Bank

1. How to determine the alkalinity or acidity using pH meter.
2. Determine the oxygen level in blood using pO_2 meter
3. Explain in detail about pCO_2 meter
4. State the principle and explain HCO_3 meter
5. Explain in detail about the enzyme electrode
6. State the principle behind the fiber optic cable and explain the working of fiber optic sensor
7. Explain in detail about the photo acoustic sensor
8. With the neat diagram of PPG sensor

SCHOOL OF BIO AND CHEMICAL ENGINEERING

DEPARTMENT OF BIOMEDICAL ENGINEERING

UNIT – 2 MEDICAL ELECTRONICS AND E-HEALTH- SBMA7004

II-DIGITAL BIOPOTENTIAL AMPLIFIER

Basic Parameters of Bio potential Amplifier design; Half cell potential, Reference electrodes, polarization effects, Polarisable and non polarisable electrodes, Micro electrodes, Equivalent Circuits Simplified recording of biopotential; Low-Polarization surface electrodes; Single ended bio potential amplifiers; Ultrahigh impedance electrode buffer arrays; Pastless bio potential electrodes.

- The basic requirements that a biopotential amplifier has to satisfy are:
- The physiological process to be monitored should not be influenced in any way by the amplifier
- The measured signal should not be distorted
- The amplifier should provide the best possible separation of signal and interferences
- The amplifier has to offer protection of the patient from any hazard of electrical shock

- The amplifier itself has to be protected against damages that might result from high input voltages as they occur during the application of defibrillators or electrosurgical instrumentation

2.1 BASIC PARAMETERS OF BIOPOTENTIAL AMPLIFIER DESIGN

Half-cell potential refers to the potential developed at the electrode of each half cell in an electrochemical cell. In an electrochemical cell, the overall potential is the total potential calculated from the potentials of two half cells. The measurement of half-cell potential is used to evaluate:

Presence of corrosion

Potential vulnerability of element surface area to corrosion

A half-cell potential measurement is a non-destructive method to assess the corrosion risk of steels in concrete. This method is cheaper and can be easily used. In reinforcing concrete, an electrode forms one half of the cell and the reinforcing steels in the concrete form the other half cell. The behavior of steel in concrete can be characterized by measuring its half-cell potential. The chances of corrosion occurring on the steel in concrete and half-cell potential are directly proportional; the higher the potential, the higher the risk of corrosion occurrence.

2.2 REFERENCE ELECTRODE

A reference electrode is an electrode which has a stable and well-known electrode potential. The high stability of the electrode potential is usually reached by employing a redox system with constant (buffered or saturated) concentrations of each participants of the redox reaction.

There are many ways reference electrodes are used. The simplest is when the reference electrode is used as a half cell to build an electrochemical cell. This allows the potential of the other half cell to be determined. An accurate and practical method to measure an electrode's potential in isolation (absolute electrode potential) has yet to be developed.

In a potentiometric electrochemical cell one half-cell provides a known reference potential and the potential of the other half-cell indicates the analyte’s concentration. The ideal reference electrode provides a stable, known potential so that any change in E_{cell} is attributed to analyte’s

effect on the potential of the indicator electrode. In addition, the ideal reference electrode should be easy to make and to use. Three common reference electrodes are portrayed here.

Although we rarely use the standard hydrogen electrode (SHE) for routine analytical work, it is the reference electrode used to establish standard-state potentials for other half-reactions. As shown here, the SHE consists of a Pt electrode immersed in a solution in which the activity of hydrogen ion is 1.00 and in which the fugacity of $H_2(g)$ is 1.00 . A conventional salt bridge connects the SHE to the indicator electrode’s half-cell. The standard-state potential for the reaction

$H^+(aq) + e^- \leftrightarrow \frac{1}{2}H_2(g)$

is, by definition, 0.00 V at all temperatures. Despite its importance as the fundamental reference electrode against which we measure all other potentials, the SHE is rarely used because it is difficult to prepare and inconvenient to use.

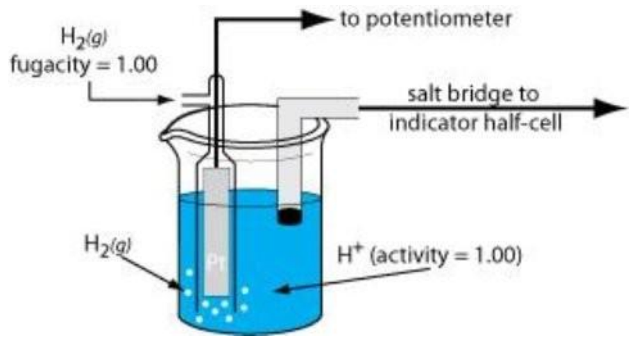

Fig:2.1 Reference electrode

A calomel electrode is based on the reduction of Hg_2Cl_2 to

Hg $Hg_2Cl_2(s) + 2e^- \leftrightarrow 2Hg(l) + 2Cl^-(aq)$

with the activity of Cl^- determining the electrode's potential. In a saturated calomel electrode (SCE), shown below, the activity of Cl^- is determined by the solubility of KCl. The electrode consists of an inner tube packed with a paste of Hg, Hg_2Cl_2, and KCl, situated within a second tube containing a saturated solution of KCl. A small hole connects the two tubes and a porous wick serves as a salt bridge to the solution in which the SCE is immersed. A stopper in the outer tube provides an opening for adding addition saturated KCl. The potential of the SCE is +0.2444 V at 25 ^oC and +0.2376 V at 35 ^oC.

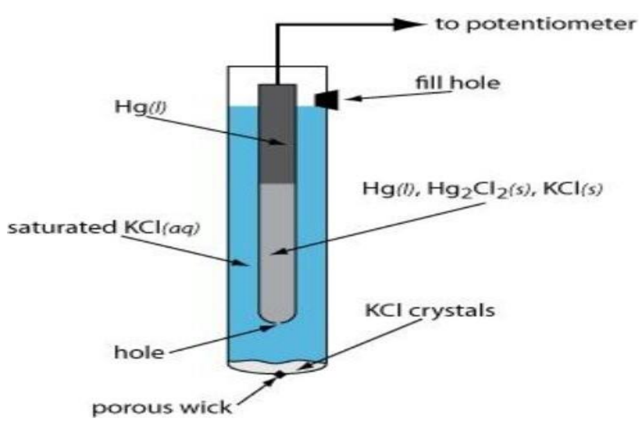

Fig:2.2 silver/silver chloride electrode

Another common reference electrode is the silver/silver chloride electrode, which is based on the following redox couple between AgCl and Ag.

$AgCl(s) + e^{-} \leftrightarrow Ag(s) + Cl^{-}(aq)$

As is the case for the calomel electrode, the activity of Cl^{-} determines the potential of the Ag/AgCl electrode. A typical Ag/AgCl electrode (shown below) consists of a silver wire, the end of which is coated with a thin film of AgCl, immersed in a solution containing the desired concentration of KCl. A porous plug serves as the salt bridge. When prepared using a saturated solution of KCl, the potential of a Ag/AgCl electrode is +0.197 V at 25 ^{o}C. Another common Ag/AgCl electrode uses a solution of 3.5 M KCl and has a potential of +0.205 V at 25 ^{o}C.

The standard state reduction potentials in most tables are reported relative to the standard hydrogen electrode's potential of +0.00 V. Because we rarely use the SHE as a reference electrode, we need to be able to convert an indicator electrode's potential to its equivalent value when using a different reference electrode. For example, the diagram below shows the relationship between the potential of an Fe^{3+}/Fe^{2+} half-cell relative to a standard hydrogen electrode (blue), a saturated silver/silver chloride electrode (red), a saturated calomel electrode (green).

2.3 POLARIZATION

In electrochemistry, polarization is a collective term for certain mechanical side-effects (of an electrochemical process) by which isolating barriers develop at the interface between electrode and electrolyte. These side-effects influence the reaction mechanisms, as well as the chemical kinetics of corrosion and metal deposition.

These mechanical side-effects are:

- Activation polarization: the accumulation of gasses (or other non-reagent products) at the interface between electrode and electrolyte.
- Concentration polarization: uneven depletion of reagents in the electrolyte cause concentration gradients in boundary layers.

Both effects isolate the electrode from the electrolyte, impeding reaction and charge transfer between the two. The immediate consequences of these barriers are:

- The reduction potential decreases, the reaction rate slows and eventually halts.
- Electric current is increasingly converted into heat rather than into desired electrochemical work.
- As predicted by Ohm's law, either electromotive force decreases and current increases, or vice-versa.

- The self-discharge rate increases in electrochemical cells.

Each of these immediate consequences has multiple secondary effects. For instance, heat affects the crystalline structure of the electrode material. This in turn can influence reaction rate, and/or accelerate dendrite formation, and/or deform the plates, and/or precipitate thermal runaway.

The mechanical side-effects can be desirable in some electrochemical processes, for example, certain types of electro polishing and electroplating take advantage of the fact that evolved gasses will first accumulate in the depressions of the plate. This feature can be used to reduce current in the depressions, and exposes ridges and edges to higher currents. Undesirable polarization can be suppressed by vigorous agitation of the electrolyte, or – when agitation is impractical (such as in a stationary battery) – with a depolarizer.

2.4 POLARISABLE AND NON POLARISABLE ELECTRODES

A polarizable electrode ideally has no DC current flow between the electrode and the surrounding electrolyte. The electrode/electrolyte interface behaves like a capacitor. Platinum in NaCl is highly polarizable, but is too expensive, stainless steel electrodes, appear to be moderately polarizable.

A non-polarizable electrode has no polarization, that is, current flows freely and the electrode-electrolyte interface behaves like a resistor. A platinum-hydrogen electrode is normally considered a good nonpolarizable electrode, but platinum in an NaCl solution (the usual case for bio-electrodes) is highly polarizable. A non-polarizable electrode that can be used in salt solutions is the silver/silver-chloride electrode.

2.5 MICROELECTRODES:

A microelectrode is an electrode of very small size, used in electrophysiology for either recording of neural signals or electrical stimulation of nervous tissue. Initially, pulled glass pipette microelectrode was used with later introduction of insulated metal wires. These microelectrodes are made from inert metals with high Young modulus such as tungsten, stainless steel, or Platinum-iridium alloy and coated with glass or polymer insulator with exposed conductive tips. More recent advances in lithography yielded to silicon based microelectrodes.

There are two main types of microelectrodes used for single-unit recordings: glass micropipettes and metal electrodes. Both are high-impedance electrodes, but glass micropipettes are highly resistive and metal electrodes have frequency-dependent impedance. Glass micropipettes are ideal for resting- and action-potential measurement, while metal electrodes are best used for extracellular spike measurements. Each type has different properties and limitations, which can be beneficial in specific applications.

Glass micropipettes

Glass micropipettes are filled with an ionic solution to make them conductive; a silver-silver chloride (Ag-AgCl) electrode is dipped into the filling solution as an electrical terminal. Ideally, the ionic solutions should have ions similar to ionic species around the electrode; the concentration inside the electrode and surrounding fluid should be the same. Additionally, the diffusive characteristics of the different ions within the electrode should be similar. The ion must also be able to "provide current carrying capacity adequate for the needs of the experiment". And importantly, it must not cause biological changes in the cell it is recording from. Ag-AgCl electrodes are primarily used with a potassium chloride (KCl) solution. With Ag- AgCl electrodes, ions react with it to produce electrical gradients at the interface, creating a voltage change with respect to time. Electrically, glass microelectrode tips have high resistance and high capacitance. They have a tip size of approximately 0.5-1.5 µm with a resistance of about 10-50 MΩ. The small tips make it easy to penetrate the cell membrane with minimal damage for intracellular recordings. Micropipettes are ideal for measurement of resting membrane potentials and with some adjustments can record action potentials. There are some issues to consider when using glass micropipettes. To offset high resistance in glass micropipettes, a cathode follower must be used as the first-stage amplifier. Additionally, high capacitance develops across the glass and conducting solution which can attenuate high- frequency responses. There is also electrical interference inherent in these electrodes and amplifiers.

Metal

Metal electrodes are made of various types of metals, typically silicon, platinum, and tungsten. They "resemble a leaky electrolytic capacitor, having a very high low-frequency impedance and low high-frequency impedance". They are more suitable for measurement of extracellular action potentials, although glass micropipettes can also be used. Metal electrodes are beneficial in some cases because they have high signal-to-noise due to lower impedance for the frequency range of spike signals. They also have better mechanical stiffness for puncturing through brain tissue. Lastly, they are more easily fabricated into different tip shapes and sizes at large quantities. Platinum electrodes are platinum black plated and insulated with glass. "They normally give stable recordings, a high signal-to-noise ratio, good isolation, and they are quite rugged in the usual tip sizes". The only limitation is that the tips are very fine and fragile. Silicon electrodes are alloy electrodes doped with silicon and an insulating glass cover layer. Silicon technology provides better mechanical stiffness and is a good supporting carrier to allow for multiple recording sites on a single electrode. Tungsten electrodes are very rugged and provide very stable recordings. This allows manufacturing of tungsten electrodes with very small tips to isolate high-frequencies. Tungsten, however, is very noisy at low frequencies. In mammalian ervous system where there are fast signals, noise can be removed with a high- pass filter. Slow signals are lost if filtered so tungsten is not a good choice for recording these signals.

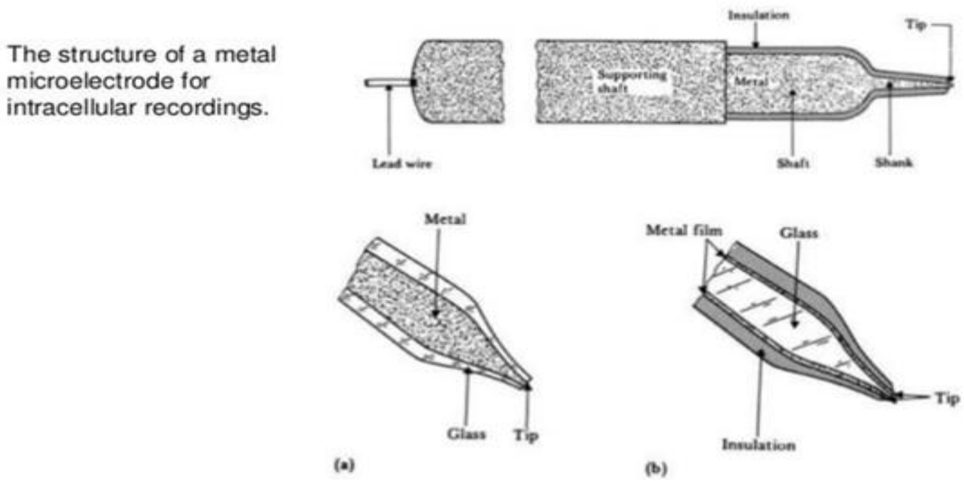

Fig:2.3- Structures of two supported metal microelectrodes (a) metal-filled glass micropipet. (b)Glass micropopet or probe, coated with metal film

EQUIVALENT CIRCUIT:

The electric characteristics of biopotential electrodes are generally nonlinear and a function of the current density at their surface. Thus, having the devices represented by linear models requires that they be operated at low potentials and currents. Under these idealized conditions, electrodes can be represented by an equivalent circuit of the form shown in the figure below. In this circuit *Rd* and *Cd* are components that represent the impedance associated with the electrode-electrolyte interface and polarization at this interface. *Rs* is the series resistance associated with interfacial effects and the resistance of the electrode materials themselves.

 Equivalent circuit

- ✓ Dc voltage source: HCP
- ✓ C_d : capacitance across the charge double layer, change with frequency, current density, electrode material, and electrolyte concentration
- ✓ R_d : leakage resistance across the charge double layer, change with frequency, current density, electrode material, and electrolyte concentration
- ✓ R_s : resistance of electrolyte, change with electrolyte concentration

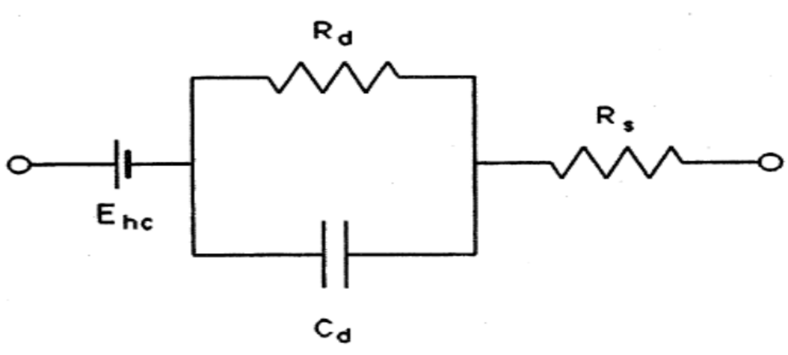

Fig:2.4- Equivalent circuit

- Skin
  - ✓ Epidermis
- Stratum corneum: outermost layer of dead cells, constantly removed

- Stratum granulosum: cells begin to die and loose nuclear material
- Stratum germinativum: cells divide and grow and displaced outward

  - ✓ Dermis
  - ✓ Subcutaneous layer

  - ✓ Vascular and nervous components, sweat glands, sweat ducts, hair follicles

- ❍ Electrode-electrolyte gel ($Cl^{?}$)-skin (Fig. 5.8)

  - ✓ Stratum corneum is the barrier

  - ✓ Rubbing or abrading the stratum corneum ? improve the stability of biopotential

- ❍ Effect of sweat ($Na^{+}, K^{+}, Cl^{?}$ ions)
- ❍ Motion artifact:

  - ✓ One electrode moved ? change in charge distribution ? change in HCP? change in the measured biopotential
  - ✓ Low frequency ? frequency components overlap with ECG, EEG, EOG, etc

  - ✓ Need better electrolyte gel

  - ✓ Skin abrasion or puncture minimize motion artifacts (skin irritation is possible)

- ❍
  - ✓

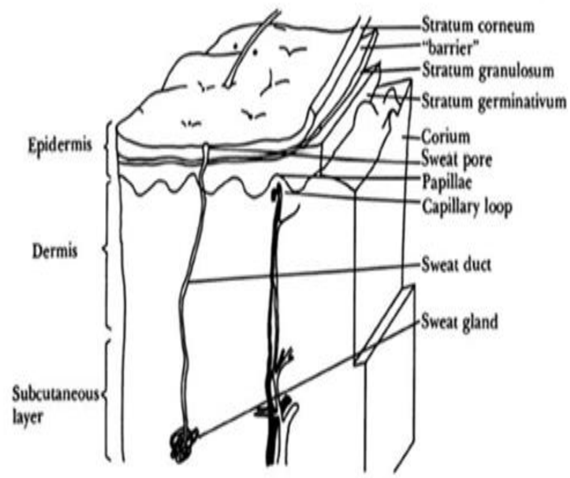

Fig:2.5 Magnified section of skin, showing the various layers

A body-surface electrode is placed against skin, showing the total electrical equivalent circuit obtained in this situation.

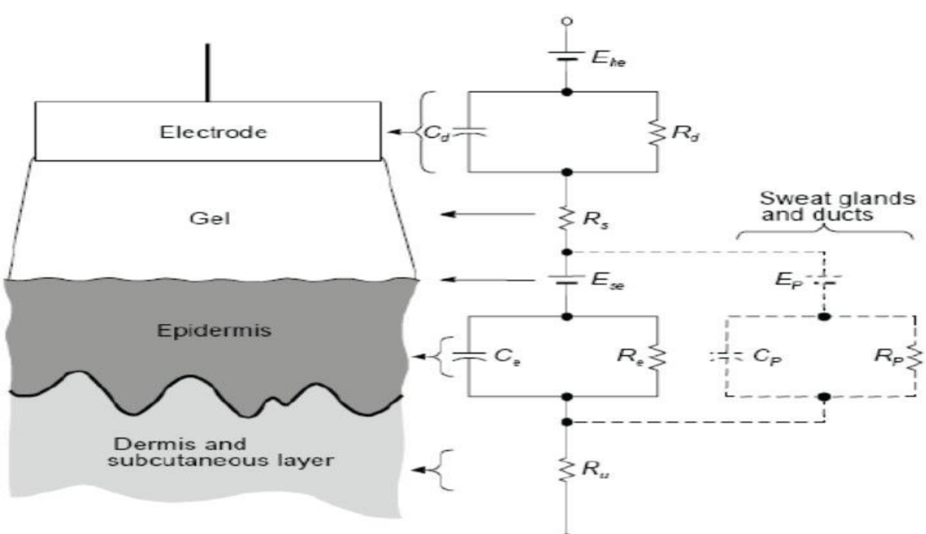

Fig:2.6 electrical equivalent circuit

Each circuit element on the right is at approximately the same level at which the physical process that it represents would be in the left hand diagram.

2.6 LOW-POLARIZATION SURFACE ELECTRODES

Silver (Ag) is a good choice for metallic skin-surface electrodes because silver forms a slightly soluble salt, silver chloride (AgCl), which quickly saturates and comes to equilibrium. A cup-shaped electrode provides enough volume to contain an electrolyte, including chlorine ions. In these electrodes, the skin never touches the electrode material directly. Rather, the interface is through an ionic solution.

One simple method to fabricate Ag/AgCl electrodes is to use electrolysis to chloride a silver base electrode (e.g., a small silver disk or silver wire). The silver substrate is immersed in a chlorine-ion-rich solution, and electrolysis is performed using a common 9-V battery connected via a series 10-kΩ potentiometer and a milliammeter. The positive terminal of the battery should be connected to the silver metal, and a plate of platinum or silver should be connected to the negative terminal and used as the opposite electrode in the solution. Our favorite electrolyte is prepared by mixing 1 part distilled water (the supermarket kind is okay), 1/2 part HCl 25%, and $FeCl_3$ at a rate of 0.5 g per milliliter of water.

If you want to make your own electrodes, use refined silver metal (99.9 to 99.99% Ag) to make the base electrode. Before chloriding, degrease and clean the silver using a concentrated aqueous ammonia solution (10 to 25%). Leave the electrodes immersed in the cleaning solution for several hours until all traces of tarnish are gone. Rinse thoroughly with deionized water (supermarket distilled water is okay) and blot-dry with clean filter paper. Don't touch the electrode surface with bare hands after cleaning. Suspend the electrodes in a suitably sized glass container so that they don't touch the sides or bottom. Pour the electrolyte into the container until the electrodes are covered, but be careful not to immerse the solder connections or leads that you will use to hook up to the electrode.

When the silver metal is immersed, the silver oxidation reaction with concomitant silver chloride precipitation occurs and the current jumps to its maximal value. As the thickness of the AgCl layer deposited increases, the reaction rate decreases and the current drops. This process continues, and the current approaches zero. Adjust the potentiometer to get an initial current density of about 2.5 mA/cm^2, making sure that no hydrogen bubbles evolve at the return electrode (large platinum or silver plate). You should remove the electrode from the solution once the current density drops to about 10 μA/cm^2. Coating should take no more than 15 to 20 minutes. Once done, remove the electrodes and rinse them thoroughly but carefully under running (tap) water.

An alternative to the electrolysis method is to immerse the silver electrode in a strong bleach solution. Yet another way of making a Ag/AgCl electrode is to coat by dipping the silver metal in molten silver chloride. To do so, heat AgCl in a small ceramic crucible with a gas flame until it melts to a dark brown liquid, then simply dip the electrode in the molten silver chloride.

If you don't want to fabricate your own electrodes, you can buy all sorts of very stable Ag/AgCl electrodes from In Vivo Metric. They make them using a very fine grained homogeneous mixture of silver and silver chloride powder, which is then compressed and sintered into various configurations. Alternatively, Ag/AgCl electrodes are cheap enough that you may get a few pregelled disposable electrodes free just by asking at the nurse's station in the emergency department or cardiology service of your local hospital.

Recording gel is available at medical supply stores (also from In Vivo Metric). However, if you really want a home brew, heat some sodium alginate (pure seaweed, commonly used to thicken food) and water with low-sodium salt (e.g., Morton Lite Salt) into a thick soup that when cooled can be applied between the electrodes and skin. Note that there is no guarantee that this concoction will be hypoallergenic! A milder paste can be made by dissolving 0.9 g of pure NaCl in 100 mL of deionized water. Add 2 g of pharmaceutical-grade Karaya gum and agitate in a magnetic stirrer for 2 hours. Add 0.09 g of methyl paraben and 0.045 g of propyl paraben as preservatives and keep in a clean capped container.

2.7 SINGLE ENDED BIOPOTENTIAL AMPLIFIERS

Most biopotential amplifiers are operational-amplifier-based circuits. As a refresher, the voltage present at the output of the operational amplifier is proportional to the differential voltage across its inputs. Thus, the noninverting input produces an in-phase output signal, while the inverting input produces an output signal that is 180° out of phase with the input.

In the circuit of Figure 1.4, an input signal V_{in} is presented through resistor R_{in} to the inverting input of an ideal operational amplifier. Resistor R_f provides feedback from the amplifier's output to its inverting input. The noninverting input is grounded, and due to the fact that in an ideal op-amp the setting conditions at one input will effectively set the same conditions at the other input, point A can be treated as it were also grounded. The power connections have been deleted for the sake of simplicity.

Ideal op-amps have an infinite input impedance, which implies that the input current i_{in} is zero. The inverting input will neither sink nor source any current. According to Kirchhoff's current law, the total current at junction A must sum to zero. Hence,

$$-i_{\text{in}} = i_f$$

But by Ohm's law, the currents are defined by

$$i_{\text{in}} = \frac{V_{\text{in}}}{R_{\text{in}}}$$

and

$$i_f = -\frac{V_{\text{out}}}{R_f}$$

Therefore, by substitution and by solving for V_{out},

$$V_{out} = \frac{R_f V_{in}}{R_{in}}$$

This equation can be rewritten as

$$V_{out} = -GV_{in}$$

where G represents the voltage gain constant R_f/R_{in}.

The circuit presented in Figure 1.5 is a noninverting voltage amplifier, also known as a *noninverting follower*, which can be analyzed in a similar manner. The setting of the noninverting input at input voltage V_{in} will force the same potential at point A. Thus,

$$i_{in} = \frac{V_{in}}{R_{in}}$$

and

$$i_f = \frac{V_{out} - V_{in}}{R_f}$$

But in the noninverting amplifier $i_{in} = i_{out}$, so by replacing and solving for V_{out}, we obtain

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

The voltage gain in this case is

$$G = 1 + \frac{R_f}{R_{in}}$$

A special case of this configuration is shown in Figure 1.6. Here $R_f = 0$, and R_{in} is unnecessary, which leads to a resistance ratio $R_f/R_{in} = 0$, which in turn results in unity gain. This configuration, termed a *unity-gain buffer* or *voltage follower*, is often used in biomedical instrumentation to couple a high-impedance signal source, through the (almost) infinite input impedance of the op-amp, to a low-impedance processing circuit connected to the very low impedance output of the op-amp.

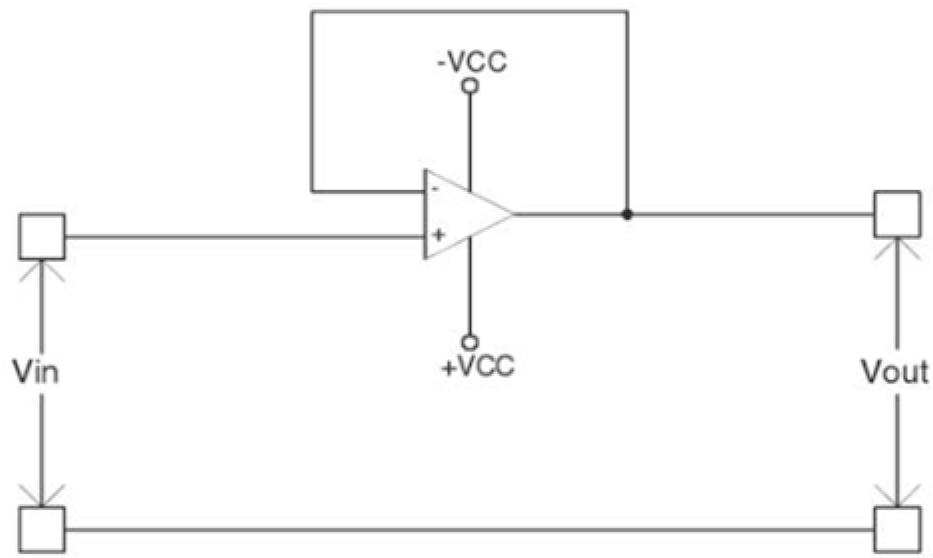

Fig:2.7 A unity-gain buffer is a special case of the non inverting voltage amplifier

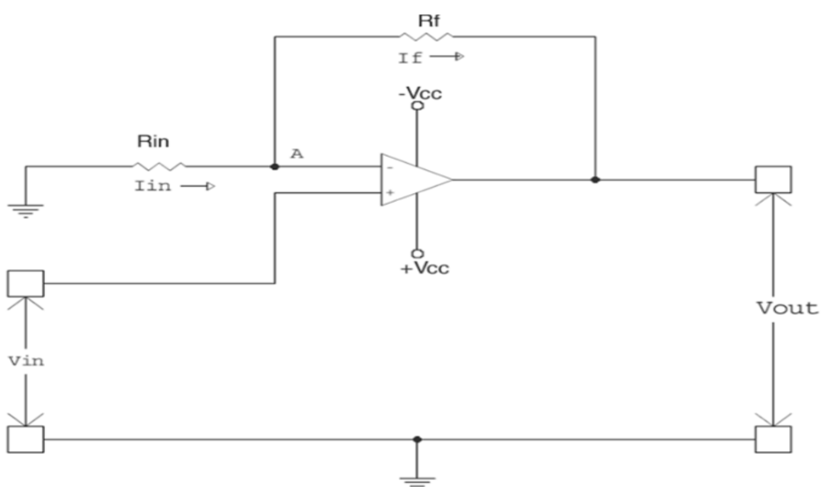

Fig:2.8 Non inverting op-amp voltage amplifier also known as a non inverting follower

2.8 ULTRAHIGH IMPEDANCE ELECTRODE BUFFER ARRAYS

A group of ultrahigh-impedance, low-power, low-noise op-amp voltage followers is commonly used as a buffer for signals collected from biopotential electrode arrays. These circuits are usually placed in close proximity to the subject or preparation to avoid contamination and degradation of biopotential signals. The circuit of Figure 1.7 comprises 32 unity-gain

buffers, which present an ultrahigh input impedance to an array of up to 32 electrodes. Each buffer in the array is implemented using a LinCMOS[1] precision op-amp operated as a unity-gain voltage follower. An output signal has the same amplitude as that of its corresponding input. The output impedance is very low, however (in the few kilohm range) and can source or sink a maximum of 25 mA. As a result of this impedance transformation, the signal at the buffer's output can be transmitted over long distances without picking up noise, despite the fact that the contact impedance of the electrodes may range into the thousands of megohms. Power for the circuit must be symmetrical ± 3 to ± 9 V dc with real or virtual ground.

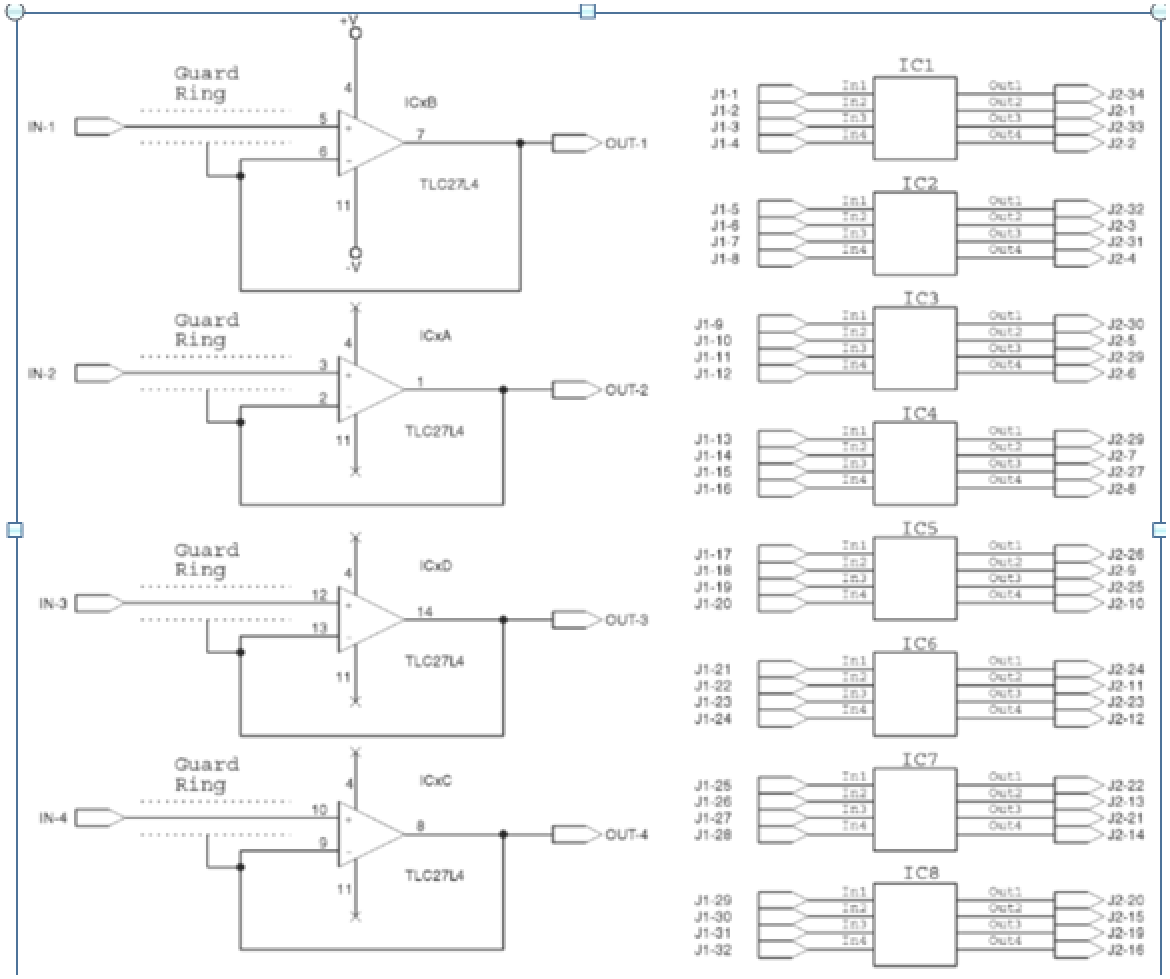

Fig:2.9 CMOS input unity gain buffers are often placed in close proximity to high-impedance electrodes to provide impedance conversion

In the circuit, input signals at J1 are buffered by eight TLC27L4 precision quad op-amp. The buffered output is available at J2. Despite its apparent simplicity, the circuit must be

laid out and constructed with care to take advantage of the op-amp's high input impedance.

2.9 PASTLESS BIOPPOTENTIAL ELECTRODES

Op-amp voltage followers are often used to buffer signals detected from biopotential sources with intrinsically high input impedance. One such application is detecting biopotential signals through capacitive bioelectrodes. One area in which these electrodes are particularly useful is in the measurement and analysis of biopotentials in humans subjected to conditions similar to those existing during flight. Knowledge regarding physiological reactions to flight maneuvers has resulted in the development of devices capable of predicting, detecting, and preventing certain conditions that might endanger the lives of crew members. For example, the detection of gravitationally induced loss of consciousness (loss of consciousness caused by extreme *g*-forces during sharp high-speed flight maneuvers in war planes) may save many pilots and their aircraft by allowing an onboard computer to take over the controls while the aviator regains consciousness [Whinnery et al., 1987]. G_{z+}-induced loss of consciousness (GLOC) detection is achieved through the analysis of various biosignals, the most important of which is the electroencephalogram (EEG).

Another new application is the use of the electrocardiography (ECG) signal to synchronize the inflation and deflation of pressure suits adaptively to gain an increase in the level of gravitational accelerations that an airman is capable of tolerating. Additional applications, such as the use of the processed electromyography (EMG) signal as a measure of muscle fatigue and pain as well as an analysis of eye blinks and eyeball movement through the detection of biopotentials around the eye as a measure of pilot alertness, constitute the promise of added safety in air operations.

One problem in making these techniques practical is that most electrodes used for the detection of bioelectric signals require skin preparation to decrease the electrical impedance of the skin–electrode interface. This preparation often involves shaving, scrubbing the skin, and applying an electrolyte paste: actions unacceptable as part of routine preflight procedures. In addition, the electrical interface characteristics deteriorate during long-term use of these electrodes as a result of skin reactions and electrolyte drying. Dry or *pasteless electrodes* can be used to get around the constraints of electrolyte–interface electrodes. Pasteless electrodes incorporate a bare or dielectric-coated metal plate, in direct contact with the skin, to form a very high impedance interface. By using an integral high-input-impedance amplifier, it is possible to record a signal through the capacitive or resistive interface.

Figure 1.10 presents the constitutive elements of a capacitive pasteless bioelectrode. In it, a highly dielectric material is used to form a capacitive interface between the skin and a conductive plate electrode. Ideally, this dielectric layer has infinite leakage resistance, but in reality this resistance is finite and decreases as the dielectric deteriorates. Signals presented to the buffer stage result from capacitive coupling of biopotentials to the network formed by series resistor R1 and the input impedance Z_{in} of the buffer amplifier. In addition, circuitry that is often used to protect the buffer stage from ESD further attenuates available signals. Shielding is usually provided in the enclosure of a bioelectrode assembly to protect it from interfering noise. The signal at the output of the buffer amplifier has low impedance and can be relayed to remotely placed processing apparatus without attenuation. External power must be supplied for operation of the active buffer circuitry.

A dielectric substance is used in capacitive biopotential electrodes to form a capacitor between the skin and the recording surface. Thin layers of aluminum anodization, pyre varnish, silicon dioxide, and other dielectrics have been used in these electrodes. For example, 17.5-μm (0.7-mil) film is easily prepared by anodic treatment, resulting in electrode plates that have a dc resistance greater than 1 GΩ and a capacitance of 5000 pF at

30 Hz. Unfortunately, standard anodization breaks down in the presence of saline (e.g., from sweat), making the electrodes unreliable for long-term use.

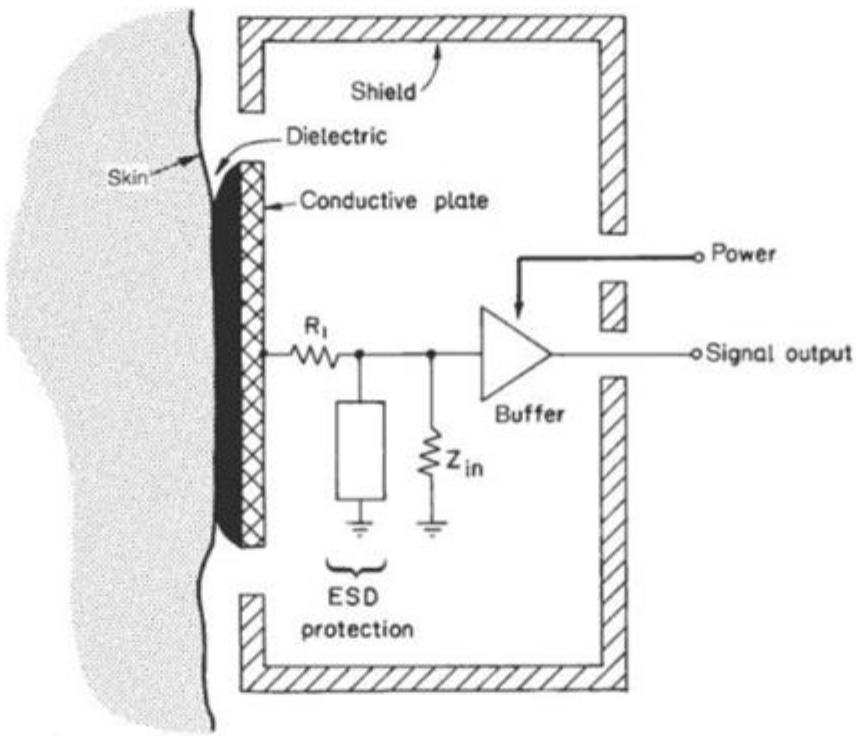

Fig:2.10 Block diagram of a typical capacitive active bio electrode

References:

1. John G.Webster, Medical Instrumentation, Application and Design, 3rd Edition, John Willey and Sons,2005.

2. Norris, A.C. Essentials of Telemedicine and Telecare. Wiley (ISBN 0-471-53151-0), 2002.

Question Bank

1. Explain detail about Half cell potential

2. With the neat diagram explain about bio potential amplifier
3. With the equivalent circuit explain about Microelectrodes.
4. Explain about the depth and needle electrodes.
5. Explain about the surface electrodes.

SCHOOL OF BIO AND CHEMICAL ENGINEERING

DEPARTMENT OF BIOMEDICAL ENGINEERING

UNIT – 3 MEDICAL ELECTRONICS AND E-HEALTH- SBMA7004

III SIGNAL CONDITIONING AND MEDICAL SAFETY.

Signal Conditioning circuits- Characteristics of Amplifiers, Differential Amplifiers, Filters, Bridge circuits, A/D Converters. Medical Safety: Electromagnetic Interference Requirements; Transient Voltage protection; Sources of Conducted and radiated interference.

3.1 SIGNAL CONDITIONING CIRCUITS

In electronics, signal conditioning means manipulating an analog signal in such a way that it meets the requirements of the next stage for further processing. Most common use is in analog-to-digital converters.

Inputs:

Signal inputs accepted by signal conditioners include DC voltage and current, AC voltage and current, frequency and electric charge. Sensor inputs can be accelerometer, thermocouple, thermistor, resistance thermometer, strain gauge or bridge, and LVDT or RVDT. Specialized inputs include encoder, counter or tachometer, timer or clock, relay or switch, and other specialized inputs. Outputs for signal conditioning equipment can be voltage, current, frequency, timer or counter, relay, resistance or potentiometer, and other specialized outputs. **Signal conditioning processes:**

Signal conditioning can include amplification, filtering, converting, range matching, isolation and any other processes required to make sensor output suitable for processing after conditioning.

Filtering

Filtering is the most common signal conditioning function, as usually not all the signal frequency spectrum contains valid data. The common example is 50/60 Hz AC power lines, present in most environments, which will produce noise if amplified.

Amplifying

Signal amplification performs two important functions: increases the resolution of the input signal, and increases its signal-to-noise ratio. For example, the output of an electronic temperature sensor, which is probably in the millivolts range is probably too low for an analog-to-digital converter (ADC) to process directly. In this case it is necessary to bring the voltage level up to that required by the ADC.

Commonly used amplifiers on signal on conditioning include sample and hold amplifiers, peak detectors, log amplifiers, antilog amplifiers, instrumentation amplifiers and programmable gain amplifiers.

Isolation

Signal isolation must be used in order to pass the signal from the source to the measuring device without a physical connection: it is often used to isolate possible sources of signal perturbations. Also notable is that it is important to isolate the potentially expensive equipment used to process the signal after conditioning from the sensor.

Magnetic or optic isolation can be used. Magnetic isolation transforms the signal from voltage to a magnetic field, allowing the signal to be transmitted without a physical connection (for example, using a transformer). Optic isolation takes an electronic signal and modulates it to a signal coded by light transmission (optical encoding), which is then used for input for the next stage of processing.

3.2 CHARACTERISTICS OF AMPLIFIERS:

The requirements for bio-potential amplifiers can often be more demanding than for a lot of electronic equipment as might be used in the entertainment or telecommunications sectors. When measuring electrical signals, such as the ECG, from the surface of the body typical requirements could be:

- Very High Input Impedance: Because the source impedance of electrodes can be quite high, this necessitates very high input impedance in the amplifier. As seen previously, this could reach to GΩ under extreme conditions. Getting very high input impedance is not that difficult with op-amp type circuits but for differential amplifiers, as would be used in ECG monitoring, it can be quite difficult to balance the impedance on both inputs.

• Moderate Bandwidth: The bandwidth requirements for most bioamplifiers are not extreme but generally need to be higher than the highest frequency component in the signal spectrum. This is in order to prevent phase distortion of the components at the higher end of the spectrum. These components are often associated with minute detail at low amplitude in biological signals and may contain diagnostically important information to be used in a clinical environment. Any band-limiting carried out by low-pass filtering must use a linear phase approach such as a BesselThompson filter design.

• Sufficient Gain-Bandwidth Product: It must be possible to obtain sufficient gain over the bandwidth of interest. Operational amplifiers can have very high open-loop gain at dc but this rapidly falls off with frequency. A measure of the closed-loop gain which can be realised is given by the Gain-Bandwidth Product parameter for the op-amps which must be sufficiently high to meet requirements.

• High Common-Mode-Rejection: Most biomedical amplifiers will have to operate in a differential manner, measuring the potential difference between two inputs. The body picks up signals, in particular mains supply interference, which are unwanted and present at both inputs. Transducers which use bridge arrangements in their construction also have large common- mode bias voltages at both outputs. Bioamplifiers in general must have a high ability to reject common-mode signals in favour of the wanted differential input signal so that only the latter contributes to the output.

3.3 DIFFERENTIAL AMPLIFIER

A differential amplifier is a type of electronic amplifier that amplifies the difference between two input voltages but suppresses any voltage common to the two inputs. It is an analog circuit with two inputs and one output in which the output is ideally proportional to the difference between the two voltages

$$V_{\text{out}} = A(V_{\text{in}}^{+} - V_{\text{in}}^{-})$$

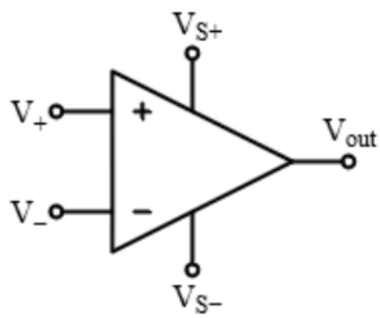

Fig:3.1 Differential Amplifier

Many electronic devices use differential amplifiers internally. The output of an ideal differential amplifier is given by:

$$V_{\text{out}} = A_{\text{d}}(V_{\text{in}}^{+} - V_{\text{in}}^{-})$$

Where V_{in}^{+} and V_{in}^{-} are the input voltages and A_{d} is the differential gain. In practice, however, the gain is not quite equal for the two inputs. This means, for instance, that if V_{in}^{+} and V_{in}^{-} are equal, the output will not be zero, as it would be in the ideal case. A more realistic expression for the output of a differential amplifier thus includes a second term.

A_{c}

$$V_{\text{out}} = A_{\text{d}}(V_{\text{in}}^{+} - V_{\text{in}}^{-}) + A_{\text{c}}\left(\frac{V_{\text{in}}^{+} + V_{\text{in}}^{-}}{2}\right)$$

is called the common-mode gain of the amplifier. As differential amplifiers are often used to null out noise or bias-voltages that appear at both inputs, a low common-mode gain is usually desired.

The common-mode rejection ratio (CMRR), usually defined as the ratio between differential-mode gain and common-mode gain, indicates the ability of the amplifier to accurately cancel voltages that are common to both inputs. The common-mode rejection ratio is defined as:

$$\mathbf{CMRR} = 10\log_{10}\left(\frac{A_d}{A_c}\right)^2 = 20\log_{10}\left(\frac{A_d}{|A_c|}\right)$$

In a perfectly symmetrical differential amplifier, A_c is zero and the CMRR is infinite. Note that a differential amplifier is a more general form of amplifier than one with a single input; by grounding one input of a differential amplifier, a single-ended amplifier results.

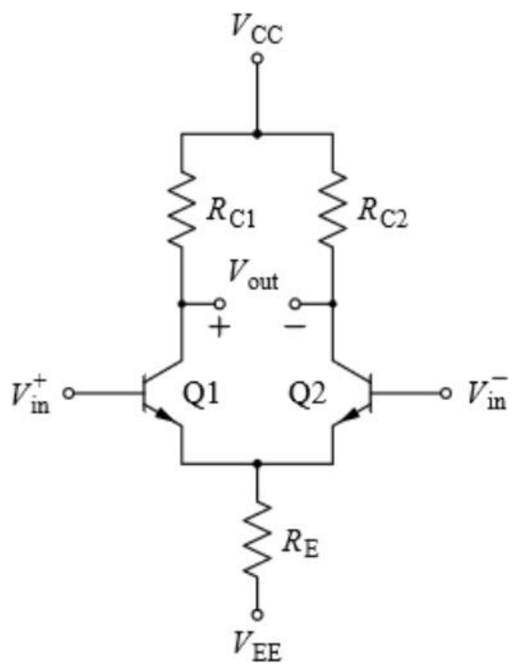

Fig:3.2 Differential Amplifier

To explain the circuit operation, four particular modes are isolated below although, in practice, some of them act simultaneously and their effects are superimposed.

Biasing

In contrast with classic amplifying stages that are biased from the side of the base (and so they are highly β-dependent), the differential pair is directly biased from the side of the emitters by sinking/injecting the total quiescent current. The series negative feedback (the emitter degeneration) makes the transistors act as voltage stabilizers; it forces them to adjust their V_{BE} voltages (base currents) to pass the quiescent current through their collector-emitter junctions. So, due to the negative feedback, the quiescent current depends only slightly on the transistor's β.

The biasing base currents needed to evoke the quiescent collector currents usually come from the ground, pass through the input sources and enter the bases. So, the sources have to be galvanic (DC) to ensure paths for the biasing current and low resistive enough to not create significant voltage drops across them. Otherwise, additional DC elements should be connected between the bases and the ground (or the positive power supply).

Common mode

At common mode (the two input voltages change in the same directions), the two voltage (emitter) followers cooperate with each other working together on the common high-resistive emitter load (the "long tail"). They all together increase or decrease the voltage of the common emitter point (figuratively speaking, they together "pull up" or "pull down" it so that it moves). In addition, the dynamic load "helps" them by changing its instant ohmic resistance in the same direction as the input voltages (it increases when the voltage increases and vice versa.) thus keeping up constant total resistance between the two supply rails. There is a full (100%) negative feedback; the two input base voltages and the emitter voltage change simultaneously while the collector currents and the total current do not change. As a result, the output collector voltages do not change as well.

Differential mode

Normal.

At differential mode (the two input voltages change in opposite directions), the two voltage (emitter) followers oppose each other - while one of them tries to increase the voltage of the common emitter point, the other tries to decrease it (figuratively speaking, one of them "pulls up" the common point while the other "pulls down" it so that it stays immovable) and v.v. So, the common point does not change its voltage; it behaves like a virtual ground with a magnitude determined by the common-mode input voltages. The high-resistive emitter element does not play any role since it is shunted by the other low-resistive emitter follower. There is no negative feedback since the emitter voltage does not change at all when the input base voltages change. The common quiescent current vigorously steers between the two transistors and the output collector voltages vigorously change. The two transistors mutually ground their emitters; so, although they are common-collector stages, they actually act as common-emitter stages with maximum gain. Bias stability and independence from variations in device parameters can be improved by negative feedback introduced via cathode/emitter resistors with relatively small resistances.

If the input differential voltage changes significantly (more than about a hundred millivolts), the transistor driven by the lower input voltage turns off and its collector voltage reaches the positive supply rail. At high overdrive the base-emitter junction gets reversed. The other transistor (driven by the higher input voltage) drives all the current. If the resistor at the collector is relatively large, the transistor will saturate. With relatively small collector resistor and moderate overdrive, the emitter can still follow the input signal without saturation. This mode is used in differential switches and ECL gates.

Breakdown.

If the input voltage continues increasing and exceeds the base-emitter breakdown voltage, the base-emitter junction of the transistor driven by the lower input voltage breaks down. If the input sources are low resistive, an unlimited current will flow directly through the "diode bridge" between the two input sources and will damage them.

At common mode, the emitter voltage follows the input voltage variations; there is a full negative feedback and the gain is minimum. At differential mode, the emitter voltage is fixed (equal to the instant common input voltage); there is no negative feedback and the gain is maximum.

3.4 FILTER

In signal processing, a filter is a device or process that removes some unwanted components or features from a signal. Filtering is a class of signal processing, the defining feature of filters being the complete or partial suppression of some aspect of the signal[*clarification needed*]. Most often, this means removing some frequencies and not others in order to suppress interfering signals and reduce background noise. However, filters do not exclusively act in the frequency domain; especially in the field of image processing many other targets for filtering exist. Correlations can be removed for certain frequency components and not for others without having to act in the frequency domain.

There are many different bases of classifying filters and these overlap in many different ways; there is no simple hierarchical classification. Filters may be:

- linear or non-linear
- in-invariant or time-variant, also known as shift invariance. If the filter operates in a spatial domain then the characterization is space invariance.
- causal or not-causal: depending if present output depends or not on "future" input; of course, for time related signals processed in real-time all the filters are causal; it is not necessarily so for filters acting on space-related signals or for deferred-time processing of time-related signals.
- analog or digital
- discrete-time (sampled) or continuous-time
- passive or active type of continuous-time filter
- infinite impulse response (IIR) or finite impulse response (FIR) type of discrete-time or digital filter.

Linear continuous-time circuit is perhaps the most common meaning for filter in the signal processing world, and simply "filter" is often taken to be synonymous. These circuits are generally designed to remove certain frequencies and allow others to pass. Circuits that perform this function are generally linear in their response, or at least approximately so. Any nonlinearity would potentially result in the output signal containing frequency components not present in the input signal.

The modern design methodology for linear continuous-time filters is called network synthesis. Some important filter families designed in this way are:

- Chebyshev filter, has the best approximation to the ideal response of any filter for a specified order and ripple.
- Butterworth filter, has a maximally flat frequency response.
- Bessel filter, has a maximally flat phase delay.
- Elliptic filter, has the steepest cutoff of any filter for a specified order and ripple.

The difference between these filter families is that they all use a different polynomial function to approximate to the ideal filter response. This results in each having a different transfer function.

Another older, less-used methodology is the image parameter method. Filters designed by this methodology are archaically called "wave filters". Some important filters designed by this method are:

- Constant k filter, the original and simplest form of wave filter.
- m-derived filter, a modification of the constant k with improved cutoff steepness and impedance matching.

Some terms used to describe and classify linear filters:

- The frequency response can be classified into a number of different bandforms describing which frequency bands the filter passes (the passband) and which it rejects (the stopband):
- Low-pass filter – low frequencies are passed, high frequencies are attenuated.
- High-pass filter – high frequencies are passed, low frequencies are attenuated.
- Band-pass filter – only frequencies in a frequency band are passed.
- Band-stop filter or band-reject filter – only frequencies in a frequency band are attenuated.
- Notch filter – rejects just one specific frequency - an extreme band-stop filter.
- Comb filter – has multiple regularly spaced narrow passbands giving the bandform the appearance of a comb.
- All-pass filter – all frequencies are passed, but the phase of the output is modified.

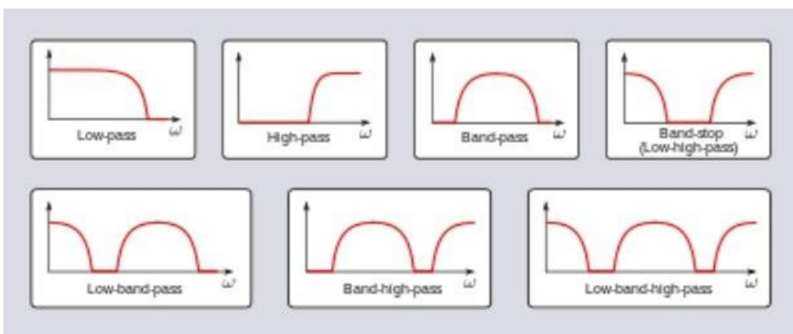

Fig:3.2 Filters

- Cutoff frequency is the frequency beyond which the filter will not pass signals. It is usually measured at a specific attenuation such as 3 dB.
- Roll-off is the rate at which attenuation increases beyond the cut-off frequency.
- Transition band, the (usually narrow) band of frequencies between a pass band and stop band.
- Ripple is the variation of the filter's insertion loss in the pass band.
- The order of a filter is the degree of the approximating polynomial and in passive filters corresponds to the number of elements required to build it. Increasing order increases roll-off and brings the filter closer to the ideal response.

One important application of filters is in telecommunication. Many telecommunication systems use frequency-division multiplexing, where the system designers divide a wide frequency band into many narrower frequency bands called "slots" or "channels", and each stream of information is allocated one of those channels. The people who design the filters at each transmitter and each receiver try to balance passing the desired signal through as accurately as possible, keeping interference to and from other cooperating transmitters and noise sources outside the system as low as possible, at reasonable cost.

Multilevel and multiphase digital modulation systems require filters that have flat phase delay—are linear phase in the pass band—to preserve pulse integrity in the time domain,[1] giving less inter symbol interference than other kinds of filters.

On the other hand, analog audio systems using analog transmission can tolerate much larger ripples in phase delay, and so designers of such systems often deliberately sacrifice linear phase to get filters that are better in other ways—better stop-band rejection, lower passband amplitude ripple, lower cost, etc.

The transfer function of a filter is most often defined in the domain of the complex frequencies. The back and forth passage to/from this domain is operated by the Laplace transform and its inverse (therefore, here below, the term "input signal" shall be understood as "the Laplace transform of" (the time representation of) the input signal, and so on).

The transfer function $H(s)$ of a filter is the ratio of the output signal $Y(s)$ to that of the input signal $X(s)$ as a function of the complex frequency : s

$$H(s) = \frac{Y(s)}{X(s)}$$

with $s = \sigma + j\omega$.

The transfer function of all linear time-invariant filters generally share certain characteristics:

- For filters which are constructed of discrete components, their transfer function must be the ratio of two polynomials in , i.e. a rational function of . The order of the transfer function will be the highest power of encountered in either the numerator or the denominator.
- The polynomials of the transfer function will all have real coefficients. Therefore, the poles and zeroes of the transfer function will either be real or occur in complex conjugate pairs.
- Since the filters are assumed to be stable, the real part of all poles (i.e. zeroes of the denominator) will be negative, i.e. they will lie in the left half-plane in complex frequency space.

Distributed element filters do not, in general, produce rational functions but can often approximate to them.

The proper construction of a transfer function involves the Laplace transform, and therefore it is needed to assume null initial conditions, because

$$\mathcal{L}\left\{\frac{df}{dt}\right\} = s \cdot \mathcal{L}\left\{f(t)\right\} - f(0),$$

And when f(0)=0 we can get rid of the constants and use the usual expression

$$\mathcal{L}\left\{\frac{df}{dt}\right\} = s \cdot \mathcal{L}\left\{f(t)\right\}$$

An alternative to transfer functions is to give the behavior of the filter as a convolution. The convolution theorem, which holds for Laplace transforms, guarantees equivalence with transfer functions.

3.5 BRIDGE CIRCUITS

A bridge circuit is a type of electrical circuit in which two circuit branches (usually in parallel with each other) are "bridged" by a third branch connected between the first two branches at some intermediate point along them. The bridge was originally developed for laboratory measurement purposes and one of the intermediate bridging points is often adjustable when so used. Bridge circuits now find many applications, both linear and non-linear, including in instrumentation, filtering and power conversion.

The best-known bridge circuit, the Wheatstone bridge, was invented by Samuel Hunter Christie and popularized by Charles Wheatstone, and is used for measuring resistance. It is constructed from four resistors, two of known values R_1 and R_3 (see diagram), one whose resistance is to be determined R_x, and one which is variable and calibrated R_2. Two opposite vertices are connected to a source of electric current, such as a battery, and a galvanometer is connected across the other two vertices. The variable resistor is adjusted until the galvanometer reads zero. It is then known that the ratio between the variable resistor and its neighbour R1 is equal to the ratio between the unknown resistor and its neighbour R3, which enables the value of the unknown resistor to be calculated.

The Wheatstone bridge has also been generalised to measure impedance in AC circuits, and to measure resistance, inductance, capacitance, and dissipation factor separately. Various arrangements are known as the Wien bridge, Maxwell bridge and Heaviside bridge. All are based on the same principle, which is to compare the output of two potentiometers sharing a common source.

In power supply design, a bridge circuit or bridge rectifier is an arrangement of diodes or similar devices used to rectify an electric current, i.e. to convert it from an unknown or alternating polarity to a direct current of known polarity.

In some motor controllers, an H-bridge is used to control the direction the motor turns.

An example for bridge circuit:

A Wheatstone bridge is an electrical circuit used to measure an unknown electrical resistance by balancing two legs of a bridge circuit, one leg of which includes the unknown component. Its operation is similar to the original potentiometer. It was invented by Samuel Hunter Christie in

1833 and improved and popularized by Sir Charles Wheatstone in 1843. One of the Wheatstone bridge's initial uses was for the purpose of soils analysis and comparison.

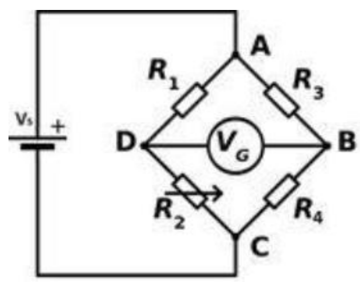

Fig:3.3 bridge circuit

Operation

In the figure, R_4 is the unknown resistance to be measured; R_1, and R_3 are resistors of known resistance and the resistance of is R_2 adjustable. If the ratio of the two resistances in the known leg (R_2/R_1) is equal to the ratio of the two in the unknown leg (R_x/R_3), then the voltage between the two midpoints (B and D) will be zero and no current (electricity) will flow through the Galvanometer (G) or V_g. No current flow through G V_g, the bridge is said to be "balanced". If the bridge is unbalanced, the direction of the current indicates whether is too R_2 high or too low. is varied R_2 until there is no current through the galvanometer, which then reads zero.

Detecting zero current with a galvanometer can be done to extremely high accuracy. Therefore, if R_1, R_2 and R_3 are known to high precision, then R_x can be measured to high precision. Very small changes in R_x disrupt the balance and are readily detected.

At the point of balance, the ratio of

$$\frac{R_2}{R_1} = \frac{R_4}{R_3}$$

$$\Rightarrow R_4 = \frac{R_2}{R_1} \cdot R_3$$

Alternatively, if R_1, R_2 , and R_3 are known, but R_2 is not adjustable, the voltage difference across or current flow through the meter can be used to calculate the value of R_4, using Kirchhoff's circuit laws (also known as Kirchhoff's rules). This setup is frequently used in strain gauge and resistance thermometer measurements, as it is usually faster to read a voltage level off a meter than to adjust a resistance to zero the voltage.

Derivation

Wheatstone Bridge current path

First, Kirchoff's current law is used to find the currents in junctions B and D:

$$I_3 - I_x + I_G = 0$$

$$I_1 - I_2 - I_G = 0$$

The potential difference between points B and D will be near zero (0 VDC) volts using a digital multimeter.

Then, Kirchhoff's voltage law (KVL) is used for finding the voltage in the loops ABD and BCD:

$$(I_3 \cdot R_3) - (I_G \cdot R_G) - (I_1 \cdot R_1) = 0$$
$$(I_x \cdot R_4) - (I_2 \cdot R_2) + (I_G \cdot R_G) = 0$$

When the bridge is balanced, then $I_G = 0$, so the second set of equations can be rewritten as:

$$I_3 \cdot R_3 = I_1 \cdot R_1$$
$$I_x \cdot R_x = I_2 \cdot R_2$$

Then, the equations are divided and rearranged, giving:

$$R_4 = \frac{R_2 \cdot I_2 \cdot I_3 \cdot R_3}{R_1 \cdot I_1 \cdot I_x}$$

From the first rule, $I_3 = I_x$ and $I_1 = I_2$. The desired value of R_4 is now known to be given as:

$$R_4 = \frac{R_3 \cdot R_2}{R_1}$$

If all four resistor values and the supply voltage (V_S) are known, and the resistance of the galvanometer is high enough that I_G is negligible, the voltage across the bridge (V_G) can be found by working out the voltage from each potential divider and subtracting one from the other. The balance state equation for this is:

$$V_G = \left(\frac{R_4}{R_4 + R_3} - \frac{R_2}{R_1 + R_2} \right) * \ V_s$$

where V_G is the voltage of node D relative to node B.

Significance

The Wheatstone bridge illustrates the concept of a difference measurement, which can be extremely accurate. Variations on the Wheatstone bridge can be used to measure capacitance, inductance, electrical impedance and other quantities, such as the amount of combustible gases in a sample, with an explosimeter. The Kelvin bridge was specially adapted from the Wheatstone bridge for measuring very low resistances. In many cases, the significance of measuring the unknown resistance is related to measuring the impact of some physical phenomenon (such as force, temperature, pressure, etc.) which thereby allows the use of Wheatstone bridge in measuring those elements indirectly.

Application

The practical applications for the Wheatstone Bridge circuit is used for many biomedical applications to include but not limited to:

- MRI/CT - table positioning, accurate movement of the CT scan imaging device and equal patient weight distribution. High accuracy is needed to perform imaging functions while preventing over-travel of the patient placed within the scanning tube.
- Infusion pumps/Syringe pumps - monitors and controls the amount of fluid flow of intravenous medication that was to be received via the tubing.
- Mammography - monitors the amount of physical force that is applied to the patients breast by the machine itself when attempting to take an image.
- Conventional radiography - monitors the amount of x-ray dose received to the AEC cells and the patient.
- Scales, weighing/Patient lifts - with the incorporation of a load cell into the bottom metal plates these scales routinely require re-zeroing which uses the above circuit.
- Remote robotic surgeries - used so physicians are able to precisely measure both depth of force and drill bit rotational force during remote hip surgeries
- Dialysis machines - ensures uniform fluid flow and circulation of proper rate, proportion and frequency according to the parameters set by its accompanying electronic controller device.
- Ventilator gas tester - ensures uniform gas flow (i.e I:E, PEEP, TV, fIO2, SIMV, AC, etc...).

Pressure Meter - ensures uniform negative and positive pressure flow (i.e. mmHg, cmH20, inH20, PSI, etc...).
Vital signs simulator

Strain gauge circuits

Pressure transducer circuits

and more

3.6 ANALOG-TO-DIGITAL CONVERTER

In electronics, an **analog-to-digital converter** (**ADC**, **A/D**, **A–D**, or **A-to-D**) is a system that converts an analog signal into a digital signal. A digital-to-analog converter (DAC) performs the reverse function.

An ADC may also provide an isolated measurement such as an electronic device that converts an input analog voltage or current to a digital number proportional to the magnitude of the voltage or current. Typically the digital output will be a two's complement binary number that is proportional to the input, but there are other possibilities.

There are several ADC architectures. Due to the complexity and the need for precisely matched components, all but the most specialized ADCs are implemented as integrated circuits (ICs).

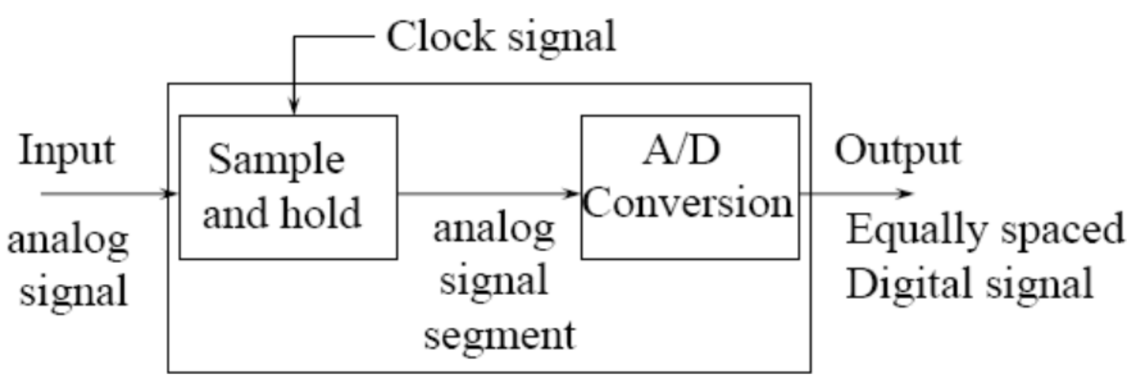

Fig:3.4 an analog-to-digital converter

Analog /Digital Conversion

2-Step Process

- Quantizing - breaking down analog value is a set of finite states
- Encoding - assigning a digital word or number to each state and matching it to the input signal

An ADC inputs an analog electrical signal such as voltage or current and outputs a binary number. In block diagram form, it can be represented as such:

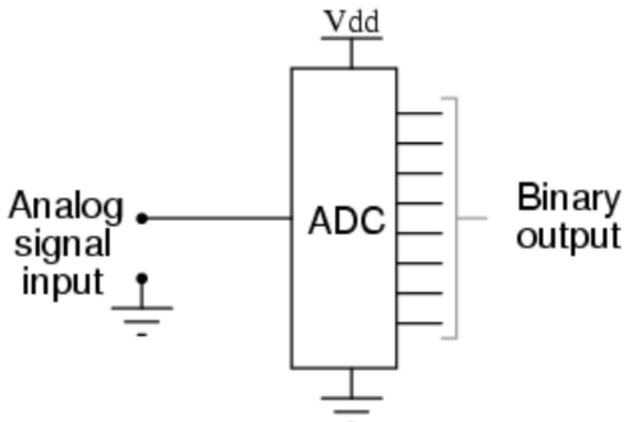

Fig:3.5 parallel A/D converter

Also called the *parallel* A/D converter, this circuit is the simplest to understand. It is formed of a series of comparators, each one comparing the input signal to a unique reference voltage. The comparator outputs connect to the inputs of a priority encoder circuit, which then produces a binary output. The following illustration shows a 3-bit flash ADC circuit:

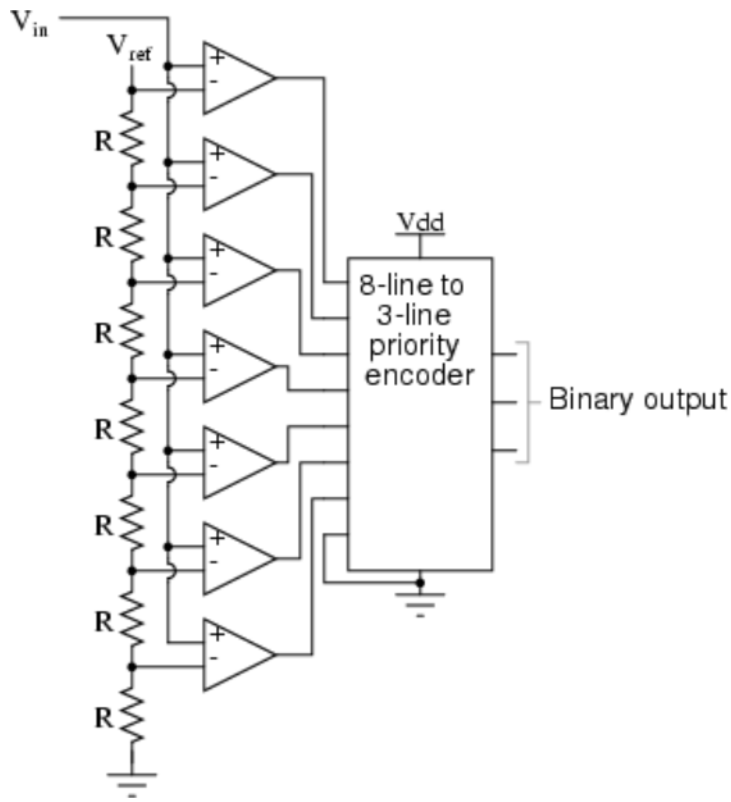

Fig:3.6 3-bit flash ADC circuit

V_{ref} is a stable reference voltage provided by a precision voltage regulator as part of the converter circuit, not shown in the schematic. As the analog input voltage exceeds the reference voltage at each comparator, the comparator outputs will sequentially saturate to a high state. The priority encoder generates a binary number based on the highest-order active input, ignoring all other active inputs.

For this particular application, a regular priority encoder with all its inherent complexity isn't necessary. Due to the nature of the sequential comparator output states (each comparator saturating "high" in sequence from lowest to highest), the same "highest-order-input selection" effect may be realized through a set of Exclusive-OR gates, allowing the use of a simpler, non-priority encoder:

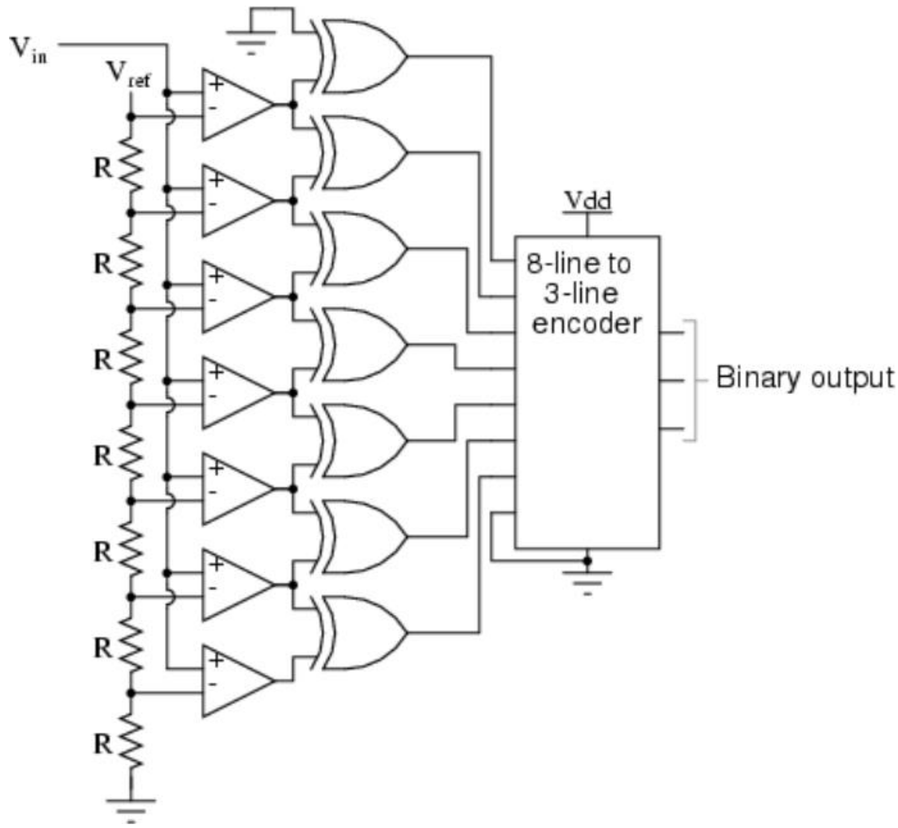

Fig:3.7 Exclusive-OR gates

And, of course, the encoder circuit itself can be made from a matrix of diodes, demonstrating just how simply this converter design may be constructed:

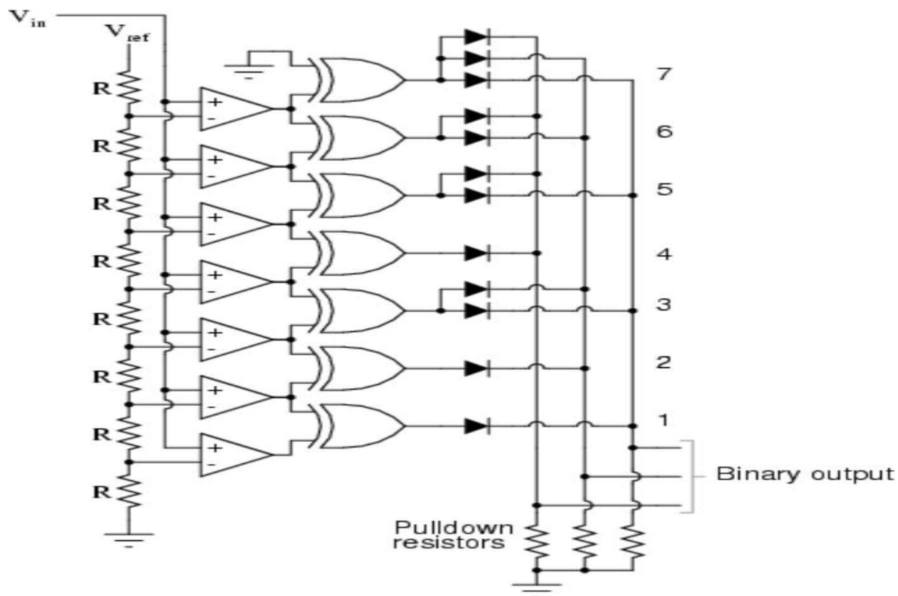

Fig:3.8: comparators

Not only is the flash converter the simplest in terms of operational theory, but it is the most efficient of the ADC technologies in terms of speed, being limited only in comparator and gate propagation delays. Unfortunately, it is the most component-intensive for any given number of output bits. This three-bit flash ADC requires seven comparators. A four-bit version would require 15 comparators. With each additional output bit, the number of required comparators doubles. Considering that eight bits is generally considered the minimum necessary for any practical ADC (255 comparators needed!), the flash methodology quickly shows its weakness.

An additional advantage of the flash converter, often overlooked, is the ability for it to produce a non-linear output. With equal-value resistors in the reference voltage divider network, each successive binary count represents the same amount of analog signal increase, providing a proportional response. For special applications, however, the resistor values in the divider network may be made non-equal. This gives the ADC a custom, nonlinear response to the analog input signal. No other ADC design is able to grant this signal-conditioning behavior with just a few component value changes.

Digital ramp ADC

Also known as the *stairstep-ramp*, or simply *counter* A/D converter, this is also fairly easy to understand but unfortunately suffers from several limitations.

The basic idea is to connect the output of a free-running binary counter to the input of a DAC, then compare the analog output of the DAC with the analog input signal to be digitized and use the comparator's output to tell the counter when to stop counting and reset. The following schematic shows the basic idea:

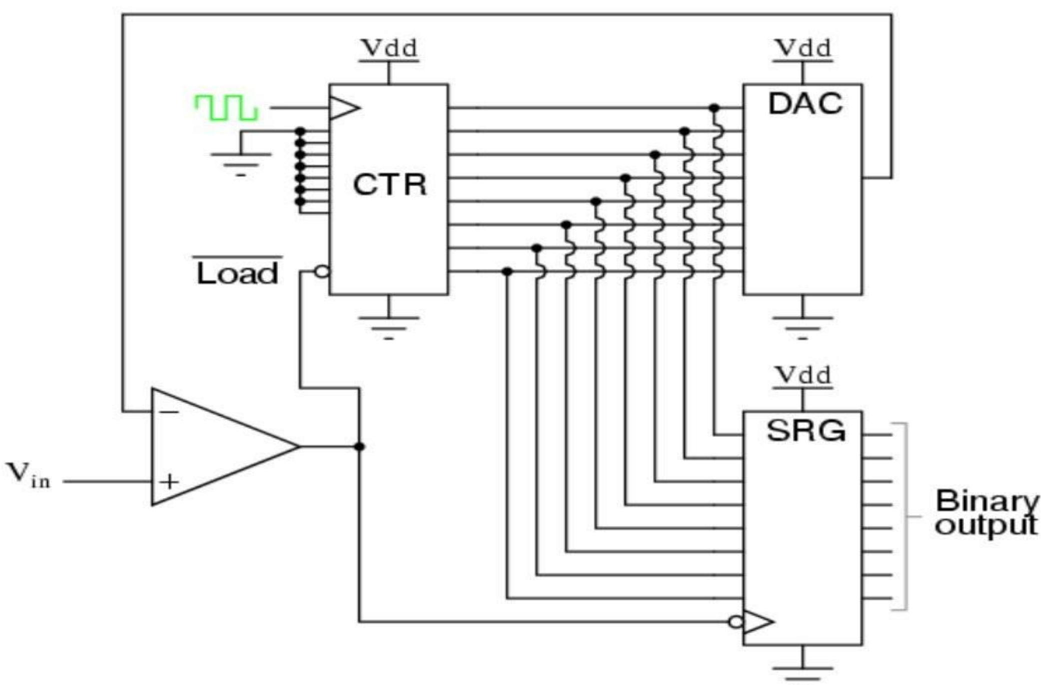

Fig:3.9: counter A/D converter

As the counter counts up with each clock pulse, the DAC outputs a slightly higher (more positive) voltage. This voltage is compared against the input voltage by the comparator. If the input voltage is greater than the DAC output, the comparator's output will be high and the counter will continue counting normally. Eventually, though, the DAC output will exceed the input voltage, causing the comparator's output to go low. This will cause two things to happen: first, the high-to-low transition of the comparator's output will cause the shift register to "load" whatever binary count is being output by the counter, thus updating the ADC circuit's output; secondly, the counter will receive a low signal on the active-low LOAD input, causing it to reset to 00000000 on the next clock pulse.

The effect of this circuit is to produce a DAC output that ramps up to whatever level the analog input signal is at, output the binary number corresponding to that level, and start over again. Plotted over time, it looks like this:

Note how the time between updates (new digital output values) changes depending on how high the input voltage is. For low signal levels, the updates are rather close-spaced. For higher signal levels, they are spaced further apart in time:

For many ADC applications, this variation in update frequency (sample time) would not be acceptable. This, and the fact that the circuit's need to count all the way from 0 at the beginning of each count cycle makes for relatively slow sampling of the analog signal, places the digital-ramp ADC at a disadvantage to other counter strategies.

Successive approximation ADC

One method of addressing the digital ramp ADC's shortcomings is the so-called *successive-approximation* ADC. The only change in this design is a very special counter circuit known as a *successive-approximation register*. Instead of counting up in binary sequence, this register counts by trying all values of bits starting with the most-significant bit and finishing at the least-significant bit. Throughout the count process, the register monitors the comparator's output to see if the binary count is less than or greater than the analog signal input, adjusting the bit values accordingly. The way the register counts is identical to the "trial-and-fit" method of decimal-to-binary conversion, whereby different values of bits are tried from MSB to LSB to get a binary number that equals the original decimal number. The advantage to this counting strategy is much faster results: the DAC output converges on the analog signal input in much larger steps than with the 0-to-full count sequence of a regular counter.

Without showing the inner workings of the successive-approximation register (SAR), the circuit looks like this:

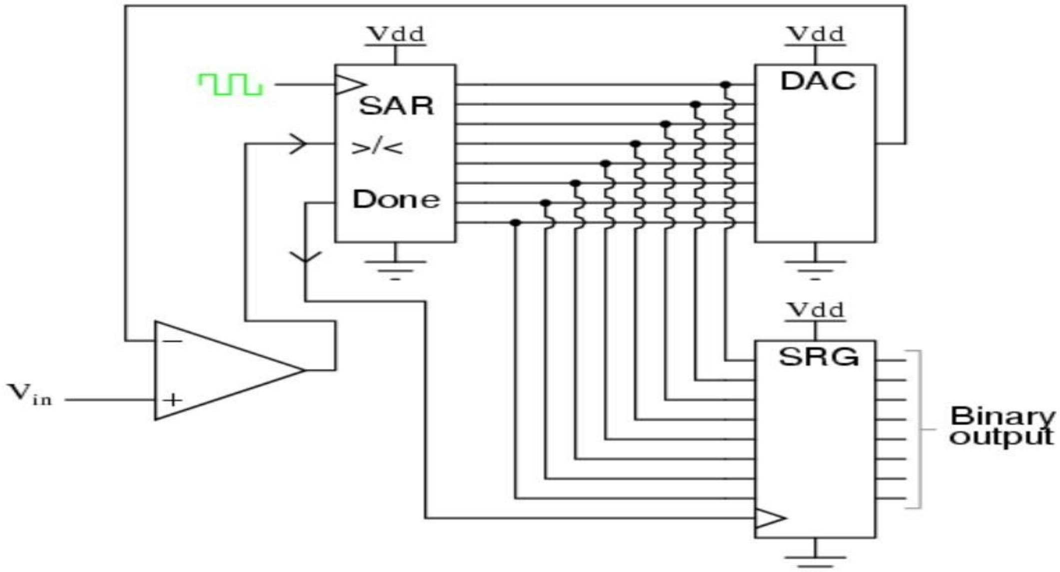

Fig:3.10 Successive approximation ADC

It should be noted that the SAR is generally capable of outputting the binary number in *serial* (one bit at a time) format, thus eliminating the need for a shift register. Plotted over time, the operation of a successive-approximation ADC looks like this:

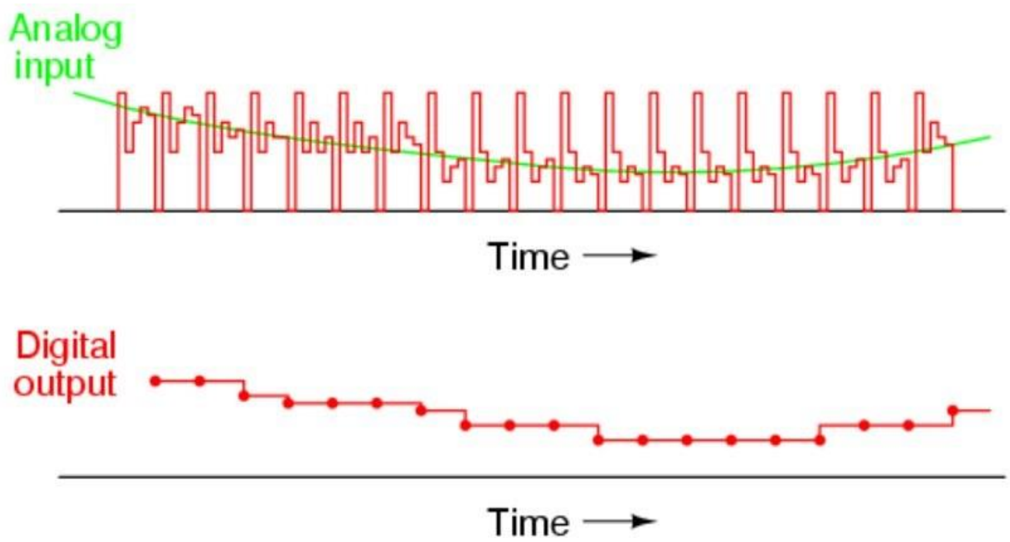

Fig:3.11 one bit at a time

Note how the updates for this ADC occur at regular intervals, unlike the digital ramp ADC circuit.

Tracking ADC

A third variation on the counter-DAC-based converter theme is, in my estimation, the most elegant. Instead of a regular "up" counter driving the DAC, this circuit uses an up/down counter. The counter is continuously clocked, and the up/down control line is driven by the output of the comparator. So, when the analog input signal exceeds the DAC output, the counter goes into the "count up" mode. When the DAC output exceeds the analog input, the counter switches into the "count down" mode. Either way, the DAC output always counts in the proper direction to *track* the input signal.

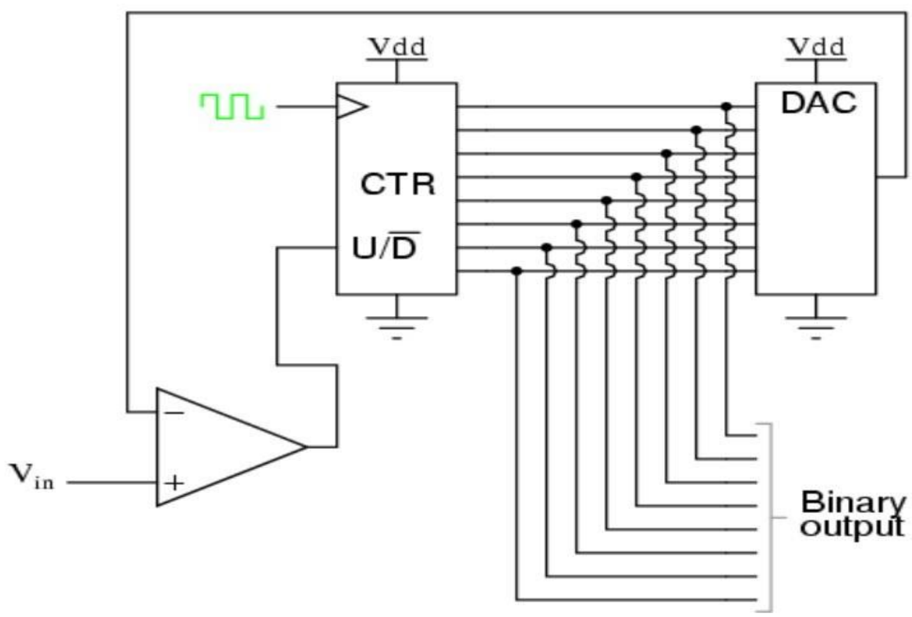

Fig:3.12 Tracking ADC

Notice how no shift register is needed to buffer the binary count at the end of a cycle. Since the counter's output continuously tracks the input (rather than counting to meet the input and then resetting back to zero), the binary output is legitimately updated with every clock pulse.

An advantage of this converter circuit is speed, since the counter never has to reset. Note the behavior of this circuit:

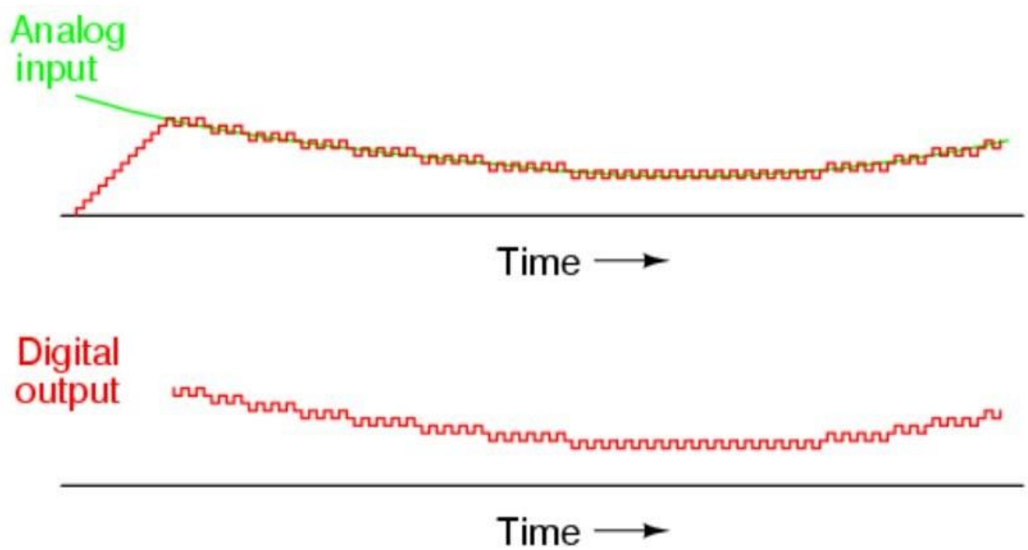

Fig:3.13 binary count

Note the much faster update time than any of the other "counting" ADC circuits. Also note how at the very beginning of the plot where the counter had to "catch up" with the analog signal, the rate of change for the output was identical to that of the first counting ADC. Also, with no shift register in this circuit, the binary output would actually ramp up rather than jump from zero to an accurate count as it did with the counter and successive approximation ADC circuits.

Perhaps the greatest drawback to this ADC design is the fact that the binary output is never stable: it always switches between counts with every clock pulse, even with a perfectly stable analog input signal. This phenomenon is informally known as *bit bobble*, and it can be problematic in some digital systems.

This tendency can be overcome, though, through the creative use of a shift register. For example, the counter's output may be latched through a parallel-in/parallel-out shift register only when the output changes by two or more steps. Building a circuit to detect two or more successive counts in the same direction takes a little ingenuity, but is worth the effort.

Slope (integrating) ADC

So far, we've only been able to escape the sheer volume of components in the flash converter by using a DAC as part of our ADC circuitry. However, this is not our only option. It is possible to avoid using a DAC if we substitute an analog ramping circuit and a digital counter with precise timing.

The is the basic idea behind the so-called *single-slope*, or *integrating* ADC. Instead of using a DAC with a ramped output, we use an op-amp circuit called an *integrator* to generate a sawtooth waveform which is then compared against the analog input by a comparator. The time it takes for the sawtooth waveform to exceed the input signal voltage level is measured by means of a digital counter clocked with a precise-frequency square wave (usually from a crystal oscillator). The basic schematic diagram is shown here:

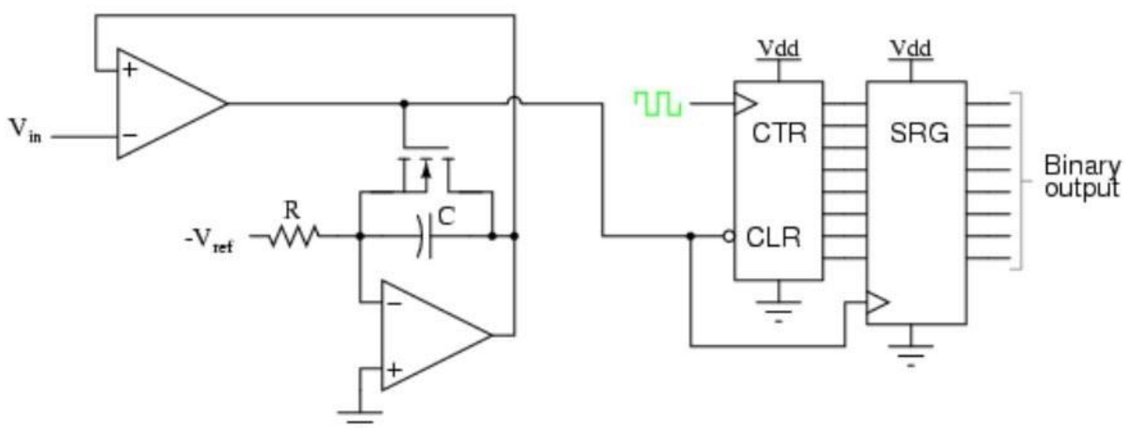

Fig:3.14 Slope (integrating) ADC

The IGFET capacitor-discharging transistor scheme shown here is a bit oversimplified. In reality, a latching circuit timed with the clock signal would most likely have to be connected to the IGFET gate to ensure full discharge of the capacitor when the comparator's output goes high. The basic idea, however, is evident in this diagram. When the comparator output is low (input voltage greater than integrator output), the integrator is allowed to charge the capacitor in a linear fashion. Meanwhile, the counter is counting up at a rate fixed by the precision clock frequency. The time it takes for the capacitor to charge up to the same voltage level as the input depends on the input signal level and the combination of $-V_{ref}$, R, and C. When the capacitor reaches that voltage level, the comparator output goes high, loading the counter's output into the shift register for a final output. The IGFET is triggered "on" by the comparator's high output, discharging the capacitor back to zero volts. When the integrator output voltage falls to zero, the comparator output switches back to a low state, clearing the counter and enabling the integrator to ramp up voltage again.

This ADC circuit behaves very much like the digital ramp ADC, except that the comparator reference voltage is a smooth saw tooth waveform rather than a "stair step:"

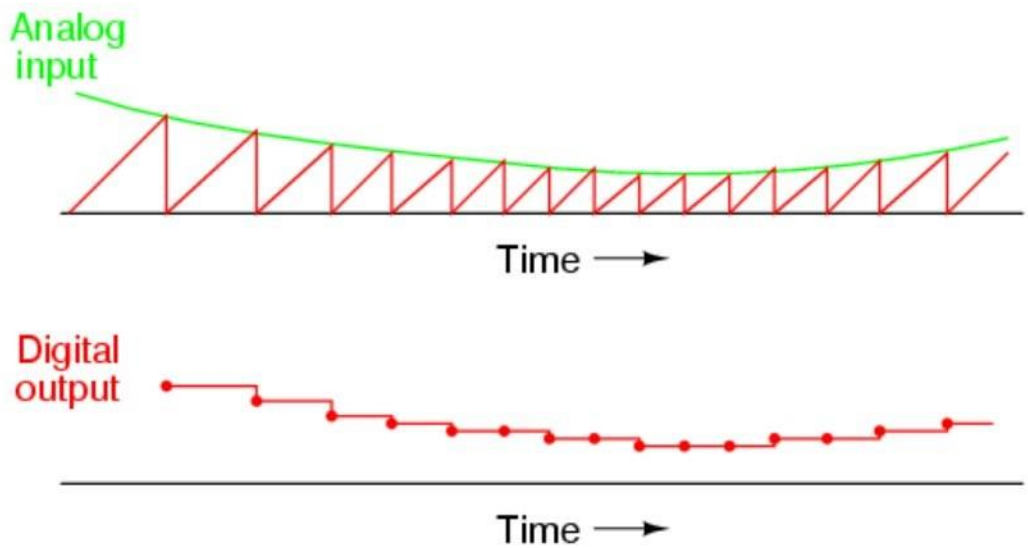

Fig:3.15 dual-slope

The single-slope ADC suffers all the disadvantages of the digital ramp ADC, with the added drawback of *calibration drift*. The accurate correspondence of this ADC's output with its input is dependent on the voltage slope of the integrator being matched to the counting rate of the counter (the clock frequency). With the digital ramp ADC, the clock frequency had no effect on conversion accuracy, only on update time. In this circuit, since the rate of integration and the rate of count are independent of each other, variation between the two is inevitable as it ages, and will result in a loss of accuracy. The only good thing to say about this circuit is that it avoids the use of a DAC, which reduces circuit complexity.

An answer to this calibration drift dilemma is found in a design variation called the *dual-slope* converter. In the dual-slope converter, an integrator circuit is driven positive and negative in alternating cycles to ramp down and then up, rather than being reset to 0 volts at the end of every cycle. In one direction of ramping, the integrator is driven by the positive analog input signal (producing a negative, variable rate of output voltage change, or output *slope*) for a fixed amount of time, as measured by a counter with a precision frequency clock. Then, in the other direction, with a fixed reference voltage (producing a fixed rate of output voltage change) with time measured by the same counter. The counter stops counting when the integrator's output reaches the same voltage as it was when it started the fixed-time portion of the cycle. The amount of time it takes for the integrator's capacitor to discharge back to its original output voltage, as measured by the magnitude accrued by the counter, becomes the digital output of the ADC circuit.

The dual-slope method can be thought of analogously in terms of a rotary spring such as that used in a mechanical clock mechanism. Imagine we were building a mechanism to measure the rotary speed of a shaft. Thus, shaft speed is our "input signal" to be measured by this device.
The measurement cycle begins with the spring in a relaxed state. The spring is then turned, or "wound up," by the rotating shaft (input signal) for a fixed amount of time. This places the spring in a certain amount of tension proportional to the shaft speed: a greater shaft speed corresponds to a faster rate of winding. and a greater amount of spring tension accumulated over that period of time. After that, the spring is uncoupled from the shaft and allowed to unwind at a fixed rate, the time for it to unwind back to a relaxed state measured by a timer device. The amount of *time* it takes for the spring to unwind at that fixed rate will be directly proportional to the *speed* at which it was wound (input signal magnitude) during the fixed-time portion of the cycle.

This technique of analog-to-digital conversion escapes the calibration drift problem of the single-slope ADC because both the integrator's integration coefficient (or "gain") and the counter's rate of speed are in effect during the entire "winding" and "unwinding" cycle portions. If the counter's clock speed were to suddenly increase, this would shorten the fixed time period where the integrator "winds up" (resulting in a lesser voltage accumulated by the integrator), but it would also mean that it would count faster during the period of time when the integrator was allowed to "unwind" at a fixed rate. The proportion that the counter is counting faster will be the same proportion as the integrator's accumulated voltage is diminished from before the clock speed change. Thus, the clock speed error would cancel itself out and the digital output would be exactly what it should be. Another important advantage of this method is that the input signal becomes averaged as it drives the integrator during the fixed-time portion of the cycle. Any changes in the analog signal during that period of time have a cumulative effect on the digital output at the end of that cycle. Other ADC strategies merely "capture" the analog signal level at a single point in time every cycle. If the analog signal is "noisy" (contains significant levels of spurious voltage spikes/dips), one of the other ADC converter technologies may occasionally convert a spike or dip because it captures the signal repeatedly at a single point in time. A dual-slope ADC, on the other hand, averages together all the spikes and dips within the integration period, thus providing an output with greater noise immunity. Dual-slope ADCs are used in applications demanding high accuracy.

Delta-Sigma (ΔΣ) ADC

One of the more advanced ADC technologies is the so-called delta-sigma, or ΔΣ (using the proper Greek letter notation). In mathematics and physics, the capital Greek letter delta (Δ) represents *difference* or *change*, while the capital letter sigma (Σ) represents *summation*: the adding of multiple terms together. Sometimes this converter is referred to by the same Greek letters in reverse order: sigma-delta, or ΣΔ.

In a ΔΣ converter, the analog input voltage signal is connected to the input of an integrator, producing a voltage rate-of-change, or slope, at the output corresponding to input magnitude. This ramping voltage is then compared against ground potential (0 volts) by a comparator. The comparator acts as a sort of 1-bit ADC, producing 1 bit of output ("high" or "low") depending on whether the integrator output is positive or negative. The comparator's output is then latched through a D-type flip-flop clocked at a high frequency, and *fed back* to another input channel on the integrator, to drive the integrator in the direction of a 0 volt output. The basic circuit looks like this:

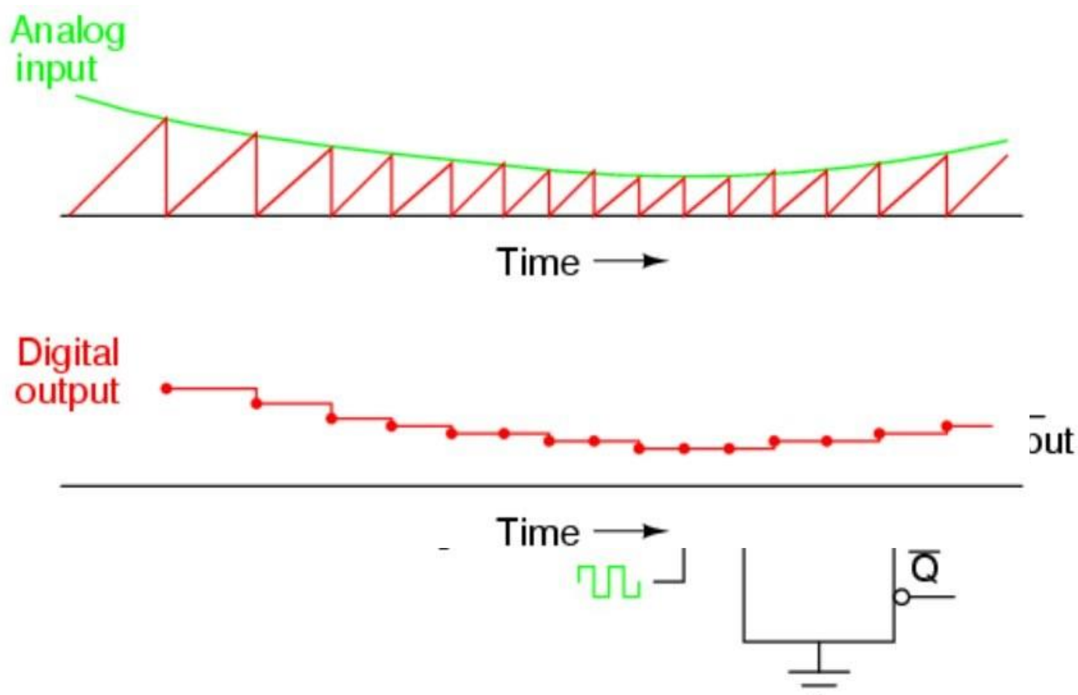

Fig:3.16 Integrator

The leftmost op-amp is the (summing) integrator. The next op-amp the integrator feeds into is the comparator, or 1-bit ADC. Next comes the D-type flip-flop, which latches the comparator's output at every clock pulse, sending either a "high" or "low" signal to the next comparator at the top of the circuit. This final comparator is necessary to convert the single-polarity 0V / 5V logic level output voltage of the flip-flop into a +V / -V voltage signal to be fed back to the integrator.

If the integrator output is positive, the first comparator will output a "high" signal to the D input of the flip-flop. At the next clock pulse, this "high" signal will be output from the Q line into the noninverting input of the last comparator. This last comparator, seeing an input voltage greater than the threshold voltage of 1/2 +V, saturates in a positive direction, sending a full +V signal to the other input of the integrator. This +V feedback signal tends to drive the integrator output in a negative direction. If that output voltage ever becomes negative, the feedback loop will send a corrective signal (-V) back around to the top input of the integrator to drive it in a positive direction. This is the delta-sigma concept in action: the first comparator senses a *difference* (Δ) between the integrator output and zero volts. The integrator *sums* (Σ) the comparator's output with the analog input signal.

Functionally, this results in a serial stream of bits output by the flip-flop. If the analog input is zero volts, the integrator will have no tendency to ramp either positive or negative, except in response to the feedback voltage. In this scenario, the flip-flop output will continually oscillate between "high" and "low," as the feedback system "hunts" back and forth, trying to maintain the integrator output at zero volts:

If, however, we apply a negative analog input voltage, the integrator will have a tendency to ramp its output in a positive direction. Feedback can only add to the integrator's ramping by a fixed voltage over a fixed time, and so the bit stream output by the flip-flop will not be quite the same:

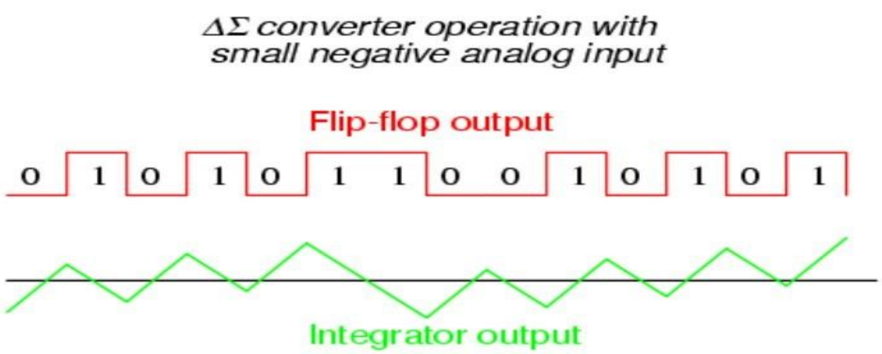

Fig:3.17 Small negative analog input

By applying a larger (negative) analog input signal to the integrator, we force its output to ramp more steeply in the positive direction. Thus, the feedback system has to output more 1's than before to bring the integrator output back to zero volts:

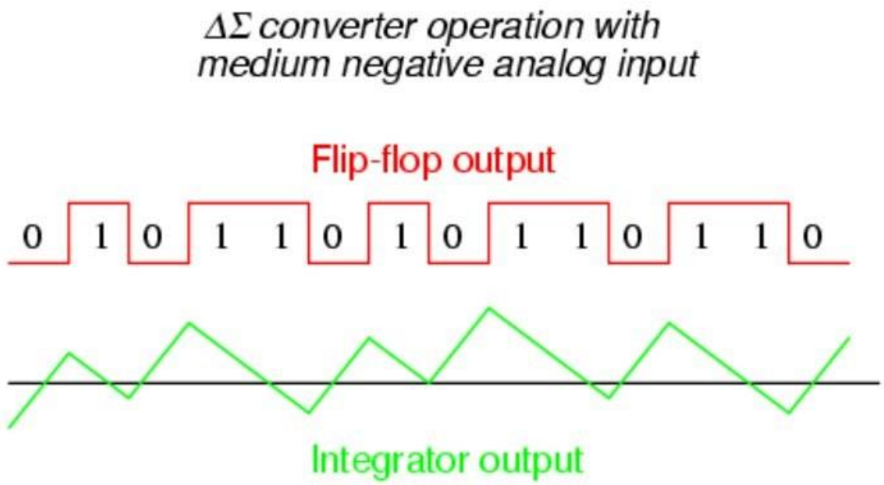

Fig:3.18 Medium negative analog input

As the analog input signal increases in magnitude, so does the occurrence of 1's in the digital output of the flip-flop:

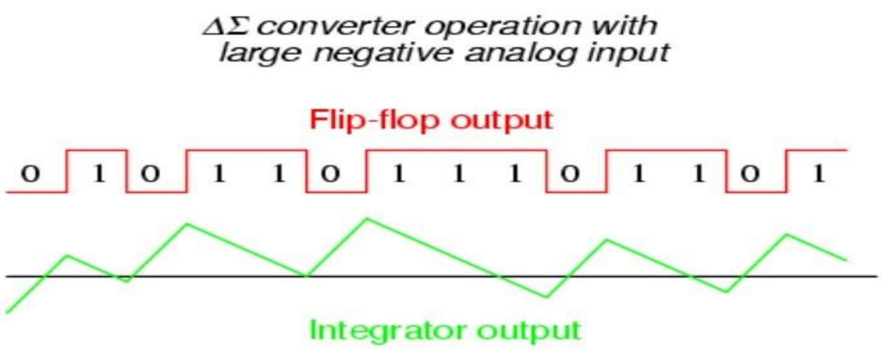

Fig:3.19 Large Negative analog input

A parallel binary number output is obtained from this circuit by averaging the serial stream of bits together. For example, a counter circuit could be designed to collect the total number of 1's output by the flip-flop in a given number of clock pulses. This count would then be indicative of the analog input voltage.

Variations on this theme exist, employing multiple integrator stages and/or comparator circuits outputting more than 1 bit, but one concept common to all ΔΣ converters is that of *oversampling*. Oversampling is when multiple samples of an analog signal are taken by an ADC (in this case, a 1-bit ADC), and those digitized samples are averaged. The end result is an effective increase in the number of bits resolved from the signal. In other words, an oversampled 1-bit ADC can do the same job as an 8-bit ADC with one-time sampling, albeit at a slower rate.

3.7 MEDICAL SAFETY

A medical device is any product used to diagnose, cure, or treat a condition, or to prevent disease. They range from small and simple, like a blood glucose meter, to large and complicated, like a ventilator. You might use one at home or at work, or you may need one in a hospital.

To use medical devices safely

- Know how your device works. Keep instructions close by
- Understand and properly respond to device alarms
- Have a back-up plan and supplies in the event of an emergency
- Keep emergency numbers available and update them as needed
- Educate your family and caregivers about your device

Electromagnetic Interference could interfere with the operation of clinical and other electronic equipment, the risk of adverse clinical effects was not well understood. Thus, it is recommended that hospitals should consider prohibiting the use of transmitting devices (e.g., cellular telephones, citizen-band radios, amateur radio transmitters) by patients and visitors throughout the facility.

The ability of devices that emit radio frequencies to interfere with the operation of electronic equipment has long been recognized. However, recent dramatic, although not necessarily representative, reports in the print and broadcast media have drawn significant attention to the effect that EMI from RF transmitting devices (e.g., cellular telephones, walkie-talkies) can have on medical devices (e.g., apnea monitors, ventilators, powered wheelchairs). These reports frequently do not convey the complexity of the issues to be considered when examining device susceptibility to EMI and the likelihood or significance of any resulting clinical effects. In reality, despite several recent technical conferences on this subject, data that can establish a reliable estimate of risk is not available.

Table3.1: medical safely

Affected Device	**Effect**	**Potential Serious Consequences?**
Monitoring system	Central station alarms and printouts suppressed	Yes—However, no adverse clinical consequence occurred in this incident, which was associated with multiple cellular telephones being used in one room
Monitoring system	Waveform artifact observed	No—Waveform interference occurred, but reverted to normal when transmission ceased
Monitoring system	High blood pressure readings displayed	No—Erroneous data was displayed, but reverted to normal when transmission ceased
ECG telemetry monitoring system	Waveform artifact observed	No—Waveform interference occurred, but reverted to normal when transmission ceased
ECG telemetry monitoring system	Unknown	No
Infusion pump	Alarm sounded (rate also may have changed)	No
Ventilator (operated in a car)	Inappropriate activation of low-pressure alarm	No

Identifying EMI Sources: High- versus Low-Power Technologies

Two prominent, and somewhat controllable, factors that affect the likelihood of electromagnetic interference (EMI) are distance and power. As the distance between an RF transmitting device and susceptible electronic equipment decreases, the likelihood of interference increases. Also, the higher the power of an RF transmitter, the more likely interference is to occur.

Many hospitals, concerned about reports of EMI from cellular telephones and walkie-talkies, have expressed reservations about incorporating lower-power wireless technologies, such as cordless telephones, personal micro-cellular telephones, and wireless LAN-based information systems. Many of these systems operate at or below 100 mW, with an operating frequency of 900 MHz and above. By comparison, conventional cellular telephones may operate at maximum powers of 600 mW or higher, and hand-held transceivers operate at powers of up to several watts.

Identifying Susceptible Devices: Ad hoc Testing

Determining which devices among a hospital's inventory are susceptible to EMI has been a general concern for many people who have contacted ECRI. (For issues related specifically to cardiac pacemakers and cellular telephones, see "Interactions between Cellular Telephones and Cardiac Pacemakers," below.) Although thorough testing to characterize a device's EMI susceptibility could be performed, doing so would require sophisticated equipment and a shielded room or other appropriate space. As an alternative, a subcommittee of the ANSI/IEEE C63 Electromagnetic Compatibility Committee is drafting a set of ad hoc testing procedures to be used to screen devices for susceptibility to transmissions from commonly used hand-held devices.

Strategies to Reduce EMI Risk

Managing EM Fields Within and Outside Hospitals

Hospital administrators and biomedical engineers or clinical engineers should understand radiofrequency and its sources. Hospital administrators should consider developing hospital policies for controlled use of radiofrequency sources and the frequency spectrum. Administrators should also consider integrating radiofrequency sources and systems with their existing communication systems to eliminate any uncontrolled communication sources in hospitals.

Managing Medical Device Immunity

Hospital administrators and clinical and technical staff should be educated about EMI-related risk. It is important for biomedical engineers, clinicians, and nurses to be aware of EMI-related incidents and report those incidents to the correct authority. If any staff member finds any sensitive medical devices susceptible to EMI, they should relocate them or separate them. Biomedical or clinical engineers should determine the immunities of their life-support and critical care medical devices. Due to financial constraints, many hospitals continue to use older medical devices that were not designed or tested for electromagnetic compatibility. Biomedical and clinical engineers should advise hospital administration that all medical devices considered for purchase must at least comply with existing EMC standards.

Recommendations for Implantable Cardiac Devices

The chances that EMI from digital cellular phones would produce life-threatening situations are low. However, data presented here support the recommendations in the Health Protection Branch's Medical Device Alert No. 108 (MDA, 1995) that cautions patients with pacemakers as follows:

1. The use of a digital cellular phone very close to the pacemaker may cause the pacemaker to malfunction.
2. It is advisable to avoid carrying the cellular phone in a breast pocket directly over the pacemaker because an incoming call will switch the phone to its transmission mode and may cause interference.
3. In using a cellular phone, patients should hold it to the ear farthest from the pacemaker.

Wireless Technology Research (WTR) and the Health Industry Manufacturers Association (HIMA) Pacemaker Interest Group recommend 15 cm as an adequate separation distance between a pacemaker and a wireless phone (WTR, 1996). This distance appears to be valid for minimum risk with most pacemaker designs.

Interference with pacemakers by security systems such as EAS systems and WTMDs may pose a potential risk, depending on the model of pacemaker, the programmed mode, and the type of the security system. The greatest risk of interference occurs when the pacemaker is close to the security system and the individual is stationary. This condition is unlikely to be encountered during normal exposure to an EAS system or walk-through metal detector, since pacemaker patients do not remain stationary within these systems.

Interference with pacemakers occurs on both sides of the transmission panel of an EAS system or a walk-through metal detector and at distances as far as 33 cm for UM, 18 cm for SM and 33 cm from WTMDs. Patients may incorrectly assume that EMI effects occur only when they are inside the gate. Data described in this chapter suggest that pacemaker patients should not stop and linger within 33 cm of either side of the transmission panel as measured from the patient's shoulder to the transmission panel of the security system. Instead, they should walk through the system at a normal pace without stopping. Manufacturers should consider placing a sign on the transmission panel to identify the presence of a security system and advise pacemaker and defibrillator patients to walk through the system at a normal pace. In addition, EAS systems should not be installed in the vicinity of cash registers in such a way that pacemaker-dependent patients must stand within 33 cm of the EAS system while completing their transactions.

EMI Awareness

Many EMI malfunctions could be prevented through education. Health Canada places a high priority on providing information to the public. *Medical Device Alert* and *It's Your Health* bulletins are sent to health care professionals to warn them of problems and provide recommendations for management of the associated risks (Tan and Hinberg, 2000b). Health Canada scientists have also published scientific papers on EMI and have given many lectures on the topic at scientific conferences and workshops (Tan et al., 1995; Tan and Hinberg, 1999; Tan and Hinberg, 1995a; Tan and Hinberg, 1995b; Tan and Hinberg, 1995c; Tan and Hinberg, 2000b).

3.8 TRANSIENT VOLTAGE PROTECTION

Transient Voltage Suppressors (TVS's) are devices used to protect vulnerable circuits from electrical overstress such as that caused by electrostatic discharge, inductive load switching and induced lightning. Within the TVS, damaging voltage spikes are limited by clamping or avalanche action of a rugged silicon pn junction which reduces the amplitude of the transient to a nondestructive level.

In a circuit, the TVS should be “invisible” until a transient appears. Electrical parameters such as breakdown voltage (VBR), standby (leakage) current (ID), and capacitance should have no effect on normal circuit performance.

The TVS breakdown voltage is usually 10 % above the reverse standoff voltage (VR), which approximates the circuit operating voltage to limit standby current and to allow for variations in VBR caused by the temperature coefficient of the TVS. When a transient occurs, the TVS clamps instantly to limit the spike voltage to a safe level, called the clamping voltage (VC), while conducting potentially damaging current away from the protected component.

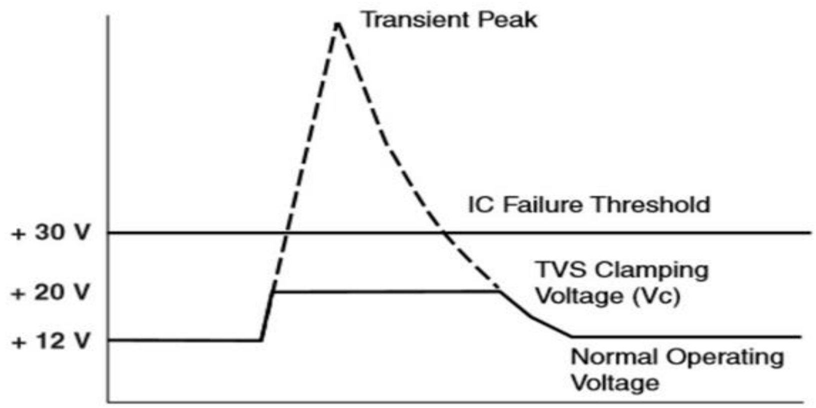

Fig:3.20 Transients of several thousand volts can be clamped a safe level by the TVS

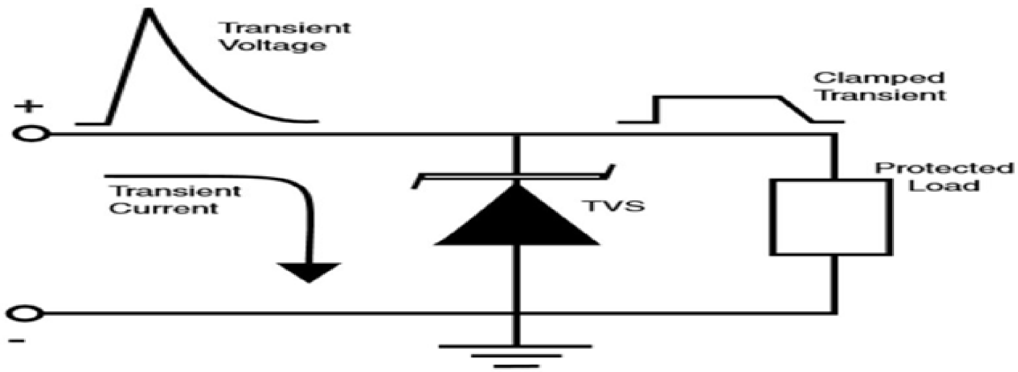

Fig:3.21 Transient current is diverted to ground through TVS

TVS’s are designed, specified and tested for transient voltage protection, while a Zener diode is designed and specified for voltage regulation. For transient protection, the designer’s choice is a TVS.

The surge power and surge current capability of the TVS are proportional to its junction area. Surge ratings for silicon TVS families are normally specified in kilowatts of peak pulse power (PPP) during a given waveform. Early devices were specified with a 10/1000 µs waveform (10 µs rise to peak and 1000 µs exponential decay to one half peak), while more recent product introductions are rated for an 8/20ìs test waveform. Power ratings range from 5 kW for 10/1000 µs, down to 400 W for 8/20 µs. This power is derived from the product of the peak voltage across the TVS and the peak current conducted through the device.

Packaging covers a broad spectrum according the need. Discrete axial leaded components are available in peak pulse power ratings of 400 W, 500 W, 600 W, 1.5 kW and 5 kW. The higher power devices are most frequently used across power buses.

For lower power, high density applications, suppressor arrays are available in both DIP and small outline surface mount configurations. Arrays are normally used across data lines for protecting I/O ports from static discharge. Specialized low capacitance TVSs are available for use in high data rate circuits to prevent signal attenuation.

TVSs have circuit operating voltages available in increments from 5 V up through 376 V for some types. Because of the broad range of voltages and power ratings available, (as well as the universal presence of transient voltages) TVSs are used in a remarkably wide variety of circuits and applications.

References:

1. Norris, A.C. Essentials of Telemedicine and Telecare. Wiley (ISBN 0-471-53151-0), 2002.

2. Wootton, R., Craig, J., Patterson, V. (Eds.), Introduction to Telemedicine. Royal Society of Medicine Press Ltd (ISBN 1853156779), 2006.

Question Bank

1. Explain the operation and working of differential amplifier
2. What is meant by filters and explain the different types of filters
3. Explain in detail about the bridge amplifier
4. How analog value of converted into digital value. Explain with different types of A/D converters
5. How electromagnetic interference can be avoided for the safety of medical instruments
6. How voltage can be protected for medical instruments

SCHOOL OF BIO AND CHEMICAL ENGINEERING
DEPARTMENT OF BIOMEDICAL ENGINEERING

UNIT – 4- MEDICAL ELECTRONICS AND E-HEALTH– SBMA7004

IV TELEMEDICINE AND HEALTH

History and Evolution of telemedicine, Functional diagram of telemedicine system, Tele health, Global and Indian scenario, Ethical and legal aspects of Telemedicine -Confidentiality, Social and legal issues, Safety and regulatory issues, Video conferencing Real -time Telemedicine integrating doctors / Hospitals. Administration of centralized medical data, security and confidentially of medical records and access control, Cyber laws related to telemedicine.

TELEMEDICINE

Telemedicine is defined as the use of telecommunications to provide medical information and services It may be as simple as two health professionals discussing a case over the telephone, or as sophisticated as using satellite technology to broadcast a consultation between providers at facilities in two countries, using videoconferencing equipment or robotic technology. The first is used daily by most health professionals, and the latter is used by the military and some large medical centers.

Telemedicine enables a physician or specialist at one site to deliver health care, diagnose patients, give intra-operative assistance, provide therapy, or consult with another physician or paramedical personnel at a remote site. Telemedicine system consists of customized medical software integrated with computer hardware, along with medical diagnostic instruments connected to the commercial VSAT (Very Small Aperture Terminal) at each location or fibre optics.

Although, telemedicine could potentially affect all medical specialties, the greatest current applications are found in radiology, pathology, cardiology and medical education. Perhaps the greatest impact of telemedicine may be in fulfilling its promise to improve the quality, increase the efficiency, and expand the access of the healthcare delivery system to the rural population and developing countries.

4.1 HISTORY AND EVOLUTION OF TELEMEDICINE

In its early manifestations, African villagers used smoke signals to warn people to stay away from the village in case of serious disease. In the early 1900s, people living in remote areas in Australia used two-way radios, powered by a dynamo driven by a set of bicycle pedals, to communicate with the Royal Flying Doctor Service of Australia.

Shortly after the invention of the telephone, attempts were made to transmit heart and lung sounds to a trained expert who could assess the state of the organs. However, poor transmission systems made the attempts a failure.

1906: ECG Transmission
Einthoven, the father of electrocardiography, first investigated on ECG transmission over telephone lines in 1906! He wrote an article "Le telecardiogramme" at the Archives Internationales Physiologie

1920s: Help for ships
Telemedicine dates back to the 1920s. During this time, radios were used to link physicians standing watch at shore stations to assist ships at sea that had medical emergencies.

1924: The first exposition of Telecare
Perhaps it was the cover showed below of "Radio News" magazine from April 1924. The article even includes a spoof electronic circuit diagram which combined all the gadgets of the day into this latest marvel!

1955: Telepsychiatry
The Nebraska Psychiatric Institute was one of the first facilities in the country to have closed-circuit television in 1955. In 1971 the Nebraska Medical Center was linked with the Omaha Veterans Administration Hospital and VA facilities in two other towns.

1967: Massachusetts General Hospital
This station was established in 1967 to provide occupational health services to airport employees and to deliver emergency care and medical attention to travelers.

1970s: Satellite telemedicine
Via ATS-6 satellites. In these projects, paramedics in remote Alaskan and Canadian villages were linked with hospitals in distant towns or cities.

The Role of NASA in Early Development
The National Aeronautics and Space Administration (NASA) played an important part in the early development of telemedicine NASA's efforts in telemedicine began in the early 1960s when humans began flying in space. Physiological parameters were telemetered from both the spacecraft and the space suits during missions. These early efforts and the enhancement in communications satellites fostered the development of telemedicine and many of the medical devices in the delivery of health care today. NASA provided much of the technology and funding for early telemedicine demonstrations, two of which are the Space Technology Applied to Rural Papago Advanced Health Care (STARPAHC) and the Nebraska Medical Center. There were several pioneering efforts not only in the US, but all over the world.

4.2 FUNCTIONAL DIAGRAM OF TELEMEDICINE SYSTEM

The *figure* shows the system-level block diagram of the medical device. The output signals from the ECG electrodes and PPG sensor connected to the patient's body are amplified and filtered initially by the analog front-end (AFE) of the system. The instrumentation amplifier, transimpedance amplifier, and filters used in the AFE are designed to meet the gain and cutoff frequencies required for ECG and PPG signal processing. The AFE also ensures that the ECG and PPG signals from the high-pass and low-pass filters respectively have sufficient amplitude and low noise levels. The output of the filters is then passed to the processing unit of the system.

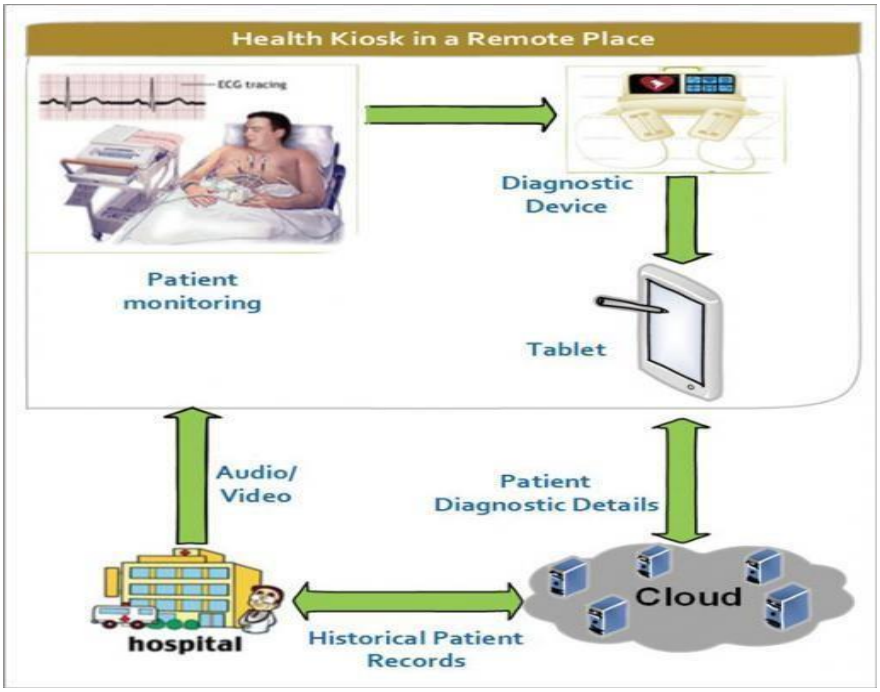

Fig:4.1 Functional Diagram of Telemedicine System

The heart of the processing unit in this example architecture is a Programmable System on Chip (PSoC3), a mixed-signal processing element based on the 8051 microcontroller architecture. PSoC3 consists of a wide variety of analog subsystems such as programmable gain amplifiers, comparators, and analog-to-digital converters, as well as digital subsystems such as timers, pulse width modulators, and various communication protocols including USB, I2C, and SPI, among others. In this design, we have incorporated different analog and digital blocks to process and calculate the patient's health parameters. The processing starts with the input buffer block, which provides proper impedance matching, and then the signal is sampled by a delta sigma ADC. The sampling rate and resolution required for ECG and PPG signals are different, but the beauty of this design lies in the fact that a single ADC is reconfigured to match the different sampling rates and resolutions required. This implementation demonstrates the features of optimum resource utilization and less power consumption. The sampled signal is used to calculate the blood pressure and heart rate of patients.

An unconvential method of measurement is used to calculate the blood pressure of the patient. Blood pressure is measured using pulse transit time (PTT), which is defined as the time taken by a pressure pulse to travel through the length of an arterial tree. PTT is measured as the time interval between the R peak of the ECG signal and the peak of the pressure pulse at the finger. Photoplethysmograph (PPG) gives a volumetric measurement of an organ for use in calculating the pulse transit time (PTT). Studies have shown that blood pressure is directly proportional to PTT. Although some of the studies question the reliability of the blood pressure measurement using PTT, it can still be used to assess blood pressure variation over a continuous period of time.

4.3 TELEHEALTH

Telehealth involves the distribution of health-related services and information. Distribution is via electronic information and telecommunication technologies. It allows long distance patient/clinician contact and care, advice, reminders, education, intervention, monitoring and remote admissions. As well as provider distance-learning; meetings, supervision, and presentations between practitioners; online information and health data management and healthcare system integration. Telehealth could include two clinicians discussing a case over video conference; a robotic surgery occurring through remote access; physical therapy done via digital monitoring instruments, live feed and application combinations; tests being forwarded between facilities for interpretation by a higher specialist; home monitoring through continuous sending of patient health data; client to practitioner online conference; or even videophone interpretation during a consult.

As the population grows and ages, and medical advances are made which prolong life, demands increase on the healthcare system. Healthcare providers are also being asked to do more, with no increase in funding, or are encouraged to move to new models of funding and care such as patient-centred or outcomes based, rather than fee-for-service. Some specific health professions already have a shortage (i.e. Speech-language pathologists). When rural settings, lack of transport, lack of mobility (i.e. In the elderly or disabled), decreased funding or lack of staffing restrict access to care, Telehealth can bridge the gap.

Telehealth requires a strong, reliable broadband connection. “Pipes”, which include wires, cables, microwaves and optic fibre are part of the infrastructure that supports signal transmission, therefore access to one or more of these is crucial to *accessing* and providing telehealth services, as well as continued maintenance of these “pipes”. The better the connection (or “bandwidth quality”), the more data can be sent and received. Historically this has priced providers or patients out of the service, but as time goes on and these parts and resources become cheaper and more accessible through higher consumer demand, the more widespread telehealth can become.

When a healthcare service decides to provide Telehealth to its patients, there are steps to consider, besides just whether the above resources are available. A needs assessment is the best way to start, which includes assessing the access the community currently has to the proposed specialists and care, whether the organisation currently has underutilized equipment which will make them useful to the area they are trying to service, and the hardships they are trying to improve by providing the access to their intended community (i.e. Travel time, costs, time off work). A service then needs to consider potential collaborators. Other services may exist in the area with similar goals that could be joined to provide a more holistic service, and/or they may already have telehealth resources available. The more services involved, the easier to spread the cost of IT, training, workflow changes and improve buy in from clients. Services need to have the patience to wait for the accrued benefits of providing their telehealth service and cannot necessarily expect community wide changes reflected straight away.

Once the need for a Telehealth service is established, delivery can come within four distinct domains. They are live video (synchronous), Store-and-forward (asynchronous), remote patient monitoring, and mobile health. Live video involves a real-time two-way interaction, such as patient/caregiver-provider or provider-provider, over a digital (i.e. broadband) connection. This often is used to substitute a face to face meeting such as consults, and saves time and cost in travel. Store-and-forward is when data is collected, recorded, and then sent on to a provider. For example, a patient's' digital health history file including x-rays and notes, being securely transmitted electronically to evaluate the current case. Remote patient monitoring includes patients' medical and health data being collected and transferred to a provider elsewhere who can continue to monitor the data and any changes that may occur. This may best suit cases that require ongoing care such as rehabilitation, chronic care, or elderly clients trying to stay in the community in their own homes as opposed to a care facility. Mobile health includes any health information, such as education, monitoring and care that is present on and supported by mobile communication devices such as cell phones or tablet computers. This might include an application, or text messaging services like appointment reminders or public health warning systems.

4.4 GLOBAL AND INDIAN SCENARIO

If the country has the dubious distinction of having one doctor for every 15,000 people, lowcost telemedicine model Remote Healthcare Delivery Solutions is set to bridge the great Indian healthcare divide. The rural healthcare scenario is a complex interplay of various parameters that include affordability, availability of healthcare personnel and medicine, infrastructure, social security/insurance and viable, sustainable and scalable business models. With a population of more than 1 billion, of which nearly 72.2 per cent reside in rural areas, the Indian healthcare industry is faced with many challenges while extending its services, particularly to those living in rural and suburban areas of the country. India, with its diverse landmass and huge population, is an ideal setting for telemedicine. Telemedicine activities were started in 1999. The Indian Space Research Organization has been deploying a SATCOM-based telemedicine network across the country since that year. Various government agencies-Department of Information Technology and Ministry of Health & Family Welfare, state governments, premier medical and technical institutions of India-have taken initiatives with the aim to provide quality healthcare facilities to the rural and remote parts of the country. The Government of India has planned and implemented various national-level projects and also extended telemedicine services to South Asian and African countries. Efforts are taking place in the field of medical e-learning by establishing digital medical libraries. Some institutions that are actively involved in telemedicine activities have started curriculum and noncurriculum telemedicine training programs. To support telemedicine activities within the country, the Department of Information Technology has defined the Standards for Telemedicine Systems and the Ministry of Health & Family Welfare has constituted the National Telemedicine Task Force. There are various government and private telemedicine solution providers and a few societies and associations actively engaged to create awareness about telemedicine within the country. With its large medical and IT manpower and expertise in these areas, India holds great promise and has emerged as a leader in the field of telemedicine. Telemedicine has been trumpeted as the great health care hope for rural India, a technology that can transform the health statistics of remote India and medical practice in the country. The advantages of telemedicine are manifold. Taking a doctor to an area where there is no doctor. Taking medical help to patients where no medical help existed before. Diagnosing a medical condition before it becomes untreatable. Tele pathology, teleradiology, tele ophthalmology – these are all ways of accurately diagnosing diseases from a distance. They have moved beyond the pilot stage to actual implementation in different parts of the world, including India. A Telemedicine consultation where a doctor remotely talks to a patient and advises (typically via a video conference link up) has had its share of controversy. If you can't touch a patient, how can you accurately diagnose his condition? Can a patient and doctor who see each other on a TV screen actually bond? Does the patient 'feel' he got a medical consultation? Is the doctor at the other end legally liable for diagnoses delivered via a telemedicine link? Several studies have been done to assess these issues. In a recent survey conducted in the United States, 85% of patients reported being satisfied with their telemedicine consultation. A similar study in Orissa reported a post-consultation satisfaction rate as high as 99%. So clearly, a telemedicine consultation can meet patient expectations. In terms of legal liability, however, this remains a grey area. In fields such as teleradiology, the radiologist who is giving the report based on the images transmitted to him is legally liable. However, in telemedicine consultations, where a doctor does not necessarily have all the clinical data available to him, the legal liability issue is fuzzier. Using newer technologies in the field – telemedicine boxes and software rather than just a video conferencing link – has increased the clinical value of the consultation. The 'tools' of telemedicine e.g. digital stethoscopes and otoscopes, oxygen saturation probes (to assess the oxygen level in the patient), blood pressure monitors etc. have made the telemedicine consultation more scientific and data based. The biggest advantage of telemedicine is that it takes the doctor to places where no doctor has been before. And in a vast country such as ours where large tracts of the country have patients but no doctor, telemedicine truly has the potential to change lives.So why is it that it has not had a huge impact in the country? Granted there have been some hospital groups making an effort in this direction and the Government of India via ISRO has been very keen to roll this out all across the country. Still, there are several reasons why telemedicine has not been able to occur on a giant scale in India:

Infrastructural issues

Infrastructural issues such as poor bandwidth in some areas; expensive bandwidth in others.

Implementation issues

Implementation issues are a major hindrance. In order to implement telemedicine, training is needed at the village end for technicians, IT staff, and local doctors. At the consulting doctor's end, a lot of pushy administration and coordination is required. The devil is in the implementation. If done effectively and consistently, telemedicine can truly be transforming.

Acceptance

For a village doctor and villager, using high end technology may be too inhibiting and radical. However, once the benefits are seen, the acceptance rate will likely be high such as has been seen with mobile telephony and rural internet services.

Viability issues

Viability issues, So far, in India, telemedicine has been largely a free offering by large hospital groups. While part of their CSR, it also has the effect of improving their bed occupancy in cases when a 'tele'patient requires hospitalization and becomes a 'real' inpatient. Large hospitals are in a position to offer these services at no extra charge because they use in-house expertise to deliver them. However, since these in-house specialists work in a busy hospital setting, making time for telemedicine consultation becomes an issue. The fact that the consultations are free also reduces the incentive to make this a high throughput service. The biggest waste in the world is the difference between what you are and what you could be. This epitomizes telemedicine in our country. In my mind, telemedicine in India can be a health innovator and affect real change in the medical scenario of our country….if done well, using multiple hospitals/centers in the country, and on a large scale. In order to do this, the Government, via ISRO connectivity, should connect up all district and village level hospitals to the closest tertiary care centers. The private sector can be used effectively – every private hospital can be connected to one remote site thus distributing the load of patients, rather than a handful of hospitals linking up to all the remote sites. Technologies for telemedicine designed by innovative majors (such as Cisco technologies) should be low cost, easy to install and use and should be able to work on low cost bandwidths.

Public interest campaigns to increase awareness of the benefits of telemedicine would help. A standardized training program for all telemedicine providers and users would be helpful in ensuring the link ups occur rapidly and the centers stay connected without the 'network going down'. A viable model wherein a small cost is paid for the telemedicine consultation would make this a long term successful model of health care delivery. An appointment system that allows patients to book their 'doctor visit' rather than have to wait for a doctor would help improve usage of the system. Also a 'pusher' is needed in every telemedicine center to ensure that after the link up is done; the telemedicine link is actually used on a daily and sustained basis. And finally, data collection on daily utilization, diagnoses made and treatment plans changed would help to measure success and impact and provide direction for the future. After the success of its telemedicine model in India and abroad, healthcare major

Apollo Hospitals is all set to start 'Telemedicine 2.0.'. "Telemedicine 2.0. is a step towards integrating healthcare delivery model with the new age technology. Through this initiative, we are trying to make telehealth a more user-friendly by providing the services on mobile phones and tablets," K Ganapathy, president, Apollo Telemedicine Networking Foundation, and director of Apollo Tele Health Services, told Business Standard."We are also integrating telehealth with the hospital management system, electronic medical records (EMR) of the hospitals, and mobile personal health records to make healthcare more affordable," he said. Apollo is currently pilot launching the 'Telemedicine 2.0.' across five tertiary hospitals in Hyderabad and Chennai. "With the success of the pilot, we will launch this service across Apollo's telemedicine centers in the country," he added. Telemedicine is a process through which patients can consult doctors located at very distant places through electronic mediums without visiting them. "Around 80 per cent from India’s population has no direct physical access to specialist healthcare.

So, we are working on new models with the help of technology to reach out to more people," he said. The telemedicine has a huge market potential. "In India alone, if we cover 10 per cent of the market, it will be more than 40 million consultations from suburban and rural India per year," he said. So far, Apollo has done 75,000 tele-consultations in 25 specialties."Apollo is also conducting a market study on attitudes and behavioral responses of the public towards accepting the mobile phone as an enabler for healthcare by taking 2,500 people from across the country," he said. Apollo Telemedicine, being the oldest multispecialty telemedicine network in South Asia, currently has 125 centers including 15 centers in international markets. It has commissioned three centers in Nigeria last week and signed a memorandum of understanding (MoU) to set up25 centers in Africa. Telemedicine is a potentially miraculous method that promises improvements to healthcare delivery systems, bettering quality and access. Interest in the field has increased dramatically in India. It is not just private healthcare institutions that are investing in creating of new telemedicine solutions, the central and the state governments are also showing interest. The Planning Commission has made numerous suggestions for using telemedicine solutions, during the 12th Five Year Plan period, for improving healthcare services in the remote parts of the country.

If Planning Commission has its way, healthcare practitioners could be using software applications such as Skype for telemedicine. The Planning Commission report says, “Computer with Internet connectivity should be ensured in every primary health centre within this Plan period; sub-centers will have extended connectivity through cell phones, depending on their state of readiness and skill set of their functionaries. The availability of Skype and other similar applications for audio-visual interaction makes telemedicine a near-universal possibility and could be used to ameliorate the professional isolation of health personnel posted in remote and rural areas.” The health ministry has also identified telemedicine as a major thrust area. Only 25 percent of India’s specialist physicians reside in semi-urban areas, and a mere three percent live in rural areas. As a result, rural areas, with a population approaching 700 million, continue to be deprived of proper healthcare facilities. Further the availability of hospital facility is very low in rural areas. Thus, the early successes of telemedicine pioneers have led to increased acceptance and proliferation of telemedicine.”

4.5 LEGAL AND ETHICAL ASPECTS OF TELEMEDICINE

There are many issues of concern regarding the legal and ethical aspects of telemedicine. These include the responsibilities and potential liabilities of the health professional, the duty to maintain the confidentiality and privacy of patient records, and the jurisdictional problems associated with cross-border consultations. There is also the issue of reimbursement for care provided using a telemedicine service. Telemedicine allows the transmission of health information across the borders of nation states. Cross-border telemedicine services have begun, particularly in specialties such as teleradiology, but questions of jurisdiction and registration have yet to be answered definitively. While this may be true of many of the legal and ethical aspects of telemedicine generally, it is also the case that health-care professionals who undertake telemedicine in a prudent manner will minimize the possibility of medico legal complications.

As with conventional healthcare, confidentiality, consent and non-malfeasance are basic principles in telemedicine. There are unforeseen medical and legal issues could arise from increased but inefficient or ineffective use of telemedicine.

Three core issues:

- Responsibilities and potential liabilities of healthcare professionals;
- Duty to maintain confidentiality and privacy of patients' records;
- Jurisdictional problems associated with cross-border consultations.

The ethical principle of beneficence to justify using technologies to increase access to care and reduce costs was discussed then. They argued that an efficient service meant a better service in terms of quality of care, mainly by increasing accessibility by minimizing traditional barriers created by time and location. However, only people with the resources to gain access benefit.

Control of data remains in the health organisation's jurisdiction. This is an advantage when coordinating a multiprofessional team as data can be readily dispersed.

There are potential ambiguities in practitioners' responsibilities, in terms of loyalty to patients or the employer. For example, if staff do not have physical/live contact with patients so are not aware of their holistic needs, this could cause them to focus on investigating the health problem rather than establishing a rapport. Staff could therefore become more committed to their employer than to patients.

Confidentiality

Confidentiality may be problematic in telemedicine. Since patients trust practitioners with personal information, it is reasonable for the onus to fall on professionals to protect the confidentiality of that data. Layman (2003), Bates et al (2001) and Briggs (2001) used the concept of non-maleficence to emphasise professional responsibility, since the legal aspect of confidentiality focuses on the relationships between individuals involved in delivering care rather than on systems used.

While security of technology is vital in safeguarding patients and care standards, individual practitioners should bear ultimate responsibility for protecting patients from emotional, spiritual, social or material harm.

The British Medical Association (2005) provided three principles to guide practice:

- Patients' right to privacy regarding medical details and records;
- Patients' privacy should be maintained unless waived in a meaningful way;
- Disclosure of information should be related to the prevailing medical condition to fulfil the immediate and specific purpose of treatment.

Data security

Rapid implementation of telemedicine, combined with major social change and mobility, means ongoing discussion is needed across international boundaries by governments and professional organizations.

Stanberry (1998) built on Bloom's (1997) foundations to offer a generic overview of legal and ethical issues (Table 1). Jones (1997) added development and security of the electronic patient record (EPR). The health and safety of staff using equipment needs to be considered (Jones, 1997).

Data transmission

Telemedicine relies on transmitting data. This means secure networks and data transmissions are critical to confidentiality and privacy.

These considerations led to a debate between the NHS Information Management Group and the BMA on access to the NHS network (NHSnet). The approach adopted is a code of connection, which sets out minimum conditions that organizations must meet if they wish to gain access to NHSnet (Asadi and Akhlaghi, 2002). The most obvious way of reducing the risk of unauthorised access to computer data across the internet is to control traffic across the interface between the NHS local area network and the external internet.

Technology offers some safeguards in firewalls and encryption protocols. However, firewalls require regular and frequent updating and are effective only against traffic that goes through them. In addition, neither firewalls nor encryption can stop people who misappropriate medical records for malicious reasons and/or economic gain.

The legal issue is not whether electronic systems can provide airtight security, but whether they can protect privacy as well as or better than paper systems. Warner (1998) said that agencies delivering care would need to ensure rigorous ways of protecting patients' electronic records.

Patient privacy

Patient privacy during telehealth consultations should be maintained as much as possible, although it is understandable that privacy might be limited when such technology is used (Mair and Whitten, 2000).

Healthcare professionals should ask patients if they have any questions that might require more privacy than provided. It is important to explain to patients that privacy and confidentiality cannot be guaranteed in telemedicine, as medical records can be shared with other practitioners involved in their care. The nature of the professional-patient relationship changes dramatically, as telemedicine challenges traditional concepts of privacy and confidentiality (Telemedicine Association of Oregon, 2004).

As Asadi and Akhlaghi (2002) pointed out, the legal aspect of confidentiality focuses on the relationships between individuals rather than the systems by which they communicate. In the UK, there are three primary pieces of legislation that are relevant to the legal and ethical aspects of telemedicine:

- Data Protection Act 1984;
- Computer Misuse Act 1990;
- Data Protection Act 1998.

Challenges

Heinzelmann et al (2006) identified several problems encountered by healthcare professionals while using telemedicine. These range from staff discomfort with new technology to those who are concerned that telemedicine threatens healthcare practice. Its future may well therefore depend on human and socioeconomic factors rather than the ability of the technology itself.

Successful integration of telemedicine into existing structures requires organisations to develop policies, procedures, guidelines and strategies to guide and govern professionals and ensure patient and staff safety.

Burmahl (2000) said that effective planning was vital for effective implementation. Hardware – the devices themselves – should be compatible with each other and suitable for their purpose. Although difficult, this is relatively straightforward compared with the culture change telemedicine demands.

Healthcare leaders and managers need to examine and, if appropriate, reconfigure entire systems of work, particularly where custom and practice may not be as efficient as is needed. Individual staff need to examine their role and activities and minimize process duplication and waste.

Most established tools/models for change management are suitable as a structure for implementing telemedicine. Perednia and Allen (1995) highlighted the need to make change elements – such as identifying goals, evaluating effectiveness, accountability, communication and periodical re-evaluation – specific to telemedicine.

Integration of telemedicine into staff development initiatives may prove useful in helping staff to accept it, leading to greater use. Burmahl (2000) said staff training should be a priority to raise awareness and expand the scope of telemedicine.

There has recently been a shift towards health promotion and illness prevention. Heinzelmann et al (2006) argued that healthcare providers are therefore less dependent on skilled and costly staff as part of a multidisciplinary approach to care delivery.
Providers are increasingly moving from financial criteria for evaluation to a more holistic analysis based on performance, in which Heinzelmann et al (2006) saw a key role for telemedicine. The technology offers a mechanism for providing cost-effective, targeted care but, before it is universally accepted, its benefits need to be demonstrated to providers, patient advocate groups and, perhaps above all, patients. Heinzelmann et al (2006) illustrated some of the major challenges facing telemedicine in an environment with increasing emphasis on self care and multidisciplinary care.
Further clinical trials to determine the effectiveness of telemedicine are needed.

Attitudes

Society is becoming increasingly dependent on advanced technology (Liederman and Morefield, 2003) and this is reflected in attitudes towards telemedicine. Adopting telemedicine generates a more open environment where defensive medicine is reduced, care is enhanced and costs are better controlled. These factors present a compelling case for easing licensing restrictions for telemedicine across international borders (Liederman and Morefield, 2003).
Technical challenges are an ever-present issue. To establish effective connectivity and to ensure that the technology works in remote communities, extensive bandwidth resources are required. Bandwidth is the transmission capacity of a system over a period of time. The term usually refers to the speeds of internet services; faster ones have a higher bandwidth than slower ones The view that bandwidth, like computer memory, doubles each year is optimistic. Regular evaluation of telemedicine modalities should be conducted to ensure that bandwith for updates and modifications is adequate.
There is some evidence that telemedicine provides adequate healthcare at a reasonable cost. In some situations the cost-effectiveness of telemedicine appears to be obvious, such as the doc@HOME system that helps patients with chronic obstructive pulmonary disease self manage their condition (Lomas, 2009). However, healthcare providers will want stronger evidence of its indirect economic benefits.

Metrics for telemedicine outcomes should be developed to demonstrate sufficient evidence of socioeconomic benefit to justify ongoing investment. Evaluation should include social, cultural, organizational and policy aspects. It is evident it will decrease the cost per contact between patient and healthcare professional (Field, 1996).

Heinzelmann et al (2006) said that behaviours related to the use of technology are influenced by culture, knowledge, attitudes, beliefs, practices and routines. Telemedicine in the future will be guided by patients' behaviour and perceptions of its applications. This is shown by the following:

- Increased use of the internet for healthcare information;
- Increased demands to access medical services more quickly;
- Growing frustration with current services;
- Greater patient involvement in decision making;
- High levels of patient satisfaction with telemedicine;
- Increased use of the internet and mobile phones (Heinzelmann et al, 2006; Nesbit et al, 2005).

It is important to examine providers' views as these will influence implementation. Providers identified several trends in relation to telemedicine including: anticipated shortages in doctor and nursing workforces; professionals such as HCAs have a larger role; and increased need for communication among various providers (Richards et al, 2005; Hibbert et al, 2004; Wood, 2003).

4.6 PRINCIPLES OF MULTIMEDIA TECHNOLOGY

1. Coherence Principle – People learn better when extraneous words, pictures and sounds are excluded rather than included.

2. Signaling Principle – People learn better when cues that highlight the organization of the essential material are added.

3. Redundancy Principle – People learn better from graphics and narration than from graphics, narration and on-screen text.

4. Spatial Contiguity Principle – People learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen.

5. Temporal Contiguity Principle – People learn better when corresponding words and pictures are presented simultaneously rather than successively.

6. Segmenting Principle – People learn better from a multimedia lesson is presented in user-paced segments rather than as a continuous unit.

7. Pre-training Principle – People learn better from a multimedia lesson when they know the names and characteristics of the main concepts.

8. Modality Principle – People learn better from graphics and narrations than from animation and on-screen text.

9. Multimedia Principle – People learn better from words and pictures than from words alone.

10. Personalization Principle – People learn better from multimedia lessons when words are in conversational style rather than formal style.

11. Voice Principle – People learn better when the narration in multimedia lessons is spoken in a friendly human voice rather than a machine voice.

12. Image Principle – People do not necessarily learn better from a multimedia lesson when the speaker's image is added to the screen.

Data communications and networks

A Computer network is a number if computers interconnected by one or more transmission paths. The transmission path often is the telephone line, due to its convenience and universal preserve.

Data Communication is the exchange of data (in the form of Os and 1s) between two devices via some form of transmission medium (such as a wire cable).

The purpose of data communication is to exchange information between two agents.

Data Communication is considered Local – if the communicating device are in the same building. Remote – if the device are farther apart.

Data: is a representation of facts, concepts and instructions presented in a formalized manner suitable for communication, interpretation or processing by human beings or by automatic means.

Information: is currently assigned to data by means by the conventions applied to those

data. The effectiveness of a data communication system depends on three characteristics.

1. Delivery: The system must deliver data to the correct destination.
2. Accuracy: The system must deliver data accurately.
3. Timeliness: The system must deliver data in a timely manner.

Components of data communication

1. Message – the message is the information to be communicated.

2. Sender – the sender is the device that sends the data message.

3. Receiver – the receiver is the device that receive the message.

4. Medium – the transmission medium is the physical path by which a message travels from sender to receiver.

5. Protocol – A protocol is a set of rules that govern data communication.

Network

A Network is a set of devices (nodes) connected by media links. A node can be a computer, printer, or any other device capable of sending and / or receiving data generated by other nodes on the network.

Advantage of distributed processing

1. Security / Encapsulation

2. Distributed database

3. Faster problem solving

4. Security through redundancy

5. Collaborative processing

Internet:

When two or more networks are connected they become an internetwork or internet.

The most notable internet is called the Internet.

The Internet is a communication system that has brought a wealth of information to out fingertips and organized it for our use.

Internet – Worldwide network.

Local Area Network (LAN) is a network that uses technology designed to span a small geographical area. For e.g. an Ethernet is a LAN technology suitable for use in a single building.

Wide Area Network (WAN) is a network that uses technology designed to span a large geographical area. For e.g. a satellite network is a WAN because a satellite can relay communication across an entire continent. WANs have higher propagation delay than LANs.

LAN, which stands for **local area network**, and **WAN**, which stands for **wide area network**, are two types of networks that allow for interconnectivity between computers. As the naming conventions suggest, LANs are for smaller, more localized networking — in a home, business, school, etc. — while WANs cover larger areas, such as cities, and even allow computers in different nations to connect. LANs are typically faster and more secure than WANs, but WANs enable more widespread connectivity. And while LANs tend to be owned, controlled and managed in-house by the organization where they are deployed, WANs typically require two or more of their constituent LANs to be connected over the public Internet or via a private connection established by a third-party telecommunications provider.

Table:4.1 Comparison chart

LAN versus WAN comparison chart		
	LAN	**WAN**
Stands For	Local Area Network	Wide Area Network
Covers	Local areas only (e.g., homes, offices, schools)	Large geographic areas (e.g., cities, states, nations)

LAN versus WAN comparison chart

	LAN	**WAN**
Definition	LAN (Local Area Network) is a computer network covering a small geographic area, like a home, office, school, or group of buildings.	WAN (Wide Area Network) is a computer network that covers a broad area (e.g., any network whose communications links cross metropolitan, regional, or national boundaries over a long distance).
Speed	High speed (1000 mbps)	Less speed (150 mbps)
Data transfer rates	LANs have a high data transfer rate.	WANs have a lower data transfer rate compared to LANs.
Example	The network in an office building can be a LAN	The Internet is a good example of a WAN
Technology	Tend to use certain connectivity technologies, primarily Ethernet and Token Ring	WANs tend to use technologies like MPLS, ATM, Frame Relay and X.25 for connectivity over longer distances
Connection	One LAN can be connected to other LANs over any distance via telephone lines and radio waves.	Computers connected to a wide-area network are often connected through public networks, such as the telephone system. They can also be connected through leased lines or satellites.
Components	Layer 2 devices like switches and bridges. Layer 1 devices like hubs and repeaters.	Layers 3 devices Routers, Multi-layer Switches and Technology specific devices like ATM or Frame-relay Switches etc.
Fault	LANs tend to have fewer problems associated with them, as there are	WANs tend to be less fault tolerant as they

LAN versus WAN comparison chart

	LAN	**WAN**
Tolerance	smaller number of systems to deal with.	consist of large number of systems.
Data Transmission Error	Experiences fewer data transmission errors	Experiences more data transmission errors as compared to LAN
Ownership	Typically owned, controlled, and managed by a single person or organization.	WANs (like the Internet) are not owned by any one organization but rather exist under collective or distributed ownership and management over long distances.
Set-up costs	If there is a need to set-up a couple of extra devices on the network, it is not very expensive to do that.	For WANs since networks in remote areas have to be connected the set-up costs are higher. However WANs using public networks can be setup very cheaply using just software (VPN etc).
Geographical Spread	Have a small geographical range and do not need any leased telecommunication lines	Have a large geographical range generally spreading across boundaries and need leased telecommunication lines
Maintenance costs	Because it covers a relatively small geographical area, LAN is easier to maintain at relatively low costs.	Maintaining WAN is difficult because of its wider geographical coverage and higher maintenance costs.
Bandwidth	High bandwidth is available for transmission.	Low bandwidth is available for transmission.
Congestion	Less congestion	More congestion

Communications satellite

A **communications satellite** is an artificial satellite that relays and amplifies radio telecommunications signals via a transponder; it creates a communication channel between a source transmitter and a receiver at different locations on Earth. Communications satellites are used for television, telephone, radio, internet, and military applications. There are over 2,000 communications satellites in Earth's orbit, used by both private and government organizations.

Wireless communication uses electromagnetic waves to carry signals. These waves require line-of-sight, and are thus obstructed by the curvature of the Earth. The purpose of communications satellites is to relay the signal around the curve of the Earth allowing communication between widely separated points. Communications satellites use a wide range of radio and microwave frequencies. To avoid signal interference, international organizations have regulations for which frequency ranges or "bands" certain organizations are allowed to use. This allocation of bands minimizes the risk of signal interference.

In satellite communication, signal transferring between the sender and receiver is done with the help of satellite. In this process, the signal which is basically a beam of modulated microwaves is sent towards the satellite. Then the satellite amplifies the signal and sent it back to the receiver's antenna present on the earth's surface. So, all the signal transferring is happening in space. Thus this type of communication is known as space communication.

Two satellites which are commonly used in satellite communication are Active and passive satellites.

Passive satellites: It is just a plastic balloon having a metal coated over it. This sphere reflects the coming microwave signals coming from one part of the earth to other part. This is also known as passive sphere. Our earth also has a passive satellite i.e. moon.

Active satellites: It basically does the work of amplifying the microwave signals coming. In active satellites an antenna system, transmitter, power supply and a receiver is used. These satellites are also called as transponders. The transmitters fitted on the earth generate the microwaves. These rays are received by the transponders attached to the satellite. Then after amplifying, these signals are transmitted back to earth. This sending can be done at the same time or after some delay. These amplified signals are stored in the memory of the satellites, when earth properly faces the satellite. Then the satellite starts sending the signals to earth. Some active satellites also have programming and recording features. Then these recording can be easily played and watched. The first active satellite was launched by Russia in 1957. The signals coming from the satellite when reach the earth, are of very low intensity. Their amplification is done by the receivers themselves. After amplification these become available for future use.

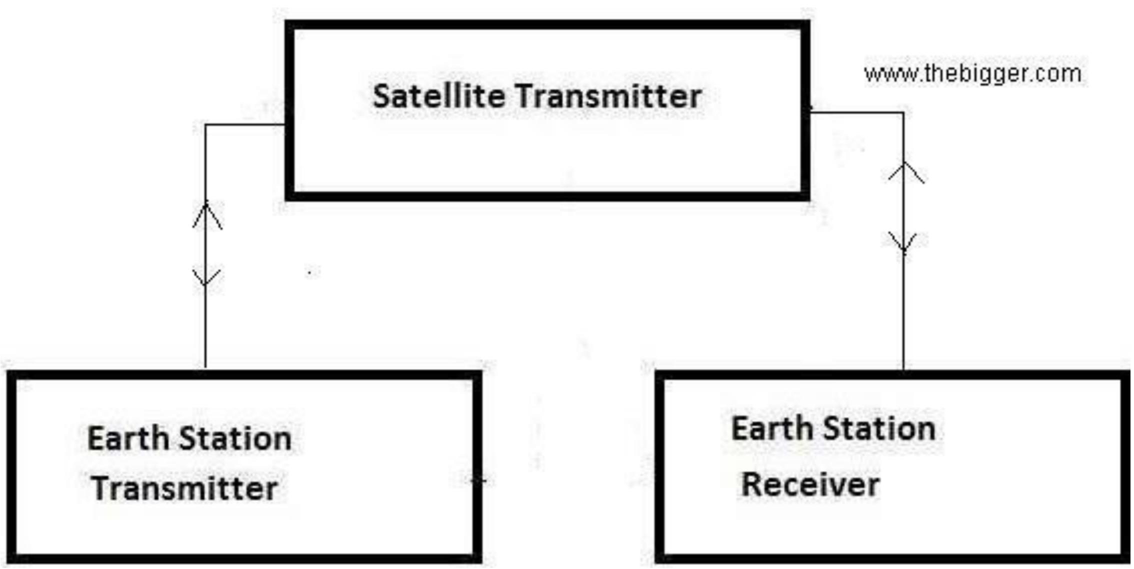

Fig4.2 Active Satellites

Microwave communication is possible only if the position of satellite becomes stationary with respect to the position of earth. So, these types of satellites are known as **geostationary satellites.**

GSM satellite - Global System for Mobile Communications

The GSM system is the most widely used cellular technology in use in the world today. It has been a particularly successful cellular phone technology for a variety of reasons including the ability to roam worldwide with the certainty of being able to be able to operate on GSM networks in exactly the same way - provided billing agreements are in place.

The GSM system was designed as a second generation (2G) cellular phone technology. One of the basic aims was to provide a system that would enable greater capacity to be achieved than the previous first generation analogue systems. GSM achieved this by using a digital TDMA (time division multiple access approach). By adopting this technique more users could be accommodated within the available bandwidth. In addition to this, ciphering of the digitally encoded speech was adopted to retain privacy. Using the earlier analogue cellular technologies it was possible for anyone with a scanner receiver to listen to calls and a number of famous personalities had been "eavesdropped" with embarrassing consequences.

GSM services

Speech or voice calls are obviously the primary function for the GSM cellular system. To achieve this the speech is digitally encoded and later decoded using a vocoder. A variety of vocoders are available for use, being aimed at different scenarios.

In addition to the voice services, GSM cellular technology supports a variety of other data services. Although their performance is nowhere near the level of those provided by 3G, they are nevertheless still important and useful. A variety of data services are supported with user data rates up to 9.6 kbps. Services including Group 3 facsimile, videotext and teletex can be supported.

One service that has grown enormously is the short message service. Developed as part of the GSM specification, it has also been incorporated into other cellular technologies. It can be thought of as being similar to the paging service but is far more comprehensive allowing bi-directional messaging, store and forward delivery, and it also allows alphanumeric messages of a reasonable length. This service has become particularly popular, initially with the young as it provided a simple, low fixed cost.

GSM basics

The GSM cellular technology had a number of design aims when the development started:

- It should offer good subjective speech quality
- It should have a low phone or terminal cost
- Terminals should be able to be handheld

- The system should support international roaming
- It should offer good spectral efficiency
- The system should offer ISDN compatibility

The resulting GSM cellular technology that was developed provided for all of these. The overall system definition for GSM describes not only the air interface but also the network or infrastructure technology. By adopting this approach it is possible to define the operation of the whole network to enable international roaming as well as enabling network elements from different manufacturers to operate alongside each other, although this last feature is not completely true, especially with older items.

GSM cellular technology uses 200 kHz RF channels. These are time division multiplexed to enable up to eight users to access each carrier. In this way it is a TDMA / FDMA system.

The base transceiver stations (BTS) are organised into small groups, controlled by a base station controller (BSC) which is typically co-located with one of the BTSs. The BSC with its associated BTSs is termed the base station subsystem (BSS).

Further into the core network is the main switching area. This is known as the mobile switching centre (MSC). Associated with it is the location registers, namely the home location register (HLR) and the visitor location register (VLR) which track the location of mobiles and enable calls to be routed to them. Additionally there is the Authentication Centre (AuC), and the Equipment Identify Register (EIR) that are used in authenticating the mobile before it is allowed onto the network and for billing. The operation of these are explained in the following pages.

Last but not least is the mobile itself. Often termed the ME or mobile equipment, this is the item that the end user sees. One important feature that was first implemented on GSM was the use of a Subscriber Identity Module. This card carried with it the users identity and other information to allow the user to upgrade a phone very easily, while retaining the same identity on the network. It was also used to store other information such as "phone book" and other items. This item alone has allowed people to change phones very easily, and this has fuelled the phone manufacturing industry and enabled new phones with additional features to be launched. This has allowed mobile operators to increase their average revenue per user (ARPU) by ensuring that users are able to access any new features that may be launched on the network requiring more sophisticated phones.

GSM system overview

The table below summarises the main points of the GSM system specification, showing some of the highlight features of technical interest.

Table 4.2: Specification summary for GSM cellular system.

SPECIFICATION SUMMARY FOR GSM CELLULAR SYSTEM	
Multiple access technology	FDMA / TDMA
Duplex technique	FDD
Uplink frequency band	890 - 915 MHz (basic 900 MHz band only)
Downlink frequency band	933 -960 MHz (basic 900 MHz band only)
Channel spacing	200 kHz
Modulation	GMSK
Speech coding	Various - original was RPE-LTP/13
Speech channels per RF channel	8
Channel data rate	270.833 kbps

Frame duration	4.615 ms

Further developments of GSM

GSM was a particularly successful mobile telecommunications system. Initially it had been intended for use within Europe, but within a relatively short while the system was being used well beyond the borders of Europe, becoming an internationally accepted system.

In addition to its success as a voice communications system, it was developed beyond the basic voice capability to be able to carry data. With the Internet becoming more widely used, GSM was developed to provide a packet data capability. The first major development was in the form of GPRS.

4.7 VIDEO AND AUDIO CONFERENCING

This is a very broad category of online tools, incorporating a range of options from free one-to-one audio conferencing all the way to more sophisticated and expensive tools such as Polycom which allow multiple sites with entire classes participating using video and audio.

1. Video and audio, or just audio connection between two computers communicating via the Internet.
    - Examples of free audio conferencing software: Gizmo, Skype (both cross platform) both enable users to speak to other Gizmo/Skype users free of charge (although users can also pay a fee and make calls to landlines using the computer). For further examples view Wikipedia list.
    - Examples of free video conferencing software: iVisit (cross platform), iChat (Mac only), NetMeeting (Windows only).
    - Breeze can also be used for video conferencing (but Breeze is more than just a video/audio conferencing tool.)
2. Transmitted to & received from any computer in any location that has Internet connection (broadband desirable for effective use). Teacher must have microphone, can have camera.

Ideally end users have microphone (camera not essential) for synchronous communication.

3. Technology requirements for video/audio conferencing:
    - Computer with access (ideally broadband) to the Internet.
    - Browser.
    - Speakers to hear audio.
    - Microphone (to contribute audio).
    - Web camera to contribute video.

Use video/audio conferencing

Enables teacher or limited number of learners need to connect from different locations at the same time when only video and/or audio connection is needed. Examples: guest speaker at remote location can talk to local class; students in one class can engage in discussion with students at another location (such as a class in another country); when a student is unable to attend face-to-face class, s/he can connect to class via VOIP; students can take virtual field trips to remote locations.

Advantages video/audio conferencing

1. Free download of easy to use software that can be used via Internet to connect student, instructor, or guest speaker to class and enable both sides to see & hear.
2. Enables individual (usually limited to one connection) to participate in synchronous learning experiences from any location worldwide. Users can connect from home, work or other location easily accessible to them.
3. Specifically useful for guest speaker who is far away from face-to-face class location, or student who cannot be in face-to-face class.
4. Enables students to take virtual field trips to remote locations (either just by viewing the video (e.g. African Voices) or engaging in an interactive lesson. This is especially an advantage to students who attend schools in isolated communities, but is an advantage to all students regardless of location or socio-economic factors.
5. Can be used to record vodcast or podcast and uploaded to course website (for web-enhanced, hybrid, or fully online classes).

6. As video conferencing technology improves, this can become a far cheaper alternative to ITV or systems like PolyCom in enabling school districts to offer specialized subjects by having one teacher teach a class to students at a number of schools.

Disadvantages of video/audio conferencing

1. Typically on free systems only one or a very limited number of users can connect to the host (instructor) computers at a time, so video/audio conferencing can be used only for individual access rather than as a larger scale tool and learning environment. However, newer systems such as Camfrog enable multi-user video conferencing.
2. Depending on the stability of the connection, users may be disconnected during the class and have to reconnect.
3. Difficult to see and hear people who are not close to the microphone/camera, especially when using one of the low-cost systems (such as iSight camera). Thus difficult to have multiple people at one site, sharing a computer to communicate with users at other sites.

Issues & problems related to video/audio conferencing

1. Works best with broad band connection, especially for video conferencing. Users report that after the initial fascination with the video component has worn off, they realize that they really only need to use the audio as this uses less bandwidth and results in higher quality audio than the video option. Because of the small video window and low quality, the video image is of limited use (as compared with ITV where image is high quality and can be used to share a variety of still and video images).
2. If used for users connecting to face-to-face class, it is important to have good quality speakers so that classroom-based students can hear the person who is calling in.
3. Students speaking from classroom must identify who they are before speaking.
4. Requires students connecting to class from remote site to be able to follow discussion relying just on audio or audio with low quality video (students report this being challenging, especially for long lectures). This is also an accessibility issue.

Emerging issues and tips

1. More often than not, once users at each site know what other users look like, more often than not it is not necessary to use video -- the audio connection is sufficient (since the video quality isn't high and you can usually only see the person's face). The video is really only necessary when users want to demonstrate something or show something to remote users.
2. If a student is connecting to a class via audio connection, handouts and visual aids can be sent to him/her via email or made available on course website ahead of time. If the instructor writes on the board or there is some other visual or interaction that happens in class, the instructor can take a digital photo or digital video and upload this to the course website. This is a benefit not only to the distant student, but also to other students who may find this useful for review.

Data security

Data security means protecting data, such as a database, from destructive forces and from the unwanted actions of unauthorized users.

The continuing evolution of computer systems has given rise to a predominantly, technologically driven work environment. The transition from physical, paper-based procedures to online processes has modified the way in which we as a society, approach the issue of data security – including secure file transfer.

Organizations that process and store what would be classed as 'Personal Information' (e.g. financial status, race, sexual orientation, health etc.) are now subject to stringent governmental controls, legislative laws and standards. If these companies choose to not comply with applicable laws, not only will they have to tackle loss of customer trust when data is leaked, severe financial penalties will be incurred.

Standardization is the process of implementing and developing technical standards based on the consensus of different parties that include firms, users, interest groups, standards organizations and governments. Standardization can help to maximize compatibility, interoperability, safety, repeatability, or quality. It can also facilitate commoditization of formerly custom processes. In social sciences, including economics, the idea of *standardization* is close to the solution for a coordination problem, a situation in which all parties can realize mutual gains, but only by making mutually consistent decisions. This view includes the case of "spontaneous standardization processes", to produce de facto standards.

Encryption

The translation of data into a secret code. Encryption is the most effective way to achieve data security. To read an encrypted file, you must have access to a secret key or password that enables you to *decrypt* it. Unencrypted data is called *plain text*; encrypted data is referred to as *cipher text.*

There are two main types of encryption: asymmetric encryption (also called public-key encryption) and symmetric encryption.

In cryptography, **encryption** is the process of encoding messages or information in such a way that only authorized parties can read it. Encryption does not of itself prevent interception, but denies the message content to the interceptor. In an encryption scheme, the intended communication information or message, referred to as plaintext, is encrypted using an encryption algorithm, generating ciphertext that can only be read if decrypted. For technical reasons, an encryption scheme usually uses a pseudo-random encryption key generated by an algorithm. It is in principle possible to decrypt the message without possessing the key, but, for a well-designed encryption scheme, large computational resources and skill are required. An authorized recipient can easily decrypt the message with the key provided by the originator to recipients, but not to unauthorized interceptors.

The purpose of encryption is to ensure that only somebody who is authorized to access data (e.g. a text message or a file), will be able to read it, using the decryption key. Somebody who is not authorized can be excluded, because he or she does not have the required key, without which it is impossible to read the encrypted information.

Encryption has long been used by military and governments to facilitate secret communication. It is now commonly used in protecting information within many kinds of civilian systems. For example, the Computer Security Institute reported that in 2007, 71% of companies surveyed utilized encryption for some of their data in transit, and 53% utilized encryption for some of their data in storage. Encryption can be used to protect data "at rest", such as information stored on computers and storage devices (e.g. USB flash drives). In recent years there have been numerous reports of confidential data such as customers' personal records being exposed through loss or theft of laptops or backup drives. Encrypting such files at rest helps protect them should physical security measures fail. Digital rights management systems, which prevent unauthorized use or reproduction of copyrighted material and protect software against reverse engineering (see also copy protection), is another somewhat different example of using encryption on data at rest.

Encryption is also used to protect data in transit, for example data being transferred via networks (e.g. the Internet, e-commerce), mobile telephones, wireless microphones, wireless intercom systems, Bluetooth devices and bank automatic teller machines. There have been numerous reports of data in transit being intercepted in recent years. Data should also be encrypted when transmitted across networks in order to protect against eavesdropping of network traffic by unauthorized users.

Data Encryption Standard:

The Data Encryption Standard is a symmetric-key algorithm for the encryption of electronic data. Although now considered insecure, it was highly influential in the advancement of modern cryptography.

Developed in the early 1970s at IBM and based on an earlier design by Horst Feistel, the algorithm was submitted to the National Bureau of Standards (NBS) following the agency's invitation to propose a candidate for the protection of sensitive, unclassified electronic government data. In 1976, after consultation with the National Security Agency (NSA), the NBS eventually selected a slightly modified version (strengthened against differential cryptanalysis, but weakened against brute force attacks), which was published as an official Federal Information Processing Standard (FIPS) for the United States in 1977.

The publication of an NSA-approved encryption standard simultaneously resulted in its quick international adoption and widespread academic scrutiny. Controversies arose out of classified design elements, a relatively short key length of the symmetric-key block cipher design, and the involvement of the NSA, nourishing suspicions about a backdoor. The intense academic scrutiny the algorithm received over time led to the modern understanding of block ciphers and their cryptanalysis.

DES is now considered to be insecure for many applications. This is mainly due to the 56-bit key size being too small; in January 1999, distributed.net and the Electronic Frontier Foundation collaborated to publicly break a DES key in 22 hours and 15 minutes (see chronology). There are also some analytical results which demonstrate theoretical weaknesses in the cipher, although they are infeasible to mount in practice. The algorithm is believed to be practically secure in the form of Triple DES, although there are theoretical attacks. In recent years, the cipher has been superseded by the Advanced Encryption Standard (AES). Furthermore, DES has been withdrawn as a standard by the National Institute of Standards and Technology (formerly the National Bureau of Standards).

Some documentation makes a distinction between DES as a standard and DES as an algorithm, referring to the algorithm as the DEA (Data Encryption Algorithm).

Cryptography

Cryptography or cryptology is the practice and study of techniques for secure communication in the presence of third parties called adversaries. More generally, cryptography is about constructing and analyzing protocols that prevent third parties or the public from reading private messages; various aspects in information security such as data confidentiality, data integrity, authentication, and non-repudiation are central to modern cryptography. Modern cryptography exists at the intersection of the disciplines of mathematics, computer science, and electrical engineering. Applications of cryptography include ATM cards, computer passwords, and electronic commerce.

Cryptography prior to the modern age was effectively synonymous with encryption, the conversion of information from a readable state to apparent nonsense. The originator of an encrypted message (Alice) shared the decoding technique needed to recover the original information only with intended recipients (Bob), thereby precluding unwanted persons (Eve) from doing the same. The cryptography literature often uses Alice ("A") for the sender, Bob ("B") for the intended recipient, and Eve ("eavesdropper") for the adversary. Since the development of rotor cipher machines in World War I and the advent of computers in World War II, the methods used to carry out cryptology have become increasingly complex and its application more widespread.

Modern cryptography is heavily based on mathematical theory and computer science practice; cryptographic algorithms are designed around computational hardness assumptions, making such algorithms hard to break in practice by any adversary. It is theoretically possible to break such a system, but it is infeasible to do so by any known practical means. These schemes are therefore termed computationally secure; theoretical advances, e.g., improvements in integer factorization algorithms, and faster computing technology require these solutions to be continually adapted. There exist information-theoretically secure schemes that provably cannot be broken even with unlimited computing power—an example is the one-time pad—but these schemes are more difficult to implement than the best theoretically breakable but computationally secure mechanisms.

The growth of cryptographic technology has raised a number of legal issues in the information age. Cryptography's potential for use as a tool for espionage and sedition has led many governments to classify it as a weapon and to limit or even prohibit its use and export. In some jurisdictions where the use of cryptography is legal, laws permit investigators to compel the disclosure of encryption keys for documents relevant to an investigation. Cryptography also plays a major role in digital rights management and copyright infringement of digital media.

PROTOCOLS:

The TCP/IP Reference Model

TCP/IP means Transmission Control Protocol and Internet Protocol. It is the network model used in the current Internet architecture as well. Protocols are set of rules which govern every possible communication over a network. These protocols describe the movement of data between the source and destination or the internet. These protocols offer simple naming and addressing schemes.

Diagram Representation of TCP/IP Model

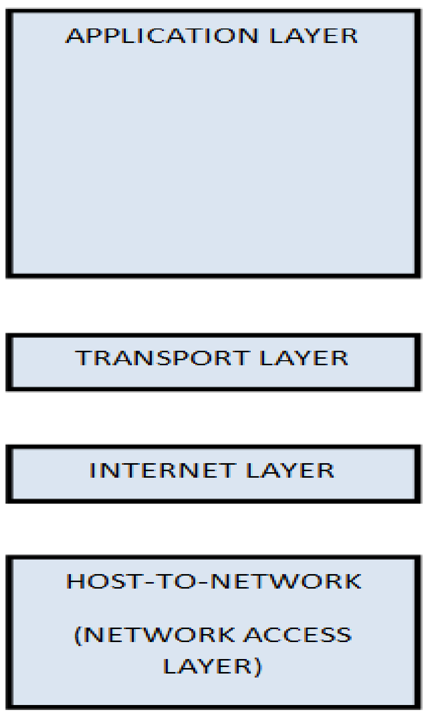

Fig4.3 TCP/IP Model

Overview of TCP/IP reference model

TCP/IP that is Transmission Control Protocol and Internet Protocol was developed by Department of Defence's Project Research Agency (ARPA, later DARPA) as a part of a research project of network interconnection to connect remote machines.

The features that stood out during the research, which led to making the TCP/IP reference model were:

- Support for a flexible architecture. Adding more machines to a network was easy.
- The network was robust, and connections remained intact untill the source and destination machines were functioning.
- The overall idea was to allow one application on one computer to talk to(send data packets) another application running on different computer.
- Description of different TCP/IP protocols

Layer 1: Host-to-network Layer

- Lowest layer of the all.
- Protocol is used to connect to the host, so that the packets can be sent over it.
- Varies from host to host and network to network.

Layer 2: Internet layer

- Selection of a packet switching network which is based on a connectionless internetwork layer is called a internet layer.
- It is the layer which holds the whole architecture together.
- It helps the packet to travel independently to the destination.
- Order in which packets are received is different from the way they are sent.
- IP (Internet Protocol) is used in this layer.

Layer 3: Transport Layer

- It decides if data transmission should be on parallel path or single path.
- Functions such as multiplexing, segmenting or splitting on the data is done by transport layer.
- The applications can read and write to the transport layer.
- Transport layer adds header information to the data.
- Transport layer breaks the message (data) into small units so that they are handled more efficiently by the network layer.
- Transport layer also arrange the packets to be sent, in sequence.

Layer 4: Application Layer

- The TCP/IP specifications described a lot of applications that were at the top of the protocol stack. Some of them were TELNET, FTP, SMTP, DNS etc.
- TELNET is a two-way communication protocol which allows connecting to a remote machine and run applications on it.
- FTP(File Transfer Protocol) is a protocol, that allows File transfer amongst computer users connected over a network. It is reliable, simple and efficient.
- SMTP(Simple Mail Transport Protocol) is a protocol, which is used to transport electronic mail between a source and destination, directed via a route.
- DNS(Domain Name Server) resolves an IP address into a textual address for Hosts connected over a network.

Merits of TCP/IP model

- It operated independently.
- It is scalable.
- Client/server architecture.
- Supports a number of routing protocols.
- Can be used to establish a connection between two computers.

Demerits of TCP/IP

- In this, the transport layer does not guarantee delivery of packets.
- The model cannot be used in any other application.
- Replacing protocol is not easy.
- It has not clearly separated its services, interfaces and protocols.

ISO/OSI Model in Communication Networks

There are n numbers of users who use computer network and are located over the world. So to ensure, national and worldwide data communication, systems must be developed which are compatible to communicate with each other. ISO has developed this. ISO stands for International organization of Standardization. This is called a model for Open System Interconnection (OSI) and is commonly known as OSI model.

The ISO-OSI model is a seven layer architecture. It defines seven layers or levels in a complete communication system.

Diagram of ISO-OSI Model

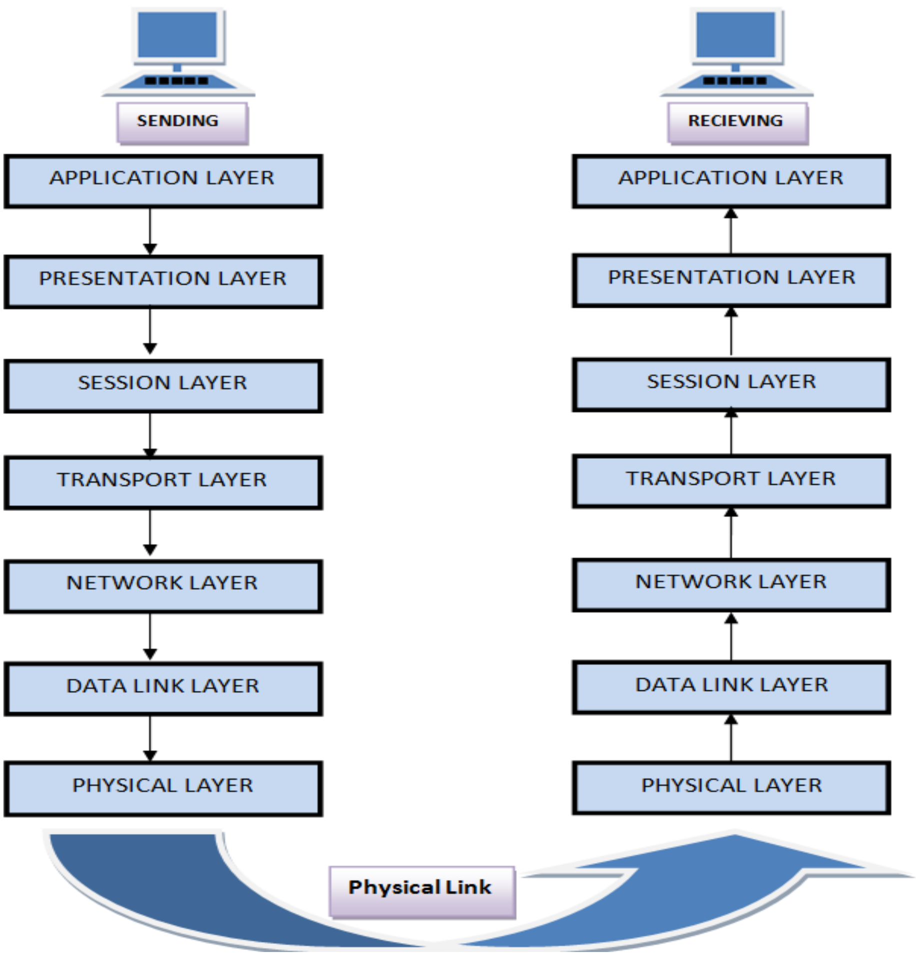

Fig4.4 ISO-OSI Model

Feature of OSI Model :

- Big picture of communication over network is understandable through this OSI model.
- We see how hardware and software work together.
- We can understand new technologies as they are developed.
- Troubleshooting is easier by separate networks.
- Can be used to compare basic functional relationships on different networks.

Functions of Different Layers :

Layer 1: The Physical Layer :

- It is the lowest layer of the OSI Model.
- It activates, maintains and deactivates the physical connection.
- It is responsible for transmission and reception of the unstructured raw data over network.
- Voltages and data rates needed for transmission is defined in the physical layer.
- It converts the digital/analog bits into electrical signal or optical signals.
- Data encoding is also done in this layer.

Layer 2: Data Link Layer :

- Data link layer synchronizes the information which is to be transmitted over the physical layer.
- The main function of this layer is to make sure data transfer is error free from one node to another, over the physical layer.
- Transmitting and receiving data frames sequentially is managed by this layer.
- This layer sends and expects acknowledgements for frames received and sent respectively. Resending of non-acknowledgement received frames is also handled by this layer.
- This layer establishes a logical layer between two nodes and also manages the Frame traffic control over the network. It signals the transmitting node to stop, when the frame buffers are full.

Layer 3: The Network Layer :

- It routes the signal through different channels from one node to other.
- It acts as a network controller. It manages the Subnet traffic.
- It decides by which route data should take.
- It divides the outgoing messages into packets and assembles the incoming packets into messages for higher levels.

- Layer 4: Transport Layer :
- It decides if data transmission should be on parallel path or single path.
- Functions such as Multiplexing, Segmenting or Splitting on the data are done by this layer
- It receives messages from the Session layer above it, convert the message into smaller units and passes it on to the Network layer.
- Transport layer can be very complex, depending upon the network requirements.
- Transport layer breaks the message (data) into small units so that they are handled more efficiently by the network layer.

Layer 5: The Session Layer:

- Session layer manages and synchronize the conversation between two different applications.
- Transfer of data from source to destination session layer streams of data are marked and are resynchronized properly, so that the ends of the messages are not cut prematurely and data loss is avoided.

Layer 6: The Presentation Layer :

- Presentation layer takes care that the data is sent in such a way that the receiver will understand the information (data) and will be able to use the data.
- While receiving the data, presentation layer transforms the data to be ready for the application layer.
- Languages(syntax) can be different of the two communicating systems. Under this condition presentation layer plays a role of translator.
- It performs Data compression, Data encryption, Data conversion etc.

Layer 7: Application Layer :

- It is the topmost layer.

- Transferring of files disturbing the results to the user is also done in this layer. Mail services, directory services, network resource etc are services provided by application layer.
- This layer mainly holds application programs to act upon the received and to be sent data.

Merits of OSI reference model:

- OSI model distinguishes well between the services, interfaces and protocols.
- Protocols of OSI model are very well hidden.
- Protocols can be replaced by new protocols as technology changes.
- Supports connection oriented services as well as connectionless service.

Demerits of OSI reference model:

- Model was devised before the invention of protocols.
- Fitting of protocols is tedious task.
- It is just used as a reference model.

Table 4.3: OSI and TCP comparison

OSI(Open System Interconnection)	**TCP/IP(Transmission Control Protocol / Internet Protocol)**
1. OSI is a generic, protocol independent standard, acting as a communication gateway between the network and end user.	1. TCP/IP model is based on standard protocols around which the Internet has developed. It is a communication protocol, which allows connection of hosts over a network.
2. In OSI model the transport layer	2. In TCP/IP model the transport layer does not guarantees delivery of

guarantees the delivery of packets.	packets. Still the TCP/IP model is more reliable.
3. Follows vertical approach.	3. Follows horizontal approach.
4. OSI model has a separate Presentation layer and Session layer.	4. TCP/IP does not have a separate Presentation layer or Session layer.
5. OSI is a reference model around which the networks are built. Generally it is used as a guidance tool.	5. TCP/IP model is, in a way implementation of the OSI model.
6. Network layer of OSI model provides both connection oriented and connectionless service.	6. The Network layer in TCP/IP model provides connectionless service.
7. OSI model has a problem of fitting the protocols into the model.	7. TCP/IP model does not fit any protocol
8. Protocols are hidden in OSI model and are easily replaced as the technology changes.	8. In TCP/IP replacing protocol is not easy.

9. OSI model defines services, interfaces and protocols very clearly and makes clear distinction between them. It is protocol independent.	9. In TCP/IP, services, interfaces and protocols are not clearly separated. It is also protocol dependent.
10. It has 7 layers	10. It has 4 layers

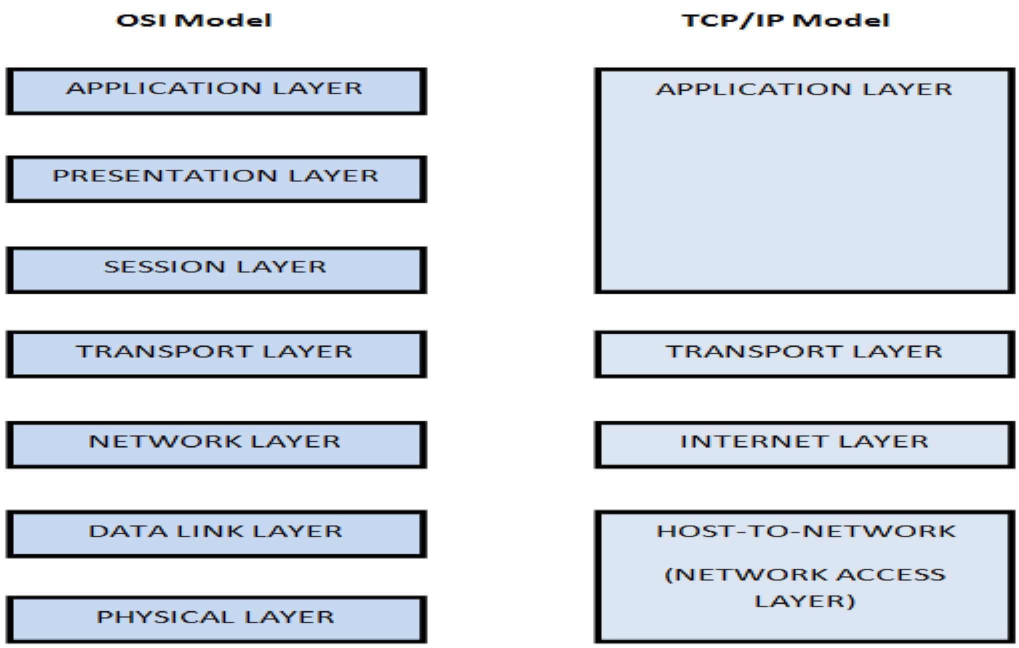

Fig4.5 Diagrammatic Comparison Of OSI And TCP/IP Model

DICOM

Digital Imaging and Communications in Medicine (DICOM) is a standard for handling, storing, printing, and transmitting information in medical imaging. It includes a file format definition and a network communications protocol. The communication protocol is an application protocol that uses TCP/IP to communicate between systems. DICOM files can be exchanged between two entities that are capable of receiving image and patient data in DICOM format. The National Electrical Manufacturers Association (NEMA) holds the copyright to this standard. It was developed by the DICOM Standards Committee, whose members are also partly members of NEMA.

DICOM enables the integration of medical imaging devices – like scanners, servers, workstations, printers, network hardware, and picture archiving and communication systems (PACS) – from multiple manufacturers. The different devices come with DICOM Conformance Statements which clearly state which DICOM classes they support. DICOM has been widely adopted by hospitals and is making inroads in smaller applications like dentists' and doctors' offices.

.

DICOM is known as NEMA standard PS3, and as ISO standard 12052:2006 "Health informatics -- Digital imaging and communication in medicine (DICOM) including workflow and data management".

HL7

Health Level-7 or HL7 refers to a set of international standards for transfer of clinical and administrative data between software applications used by various healthcare providers. These standards focus on the application layer, which is "layer 7" in the OSI model. The HL7 standards are produced by the Health Level Seven International, an international standards organization, and are adopted by other standards issuing bodies such as American National Standards Institute and International Organization for Standardization.

Hospitals and other healthcare provider organizations typically have many different computer systems used for everything from billing records to patient tracking. All of these systems should communicate with each other (or "interface") when they receive new information, or when they wish to retrieve information, but not all do so.

HL7 International specifies a number of flexible standards, guidelines, and methodologies by which various healthcare systems can communicate with each other. Such guidelines or data standards are a set of rules that allow information to be shared and processed in a uniform and consistent manner. These data standards are meant to allow healthcare organizations to easily share clinical information. Theoretically, this ability to exchange information should help to minimize the tendency for medical care to be geographically isolated and highly variable.

HL7 International considers the following standards to be its primary standards - those standards that are most commonly used and implemented:

Version 2.x Messaging Standard – an interoperability specification for health and medical transactions

Version 3 Messaging Standard – an interoperability specification for health and medical transactions

Clinical Document Architecture (CDA) – an exchange model for clinical documents, based on HL7 Version 3

Continuity of Care Document (CCD) - a US specification for the exchange of medical summaries, based on CDA.

Structured Product Labeling (SPL) – the published information that accompanies a medicine, based on HL7 Version 3

Clinical Context Object Workgroup (CCOW) – an interoperability specification for the visual integration of user applications

Other HL7 standards/methodologies include:

Fast Healthcare Interoperability Resources (FHIR) - a draft standard for the exchange of resources

Arden Syntax – a grammar for representing medical conditions and recommendations as a Medical Logic Module (MLM)

Claims Attachments – a Standard Healthcare Attachment to augment another healthcare transaction

Functional Specification of Electronic Health Record (EHR) / Personal Health Record (PHR) systems - a standardized description of health and medical functions sought for or available in such software applications

GELLO – a standard expression language used for clinical decision support.

H.320

H.320 is an umbrella recommendation by the ITU-T for running Multimedia (Audio/Video/Data) over ISDN based networks. The main protocols in this suite are H.221, H.230, H.242, audio codecs such as G.711, and video codecs such as

H.261 and H.263.

It is formally named as Narrow-band visual telephone systems and terminal equipment. It specifies technical requirements for narrow-band visual telephone systems and terminal equipment, typically for videoconferencing and videophone services. It describes a generic system configuration consisting of a number of elements which are specified by respective ITU-T Recommendations, definition of communication modes and terminal types, call control arrangements, terminal aspects and interworking requirements.

The service requirements for visual telephone services are presented in ITU- T Recs F.720 for video telephony and F.702 for videoconference. Video and audio coding systems and other technical aspects common to audiovisual services are covered in other Recommendations in the H.200/F.700-series.

Narrow-band for this specification is defined as bit rates ranging from 64 kbit/s to 1920 kbit/s. This channel capacity may be provided as a single B/H0/H11/H12-channel or multiple B/H0-channels in ISDN.

Used video codecs: H.261, and optionally H.262, H.263, H.264 according to the video hierarchy specified in specification, and in ITU-T Recs H.241 and H.242.

H.261 is mandatory in any enhanced H.320 system with video capability. Baseline H.263 capability shall be required in systems that use enhanced video modes.[1]

Used audio codecs: G.711, and optionally G.722, G.728, G.723.1, G.729. (Example of usage: If a visual telephone interworks with a wideband speech terminal, G.722 audio may be used instead of G.711 audio.)

T.120

T.120 is an ITU-T recommendation that describes a series of communication and application protocols and services that provide support for real-time, multipoint data communications. It is used by products such as Cisco WebEx's MeetingCenter, Microsoft NetMeeting, Nortel CS 2100 and Lotus Sametime to support application sharing, real-time text conferencing and other functions.

The set of "T.120" recommendations includes:

- T.120 - Data protocols for multimedia conferencing
- T.Imp120 - Revised Implementor's Guide for the ITU-T T.120 Recommendation series
- T.121 - Generic application template
- T.122 - Multipoint communication service - Service definition
- T.123 - Network-specific data protocol stacks for multimedia conferencing
- T.124 - Generic Conference Control
- T.125 - Multipoint communication service protocol specification
- T.126 - Multipoint still image and annotation protocol
- T.127 - Multipoint binary file transfer protocol
- T.128 - Multipoint application sharing (was known as T.share while in draft stage)
- T.134 - Text chat application entity
- T.135 - User-to-reservation system transactions within T.120 conferences
- T.136 - Remote device control application protocol
- T.137 - Virtual meeting room management - services and protocol

The PCMag encyclopedia of computer terms lists the following additional recommendations:

- T.130 - Realtime Architecture - Interaction between T.120 and H.320.
- T.131 - Network Specific Mappings - Transport of realtime data used with T.120 over LANs.
- T.132 - Realtime Link Management - Creation and routing of realtime data streams.
- T.133 - Audio Visual Control Services - Controls for realtime data streams.
- T.RES - Reservation Services - Interaction between devices and reservation systems
- T.Share - this is now known as T.128
- T.TUD - User Reservation - Transporting user-defined data

H.324

H.324 is an ITU-T recommendation for voice, video and data transmission over regular analog phone lines. It uses a regular 33,600 bit/s modem for transmission, the H.263 codec for video encoding and G.723.1 for audio.

H.324 standard is formally known as Terminal for low bit-rate multimedia communication. H.324 covers the technical requirements for very low bit-rate multimedia telephone terminals operating over the General Switched Telephone Network (GSTN). H.324 terminals provide real-time video, audio, or data, or any combination, between two multimedia telephone terminals over a GSTN voice band network connection.

H.324 terminals offering audio communication shall support the G.723.1 audio codec. H.324 terminals offering video communication shall support the
H.263 and H.261 video codecs. G.722.1 may be used for wideband audio applications. Annex G of H.324 specification defines usage of ISO/IEC 14496-1 (MPEG-4 Systems) generic capabilities in H.324 terminals. H.324/I terminals shall support interoperation with voice telephones using G.711 speech coding, if the connected network supports transmission and reception of G.711. Other modes such as G.722 audio may optionally be supported as well.

H.324 was adapted by 3GPP to form 3G-324M.

It is for example used in the Vialta Beamer BM-80 Phone Video Station, the MINX system from Datapoint Corporation, and in several other videophones.

4.8 VIDEOCONFERENCING

Videoconferencing (VC) is the conduct of a videoconference (also known as a video conference or videoteleconference) by a set of telecommunication technologies which allow two or more locations to communicate by simultaneous two-way video and audio transmissions. It has also been called 'visual collaboration' and is a type of groupware.

Videoconferencing differs from videophone calls in that it's designed to serve a conference or multiple locations rather than individuals. It is an intermediate form of video telephony, first used commercially in Germany during the late-1930s and later in the United States during the early 1970s as part of AT&T's development of Picture phone technology.

With the introduction of relatively low cost, high capacity broadband telecommunication services in the late 1990s, coupled with powerful computing processors and video compression techniques, videoconferencing has made significant inroads in business, education, medicine and media.

The core technology used in a videoconferencing system is digital compression of audio and video streams in real time. The hardware or software that performs compression is called a codec (coder/decoder). Compression rates of up to 1:500 can be achieved. The resulting digital stream of 1s and 0s is subdivided into labeled packets, which are then transmitted through a digital network of some kind (usually ISDN or IP). The use of audio modems in the transmission line allow for the use of POTS, or the Plain Old Telephone System, in some low-speed applications, such as videotelephony, because they convert the digital pulses to/from analog waves in the audio spectrum range.

The other components required for a videoconferencing system include:

- Video input: (PTZ / 360° / Fisheye) video camera or webcam
- Video output: computer monitor, television or projector
- Audio input: microphones, CD/DVD player, cassette player, or any other source of PreAmp audio outlet.
- Audio output: usually loudspeakers associated with the display device or telephone
- Data transfer: analog or digital telephone network, LAN or Internet
- Computer: a data processing unit that ties together the other components, does the compressing and decompressing, and initiates and maintains the data linkage via the network.

There are basically two kinds of videoconferencing systems:

Dedicated systems have all required components packaged into a single piece of equipment, usually a console with a high quality remote controlled video camera. These cameras can be controlled at a distance to pan left and right, tilt up and down, and zoom. They became known as PTZ cameras. The console contains all electrical interfaces, the control computer, and the software or hardware-based codec. Omnidirectional microphones are connected to the console, as well as a TV monitor with loudspeakers and/or a video projector. There are several types of dedicated videoconferencing devices:

- Large group videoconferencing are non-portable, large, more expensive devices used for large rooms and auditoriums.
- Small group videoconferencing are non-portable or portable, smaller, less expensive devices used for small meeting rooms.
- Individual videoconferencing are usually portable devices, meant for single users, have fixed cameras, microphones and loudspeakers integrated into the console.
- Desktop systems are add-ons (hardware boards or software codec) to normal PCs and laptops, transforming them into videoconferencing devices. A range of different cameras and microphones can be used with the codec, which contains the necessary codec and transmission interfaces. Most of the desktops systems work with the H.323 standard. Videoconferences carried out via dispersed PCs are also known as e-meetings. These can also be nonstandard, Microsoft Lync, Skype for Business, Google Hangouts, or Yahoo Messenger or standards based, Cisco Jabber.

- WebRTC Platforms are video conferencing solutions that are not resident by using a software application but is available through the standard web browser. Solutions such as Adobe Connect and Cisco WebEX can be accessed by going to a URL sent by the meeting organizer and various degrees of security can be attached to the virtual "room". Often the user will be required to download a piece of software, called an "Add In" to enable the browser to access the local camera, microphone and establish a connection to the meeting.

Conferencing layers:

The components within a Conferencing System can be divided up into several different layers: User Interface, Conference Control, Control or Signal Plane, and Media Plane.

Videoconferencing User Interfaces (VUI) can be either graphical or voice responsive. Many in the industry have encountered both types of interfaces, and normally graphical interfaces are encountered on a computer. User interfaces for conferencing have a number of different uses; they can be used for scheduling, setup, and making a videocall. Through the user interface the administrator is able to control the other three layers of the system.

Conference Control performs resource allocation, management and routing. This layer along with the User Interface creates meetings (scheduled or unscheduled) or adds and removes participants from a conference.

Control (Signaling) Plane contains the stacks that signal different endpoints to create a call and/or a conference. Signals can be, but aren't limited to, H.323 and Session Initiation Protocol (SIP) Protocols. These signals control incoming and outgoing connections as well as session parameters.

The Media Plane controls the audio and video mixing and streaming. This layer manages Real-Time Transport Protocols, User Datagram Packets (UDP) and Real-Time Transport Control Protocol (RTCP). The RTP and UDP normally carry information such the payload type which is the type of codec, frame rate, video size and many others. RTCP on the other hand acts as a quality control Protocol for detecting errors during streaming.

Multipoint videoconferencing:

Main article: Multipoint Control Unit

Simultaneous videoconferencing among three or more remote points is possible by means of a Multipoint Control Unit (MCU). This is a bridge that interconnects calls from several sources (in a similar way to the audio conference call). All parties call the MCU, or the MCU can also call the parties which are going to participate, in sequence. There are MCU bridges for IP and ISDN-based videoconferencing. There are MCUs which are pure software, and others which are a combination of hardware and software. An MCU is characterised according to the number of simultaneous calls it can handle, its ability to conduct transposing of data rates and protocols, and features such as Continuous Presence, in which multiple parties can be seen on-screen at once. MCUs can be stand-alone hardware devices, or they can be embedded into dedicated videoconferencing units.

The MCU consists of two logical components:

- A single multipoint controller (MC), and
- Multipoint Processors (MP), sometimes referred to as the mixer.

The MC controls the conferencing while it is active on the signaling plane, which is simply where the system manages conferencing creation, endpoint signaling and in-conferencing controls. This component negotiates parameters with every endpoint in the network and controls conferencing resources. While the MC controls resources and signaling negotiations, the MP operates on the media plane and receives media from each endpoint. The MP generates output streams from each endpoint and redirects the information to other endpoints in the conference.

Some systems are capable of multipoint conferencing with no MCU, stand- alone, embedded or otherwise. These use a standards-based H.323 technique known as "decentralized multipoint", where each station in a multipoint call exchanges video and audio directly with the other stations with no central "manager" or other bottleneck. The advantages of this technique are that the video and audio will generally be of higher quality because they don't have to be relayed through a central point. Also, users can make ad-hoc multipoint calls without any concern for the availability or control of an MCU. This added convenience and quality comes at the expense of some increased network bandwidth, because every station must transmit to every other station directly.

Videoconferencing modes:

Videoconferencing systems use several common operating modes:

- Voice-Activated Switch (VAS);
- Continuous Presence.

In VAS mode, the MCU switches which endpoint can be seen by the other endpoints by the levels of one's voice. If there are four people in a conference, the only one that will be seen in the conference is the site which is talking; the location with the loudest voice will be seen by the other participants.

Continuous Presence mode, displays multiple participants at the same time. The MP in this mode takes the streams from the different endpoints and puts them all together into a single video image. In this mode, the MCU normally sends the same type of images to all participants. Typically these types of images are called "layouts" and can vary depending on the number of participants in a conference.

Echo cancellation

A fundamental feature of professional videoconferencing systems is Acoustic Echo Cancellation (AEC). Echo can be defined as the reflected source wave interference with new wave created by source. AEC is an algorithm which is able to detect when sounds or utterances reenter the audio input of the videoconferencing codec, which came from the audio output of the same system, after some time delay. If unchecked, this can lead to several problems including:

- the remote party hearing their own voice coming back at them (usually significantly delayed)
- strong reverberation, which makes the voice channel useless, and
- howling created by feedback.

- Echo cancellation is a processor-intensive task that usually works over a narrow range of sound delays.

Cloud-based video conferencing

Cloud-based video conferencing can be used without the hardware generally required by other video conferencing systems, and can be designed for use by SMEs, or larger international companies like Facebook. Cloud-based systems can handle either 2D or 3D video broadcasting. Cloud-based systems can also implement mobile calls, VOIP, and other forms of video calling. They can also come with a video recording function to archive past meetings.

4.9 ADMINISTRATION OF CENTRALIZED MEDICAL DATA

An electronic health record (EHR), or electronic medical record (EMR), refers to the systematized collection of patient and population electronically-stored health information in a digital format. These records can be shared across different health care settings. Records are shared through network-connected, enterprise- wide information systems or other information networks and exchanges. EHRs may include a range of data, including demographics, medical history, medication and allergies, immunization status, laboratory test results, radiology images, vital signs, personal statistics like age and weight, and billing information.

EHR systems are designed to store data accurately and to capture the state of a patient across time. It eliminates the need to track down a patient's previous paper medical records and assists in ensuring data is accurate and legible. It can reduce risk of data replication as there is only one modifiable file, which means the file is more likely up to date, and decreases risk of lost paperwork. Due to the digital information being searchable and in a single file, EMR's are more effective when extracting medical data for the examination of possible trends and long term changes in a patient. Population-based studies of medical records may also be facilitated by the widespread adoption of EHR's and EMR's.

SECURITY AND CONFIDENTIALLY OF MEDICAL RECORDS

Health care is changing and so are the tools used to coordinate better care for patients like you and me. During your most recent visit to the doctor, you may have noticed your physician entering notes on a computer or laptop into an electronic health record (EHR). With EHRs comes the opportunity for patients to receive improved coordinated care from providers and easier access to their health information. It’s a way to make it easier for everyone to be better informed and more involved in the patient’s health care. However for many of us, EHRs also come with questions and concerns about the privacy and security of our health information. Who can access the information on my EHR? How can I see the information in my record and make sure it’s correct? How is it protected from loss, theft and hacking? What should I do if I think my information has been compromised?

Many of you have heard of HIPAA– the Health Insurance Portability and Accountability Act. The HHS Office for Civil Rights (OCR) enforces the HIPAA Privacy and Security Rules, which help keep entities covered under HIPAA accountable for the privacy and security of patients’ health information. As a former health care lawyer, I know that many health care providers understand and abide by their obligations under the Privacy and Security Rules. Although EHRs allow providers to use information more effectively to improve the quality and efficiency of your care, they do not change the obligations providers have to keep your protected health information private and secure.

Following my recent appointment as OCR's Director, I had a number of conversations that made it apparent to me that many patients recognize some of the health privacy jargon such as "HIPAA" or "the Notice of Privacy Practices," but often do not know their rights under the HIPAA Privacy and Security Rules — especially in terms of how these rules relate to EHRs.

The HIPAA Privacy Rule gives you rights over your own health information, regardless of its form. Whether your record is in paper or electronic form, under the Privacy Rule you have the right:

- To see or get a copy of your medical record;
- To request to have any mistakes corrected;
- To get a notice about how your health information is used and shared;
- To say how and where you want to be contacted by your health care provider; and
- To file a complaint if you think any of these rights have been violated. One way to do this is through OCR's website: www.hhs.gov/ocr.

These rights are spelled out in the Notice of Privacy Practices that is given to you at your doctor's office or hospital. Your health plan may also send this notice to you in the mail.

Specific to protecting the information stored in EHRs, the HIPAA Security Rule requires that health care providers set up physical, administrative, and technical safeguards to protect your electronic health information. Some safety measures that may be built in to EHR systems include:

“Access controls” like passwords and PIN numbers, to help limit access to your information;

“Encrypting” your stored information. This means your health information cannot be read or understood except by someone who can “decrypt” it, using a special “key” made available only to authorized individuals;

An “audit trail,” which records who accessed your information, what changes were made and when.

In certain circumstances, if your data is seen by someone who should not see it, federal law requires doctors, hospitals, and other health care providers to notify you of a “breach” of your health information. This requirement helps patients know if something has gone wrong with the protection of their information and helps keep providers accountable.

OCR works to help make sure your health information is kept private and secure by your health professionals. We are here to help you understand these rights, how you can take action if your rights are violated and how your health information is required to be safeguarded under the law. The first step is to know your rights.

Access control

Access control is of vital importance in healthcare. Confidentiality is a main concern when it is related to patient clinical information that needs to be private. It is essential to protect this information from unauthorized access and, therefore, misuse or legal liability.

The introduction of the EMR within healthcare organizations has the main goal of integrating heterogeneous patient information that is usually scattered throughout different locations. This is why the EMR is becoming an essential source of information and an important support tool for the healthcare professional. There is also an increasing need to access healthcare information at remote locations. This and the distributed nature of the information stress the need for access control requirements to be taken seriously.

Although the EMR is an essential tool for the healthcare professional, the reality is that it still does not integrate easily and effectively with healthcare professionals' daily workflow and processes. Several obstacles are mentioned by healthcare professionals concerning the use of EMR. The obstacles are associated with a concern for patient privacy and other security vulnerabilities related to the easy distribution, sharing and wider online access of the information.

Other barriers that prevent the successful integration and use of EMR are mostly related to human interactions with the system. These include the time taken by healthcare professionals to learn and to use the system, and the consequent extra time and costs the patients may incur if they have to wait longer to be seen and treated. In addition, relational and educational barriers also hinder the right use of the EMR. Relational barriers include the perceptions that the physician and the patient have about the use of the EMR and how their relationship can be affected by it. Educational barriers comprise the lack of proficiency and difficulties that healthcare professionals have whilst interacting with the EMR to perform their daily tasks.

Taking into account the problems mentioned above and considering that the main factor that is driving the integration of EMR systems is the need to improve clinical processes and workflow efficiency, a deeper understanding of how access control systems can affect this integration and how they are being developed within the EMR is required.

4.10 CYBER LAWS RELATED TO TELEMEDICINE

The concept of telemedicine and web based medical services are derived from the exchange or transmission of medical knowledge or information through electronic formats and mediums, so as to cut across time and space across the world for the benefit of medical advancement. Medical information is communicated through electronic media in interactive formats such as audiovisual media, telephonic conferences, satellite communication, internet etc. for medical consultation, examination or remote monitoring / medical procedure purposes. The model has popularized since it links isolated communities to advanced medical services and provides speedy delivery of medical expertise. Certified medical practitioners world over have started taking advantage of the telemedicine concept, expanding their services. The United States licensure laws promote the model while requiring a practitioner following the format to obtain a full license across states to deliver telemedicine healthcare services across state lines. In India practitioners and medical societies have been tele-transmitting medical information and remote monitoring health services since as early as 1975. The medium of tele applications and web-interface based systems linking patients and medical practitioners through telemedicine services may use wireless diagnostics tools like stethoscopes; blood pressure, temperature and insulin monitors, and ultrasounds enabling remote diagnosis, treatment, advanced healthcare and medical services.

The Medical Council of India regulates uniform standards of higher qualifications in medicine and recognition of medical qualifications in India and abroad. Official registration of doctors with recognized medical qualifications is controlled by the council, and procedures have been laid out under the Indian Medical Council Act 1956 and Indian Medical Degree Act 1916. Although there are no legal constraints specifically dealing with methodology of executing or dispensing medical services in India, various laws including the Drugs and Cosmetics Act, 1940 define negligence; criminal intent; sale, manufacture and distribution of drugs etc., while judicial precedent and case laws determine medical negligence on a case by case basis. The healthcare service provider adopting telemedicine methods of medical practice must ensure that medical consultation, prescriptions, treatment and drugs are dispensed only in accordance with legal provisions and guidelines regulating the medical and healthcare sector in India.

Under the present laws relating to the above in India, a fully automated process solely based on an artificial intelligence program may not be legally feasible, as it is a basic requirement that only medical practitioners registered before the Medical Council of India and other relevant lawsare allowed to provide medical consultation, prescriptions and treatment. For understanding the legal proposition in regard to the telemedicine in India, one has to understand the implications of some important legal provisions relating to medical healthcare and drugs in India, as under:

A *"Registered medical practitioner" has been defined under Section 2 (ee) of the Drugs and Cosmetics Rules, 1945 of India as a person-

- holding a qualification granted by an authority specified or notified under Section 3 of the Indian Medical Degrees Act, 1916 (7 of 1916), or specified In the Schedules to the Indian Medical Council Act, 1956 (102 of 1956); or

- registered or eligible for registration in a medical register of a State meant for the registration of persons practicing the modern scientific system of medicine excluding the Homoeopathic system of medicine; or
- registered in a medical register, other than a register for the registration of Homoeopathic practitioner, of a State, who although not falling within subclause (i) or sub-clause (ii) declared by a general or special order made by the State Government in this behalf as a person practicing the modern scientific system of medicine for the purposes of this Act; or
- registered or eligible for registration in the register of dentists for a State under the Dentists Act, 1948 (16 of 1948); or
- who is engaged in the practice of veterinary medicine and who possesses qualification approved by the State Government.

A "Drug" has been defined under Section 3 (b) of the Drugs and Cosmetics Act, 1940 and includes-

- all medicines for internal or external use of human beings or animals and all substances intended to be used for or in the diagnosis, treatment, mitigation or prevention of any disease or disorder in human beings or animals, including preparations applied on human body for the purpose of repelling insects like mosquitoes;
- such substances (other than food) intended to affect the structure or any function of human body or intended to be used for the destruction

of (vermin) or insects which cause disease in human beings or animals, as may be specified from time to time by the Central Government by notification in the Official Gazette;

- all substances intended for use as components of a drug including empty gelatin capsules; and
- such devices intended for internal or external use in the diagnosis, treatment, mitigation or prevention of disease or disorder in human beings or animals, as may be specified from time to time by the Central Government by notification in the Official Gazette, after consultation with the Board.

The term "prescribed" as per Section 3 [(i)] of the Drugs and Cosmetics Act, 1940 means prescribed by rules made under the Act.

Prescriptions made against medical consultation and diagnosis services under telemedicine formats should satisfy legal requirements given below so as to be a valid legal prescriptionunder the laws of India. The Drugs and Cosmetics Rules, 1945 specify the type of drugs that require valid medical prescriptions for retail purchase, classifying them under Schedules appended to the Rules.

"Prescription only drugs" are defined under Section 65(9) of the Drugs and Cosmetics Rules, 1945, which states that –

- Substances specified in Schedule H or Schedule X shall not be sold by retail except on and in accordance with the prescription of a Registered Medical Practitioner only.

- Further, in the case of substances specified in schedule X, the prescriptions shall be in duplicate, one copy of which shall be retained by the licensee for a period of two years.
- The supply of drugs specified in Schedule H or Schedule X to Registered Medical Practitioners, Hospitals, Dispensaries and Nursing Homes shall be made only against the signed order in writing which shall be preserved by the licensee for a period of two years;
- The above provision deals only with the dispensing of medicine and supply of a certain category of medicine. However, irrespective of the schedule in which a medicine may fall, prescription of a medicine can be made only by a registered medical professional as per the Rules. Since there are no legal provisions describing the manner of treating a patient, prescriptions instructing a patient to consume any drugs are very important documentary evidence of negligence or lack of it on the part of a medical practitioner while treating a patient. The Rules have defined the important components that constitute a valid legal prescription, for all medical practice purposes.

A "prescription" has been defined under Section 65(10) of the Drugs and Cosmetics Rules, 1945 so as to have the following components-

- be in writing** and be signed*** by the person giving it with his usual signature and be dated by him;
- specify the name and address of the person for whose treatment it is given, or the name and address of the owner of the animal if the drug is meant for veterinary use;

- indicate the total amount of the medicine to be supplied and the dose to be taken.

For all medical treatments through telemedicine or web-interface format, it is important to ensure that the prescriptions must satisfy the above requirements of being in writing and signed by a registered medical practitioner, without which the prescription will be invalid in the eyes of the law.

Due to the recognition of electronic documents under the Information Technology Act, 2000, a prescription in an electronic format may be validated as a legal prescription if it is a secure electronic record affixed with a secure digital signature as prescribed under the Information Technology Act, 2000 of India. The Information Technology Act, 2000 provides for authentication of secure electronic records and affixing of digital signatures so as to ensure the legal validity of the same. Section 4, mentioned herein below, of the Information Technology Act, 2000 which recognizes electronic records is important for understanding above:-

"where any law provides that information or any other matter shall be in writing or in the typewritten or printed form, then, notwithstanding anything contained in such law, such requirement shall be deemed to have been satisfied if such information or matter is –

Rendered or made available in electronic form, and Accessible

so as to be usable for subsequent reference."

Section 3 of the Information Technology Act, 2000 deals with authentication of electronic records as under:

- Subject to the provisions of the section any subscriber may authenticate an electronic record by affixing his digital signature.
- The authentication of the electronic record shall be effected by the use of asymmetric crypto system and hash function which envelop and transform the initial electronic record into another electronic record.

Digital signatures are legally recognized under Section 5 of the Information Technology Act, 2000, which states as under:-

"where any law provides that information or any other matter shall be authenticated by affixing the signature or any document shall be signed or bear the signature of any person then; notwithstanding anything contained in such law, such requirement shall be deemed to have been satisfied, if such information or matter is authenticated by means of digital signature affixed in such manner as may be prescribed by the Central Government."

Automated artificial intelligence based telemedicine formats controlled by a registered medical practitioner can formulate legal prescriptions in the form of an electronic record, provided the same can be attributed, under Section 11 of the Information Technology Act, 2000, to the originator-

- if it was sent by the originator himself;
- by a person who had the authority to act on behalf of the originator in respect of that electronic record; or
- by an information system programmed by or on behalf of the originator to operate automatically.

Section 14 of the Information Technology Act, 2000, defines a secure electronic record, wherein any security procedure has been applied to it at a specific point of time, after which such record shall be deemed to be a secure electronic record from such point of time to the time of verification.

Under Section 15 of the Information Technology Act, 2000, a secure digital signature by application of a security procedure agreed to by the parties concerned, can be verified to be a digital signature, at the time it was affixed, if it was—

- unique to the subscriber affixing it;
- capable of identifying such subscriber;
- created in a manner or using a means under the exclusive control of the subscriber and is linked to the electronic record to which it relates in such a manner that if the electronic record was altered the digital signature would be invalidated.

Since telemedicine formats of medical practice are essentially based on mediums of technology, the medical practice model may use the above legal provisions to their advantage with respect to preparation of valid legal electronic prescriptions.

'OTC Drugs' (Over The Counter drugs) are drugs legally allowed to be sold 'Over The Counter', i.e. without the prescription of a Registered Medical Practitioner. In India, though the phrase has no legal recognition, all drugs not included in the list of 'prescription only drugs' under the Drugs and Cosmetics Act, 1940 may be considered as non-prescription drugs (or OTC drugs). A proposal for a list of over the counter (OTC) drugs has been under the government's consideration and a committee appointed for the purpose has been working on it.

The Drug Controller General of India is expected to lay down a separate set of rules or guidelines for OTC marketing once the list is official.

In a scenario where advice is provided electronically through a telemedicine interface, which is manned by certified medical practitioners and/ or an artificial intelligence system validated by medical practitioners entitled to practice medicine in India, the guidelines issued by the Medical Council of India under the Code of Ethics Regulations, 2002 also apply. Some guidelines, which may apply to a telemedicine system are listed below:-

Section 1.4 of the Code of Ethics Regulations, 2002 states that registration numbers of medical practitioners/ doctors accorded by the State Medical Council / Medical Council of India must be displayed in the clinic and in all prescriptions, certificates, money receipts given to patients.

Under Section 6.1.1 of the Code of Ethics Regulations, 2002 the act of soliciting patients directly or indirectly is unethical, by a physician or a group of physicians, institutions or organizations. Although no legal provision deals with the manner of communication between a medical practitioner and patient with respect to diagnosis and treatment, there are numerous legal provisions dealing with ethical conduct to be followed by medical practitioners while dispensing specialized medical services.

References:

1. Wootton, R., Craig, J., Patterson, V. (Eds.), Introduction to Telemedicine. Royal Society of Medicine Press Ltd (ISBN 1853156779), 2006.

2. O'Carroll, P.W., Yasnoff, W.A., Ward, E., Ripp, L.H., Martin, E.L. (Eds), Public Health Informatics and Information Systems. Springer (ISBN 0-387-95474-0), 2003.

Question Bank

1. Draw the functional block diagram of telemedicine and explain
2. State the ethical and legal aspects of telemedicine
3. How video conferencing is used for the treatment of patients using telemedicine.
4. How the medical data's are centralized for the administration purpose
5. State the cyber laws related to telemedicine

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF BIO AND CHEMICAL ENGINEERING
DEPARTMENT OF BIOMEDICAL ENGINEERING

UNIT – 5- MEDICAL ELECTRONICS AND E-HEALTH– SBMA7004

V- TELERADIOLOGY, HEALTH INFORMATICS AND APPLICATIONS

Tele radiology: Definition, Basic parts of teleradiology system: Image Acquisition system Display system, Tele pathology, multimedia databases, color images of sufficient resolution, Dynamic range, spatial resolution, compression methods, Interactive control of color, Medical information storage and management for telemedicine - patient information medical history, test reports, medical images diagnosis and treatment. Hospital information system - Doctors, paramedics, facilities available. Pharmaceutical information system. Introduction to robotics surgery, telesurgery. Telecardiology, Teleoncology, Telemedicine access to health education and self care.

5.1 TELE RADIOLOGY:

Teleradiology is the transmission of radiological patient images, such as x-rays, CTs, and MRIs, from one location to another for the purposes of sharing studies with other radiologists and physicians. Teleradiology is a growth technology given that imaging procedures are growing approximately 15% annually against an increase of only 2% in the radiologist population.

Teleradiology improves patient care by allowing radiologists to provide services without actually having to be at the location of the patient. This is particularly important when a sub-specialist such as a MRI radiologist, neuroradiologist, pediatric radiologist, or musculoskeletal radiologist is needed, since these professionals are generally only located in large metropolitan areas working during daytime hours. Teleradiology allows for trained specialists to be available 24/7.

Teleradiology utilizes standard network technologies such as the internet, telephone lines, wide area network, local area network (LAN) and the latest high tech being computer clouds. Specialized software is used to transmit the images and enable the radiologist to effectively analyze what can be hundreds of images for a given study. Technologies such as advanced graphics processing, voice recognition, and image compression are often used in teleradiology. Through tele radiology and mobile DICOM viewers, images can be sent to another part of the hospital, or to other locations around the world.

BASIC PARTS OF TELERADIOLOGY SYSTEM

A basic teleradiology system consists of three major components:

1. An image sending station

2. A transmission network

3. A receiving/image review station.

These components are interconnected as shown.

Fig:5.1 basic teleradiology system

The image sending station is usually a personal computer that consists of two main parts: the image digitizer and the network interface. As stated before, the digitizer is the flatbed scanner used to transform the film image to a computer image file format. The network interface includes a network card or a modem. The personal computer will also consist of specialized software that will help in the process of transforming the digitized image into a file format that can be recognized on the other end.

The transmission network is the medium by which the file can be sent between the sending and receiving station. This medium could include the use of fiber optics, ISDN, wireless, and generic phone lines.

The receiving/image review station consists of:

1. network interface (modem or network interface card)

2. personal computer with storage medium (hard disk drive, DVD- recorder, or optical drive)

3. one or two TV monitors

4. optional hard copy device printer

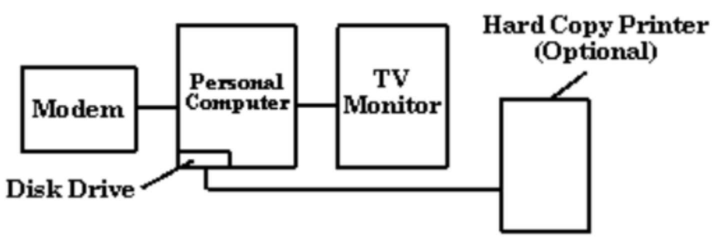

Fig:5.2 transmission network

The network interface on the review station receives the electrical impulses from the transmission network and converts them back to digital image data. This data are then sent via the network interface to the computer disk for storage. Once the image data are stored, the radiologist at the station can access the image via the computer and display it on the TV monitor. The radiologist then performs software manipulation functions to "enhance" the image. Software functions vary, but almost all have the capability of window/level and magnification. If a printer is attached to the receiving station, hard copy images can be printed.

5.2 TELEPATHOLOGY

Telepathology is the practice of pathology at a distance. It uses telecommunications technology to facilitate the transfer of image-rich pathology data between distant locations for the purposes of diagnosis, education, and research. Performance of telepathology requires that a pathologist selects the video images for analysis and the rendering of diagnoses. The use of “television microscopy”, the forerunner of telepathology, did not require that a pathologist have physical or virtual “hands-on” involvement in the selection of microscopic fields-of-view for analysis and diagnosis.

An academic pathologist, Ronald S. Weinstein, M.D., coined the term “telepathology” in 1986. In a medical journal editorial, Weinstein outlined the actions that would be needed to create remote pathology diagnostic services. He and his collaborators published the first scientific paper on robotic telepathology. Weinstein was also granted the first U.S. patents for robotic telepathology systems and telepathology diagnostic networks. Dr. Weinstein is known to many as the "father of telepathology". In Norway, Eide and Nordrum implemented the first sustainable clinical telepathology service in 1989; this is still in operation decades later. A number of clinical telepathology services have benefited many thousands of patients in North America, Europe, and Asia.

Telepathology has been successfully used for many applications, including the rendering of histopathology tissue diagnoses at a distance. Although digital pathology imaging, including virtual microscopy, is the mode of choice for telepathology services in developed countries, analog telepathology imaging is still used for patient services in some developing countries.

MULTIMEDIA DATABASE

A Multimedia database (MMDB) is a collection of related multimedia data. The multimedia data include one or more primary media data types such as text, images, graphic objects (including drawings, sketches and illustrations) animation sequences, audio and video.

A Multimedia Database Management System (MMDBMS) is a framework that manages different types of data potentially represented in a wide diversity of formats on a wide array of media sources. It provides support for multimedia data types, and facilitate for creation, storage, access, query and control of a multimedia database.

A Multimedia Database (MMDB) hosts one or more multimedia data types[3] (i.e. text, images, graphic objects, audio, video, animation sequences. These data types are broadly categorized into three classes:

Static media (time-independent: image and graphic object).

Dynamic media (time-dependent: audio, video and animation).

Dimensional media(3D game and computer aided drafting programs).

Comparison of multimedia data types

Medium	Elements	Time-dependence
Text	Printable characters	No
Graphic	Vectors, regions	No
Image	Pixels	No
Audio	Sound, Volume	Yes
Video	Raster images, graphics	Yes

Additionally, a Multimedia Database (MMDB) needs to manage additional information pertaining to the actual multimedia data. The information is about the following:

Media data: the actual data representing an object.

Media format data: information about the format of the media data after it goes through the acquisition, processing, and encoding phases.

Media keyword data: the keyword descriptions, usually relating to the generation of the media data.

Media feature data: content dependent data such as contain information about the distribution of colors, the kinds of textures and the different shapes present in an image.

The last three types are called metadata as they describe several different aspects of the media data. The media keyword data and media feature data are used as indices for searching purpose. The media format data is used to present the retrieved information.

Requirements of Multimedia databases

Like the traditional databases, Multimedia databases should address the following requirements:

- Integration
- Data items do not need to be duplicated for different programs invocations
- Data independence
- Separate the database and the management from the application programs
- Concurrency control
- Allows concurrent transactions
- Persistence
- Data objects can be saved and re-used by different transactions and program invocations
- Privacy
- Access and authorization control

- Integrity control
- Ensures database consistency between transactions
- Recovery
- Failures of transactions should not affect the persistent data storage
- Query support
- Allows easy querying of multimedia data

Multimedia databases should have the ability to uniformly query data (media data, textual data) represented in different formats and have the ability to simultaneously query different media sources and conduct classical database operations across them. (Query support)

They should have the ability to retrieve media objects from a local storage device in a good manner. (Storage support)

They should have the ability to take the response generated by a query and develop a presentation of that response in terms of audio-visual media and have the ability to deliver this presentation. (Presentation and delivery support)

Issues and challenges

Multimedia data consists of a variety of media formats or file representations including TIFF, BMP, PPT, IVUE, FPX, JPEG, MPEG, AVI, MID, WAV, DOC, GIF, EPS, PNG, etc. Because of restrictions on the conversion from one format to the other, the use of the data in a specific format has been limited as well. Usually, the data size of multimedia is large such as video; therefore, multimedia data often require a large storage.

Multimedia database consume a lot of processing time, as well as bandwidth.

Some multimedia data types such as video, audio, and animation sequences have temporal requirements that have implications on their storage, manipulation and presentation, but images, video and graphics data have spatial constraints in terms of their content.

5.3 COLOUR IMAGES OF SUFFICIENT RESOLUTION

DYNAMIC RANGE:

The dynamic range of a system is defined as the difference between the saturation level and minimum detection level divided by the capability of detection change.

$$Dr=(Smax - Smin) / Sdif$$

Smax = Saturation level or maximum value able to be detected without saturation

Smin = Detection level or minimum value able to be detected

Sdif = Minimum magnitude change able to be detected by the system

Diagnosis in pathology is based on colour images, that with 8 bits of dynamic range may produce sufficient information; the reason being that pathology acquisition systems are based in Video cameras that have an implicit non-linear response (gamma correction). This effect on the image is inverted by the gamma correction of the display systems obtaining a final effect of an adequate visual perception.

SPATIAL RESOLUTION

In the digital sampling of pathology images the spatial resolution is of paramount importance in the accurate visual perception, and therefore:

a) Introduces variability on the minimum requirements for a capture system depending on the microscopic power used.

b) Is directly related to the capability of pathology images to support lossy compression algorithms.

If minimum requirements on spatial resolution are respected, compression may be possible without sensible loss of visual information.

SAMPLING THEORY

Spatial frequency is related with the amount of information present in a one spatial dimension unit. Spatial resolution is defined as the maximum spatial frequency (Fmax) able to be detected or transmitted; for digital images, it represents the two dimensional pixel matrix of the acquisition device.

Sampling theory tries to reduce to a reasonable limit the amount of information that should be stored and processed to avoid heavy computation processes and optimize transmission times over the network. According to the Sampling Theory of Shannon (sample frequency or distance), in order to assure that a discrete sample (digital) will unequivocally reproduce an analogue (continuous) image, samples should be taken at 2Fmax. It means double of the maximum spatial frequency; all this limited by the signal to noise ratio of the system.

Shannon Theorem: $R<2*B$

$n=sqr(1+ signal/noise)$

$D=R*log2(n)$

$D<B*log2(1+signal/noise)$

That is the reason why 512 x 512 image sampling at higher power can provide good quality images (some investigators mention 1Kx1K requirements (Black-Schaffer et al. 1995)) but that could be unacceptable for lower power sampling, where the amount of information per space unit is much higher.

The high resolution cameras present in the market (still too expensive) directly record digital images with very high spatial resolution. Two main consequences are derived from this technique:

A recorded image at low microscopic power can support digital zooming simulating higher power microscopic views.

The image is not affected by the gamma correction of the Video-systems (Annex IV) - depending on the cameras- and is therefore comparable to a digital radiography from the visual perception and image processing point of view.

All principles pointed out previously are true considering that the optics of the microscope is optimal.

COMPRESSION METHODS

Address the issue of how to reduce the amount of data without a sensible loss of important information. The term "important", varies according to the subsequent analysis procedure:

a) Visual inspection or diagnosis;

b) Image analysis;

c) Image quantization

Include techniques of luminance and colour reduction (reduction of dynamic range), spatial resolution reduction, or both, as well as simple data reduction.

A.- Colour reduction techniques:

A.1. Through YUV encoding. Usually based on colour sampling reduction of 8,4,4 bits information acquisition.

A.2. Reduction of colour palette to 256 colours (8 bits).

A.3. Median cut colour quantization. It is a colour quantization technique that optimises the representation of the original colour in the final palette.

This technique produces median cuts of colours on RGB to arrive to 256 colours (8 bits), therefore, the most frequent colours have greater range of colour gradation than infrequent colours. This can be further optimised by dithering that appears to expand the available colour palette by juxtaposing pixels of different colours to create the illusion of additional colours by visual blending. In diffusion dithering this is accomplished by calculating the numeric difference between the original and final colour of each pixel and distributing this difference among the neighbouring pixels.

B.- Spatial resolution reduction: Includes sampling reduction with or without inter-pixel interpolation e.g. digitise one every two pixels.

C.- Data reduction as the ones included in the classical lossy and lossless compression methods.

INTERACTIVE CONTROL OF COLOUR

Misinterpretation of colour at the receiving point is an important cause of error together with sampling problems. Therefore, a minimum knowledge of the Colour Theory is required to understand the need for interactivity due to the various *Spectral Responses* as well as built-in *Gamma Correction* that cameras and displays provide, and that affects the colours and visual perception.

5.4 MEDICAL INFORMATION STORAGE AND MANAGEMENT FOR TELEMEDICINE

Implications of Health Care

Organization for Informatics

• Money determines much

– Medicine spends 1-2% on IT, vs. 6-7% for business overall, vs. 10-12% for banking

– "Bottom line" rules, therefore emphasis on

• Billing

• Cost control

• Quality control, especially if demonstrable cost savings

• Retention and satisfaction (maybe)

– Management by accountants

Basis for historical record

• Communication among providers

• Anticipate future health problems

• Record standard preventive measures

• Identify deviations from the expected

- Legal record
- Basis for clinical research

Who Keeps Records?

- Doctor
- Nurse
- radiologist
- Office staff,
- pharmacist admissions
- patient
- Administrator
- physical therapist
- lab personnel

Forms of Clinical Data

- Numerical Measurements
- Coded (?) discrete data
  - – Lab data – Family history
  - – Bedside measurements – Patient's medical history
  - – Home instrumentation – Current complaint
- Recorded signals (e.g., • Symptoms (patient) ECG, EEG, EMG) • Signs (doc)
  - – Physical examination
- Images (X-ray, MRI, CAT, – MedicationsUltrasound, Pathology, • Narrative text…)
- Genes (SNPs, – Doctor's, nurse's notes expression arrays, – Discharge summaries pedigrees, …) – Referring letters

Organization of Data

- Doctor's journal (traditional)

• Time order of collection, per patient (Mayo)

• Source of data

• Problem-Oriented Medical Record (POMR) (L. Weed, 1969)

– Notes organized by problems

– SOAP: subjective, objective, assessment, plans

5.5 PATIENT INFORMATION MEDICAL HISTORY

The content of the history required in primary care consultations is very variable and will depend on the presenting symptoms, patient concerns and the past medical, psychological and social history. However the general framework for history taking is as follows:

Presenting complaint.

History of presenting complaint, including investigations, treatment and referrals already arranged and provided..

Past medical history: significant past diseases/illnesses, surgery, including complications, trauma.

Drug history: now and past, prescribed and over-the-counter, allergies.

Family history: especially parents, siblings and children.

Social history: smoking, alcohol, drugs, accommodation and living arrangements, marital status, baseline functioning, occupation, pets and hobbies.

Systems review: cardiovascular system, respiratory system, gastrointestinal system, nervous system, musculoskeletal system, genitourinary system.

The art of history taking

It is widely taught that diagnosis is revealed in the patient's history. 'Listen to your patient; they are telling you the diagnosis' is a much quoted aphorism.

The basis of a true history is good communication between doctor and patient. The patient may not be looking for a diagnosis when giving their history and the doctor's search for one under such circumstances is likely to be fruitless. The patient's problem, whether it has a medical diagnosis attached or not, needs to be identified.

It is important for doctors to acquire good consultation skills which go beyond prescriptive history taking learned as part of the comprehensive and systematic clerking process outlined in textbooks. A good history is one which reveals the patient's ideas, concerns and expectations as well as any accompanying diagnosis. The doctor's agenda, incorporating lists of detailed questions, should not dominate the history taking. Listening is at the heart of good history taking. Without the patient's perspective, the history is likely to be much less revealing and less useful to the doctor who is attempting to help the patient.

Often the history alone does reveal a diagnosis. Sometimes it is all that is required to make the diagnosis. A good example is with the complaint of headache where the diagnosis can be made from the description of the headache and perhaps some further questions. For example, in cluster headache the history is very characteristic and reveals the diagnosis without the need for examination or investigations.

To obtain a true, representative account of what is troubling a patient and how it has evolved over time, is not an easy task. It takes practice, patience, understanding and concentration. The history is a sharing of experience between patient and doctor. A consultation can allow a patient to unburden himself or herself. They may be upset about their condition or with the frustrations of life and it is important to allow patients to give vent to these feelings. The importance of the lament and how it may be transformed from the grumbles of a heartsink patient, to a useful diagnostic and therapeutic tool for both patient and physician, has been discussed in an excellent paper.

5.6 MEDICAL IMAGES DIAGNOSIS AND TREATMENT

Medical imaging is the technique and process of creating visual representations of the interior of a body for clinical analysis and medical intervention, as well as visual representation of the function of some organs or tissues (physiology). Medical imaging seeks to reveal internal structures hidden by the skin and bones, as well as to diagnose and treat disease. Medical imaging also establishes a database of normal anatomy and physiology to make it possible to identify abnormalities. Although imaging of removed organs and tissues can be performed for medical reasons, such procedures are usually considered part of pathology instead of medical imaging.

As a discipline and in its widest sense, it is part of biological imaging and incorporates radiology which uses the imaging technologies of X-ray radiography, magnetic resonance imaging, medical ultrasonography or ultrasound, endoscopy, elastography, tactile imaging, thermography, medical photography and nuclear medicine functional imaging techniques as positron emission tomography (PET) and Single-photon emission computed tomography (SPECT).

Measurement and recording techniques which are not primarily designed to produce images, such as electroencephalography (EEG), magnetoencephalography (MEG), electrocardiography (ECG), and others represent other technologies which produce data susceptible to representation as a parameter graph vs. time or maps which contain data about the measurement locations. In a limited comparison these technologies can be considered as forms of medical imaging in another discipline.

Up until 2010, 5 billion medical imaging studies had been conducted worldwide. Radiation exposure from medical imaging in 2006 made up about 50% of total ionizing radiation exposure in the United States.

Medical imaging is often perceived to designate the set of techniques that noninvasively produce images of the internal aspect of the body. In this restricted sense, medical imaging can be seen as the solution of mathematical inverse problems. This means that cause (the properties of living tissue) is inferred from effect (the observed signal). In the case of medical ultrasonography, the probe consists of ultrasonic pressure waves and echoes that go inside the tissue to show the internal structure. In the case of projectional radiography, the probe uses X-ray radiation, which is absorbed at different rates by different tissue types such as bone, muscle and fat.`

The term noninvasive is used to denote a procedure where no instrument is introduced into a patient's body which is the case for most imaging techniques used.

5.7 HOSPITAL INFORMATION SYSTEM

A hospital information system (HIS) is an element of health informatics that focuses mainly on the administrational needs of hospitals. In many implementations, a HIS is a comprehensive, integrated information system designed to manage all the aspects of a hospital's operation, such as medical, administrative, financial, and legal issues and the corresponding processing of services.

Architecture

Hospital Information System architecture has three main levels, Central Government Level, Territory Level, and Patient Carrying Level. Generally, all types of hospital information system (HIS) are supported in client-server architectures for networking and processing. Most work positions for HIS are currently resident types. Mobile computing began with wheeled PC stands. Now tablet computers and smartphone applications are used.

Enterprise HIS with Internet architectures have been successfully deployed in Public Healthcare Territories and have been widely adopted by further entities. The Hospital Information System (HIS) is a province-wide initiative designed to improve access to patient information through a central electronic information system. HIS's goal is to streamline patient information flow and its accessibility for doctors and other health care providers. These changes in service will improve patient care quality and patient safety over time.

The patient carries system record patient information, patient laboratory test results, and patient's doctor information. Doctors can access easily person information, test results, and previous prescriptions. Patient schedule organization and early warning systems can provide by related systems.

Functional split

HIS has data warehousing as the main topic, hence a more static model of information management. HIS is often composed of one or several software components with specialty-specific extensions, as well as of a large variety of sub-systems in medical specialties from a multi-vendor market. Specialized implementations name for example Laboratory Information System (LIS), Policy and Procedure Management System, Radiology Information System (RIS) or Picture archiving and communication system (PACS).

Architecture is based on a distributed approach and on the utilization of standard software products complying with the industrial and market standards must be utilized (such as: UNIX operating systems, MS-Windows, local area network based on Ethernet and TCP/IP protocols, relational database management systems based on SQL language or Oracle databases, C programming language).

Portable devices such as smartphones and table computers may be used at the bedside.

PHARMACY INFORMATION SYSTEMS

Pharmacy information systems (PIS) are complex computer systems that have been designed to meet the needs of a pharmacy department. Through the use of such systems, pharmacists can supervise and have inputs on how medication is used in a hospital.

Some of the activities which Pharmacy Information Systems have been employed in pharmacy departments include:

Clinical Screening: The Pharmacy Information System can assist in patient care by the monitoring of drug interactions, drug allergies and other possible medication-related complications.

When a prescription order is entered, the system can check to see if there are any interactions between two or more drugs taken by the patient simultaneously or with any typical food, any known allergies to the drug, and if the appropriate dosage has been given based on the patient’s age, weight and other physiologic factors. Alerts and flags come up when the system picks up any of these.

Prescription Management: The Pharmacy Information System can also be use to mange prescription for inpatients and/or outpatients. When prescription orders are received, the orders are matched to available pharmaceutical products and then dispensed accordingly depending on whether the patient is an inpatient or outpatient.

It is possible to track all prescriptions passed through the system from who prescribed the drug, when it was prescribed to when it was dispensed.

It is also possible to print out prescription labels and instructions on how medication should be taken based on the prescription.

Inventory Management: Pharmacies require a continuous inventory culture in order to ensure that drugs do not go out of stock. This is made even more difficult when there are multiple dispensing points. When don manually it is very difficult to maintain an accurate inventory.

Pharmacy Information Systems aid inventory management by maintaining an internal inventory of all pharmaceutical products, providing alerts when the quantity of an item is below a set quantity and providing an electronic ordering system that recommends the ordering of the affected item and with the appropriate quantity from approved suppliers.

Patient Drug Profiles: These are patient profiles managed by the Pharmacy Information System and contain details of their current and past medications, known allergies and physiological parameters. These profiles are used for used for clinical screening anytime a prescription is ordered for the patient.

Report Generation: Most Pharmacy Information Systems can generate reports which range from determining medication usage patterns in the hospital to the cost of drugs purchased and /or dispensed.

Interactivity with other systems: It is important that Pharmacy Information Systems should be able to interact with other available systems such as the clinical information systems to receive prescription orders and financial information system for billing and charging.

5.8 INTRODUCTION TO ROBOTICS SURGERY

Robotic surgery, computer-assisted surgery, and robotically-assisted surgery are terms for technological developments that use robotic systems to aid in surgical procedures. Robotically-assisted surgery was developed to overcome the limitations of pre-existing minimally-invasive surgical procedures and to enhance the capabilities of surgeons performing open surgery.

In the case of robotically-assisted minimally-invasive surgery, instead of directly moving the instruments, the surgeon uses one of two methods to control the instruments; either a direct telemanipulator or through computer control. A telemanipulator is a remote manipulator that allows the surgeon to perform the normal movements associated with the surgery whilst the robotic arms carry out those movements using end-effectors and manipulators to perform the actual surgery on the patient. In computer-controlled systems the surgeon uses a computer to control the robotic arms and its end-effectors, though these systems can also still use telemanipulators for their input. One advantage of using the computerised method is that the surgeon does not have to be present, but can be anywhere in the world, leading to the possibility for remote surgery.

In the case of enhanced open surgery, autonomous instruments (in familiar configurations) replace traditional steel tools, performing certain actions (such as rib spreading) with much smoother, feedback-controlled motions than could be achieved by a human hand. The main object of such smart instruments is to reduce or eliminate the tissue trauma traditionally associated with open surgery without requiring more than a few minutes' training on the part of surgeons. This approach seeks to improve open surgeries, particularly cardio-thoracic, that have so far not benefited from minimally-invasive techniques.

Robotic surgery has been criticized for its expense, by one estimate costing $1,500 to $2000 more per patient.

5.9 TELESURGERY

Telesurgery, also called remote surgery, is performed by a surgeon at a site removed from the patient. Surgical tasks are directly performed by a robotic system controlled by the surgeon at the remote site. The word "telesurgery" is derived from the Greek words tele , meaning "far off," and cheirourgia , meaning "working by hand."

Remote surgery (also known as telesurgery) is the ability for a doctor to perform surgery on a patient even though they are not physically in the same location. It is a form of telepresence. A robot surgical system generally consists of one or more arms (controlled by the surgeon), a master controller (console), and a sensory system giving feedback to the user. Remote surgery combines elements of robotics, cutting edge communication technology such as high-speed data connections and elements of management information systems. While the field of robotic surgery is fairly well established, most of these robots are controlled by surgeons at the location of the surgery. Remote surgery is essentially advanced telecommuting for surgeons, where the physical distance between the surgeon and the patient is immaterial. It promises to allow the expertise of specialized surgeons to be available to patients worldwide, without the need for patients to travel beyond their local hospital.

5.10 TELECARDIOLOGY

Telecardiology is a modern medical practice, which uses the power of telecommunications to achieve remote diagnosis and treatment of heart disease. This includes coronary heart disease, chronic and acute, as well as arrhythmias, congestive cardiac failure and sudden cardiac arrest.

In this situation, doctors and other healthcare providers use electrocardiographic data, which is transmitted remotely, in real time, for interpretation by a specialist. It enables specialist care to be accessed by people in remote locations. Advancing technology is making it easier and less expensive to set up wireless or satellite networks for this purpose, increasing their effectiveness and ease.

The practice of telecardiology depends upon the availability of a specialized device, which not only takes and records a 12-lead ECG in the primary care setting, but also transmits the ECG image in the form of a sound signal over the telephone line.

At the other end, namely, the telecardiology facility, it is converted back into an image on screen. After specialists interpret it, an oral report is quickly sent, while a written summary is emailed or faxed to the patient hub. All ECGs are stored in an electronic database at the telecardiology center, to enable future comparison of ECGs for the same patient over time.

Single-lead ECG machines are available in the form of a watch-like device to enable quick monitoring when the patient needs it, while still at home. This allows for a better interpretation and diagnosis of the disease condition.

This device can store the images recorded, and transmit them once the patient reaches the GP's office. The advantage is that the patient need not wait to reach the doctor's office, but can record the ECG as and when symptoms are present.

5.11 TELEONCOLOGY

Access to quality cancer care is often unavailable not only in low- and middle-income countries but also in rural or remote areas of high-income countries. Teleoncology (oncology applications of medical telecommunications, including pathology, radiology, and other related disciplines) has the potential to enhance both access to and the quality of clinical cancer care as well as education and training. Its implementation in the developing world requires an approach tailored to priorities, resources, and needs.

References:

1. O'Carroll, P.W., Yasnoff, W.A., Ward, E., Ripp, L.H., Martin, E.L. (Eds), Public Health Informatics and Information Systems. Springer (ISBN 0-387-95474-0), 2003

2. Reinaldo Perez, Design of Medical Electronic Device, Elsevier Science, 1st Edition

Question bank

1. Explain the various parts of teleradiology system
2. With an example explain the image acquisition system
3. With a neat diagram explain the patient information system
4. Explain how the robotics surgery is taking place in hospital.
5. How telecardiology is used in hospitals. Explain in detail
6. What is meant by resolution? how resolution is helpful in image processing technique.