ADVANCE COMMUNICATION ENGINEERING

Prof. Satya Prakash Rout
Assistant Professor
TempleCity Institute of Technology & Engineering (Diploma)
Taraboi, Khordha

SATELLITE COMMUNICATION

UNIT I SATELLITE ORBITS

Prepared by

Prof. Satya Prakash Rout

ASSISTANT PROFESSOR

TEMPLECITY INSTITUTE OF TECHNOLOGY & ENGINEERING

UNIT I SATELLITE ORBITS

AIM & OBJECTIVE

- To understand the basics of satellite orbits.
- To analyze the geo stationary and non geo stationary orbits.
- To acquire the knowledge about launching procedures.

PRE-TEST MCQ

1. A television (TV) transmission is an example of which type of transmission?

   a) Simplex
   b) Half duplex
   c) Full duplex
   d) None of the above

2. What is application of satellite systems?

   a) Weather forecasting
   b) Terrestrial communication
   c) point to point communication
   d) None of the above

3. The down link frequency in the C band transponder is

   (a) 6 GHz
   (b) 4 GHz
   (c) 14 GHz
   (d) 11 GHz

THEORY

1.1 Introduction to satellite communication

Satellites are specifically made for telecommunication purpose. They are used for mobile applications such as communication to ships, vehicles, planes, hand -held terminals and for TV and radio broadcasting.

They are responsible for providing these services to an assigned region (area) on the earth. The power and bandwidth of these satellites depend upon the preferred size of the footprint, complexity of the traffic control protocol schemes and the cost of ground stations.

A satellite works most efficiently when the transmissions are focused with a desired area.

When the area is focused, then the emissions don‟t go outside that designated area and thus minimizing the interference to the other systems. This leads more efficient spectrum usage.

Satellites antenna patterns play an important role and must be designed to best cover the designated geographical area (which is generally irregular in shape).

Satellites should be designed by keeping in mind its usability for short and long term effects throughout its life time.

The earth station should be in a position to control the satellite if it drifts from its orbit it is subjected to any kind of drag from the external forces.

Applications of Satellites:

- Weather Forecasting
- Radio and TV Broadcast
- Military Satellites
- Navigation Satellites
- Global Telephone
- Connecting Remote Area
- Global Mobile Communication

1.2 Kepler's laws

1.2.1 Kepler's law Introduction

Satellites (spacecraft) orbiting the earth follow the same laws that govern the motion of the planets around the sun.

Kepler's laws apply quite generally to any two bodies in space which interact through gravitation. The more massive of the two bodies is referred to as the *primary,* the other, the *secondary* or *satellite.*

1.2.2 Kepler's First Law

Kepler's first law states that the path followed by a satellite around the primary will be an ellipse. An ellipse hast Two focal points shown as $F1$ and $F2$in Fig. 2.1. The center of mass of the two-body system, termed the *bary center,* is always center of the foci.

The semi major axis of the ellipse is denoted by a, and the semi minor axis, by b. The eccentricity e is given by

$$e = \frac{\sqrt{a^2 - b^2}}{a}$$

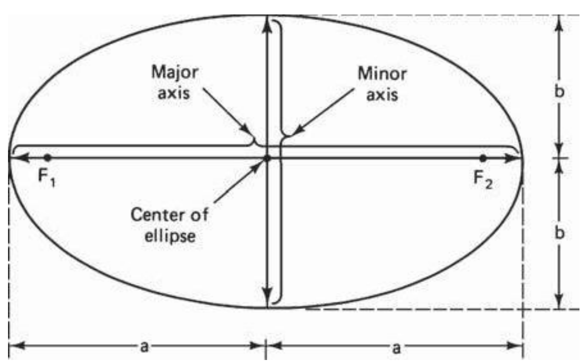


Figure 1.1 The foci $F1$ and $F2$, the semi major axis a, and the semi minor axis b of an ellipse.

1.2.3 Kepler's Second Law

Kepler's second law states that, for equal time intervals, a satellite will sweep out equal areas in its orbital plane, focused at the barycenter. Referring to Fig. 2.2, assuming the satellite travels distances $S1$ and $S2$ meters in 1 s, then the areas $A1$ and $A2$ will be equal. The average velocity in each case is $S1$ and $S2$ m/s, and because of the equal area law, it follows that the velocity at $S2$ is less than that at $S1.$

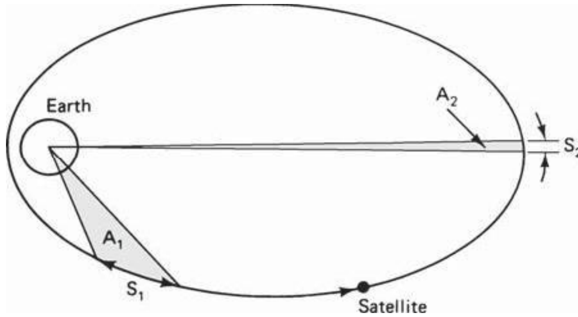


Figure 1.2 Kepler's second law. The areas A_1 and A_2 swept out in unit time are equal.

1.2.4 Kepler's Third Law

Kepler's third law states that the square of the periodic time of orbit is proportional to the cube of the mean distance between the two bodies. The mean distance is equal to the semi major axis *a*.

For the artificial satellites orbiting the earth, Kepler's third law can be written in the form

$$a^3 = \mu/n^2$$

Where n is the mean motion of the satellite in radians per second and is the earth's geocentric gravitational constant $\mu = 3.986005 \times 10^{14}\, m^3/s^2$

1.3. Newton's law:

1.3.1 Newton's first law

An object at rest will remain at rest unless acted on by an unbalanced force. An object in motion continues in motion with the same speed and in the same direction unless acted upon by an unbalanced force. This law is often called "the law of inertia".

1.3.2 Newton's second law

Acceleration is produced when a force acts on a mass. The greater the mass (of the object being accelerated) the greater the amount of force needed (to accelerate the object).

1.3.3 Newton's first law

For every action there is an equal and opposite re-action. This means that for every force there is a reaction force that is equal in size, but opposite in direction. That is to say that whenever an object pushes another object it gets pushed back in the opposite direction equally hard.

1.4. orbital parameters

Apogee: A point for a satellite farthest from the Earth. It is denoted as **ha.**

Perigee**:** A point for a satellite closest from the Earth. It is denoted as **hp.**

Line of Apsides: Line joining perigee and apogee through centre of the Earth. It is the major axis of the orbit. One-half of this line‟s length is the semi-major axis equivalents to satellite‟s mean distance from the Earth.

Ascending Node: The point where the orbit crosses the equatorial plane going from north to south.

Descending Node: The point where the orbit crosses the equatorial plane going from south to north.

Inclination: the angle between the orbital plane and the Earth‟s equatorial plane. Its measured at the ascending node from the equator to the orbit, going from East to North. Also, this angle is commonly denoted as **i.**

Line of Nodes: the line joining the ascending and descending nodes through the centre of Earth.

Prograde Orbit: an orbit in which satellite moves in the same direction as the Earth‟s rotation. Its inclination is always between 00 to 900. Many satellites follow this path as Earth‟s velocity makes it easier to lunch these satellites.

Retrograde Orbit: an orbit in which satellite moves in the same direction counter to the Earth‟s rotation.

Argument of Perigee: An angle from the point of perigee measure in the orbital plane at the Earth‟s centre, in the direction of the satellite motion.

Right ascension of ascending node: The definition of an orbit in space, the position of ascending node is specified. But as the Earth spins, the longitude of ascending node changes and cannot be used for reference. Thus for practical determination of an orbit, the longitude and time of crossing the ascending node is used.For absolute measurement, a fixed reference point in space is required.

It could also be defined as "*right ascension of the ascending node; right ascension is the angular position measured eastward along the celestial equator from the vernal equinox vector to the hour circle of the object*".

Mean anamoly: It gives the average value to the angular position of the satellite with reference to the perigee.

True anamoly: It is the angle from point of perigee to the satellite"s position, measure at the Earth"s centre.

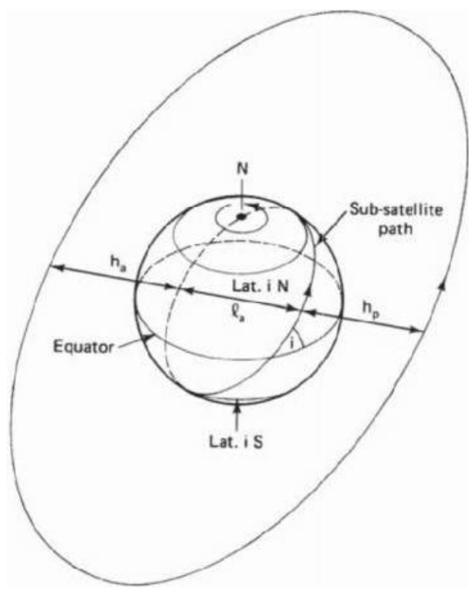


Figure1.2 Apogee height h_a, perigee height h_p, and inclination i. L_a is the line of apsides.

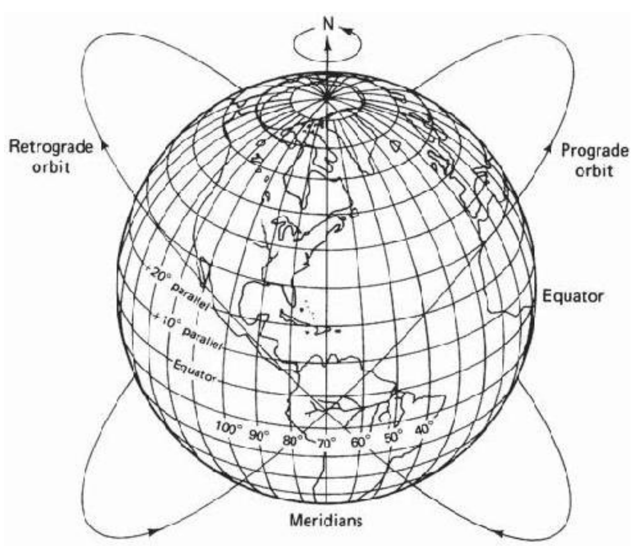


Figure 1.3(a) Prograde and retrograde orbits.

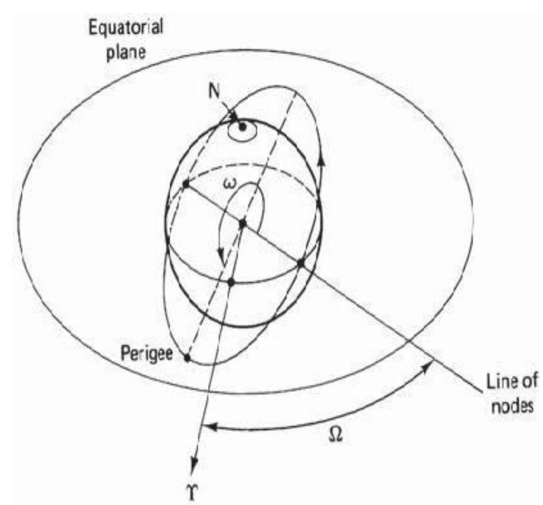


Figure.1.4 The argument of perigee w & right ascension of the ascending node Ω.

1.5. Orbital Perturbations

Theoretically, an orbit described by Kepler is ideal as Earth is considered to be a perfect sphere and the force acting around the Earth is the centrifugal force. This force is supposed to balance the gravitational pull of the earth.

In reality, other forces also play an important role and affect the motion of the satellite. These forces are the gravitational forces of Sun and Moon along with the atmospheric drag.

Effect of Sun and Moon is more pronounced on geostationary earth satellites where as the atmospheric drag effect is more pronounced for low earth orbit satellites.

1.5.1 Effects of non-Spherical Earth

As the shape of Earth is not a perfect sphere, it causes some variations in the path followed by the satellites around the primary. As the Earth is bulging from the equatorial belt, and keeping in mind that an orbit is not a physical entity, and it is the forces resulting from an oblate Earth which act on the satellite produce a change in the orbital parameters.

This causes the satellite to drift as a result of regression of the nodes and the latitude of the point of perigee (point closest to the Earth). This leads to rotation of the line of apsides. As the orbit itself is moving with respect to the Earth, the resultant changes are seen in the values of argument of perigee and right ascension of ascending node.

Due to the non-spherical shape of Earth, one more effect called as the "Satellite Graveyard" is seen. The non-spherical shape leads to the small value of eccentricity (10-5) at the equatorial plane. This causes a gravity gradient on GEO satellite and makes them drift to one of the two stable points which coincide with minor axis of the equatorial ellipse.

1.5.2 Atmospheric Drag

For Low Earth orbiting satellites, the effect of atmospheric drag is more pronounces. The impact of this drag is maximum at the point of perigee. Drag (pull towards the Earth) has an effect on velocity of Satellite (velocity reduces).

This causes the satellite to not reach the apogee height successive revolutions. This leads to a change in value of semi-major axis and eccentricity. Satellites in service are maneuvered by the earth station back to their original orbital position.

1.6 Station Keeping

In addition to having its attitude controlled, it is important that a geo-stationary satellite be kept in its correct orbital slot. The equatorial ellipticity of the earth causes geostationary satel- lites to drift slowly along the orbit, to one of two stable points, at 75°E and 105°W.

To counter this drift, an oppositely directed velocity com-ponent is imparted to the satellite by means of jets, which are pulsed once every 2 or 3 weeks.

These maneuvers are termed *east-west station-keeping maneuvers*. Satellites in the 6/4-GHz band must be kept within 0.1° of the desig- nated longitude, and in the 14/12-GHz band, within 0.05°.

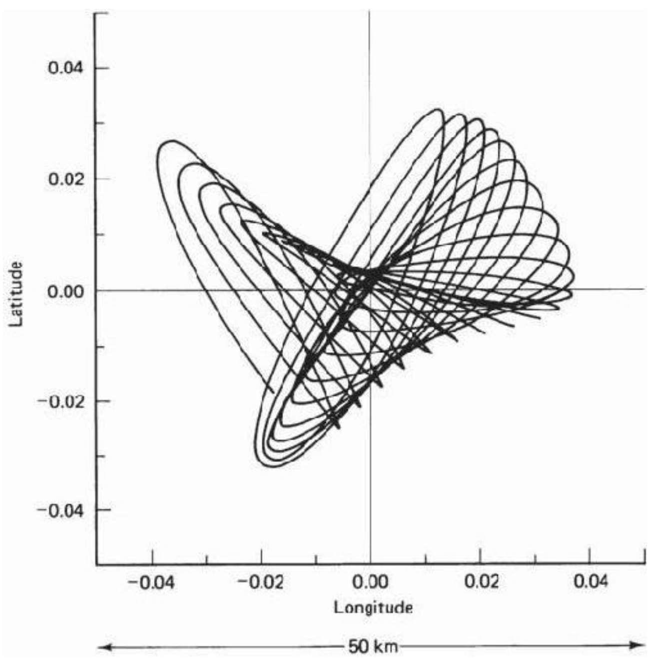


Figure 1.5 Typical satellite motion.(*CourtesyofTelesat,Canada,1983.*)

1.7. Geo stationary and Non Geo-stationary orbits

1.7.1 Geo stationary

A **geostationary** orbit is one in which a satellite orbits the earth at exactly the same speed as the earth turns and at the same latitude, specifically zero, the latitude of the equator. A satellite orbiting in a geostationary orbit appears to be hovering in the same spot in the sky, and is directly over the same patch of ground at all times.

A **geosynchronous** orbit is one in which the satellite is synchronized with the earth's rotation, but the orbit is tilted with respect to the plane of the equator. A satellite in a geosynchronous orbit will wander up and down in latitude, although it will stay over the same line of longitude. Although the terms 'geostationary' and 'geosynchronous' are sometimes used interchangeably, they are not the same technically; geostationary orbit is a subset of all possible geosynchronous orbits.

The person most widely credited with developing the concept of geostationary orbits is noted science fiction author Arthur C. Clarke (Islands in the Sky, Childhood's End, Rendezvous with Rama, and the movie 2001: a Space Odyssey). Others had earlier pointed out that bodies traveling a certain distance above the earth on the equatorial plane would remain motionless with respect to the earth's surface. But Clarke published an article in 1945's Wireless World that made the leap from the Germans' rocket research to suggest permanent manmade satellites that could serve as communication relays.

Geostationary objects in orbit must be at a certain distance above the earth; any closer and the orbit would decay, and farther out they would escape the earth's gravity altogether. This distance is 35,786 kilometers (22,236 miles) from the surface.

The first geosynchrous satellite was orbited in 1963, and the first geostationary one the following year. Since the only geostationary orbit is in a plane with the equator at 35,786 kilometers, there is only one circle around the world where these conditions obtain.

This means that geostationary 'real estate' is finite. While satellites are in no danger of bumping in to one another yet, they must be spaced around the circle so that their frequencies do not interfere with the functioning of their nearest neighbors.

Geostationary Satellites

There are 2 kinds of manmade satellites in the heavens above: One kind of satellite ORBITS the earth once or twice a day, and the other kind is called a communications satellite and it is PARKED in a STATIONARY position 22,300 miles (35,900 km) above the equator of the STATIONARY earth.

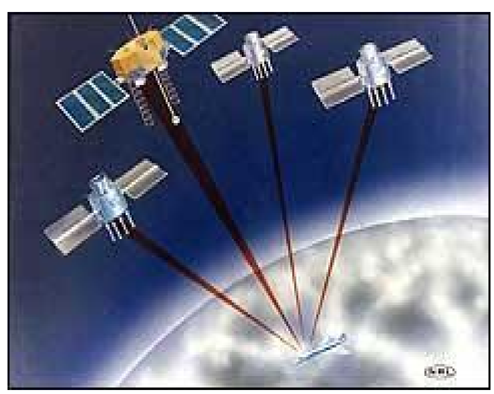
A type of the orbiting satellite includes the space shuttle and the international space station which keep a low earth orbit (LEO) to avoid the deadly Van Allen radiation belts.

The most prominent satellites in medium earth orbit (MEO) are the satellites which comprise the GLOBAL POSITIONING SYSTEM or GPS as it is called.

The Global Positioning System

The global positioning system was developed by the U.S. military and then opened to civilian use. It is used today to track planes, ships, trains, cars or literally anything that moves. Anyone can buy a receiver and track their exact location by using a GPS receiver.

GPS satellites orbit at a height of about 12,000 miles (19,300 km) and orbit the earth once every 12 hours.

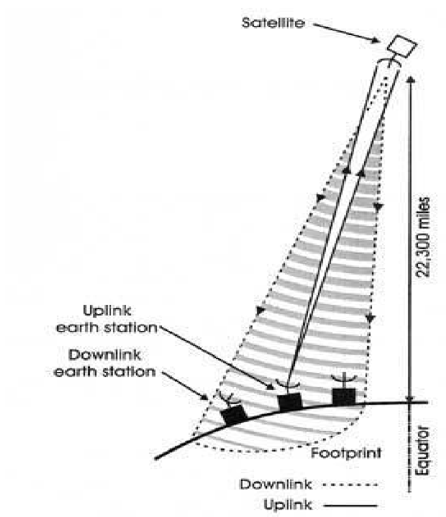


About 24 GPS satellites orbit the earth every 12 hours.

These satellites are traveling around the earth at speeds of about 7,000 mph (11,200 kph). GPS satellites are powered by solar energy. They have backup batteries onboard to keep them running in the event of a solar eclipse, when there's no solar power.

Small rocket boosters on each satellite keep them flying in the correct path. The satellites have a lifetime of about 10 years until all their fuel runs out.

At exactly 22,300 miles above the equator, the force of gravity is cancelled by the centrifugal force of the rotating universe. This is the ideal spot to park a stationary satellite.



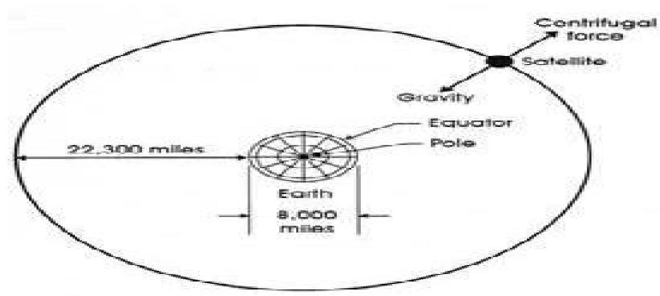


Figure. 1.6 & 1.7 At exactly 22,000 miles (35,900 km) above the equator, the earth's force of gravity is canceled by the centrifugal force of the rotating universe. .

1.7.2 Non Geo-Stationary Orbit

For the geo- stationary case, the most important of these are the gravitational fields of the moon and the sun, and the nonspherical shape of the earth.

Other significant forces are solar radiation pressure and reaction of the satellite itself to motor movement within the satellite. As a result, station-keeping maneuvers must be carried out to maintain the satel- lite within set limits of its nominal geostationary position.

An exact geostationary orbit therefore is not attainable in practice, and the orbital parameters vary with time. The two-line orbital elements are published at regular intervals.

The period for a geostationary satellite is 23 h, 56 min, 4 s, or 86,164 s. The reciprocal of this is 1.00273896 rev/day, which is about the value tabulated for most of the satellites in Fig.

Thus these satellites are *geo- synchronous,* in that they rotate in synchronism with the rotation of the earth. However, they are not geostationary. The term *geosynchronous satellite* is used in many cases instead of *geostationary* to describe these near-geostationary satellites.

It should be noted, however, that in gen- eral a geosynchronous satellite does not have to be near-geostationary, and there are a number of geosynchronous satellites that are in highly elliptical orbits with comparatively large inclinations (e.g., the Tundra satellites).

The small inclination makes it difficult to locate the position of the ascending node, and the small eccentricity makes it difficult to locate the position of the perigee.

However, because of the small inclination, the angles w and Ω can be assumed to be in the same plane.The longitude of the subsatellite point (thesatellitelongitude) is the east early rotation from the Greenwich meridian.

$$\phi_{SS} = \omega + \Omega + v - \mathrm{GST}$$

The *Greenwich sidereal time* (GST) gives the eastward position of the Greenwich meridian relative to the line of Aries, and hence the subsatellite point is at longitudeand the mean longitude of the satellite is given by

$$\phi_{SSmean} = \omega + \Omega + M - \mathrm{GST}$$

Equation(2.31)can be used to calculate the trueanomaly, and because of the small eccentricity, this can be approximated as v= $M + 2e\sin M$.

1.8 Look Angle Determination

The look angles for the ground station antenna are Azimuth and Elevation angles. They are required at the antenna so that it points directly at the satellite. Look angles are calculated by considering the elliptical orbit. These angles change in order to track the satellite.

For geostationary orbit, these angels values does not change as the satellites are stationary with respect to earth. Thus large earth stations are used for commercial communications.

For home antennas, antenna beamwidth is quite broad and hence no tracking is essential. This leads to a fixed position for these antennas.

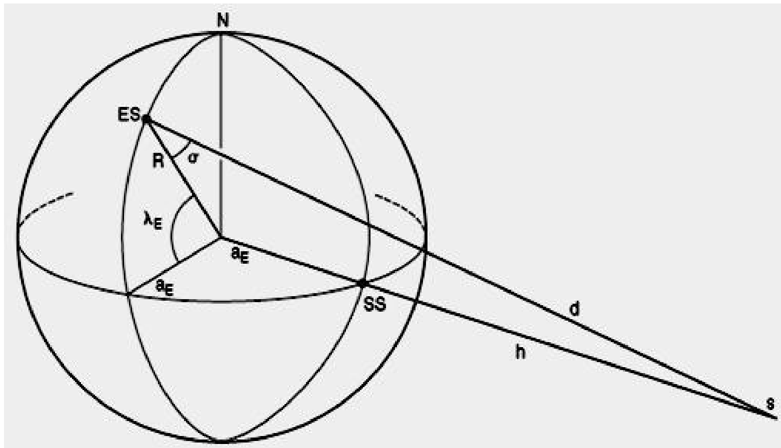


Figure 1.8: The geometry used in determining the look angles for Geostationary Satellites.

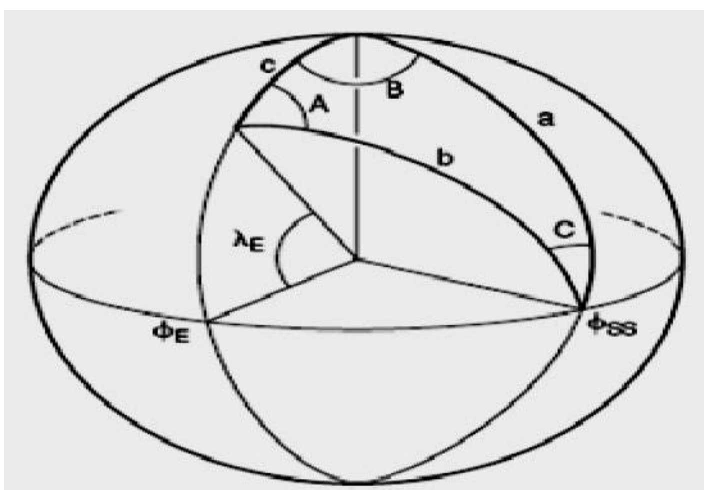


Figure 1.9: The spherical geometry related to figure 1.8

With respect to the figure 1.8 and 1.9, the following information is needed to determine the look angles of geostationary orbit.

1. Earth Station Latitude: λE
2. Earth Station Longitude: ΦE

3. Sub-Satellite Point"s Longitude: ΦSS
4. ES: Position of Earth Station
5. SS: Sub-Satellite Point
6. S: Satellite
7. d: Range from ES to S
8. ζ: angle to be determined

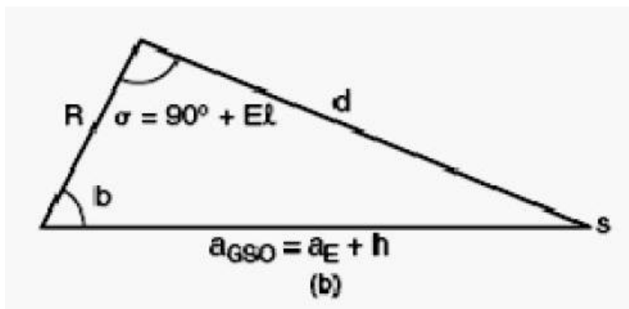


Figure 1.10: A plane triangle obtained from figure 1.8

Considering figure 3.3, it"s a spherical triangle. All sides are the arcs of a great circle. Three sides of this triangle are defined by the angles subtended by the centre of the earth.

o Side a: angle between North Pole and radius of the sub-satellite point.

o Side b: angle between radius of Earth and radius of the sub-satellite point.

o Side c: angle between radius of Earth and the North Pole.

a $=90^0$ and such a spherical triangle is called quadrantal triangle. c = $90^0 - \lambda$

Angle B is the angle between the plane containing c and the plane containing a.

Thus, B = ΦE-ΦSS

Angle A is the angle between the plane containing b and the plane containing c.

Angle C is the angle between the plane containing a and the plane containing b.

Thus, $a = 90^0$

$c = 90^0 - \lambda E$

$B = \Phi E - \Phi SS$

Thus, $b = \arccos(\cos B \cos \lambda E)$

And $A = \arcsin(\sin |B| / \sin b)$

Applying the cosine rule for plane triangle to the triangle of figure

$$d = \sqrt{R^2 + a_{GSO}{}^2 - 2Ra_{GSO} \cos b}$$

Applying the sine rule for plane triangles to the triangle of figure 3.3 allows the angle of elevation to be found:

$$El = \arccos\left(\frac{a_{GSO}}{d} \sin b\right)$$

1.9. Limits of visibility

The east and west limits of geostationary are visible from any given Earth station. These limits are set by the geographic coordinates of the Earth station and antenna elevation.

The lowest elevation is zero (in theory) but in practice, to avoid reception of excess noise from Earth. Some finite minimum value of elevation is issued. The earth station can see a satellite over a geostationary arc bounded by **+- (81.30)** about the earth station"s longitude.

1.10. Eclipse

It occurs when Earth‟s equatorial plane coincides with the plane f he Earth‟s orbit around the sun.

Near the time of spring and autumnal equinoxes, when the sun is crossing the equator, the satellite passes into sun‟s shadow. This happens for some duration of time every day.

These eclipses begin 23 days before the equinox and end 23 days after the equinox. They last for almost 10 minutes at the beginning and end of equinox and increase for a maximum period of 72 minutes at a full eclipse.

The solar cells of the satellite become non-functional during the eclipse period and the satellite is made to operate with the help of power supplied from the batteries.

A satellite will have the eclipse duration symmetric around the time t=Satellite Longitude/15 + 12 hours. A satellite at Greenwich longitude 0 will have the eclipse duration symmetric around 0/15

UTC +12hours = 00:00 UTC.

The eclipse will happen at night but for satellites in the east it will happen late evening local time.

For satellites in the west eclipse will happen in the early morning hour's local time.

An earth caused eclipse will normally not happen during peak viewing hours if the satellite is located near the longitude of the coverage area. Modern satellites are well equipped with batteries for operation during eclipse.

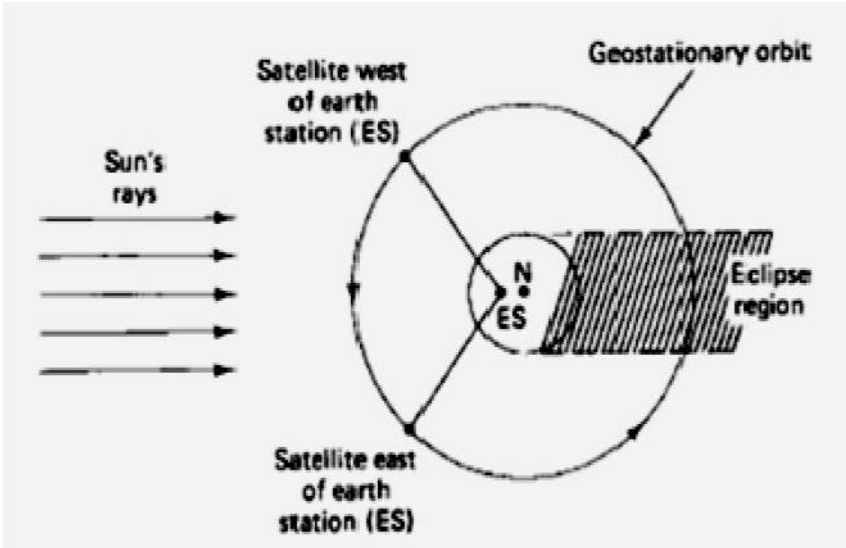


Figure 1.11(i): A satellite east of the earth station enters eclipse during daylight busy) hours at the earth station. A Satellite west of earth station enters eclipse during night and early morning hours (non busy time).

1.11. Sub satellite Point

Point at which a line between the satellite and the center of the Earth intersects the Earth's surface

Location of the point expressed in terms of latitude and longitude

If one is in the US it is common to use

- o Latitude – degrees north from equator
- o Longitude – degrees west of the Greenwich meridian

Location of the sub satellite point may be calculated from coordinates of the rotating system as:

$$L_s \quad - \quad \cos^{-1} \frac{z_r}{\sqrt{x_r^2 + y_r^2 + z_r^2}}$$

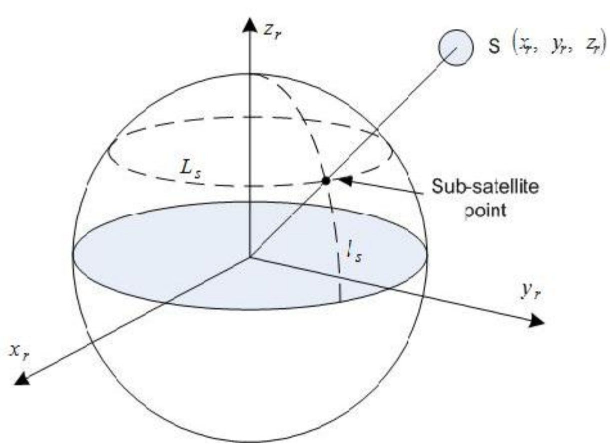


Figure 1.11(ii) **Sub satellite Point**

1.12. Sun Transit Outage

Sun transit outage is an interruption in or distortion of geostationary satellite signals caused by interference from solar radiation.

Sun appears to be an extremely noisy source which completely blanks out the signal from satellite. This effect lasts for 6 days around the equinoxes. They occur for a maximum period of 10 minutes.

Generally, sun outages occur in February, March, September and October, that is, around the time of the equinoxes.

At these times, the apparent path of the sun across the sky takes it directly behind the line of sight between an earth station and a satellite.

As the sun radiates strongly at the microwave frequencies used to communicate with satellites (C-band, Ka band and Ku band) the sun swamps the signal from the satellite.

The effects of a sun outage can include partial degradation, that is, an increase in the error rate, or total destruction of the signal.

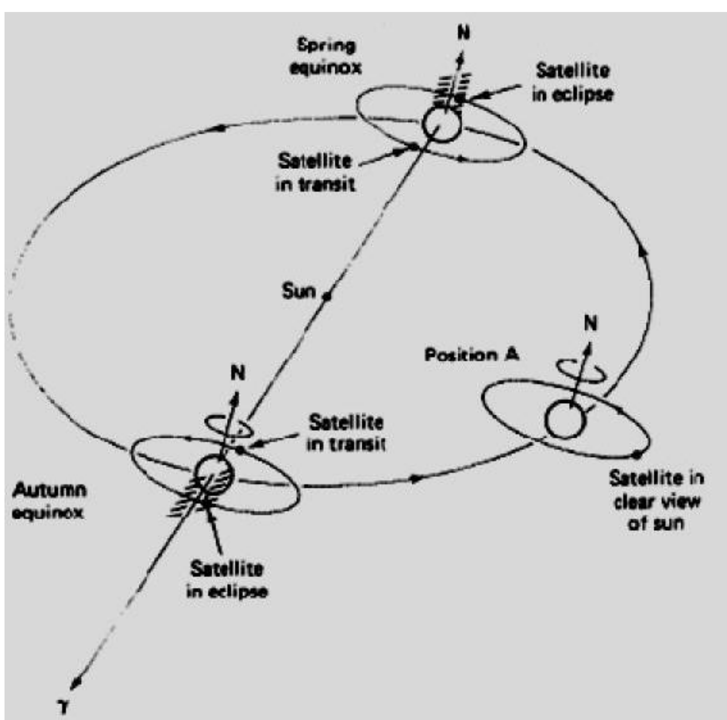


Figure 1.12 : Earth Eclipse of a Satellite and Sun transit Outage

1.13. Launching Procedures

1.13.1 Intoduction

Low Earth Orbiting satellites are directly injected into their orbits. This cannot be done incase of GEOs as they have to be positioned 36,000kms above the Earth‟s surface.

Launch vehicles are hence used to set these satellites in their orbits. These vehicles are reusable. They are also known as „Space Transportation System‟(STS).

When the orbital altitude is greater than 1,200 km it becomes expensive to directly inject the satellite in its orbit.

For this purpose, a satellite must be placed in to a transfer orbit between the initial lower orbit and destination orbit. The transfer orbit is commonly known as *Hohmann-Transfer Orbit.

1.13.2 Orbit Transfer

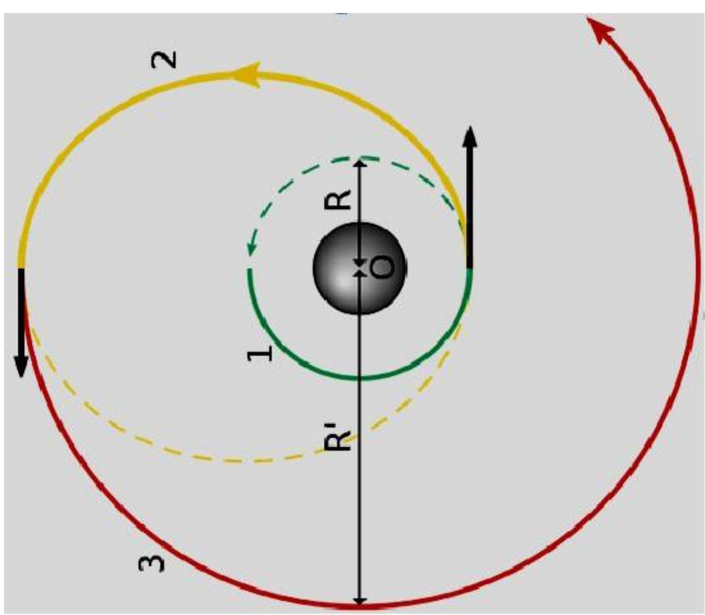


Figure 1.13*: Orbit Transfer positions*

(*About Hohmann Transfer Orbit: This manoeuvre is named for the German civil engineer who first proposed it, Walter Hohmann, who was born in 1880. He didn't work in rocketry professionally (and wasn't associated with military rocketry), but was a key member of Germany's pioneering Society for Space

Travel that included people such as Willy Ley, Hermann, and Werner von Braun. He published his concept of how to transfer between orbits in his 1925 book, The Attainability of Celestial Bodies.)

The transfer orbit is selected to minimize the energy required for the transfer. This orbit forms a tangent to the low attitude orbit at the point of its perigee and tangent to high altitude orbit at the point of its apogee.

1.14 Launch vehicles and propulsion

The rocket injects the satellite with the required thrust** into the transfer orbit. With the STS, the satellite carries a perigee kick motor*** which imparts the required thrust to inject the satellite in its transfer orbit. Similarly, an apogee kick motor (AKM) is used to inject the satellite in its destination orbit.

Generally it takes 1-2 months for the satellite to become fully functional. The Earth Station performs the Telemetry Tracking and Command**** function to control the satellite transits and functionalities.

(**Thrust: It is a reaction force described quantitatively by Newton's second and third laws. When a system expels or accelerates mass in one direction the accelerated mass will cause a force of equal magnitude but opposite direction on that system.)

Kick Motor refers to a rocket motor that is regularly employed on artificial satellites destined for a geostationary orbit. As the vast majority of geostationary satellite launches are carried out from spaceports at a significant distance away from Earth's equator.

The carrier rocket would only be able to launch the satellite into an elliptical orbit of maximum apogee 35,784-kilometres and with a non-zero inclination approximately equal to the latitude of the launch site.

TT&C: it"s a sub-system where the functions performed by the satellite control network to maintain health and status, measure specific mission parameters and processing over time a sequence of these measurement to refine parameter knowledge, and transmit mission commands to the satellite. Detailed study of TT&C in the upcoming units.

1.14.1 Transfer Orbit

It is better to launch rockets closer to the equator because the Earth rotates at a greater speed here than that at either pole. This extra speed at the equator means a rocket needs less thrust (and therefore less fuel) to launch into orbit.

In addition, launching at the equator provides an additional 1,036 mph (1,667 km/h) of speed once the vehicle reaches orbit. This speed bonus means the vehicle needs less fuel, and that freed space can be used to carry more pay load.

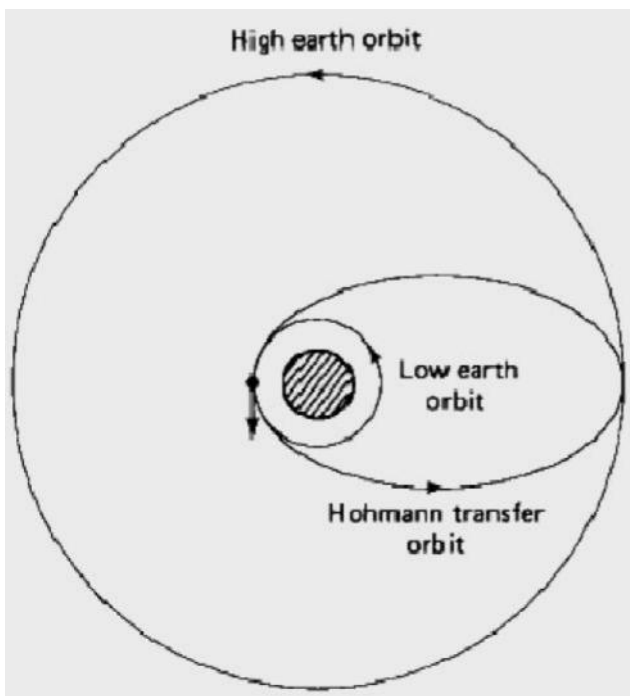


Figure 1.14: *Hohmann Transfer Orbit*

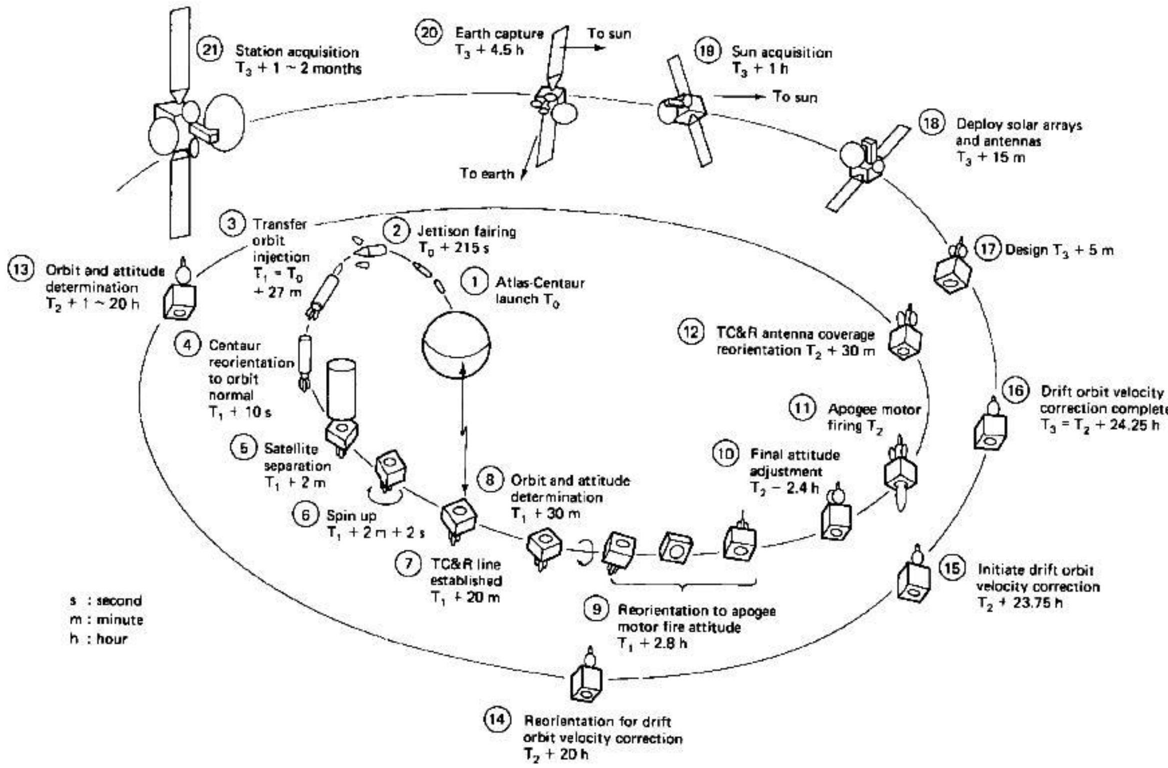


***Figure 1.15**: Launching stages of a GEO (example INTELSAT)*

Rocket launch

A **rocket launch** is the takeoff phase of the flight of a rocket. Launches for orbital spaceflights, or launches into interplanetary space, are usually from a fixed location on the ground, but may also be from a floating platform (such as the Sea Launch vessel) or, potentially, from a super heavy An-225-class airplane

Launches of suborbital flights (including missile launches), can also be from:

- a missile silo
- a mobile launcher vehicle
- a submarine
- air launch:
  - from a plane (e.g. Scaled Composites Space Ship One, Pegasus Rocket, X-15)
  - from a balloon (Rockoon, da Vinci Project (under development))
  - a surface ship (Aegis Ballistic Missile Defense System)
  - an inclined rail (e.g. rocket sled launch)

"Rocket launch technologies" generally refers to the entire set of systems needed to successfully launch a vehicle, not just the vehicle itself, but also the firing control systems, ground control station, launch pad, and tracking stations needed for a successful launch and/or recovery.

Orbital launch vehicles commonly take off vertically, and then begin to progressively lean over, usually following a gravity turn trajectory.

Once above the majority of the atmosphere, the vehicle then angles the rocket jet, pointing it largely horizontally but somewhat downwards, which permits the vehicle to gain and then maintain altitude while increasing horizontal speed. As the speed grows, the vehicle will become more and more horizontal until at orbital speed, the engine will cut off.

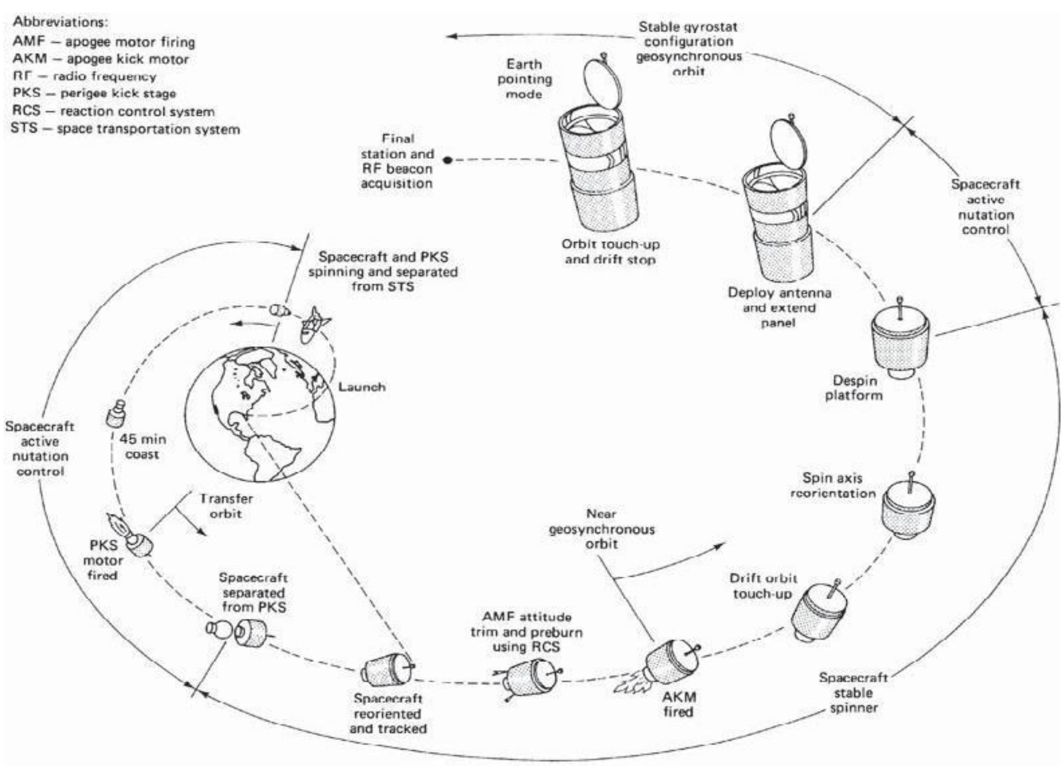


Figure 1.16 STS-7/Anik C2 mission scenario. (*From Anik C2 Launch Handbook; courtesy of Telesat, Canada.*)

APPLICATIONS

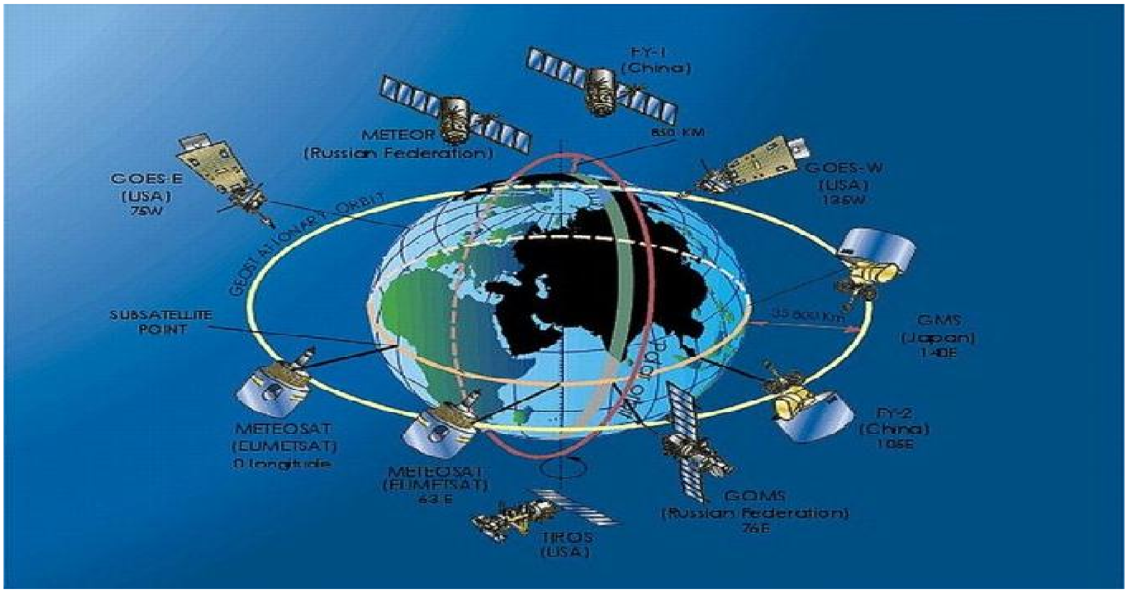


Figure example of geostationary satellites

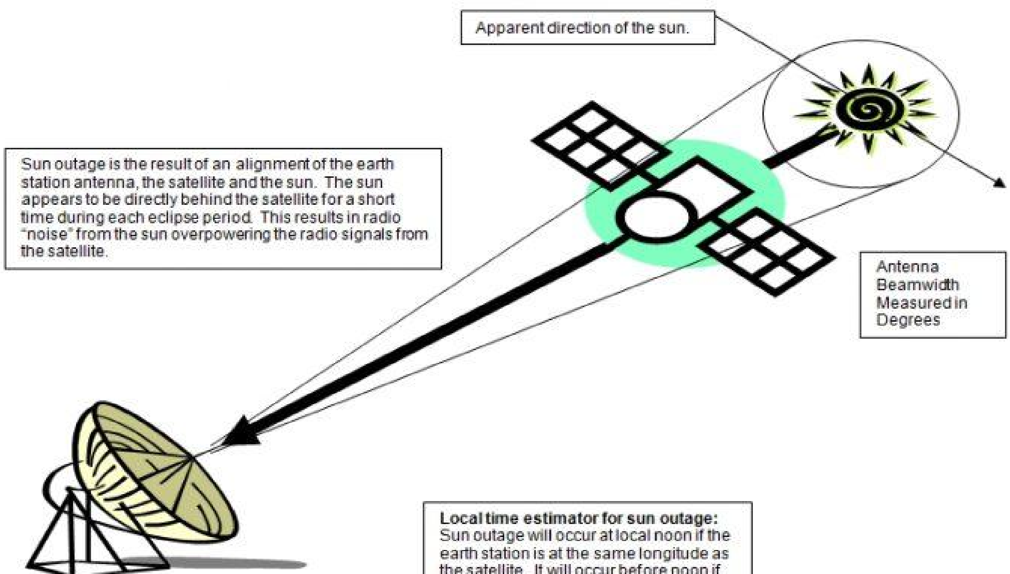


Figure example of sun transit outage

POST TEST MCQ

1. Kepler's first law states

a) The path followed by a satellite around the primary will be an ellipse.

b) The path followed by a satellite around the primary will be an circle.

c) The path followed by a satellite around the primary will be an sphere

d) None of the above

2. INTELSAT stands?

a) **International Telecommunications Satellite**

b) India Telecommunications Satellite

c) Inter Telecommunications Satellite

d) None of the above

3. The carrier to noise ratio for a satellite depends upon

(a) Effective Isotropic Radiated power

(b) Bandwidth.

(c) Free space path losses

(d) All of them

4. Mention the different services of satellite systems.

a) Broadcasting satellite services

b) Signal transmission

c) Information transmission

d) None of the above

5. Calculate the radius of a circular orbit for which the period is 1 day?

a) 42.241Km

b) 42.241m

c) 4.241Km

d) 2.241K

6. The period of a satellite, the time required for a satellite to make a complete trip around the Earth, is determined by _________ law.

a) Kepler's

b) Newton's

c) Ohm's

d) none of the above

7. Kepler's second law states

a) If t2-t1=t4-t3, then A12=A34.

b) If t2+t1 = t4+t3, then A12=A34.

c) If t2/t1=t4/t3, then A12=A34.

d) The path followed by a satellite around the primary will be an ellipse

8. Apogee is

a) The point farthest from earth

b) The point nearest from earth

c) The point smallest from earth

d) None of the above

9. Perigee is

a) The point farthest from earth

b) The point longest from earth

c) The point closest approach to earth

d) None of the above

10. True anomaly is

a) the angle from perigee to the satellite position, measured at the earth's center.

b) The point longest from earth

c) The point closest approach to earth

d) None of the above

11. ____________ is s a loss of power of a satellite downlink signal due to earth's atmosphere.

a) Atmospheric loss

b) Path loss

c) Radiation loss

d) RFI

12. Collects very weak signals from a broadcast satellite

a) Helical antenna

b) Satellite dish

c) LNA

d) TWT

13. As the height of a satellite orbit gets lower, the speed of the satellite ___________.

a) Increases

b) Decreases

c) Remains the same

d) None of the above

14. The term Eclipse is defined as

a) During equinox periods, the earth the sun & the satellite are in alignment with the result that earth' s shadow eclipses that satellite & the sunlight fails to reach the satellite solar cells.

b) During equinox periods, the earth the sun & the satellite are in alignment with the result that earth' s shadow eclipses that satellite & the sunlight success to reach the satellite solar cells.

c) a & b

d) None of above

15. Why are VHF, UHF, and microwave signals used in satellite communication?

a) More bandwidth

b) More spectrum space

c) Are not diffracted by the ionosphere

d) Economically viable

16. What is the reason for shifting from c band to ku band in satellite communication?

a) Lesser attenuation

b) Less power requirements

c) More bandwidth

d) Overcrowding

17. Which of the following bands cannot be used for satellite communication?

a) MF

b) Ku

c) X

d) C

18. The eclipse effect is noticeable for periods of ____________________

(a) about four weeks & the maximum daily eclipse duration is about 1.20 hours.

(b) about one weeks & the maximum daily eclipse duration is about 12 hours.

(c) **about five weeks & the maximum daily eclipse duration is about 1.20 hours.**

(d) about two weeks & the maximum daily eclipse duration is about 1.20 hours.

19. Which technique uses spot beam antennas to divide the area covered by the satellite into smaller segments?

a) Spatial isolation

b) Frequency reuse

c) Multiplexing

d) Modulation

CONCLUSION

- ❖ In this unit, analysis of the satellite orbits, geo stationary and non Geo-stationary orbits was discussed.
- ❖ The Kepler's Laws, Newton's law were elaborated.
- ❖ The Sub satellite point –Sun transit outage-Launching Procedures - launch vehicles and propulsion was discussed.

REFERENCES

1. Wilbur L.Pritchard, Hendri G. Suyderhoud, Robert A. Nelson, "Satellite Communication Systems Engineering", Prentice Hall/Pearson, 2007.
2. N. Agarwal, "Design of Geosynchronous Space Craft", Prentice Hall, 1986.
3. Bruce R. Elbert, "Satellite Communication Applications", Hand Book, Artech House Bostan London,1997.
4. Tri T. Ha, "Digital Satellite Communication", II nd edition, 1990.
5. Robert G. Winch, "Telecommunication Transmission Systems", Mc Graw-Hill Book Co., 1983.
6. Brian Ackroyd, "World Satellite Communication and earth station Design", BSP professional Books, 1990.
7. G.B.Bleazard, "Introducing Satellite communications", NCC Publication, 1985.
8. M.Richharia, "Satellite Communication Systems-Design Principles", Macmillan 2003.
9. Emanuel Fthenakis, "Manual of Satellite Communications", Mc Graw Hill Book Co., 1984

ASSIGNMENT

1. Explain the Kepler's law of planetary motion and how are they applicable to the geostationary satellite.
2. What is orbit ? Derive an expression for the equation of satellite orbit.
3. What is meant by look angles ? Explain them with reference to a geostationary satellite and earth station.
4. Explain briefly the orbital parameters required to determine a satellite's orbit.
5. Explain the concept of earth coverage and slant range for geostationary satellite? What are the maximum values of these parameter.
6. Explain the concept and significance of station keeping. What are N-S and E-W station keeping.
7. Explain how a satellite is placed into geostationary orbit from earth ?
8. What is meant by /orbit perturbations? Explain in brief.
9. Differentiate between geosynchronous and geostationary orbit.
10. Briefly discuss the various types of orbits.

SATELLITE COMMUNICATION

UNIT II SPACE SEGMENT AND SATELLITE LINK DESIGN

Prepared by

Dr.M.A.ARCHANA

AP/ECE,SCSVMV

UNIT II SPACE SEGMENT AND SATELLITE LINK DESIGN

AIM &OBJECTIVE

- To understand the satellite segment and earth segment.
- To analyze the Satellite Uplink and Downlink.
- To understand the G/T ratio-Performance Impairments-System noise.

PRE -TEST MCQ

1. For satellite communication, standard Earth stations have antenna diameters in the range of ____________________ metre.

a) 27.5 to 30

b) 10 to 15

c) 30 to 50

d) 5 to 10

2. In satellite communication, frequency modulation is used because satellite channel has

a) small bandwidth and negligible noise

b) large bandwidth and severe noise

c)maximum bandwidth and minimum noise

d) high modulation index

3. The noise temperature of sky is about __________ °K.

a) 100

b) 273

c) 0

d) 30

2.1 Spacecraft Technology- Structure

A satellite communications system can be broadly divided into two segments—a ground segment and a space segment.

The space segment will obviously include the satellites, but it also includes the ground facilities needed to keep the satellites operational, these being referred to as the *tracking, telemetry, and command* (TT&C) facilities. In many networks it is common practice to employ a ground station solely for the purpose of TT&C.

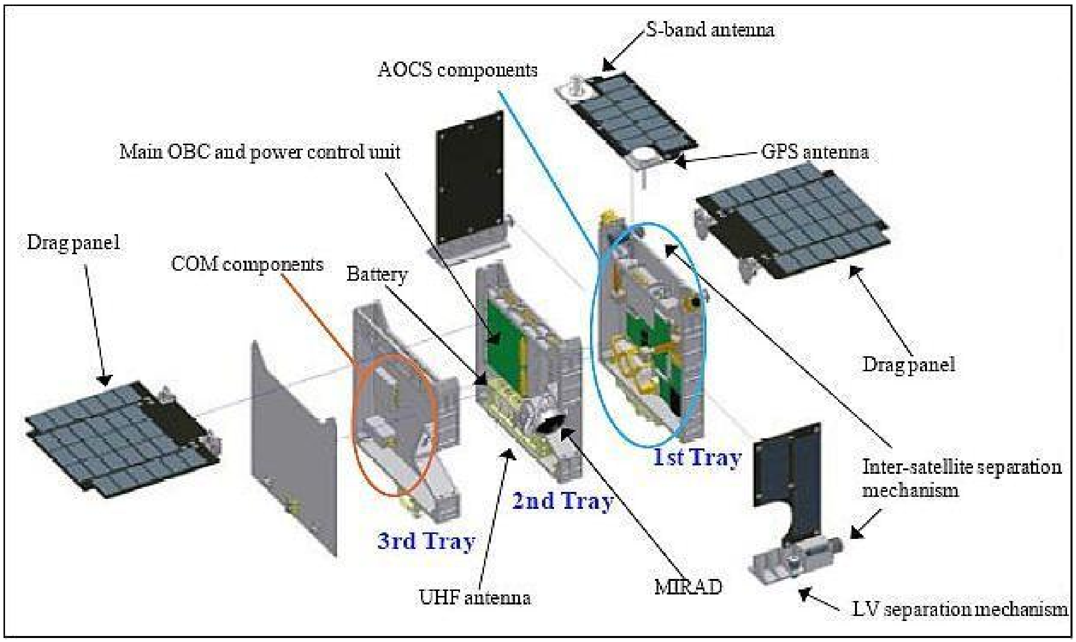


Figure 2.1 (a) Satellite Structure

The equipment carried aboard the satellite also can be classified according to function. The *payload* refers to the equipment used to pro- vide the service for which the satellite has been launched.

In a communications satellite, the equipment which provides the con- necting link between the satellite's transmit and receive antennas is referred to as the *transponder*. The transponder forms one of the main sections of the payload, the other being the antenna subsystems.In this chapter the main characteristics of certain bus systems and payloads are described.

2.2 The Power Supply

The primary electrical power for operating the electronic equipment is obtained from solar cells. Individual cells can generate only small amounts of power, and therefore, arrays of cells in series-parallel connection are required.

Figure shows the solar cell panels for the HS 376 satellite manufactured by Hughes Space and Communications Company.

In geostationary orbit the telescoped panel is fully extended so that both are exposed to sun- light. At the beginning of life, the panels produce 940 W dc power, which may drop to 760 W at the end of 10 years.

During eclipse, power is provided by two nickel-cadmium (Ni-Cd) long-life batteries, which will deliver 830 W. At the end of life, battery recharge time is less than 16 h.

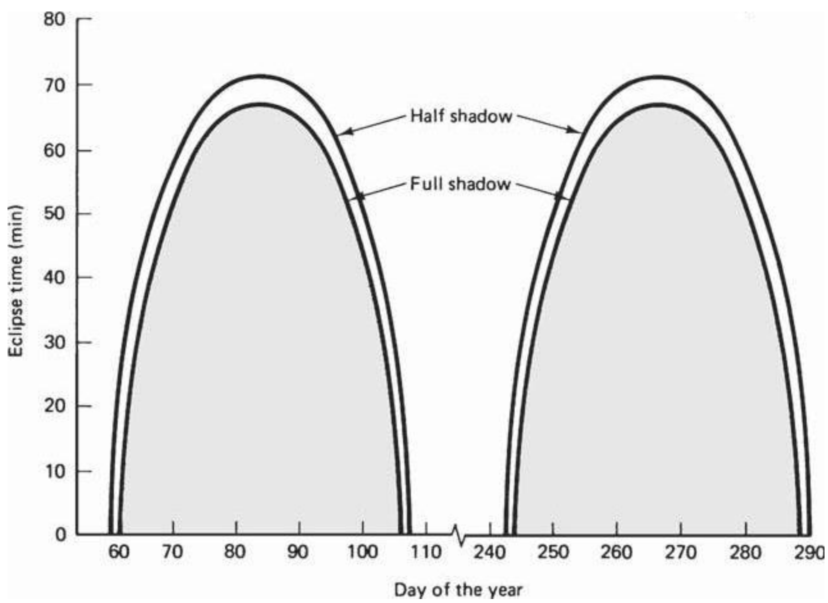


Figure 2.1.(b) Satellite eclipse time as a function of the current day of the year. (*Courtesy of Spilker, 1977. Reprinted by permission of Prentice-Hall, Englewood Cliffs, NJ.*)

capacity of cylindrical and solar-sail satellites, the cross-over point is esti- mated to be about 2 kW, where the solar-sail type is more economical than the cylindrical type (Hyndman, 1991).

2.3 Attitude Control & Orbit Control

The *attitude* of a satellite refers to its orientation in space. Much of the equipment carried aboard a satellite is there for the purpose of control- ling its attitude. Attitude control is necessary, for example, to ensure that directional antennas point in the proper directions.

In the case of earth environmental satellites, the earth-sensing instruments must cover the required regions of the earth, which also requires attitude control. A number of forces, referred to as *disturbance torques*, can alter the attitude, some examples being the gravitational fields of the earth and the moon, solar radiation, and meteorite impacts.

Attitude control must not be con- fused with station keeping, which is the term used for maintaining a satellite in its correct orbital position, although the two are closely related.

To exercise attitude control, there must be available some measure of a satellite's orientation in space and of any tendency for this to shift. In one method, infrared sensors, referred to as *horizon detectors*, are used to detect the rim of the earth against the background of space.

With the use of four such sensors, one for each quadrant, the center of the earth can be readily established as a reference point.

Usually, the attitude-control process takes place aboard the satellite, but it is also possible for control signals to be transmitted from earth, based on attitude data obtained from the satellite.

Also, where a shift in attitude is desired, an *attitude maneuver* is executed. The control signals needed to achieve this maneuver may be transmitted from an earth station.

Controlling torques may be generated in a number of ways. *Passive attitude control* refers to the use of mechanisms which stabilize the satellite without putting a drain on the satellite's energy supplies; at most, infrequent use is made of these supplies, for example, when thruster jets are impulsed to provide corrective torque. Examples of passive attitude control are *spin stabilization* and *gravity gradient sta- bilization*.

The other form of attitude control is *active control*. With active attitude control, there is no overall stabilizing torque present to resist the disturbance torques. Instead, corrective torques are applied as required in response to disturbance torques. Methods used to generate active control torques include momentum wheels, electromagnetic coils, and mass expulsion devices, such as gas jets and ion thrusters.

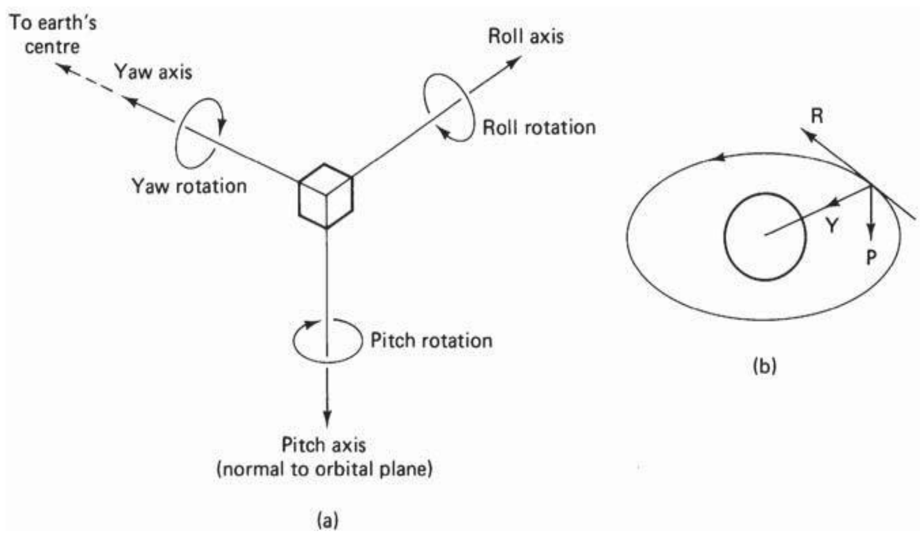


Figure 2.2 (*a*) Roll, pitch, and yaw axes. The yaw axis is directed toward the earth's center, the pitch axis is normal to the orbital plane, and the roll axis is perpendicular to the other two. (*b*) RPY axes for the geostationary orbit. Here, the roll axis is tangential to the orbit and lies along the satellite velocity vector.

The three axes which define a satellite's attitude are its *roll, pitch*, and *yaw* (RPY) axes. These are shown relative to the earth in Fig. 7.4. All three axes pass through the center of gravity of the satellite. For an equatorial orbit, movement of the satellite about the roll axis moves the antenna footprint north and south; movement about the pitch axis moves the footprint east and west; and movement about the yaw axis rotates the antenna footprint.

2.3.1 Spinning satellite stabilization

Spin stabilization may be achieved with cylindrical satellites. The satellite is constructed so that it is mechanically balanced about one partic- ular axis and is then set spinning around this axis. For geostationary satellites, the spin axis is adjusted to be parallel to the N-S axis of the earth, as illustrated in Fig. 7.5. Spin rate is typically in the range of 50 to 100 rev/min. Spin is initiated during the launch phase by means of small gas jets.

In the absence of disturbance torques, the spinning satellite would maintain its correct attitude relative to the earth. Disturbance torques are generated in a number of ways, both external and internal to the satellite.

Solar radiation, gravitational gradients, and meteorite impacts are all examples of external forces which can give rise to disturbance torques. Motor-bearing friction and the movement of satellite elements such as the antennas also can give rise to disturbance torques. The

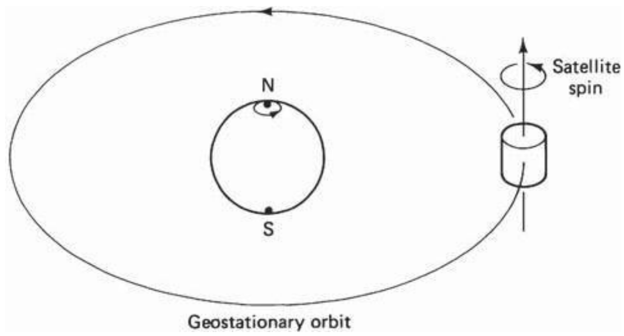


Figure 2.3 Spin stabilization in the geostationary orbit. The spin axis lies along the pitch axis, parallel to the earth's N-S axis.

overall effect is that the spin rate will decrease, and the direction of the angular spin axis will change. Impulse-type thrusters, or jets, can be used to increase the spin rate again and to shift the axis back to its cor- rect N-S orientation.

Nutation, which is a form of wobbling, can occur as a result of the disturbance torques and/or from misalignment or unbalance of the control jets. This nutation must be damped out by means of energy absorbers known as *nutation dampers*.

The antenna feeds can therefore be connected directly to the transponders without the need for radiofrequency (rf) rotary joints, while the complete platform is despun. Of course, control signals and power must be transferred to the despun section, and a mechanical bearing must be provided.

The complete assembly for this is known as the *bearing and power transfer assembly* (BAPTA). Figure 2.4 shows a photograph of the internal structure of the HS 376.

Certain dual-spin spacecraft obtain spin stabilization from a spinning fly-wheel rather than by spinning the satellite itself. These flywheels are termed *momentum wheels*, and their average momentum is referred to as *momentum bias*

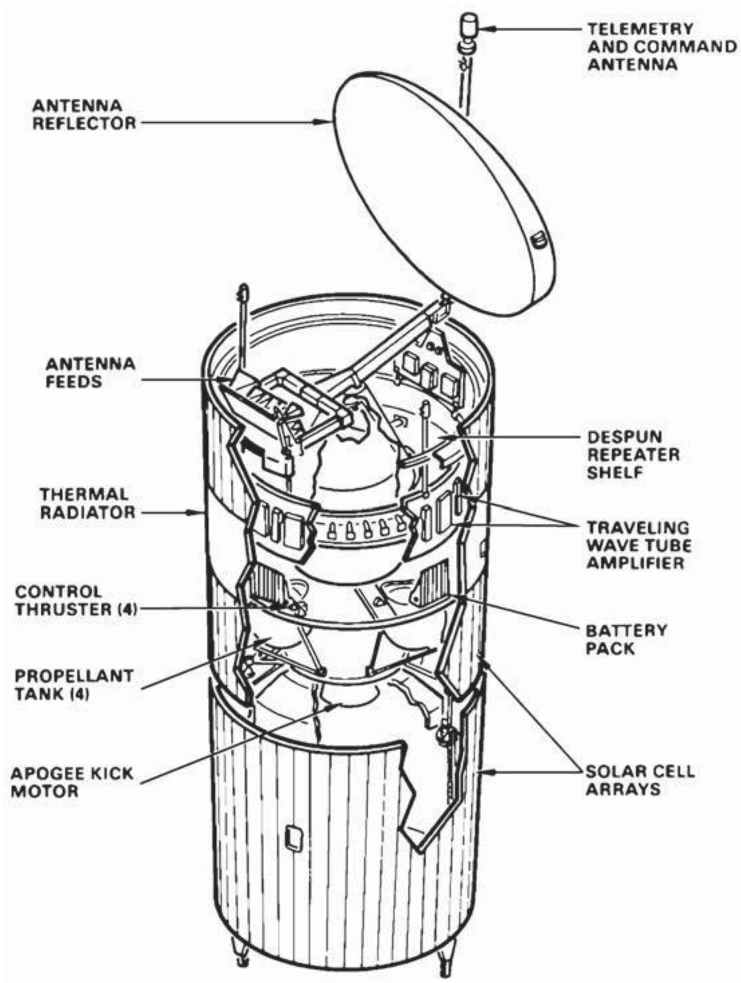


Figure 2.4 HS 376 spacecraft. (*Courtesy of Hughes Aircraft Company Space and Communication Group.*)

2.3.2 Momentum wheel stabilization

In the previous section the gyroscopic effect of a spinning satellite was shown to provide stability for the satellite attitude.

Stability also can be achieved by utilizing the gyroscopic effect of a spinning flywheel, and this approach is used in satellites with cube-like bodies (such as shown in Fig. and the INTELSAT V type satellites shown in Fig. These are known as *body-stabilized* satellites.

The complete unit, termed a momentum wheel, consists of a flywheel, the bearing assembly, the casing, and an electric drive motor with associated electronic con- trol circuitry.

The flywheel is attached to the rotor, which consists of a permanent magnet providing the magnetic field for motor action. The stator of the motor is attached to the body of the satellite.

Thus the motor provides the coupling between the flywheel and the satellite structure. Speed and torque control of the motor is exercised through the currents fed to the stator.

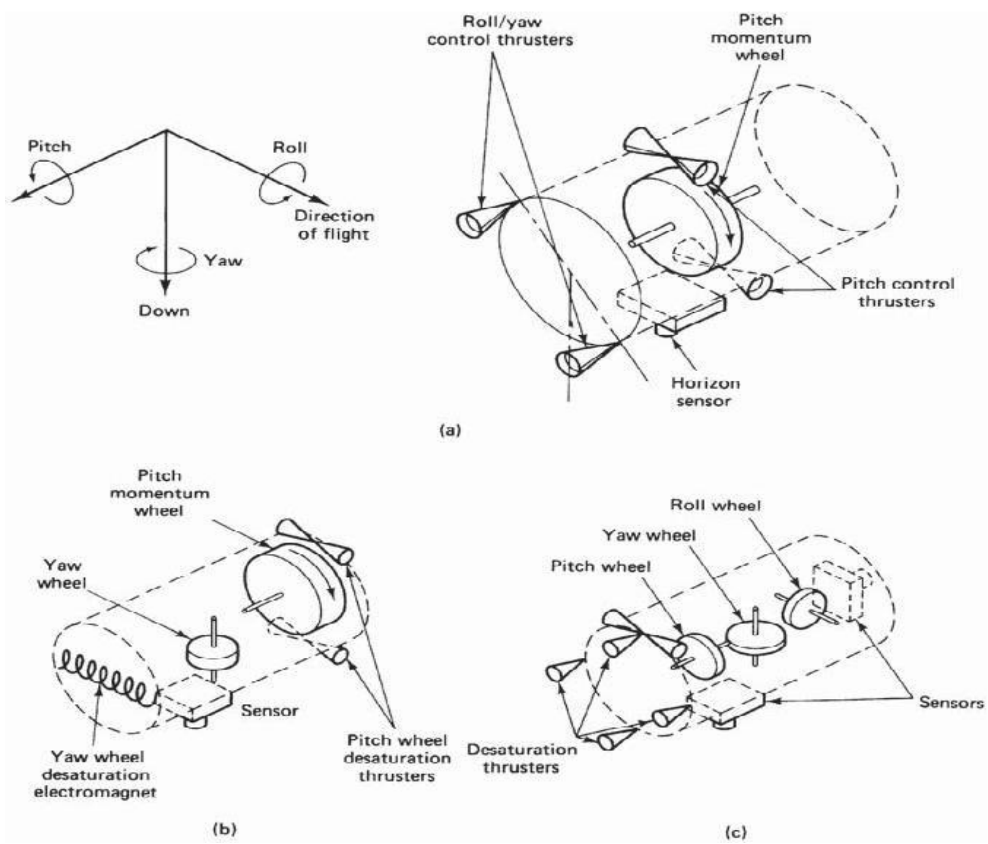


Figure 2.5 Alternative momentum wheel stabilization systems: (*a*) one-wheel, (*b*) two- wheel, (*c*) three-wheel.

When a momentum wheel is operated with zero momentum bias, it is generally referred to as a *reaction wheel*. Reaction wheels are used in three-axis stabilized systems. Here, as the name suggests, each axis is stabilized by a reaction wheel, as shown in Fig. 7.8*c*. Reaction wheels can also be combined with a momentum wheel to provide the control needed (Chetty, 1991).

Random and cyclic disturbance torques tends to produce zero momentum on average. However, there will always be some disturbance torques that causes a cumulative increase in wheel momentum, and eventually at some point the wheel *saturates*.

In effect, it reaches its maximum allowable angular velocity and can no longer take in any more momentum. Mass expulsion devices are then used to unload the wheel, that is, remove momentum from it (in the same way a brake removes energy from a moving vehicle). Of course, operation of the mass expulsion devices consumes part of the satellite's fuel supply.

2.4 Thermal Control and Propulsion

Satellites are subject to large thermal gradients, receiving the sun's radiation on one side while the other side faces into space. In addition, thermal radiation from the earth and the earth's *albedo*, which is the fraction of the radiation falling on earth which is reflected, can be sig- nificant for low-altitude earth-orbiting satellites, although it is negligi- ble for geostationary satellites.

Equipment in the satellite also generates heat which has to be removed. The most important consideration is that the satellite's equipment should operate as nearly as possible in a stable temperature environment. Various steps are taken to achieve this. Thermal blankets and shields may be used to provide insulation. Radiation mirrors are often used to remove heat from the communications payload.

The mirrored thermal radiator for the Hughes HS 376 satellite can be seen in Fig. These mirrored drums surround the communications equipment shelves in each case and pro- vide good radiation paths for the generated heat to escape into the surrounding space.

One advantage of spinning satellites compared with body-stabilized is that the spinning body provides an averaging of the temperature extremes experienced from solar flux and the cold back- ground of deep space.

In order to maintain constant temperature conditions, heaters may be switched on (usually on command from ground) to make up for the heat reduction which occurs when transponders are switched off. The INTELSAT VI satellite used heaters to maintain propulsion thrusters and line temperatures (Pilcher, 1982).

2.5 Communication Payload & Supporting Subsystems

The physical principle of establishing communication connections between remote communication devices dates back to the late 1800s when scientists were beginning to understand electromagnetism and discovered that electromagnetic (EM) radiation (also called EM waves) generated by one device can be detected by another located at some distance away.

By controlling certain aspect s of the radiation (through a process called modulation , explained in Section 4.4), useful information can be embedded in the EM waves and transmitted from one device to another.

The second major module is the communication payload, which is made up of transponders. A transponder is capable of :

Receiving uplinked radio signals from earth satellite transmission stations (antennas).
Amplifying received radio signals
Sorting the input signals and directing the output signals through input/output signal multiplexers to the proper downlink antennas for retransmission to earth satellite receiving stations (antennas).

2.6 TT&C Subsystem

The TT&C subsystem performs several routine functions aboard the spacecraft. The telemetry, or telemetering, function could be interpreted as *measurement at a distance*. Specifically, it refers to the overall oper- ation of generating an electrical signal proportional to the quantity being measured and encoding and transmitting this to a distant station, which for the satellite is one of the earth stations.

Data which are trans- mitted as telemetry signals include attitude information such as that obtained from sun and earth sensors; environmental information such as the magnetic field intensity and direction, the frequency of meteorite impact, and so on; and spacecraft information such as temperatures, power supply voltages, and stored-fuel pressure.

Telemetry and command may be thought of as complementary functions. The telemetry subsystem transmits information about the satellite to the earth station, while the command subsystem receives command signals from the earth station, often in response to telemetered information. The command subsystem demodulates and, if necessary, decodes the command signals and routes these to the appropriate equipment needed to exe- cute the necessary action.

Thus attitude changes may be made, communication transponders switched in and out of circuits, antennas redirected, and station-keeping maneuvers carried out on command. It is clearly important to prevent unauthorized commands from being received and decoded, and for this reason, the command signals are often encrypted.

Encrypt is derived from a Greek word *kryptein*, meaning *to hide*, and represents the process of concealing the command signals in a secure code. This differs from the normal process of encoding which converts characters in the command signal into a code suitable for transmission.

Tracking of the satellite is accomplished by having the satellite transmit beacon signals which are received at the TT&C earth stations.

Tracking is obviously important during the transfer and drift orbital phases of the satellite launch. Once it is on station, the position of a geostationary satellite will tend to be shifted as a result of the various disturbing forces, as described previously.

Therefore, it is necessary to be able to track the satellite's movement and send correction signals as required.

2.6.1 Transponders

A transponder is the series of interconnected units which forms a single communications channel between the receive and transmit antennas in a communications satellite.

Some of the units utilized by a transponder in a given channel may be common to a number of transponders. Thus, although reference may be made to a specific transponder, this must be thought of as an equipment *channel* rather than a single item of equipment.

Before describing in detail the various units of a transponder, the overall frequency arrangement of a typical C-band communications satellite will be examined briefly. The bandwidth allocated for C-band service is 500 MHz, and this is divided into subbands, one transponder.

A typical transponder bandwidth is 36 MHz, and allowing for a 4-MHz guardband between transponders, 12 such transponders can be accommodated in the 500-MHz bandwidth.

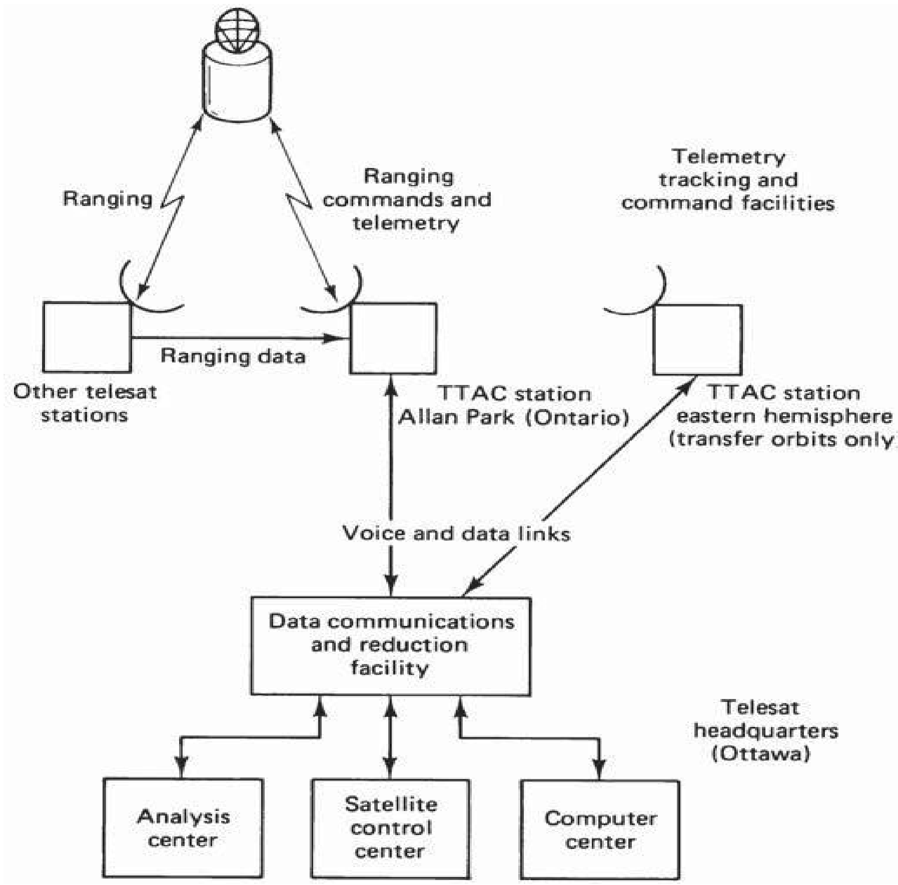
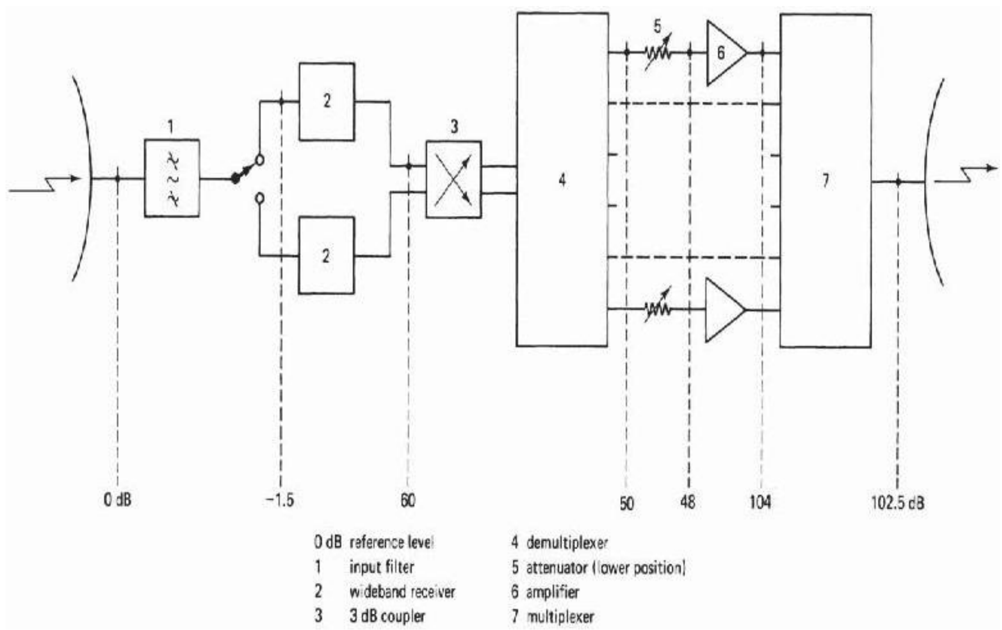


Figure 2.8 Satellite control system. (*Courtesy of Telesat Canada, 1983.*)

By making use of *polar- ization isolation*, this number can be doubled. Polarization isolation refers to the fact that carriers, which may be on the same frequency but with opposite senses of polarization, can be isolated from one another by receiving antennas matched to the incoming polarization.

With linear polarization, vertically and horizontally polarized carriers can be separated in this way, and with circular polarization, left-hand circular and right-hand circular polarizations can be separated.

Because the carriers with opposite senses of polarization may overlap in frequency, this technique is referred to as *frequency reuse*. Figure 2.9 shows part of the frequency and polarization plan for a C-band communications satellite.



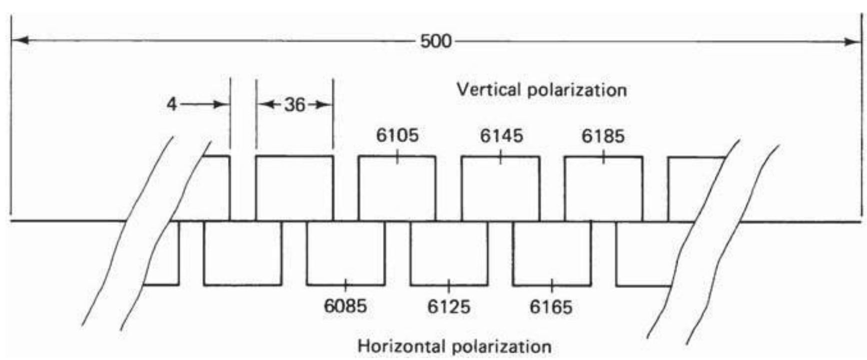


Figure 2.9 Section of an uplink frequency and polarization plan. Numbers refer to frequency in megahertz.

Frequency reuse also may be achieved with spot-beam antennas, and these may be combined with polarization reuse to provide an effective bandwidth of 2000 MHz from the actual bandwidth of 500 MHz.

For one of the polarization groups, Fig. 2.9 shows the channeling scheme for the 12 transponders in more detail. The incoming, or uplink, frequency range is 5.925 to 6.425 GHz.

The frequency conversion shifts the carriers to the downlink frequency band, which is also 500 MHz wide, extending from 3.7 to 4.2 GHz. At this point the signals are channelized into frequency bands which represent the individual transponder bandwidths.

2.6.2 The wideband receiver

The wideband receiver is shown in more detail in Fig. 2.10. A duplicate receiver is provided so that if one fails, the other is automatically switched in. The combination is referred to as a *redundant receiver*, meaning that although two are provided, only one is in use at a given time.

The first stage in the receiver is a *low-noise amplifier* (LNA). This amplifier adds little noise to the carrier being amplified, and at the same time it provides sufficient amplification for the carrier to override the higher noise level present in the following mixer stage.

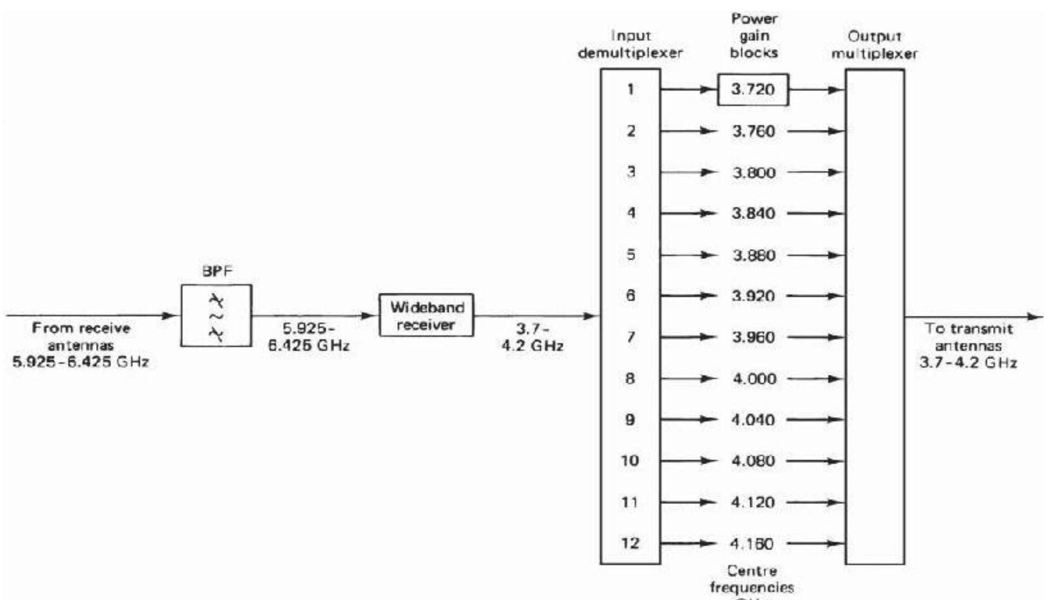


Figure 2.10 Satellite transponder channels

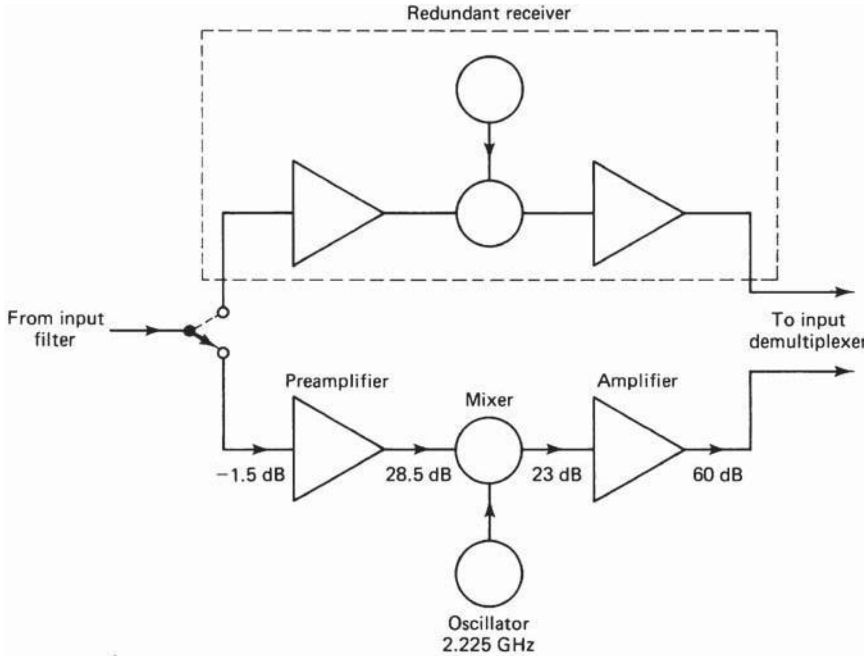


Figure 2.11 Satellite wideband receiver. (*Courtesy of CCIR, CCIR Fixed Satellite Services Handbook, final draft 1984.*)

involving noise, it is usually more convenient to refer all noise levels to the LNA input, where the total receiver noise may be expressed in terms of an equivalent noise temperature.

In a well-designed receiver, the equivalent noise temperature referred to the LNA input is basically that of the LNA alone. The overall noise temperature must take into account the noise added from the antenna, and these calculations are presented in detail in Chap. 12. The equivalent noise temperature of a satellite receiver may be on the order of a few hundred kelvins.

The LNA feeds into a mixer stage, which also requires a *local oscillator* (LO) signal for the frequency-conversion process.

With advances in *field-effect transistor* (FET) technology, FET amplifiers, which offer equal or better performance, are now available for both bands. Diode mixer stages are used.

The amplifier following the mixer may utilize *bipolar junction transistors* (BJTs) at 4 GHz and FETs at 12 GHz, or FETs may in fact be used in both bands.

2.6.3 The input demultiplexer

The input demultiplexer separates the broadband input, covering the frequency range 3.7 to 4.2 GHz, into the transponder frequency channels.

This provides greater frequency separation between adjacent channels in a group, which reduces adjacent channel interference.

The output from the receiver is fed to a power splitter, which in turn feeds the two separate chains of circulators.

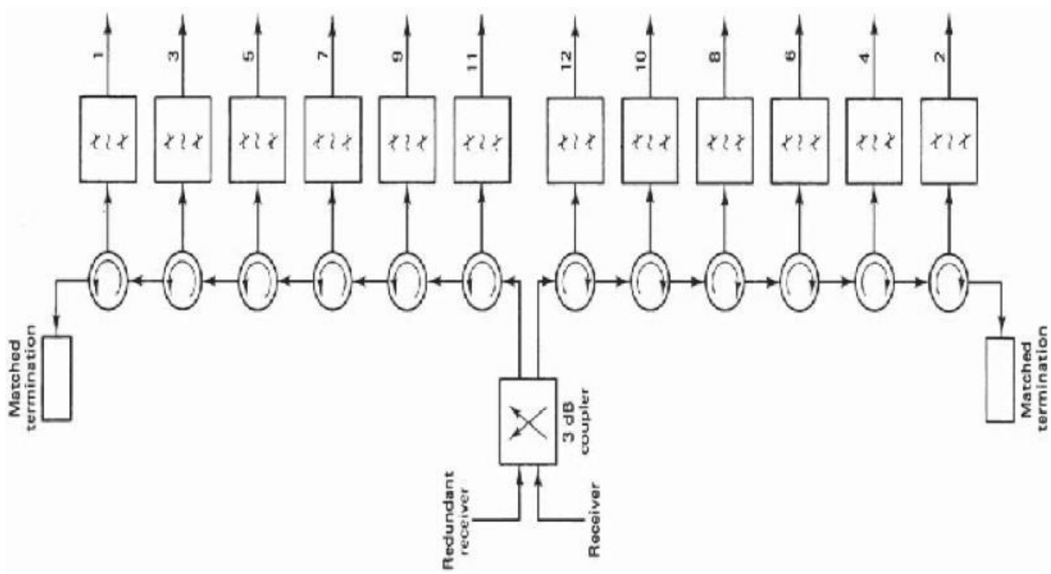


Figure 2.12 Satellite input multiplexer

The full broadband signal is transmitted along each chain, and the channelizing is achieved by means of channel filters con- nected to each circulator,

Each filter has a bandwidth of 36 MHz and is tuned to the appropriate center frequency, as shown in Fig. 2.11.

Although there are considerable losses in the demultiplexer, these are easily made up in the overall gain for the transponder channels.

2.6.4 The power amplifier

The fixed attenuation is needed to balance out variations in the input attenuation so that each transpon- der channel has the same nominal attenuation, the necessary adjust- ments being made during assembly.

The variable attenuation is needed to set the level as required for different types of service (an example being the requirement for input power backoff discussed later). Because this variable attenuator adjustment is an operational requirement, it must be under the control of the ground TT&C station.

Traveling-wave tube amplifiers (TWTAs) are widely used in transpon- ders to provide the final output power required to the transmit antenna. Figure 2.13 shows the schematic of a *traveling wave tube* (TWT) and its power supplies.

In the TWT, an electron-beam gun assembly consisting of a heater, a cathode, and focusing electrodes is used to form an elec- tron beam. A magnetic field is required to confine the beam to travel along the inside of a wire helix.

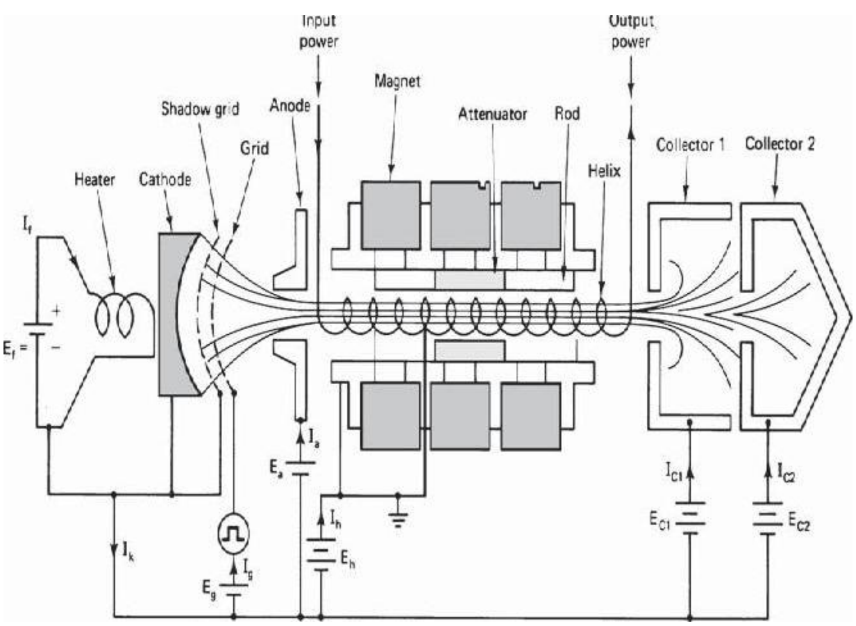


Figure 2.13 Satellite TWTA

used in ground stations, the magnetic field can be provided by means of a solenoid and dc power supply. The comparatively large size and high power consumption of solenoids make them unsuitable for use aboard satellites, and lower-power TWTs are used which employ permanent- magnet focusing.

The wave actually will travel around the helical path at close to the speed of light, but it is the axial component of wave velocity which interacts with the electron beam.

This component is less than the velocity of light approximately in the ratio of helix pitch to circumference. Because of this effective reduction in phase velocity, the helix is referred to as a *slowwave structure*.

The advantage of the TWT over other types of tube amplifiers is that it can provide amplification over a very wide bandwidth. Input levels to the TWT must be carefully controlled, however, to minimize the effects of certain forms of distortion.

The worst of these result from the nonlinear transfer characteristic of the TWT, illustrated in Fig. 2.14.

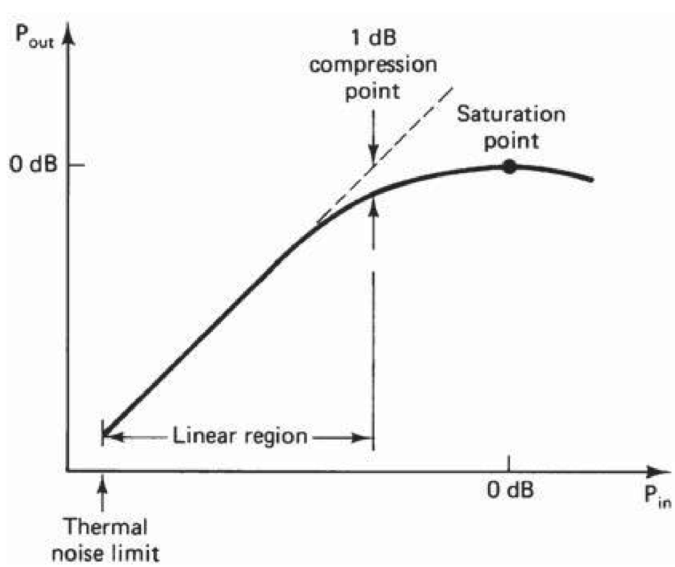


Figure 2.14 Power transfer characteristics of a TWT. The saturation point is used as 0-dB reference for both input and output.

At low-input powers, the output-input power relationship is linear; that is, a given decibel change in input power will produce the same decibel change in output power. At higher power inputs, the output power sat- urates, the point of maximum power output being known as the *satu- ration point*.

The saturation point is a very convenient reference point, and input and output quantities are usually referred to it. The linear region of the TWT is defined as the region bound by the thermal noise limit at the low end and by what is termed the *1-dB compression point* at the upper end. This is the point where the actual transfer curve drops

2.7. Satellite uplink and downlink Analysis and Design

2.7.1 Introduction

This chapter describes how the link-power budget calculations are made. These calculations basically relate two quantities, the transmit power and the receive power, and show in detail how the difference between these two powers is accounted for.

Link-budget calculations are usually made using decibel or decilog quantities. These are explained in App. G. In this text [square] brackets are used to denote decibel quantities using the basic power definition.

Where no ambiguity arises regarding the units, the abbreviation dB is used. For example, Boltzmann's constant is given as 228.6 dB, although, strictly speaking, this should be given as 228.6 deci logs relative to 1 J/K.

2.7.2 Equivalent Isotropic Radiated Power

A key parameter in link-budget calculations is the *equivalent isotropic radiated power*, conventionally denoted as EIRP. From Eqs, the maximum power flux density at some distance *r* from a transmitting antenna of gain *G* i

$$Pr= \frac{GP}{4\pi^2}$$

An isotropic radiator with an input power equal to GP$_S$ would produce the same flux density. Hence, this product is referred to as the EIRP, or EIRP is often expressed in decibels relative to 1 W, or dBW. Let P_S be in watts; then [EIRP] = [P_S] x [G] dB ,where [P_S] is also in dBW and [G] is in dB.

2.7.3 Transmission Losses

The [EIRP] may be thought of as the power input to one end of the transmission link, and the problem is to find the power received at the other end. Losses will occur along the way, some of which are constant.

Other losses can only be estimated from statistical data, and some of these are dependent on weather conditions, especially on rainfall.

The first step in the calculations is to determine the losses for *clear- weather* or *clear-sky conditions*. These calculations take into account the losses, including those calculated on a statistical basis, which do not vary significantly with time. Losses which are weather-related, and other losses which fluctuate with time, are then allowed for by introducing appropriate *fade margins* into the transmission equation.

Free-space transmission:

As a first step in the loss calculations, the power loss resulting from the spreading of the signal in space must be determined.

Feeder losses:

Losses will occur in the connection between the receive antenna and the receiver proper. Such losses will occur in the connecting waveguides, filters, and couplers. These will be denoted by RFL, or [RFL] dB, for *receiver feeder losses*.

Antenna misalignment losses

When a satellite link is established, the ideal situation is to have the earth station and satellite antennas aligned for maximum gain, as shown in Fig. There are two possible sources of off-axis loss, one at the satellite and one at the earth station, as shown in Fig.

The off-axis loss at the satellite is taken into account by designing the link for operation on the actual satellite antenna contour; this is described in more detail in later sections. The off-axis loss at the earth station is referred to as the *antenna pointing loss*. Antenna pointing losses are usually only a few tenths of a decibel;

In addition to pointing losses, losses may result at the antenna from misalignment of the polarization direction (these are in addition to the polarization losses described in Chap. 5). The polarization misalign- ment losses are usually small, and it will be assumed that the antenna misalignment losses, denoted by [AML], include both pointing and polar- ization losses resulting from antenna misalignment. It should be noted

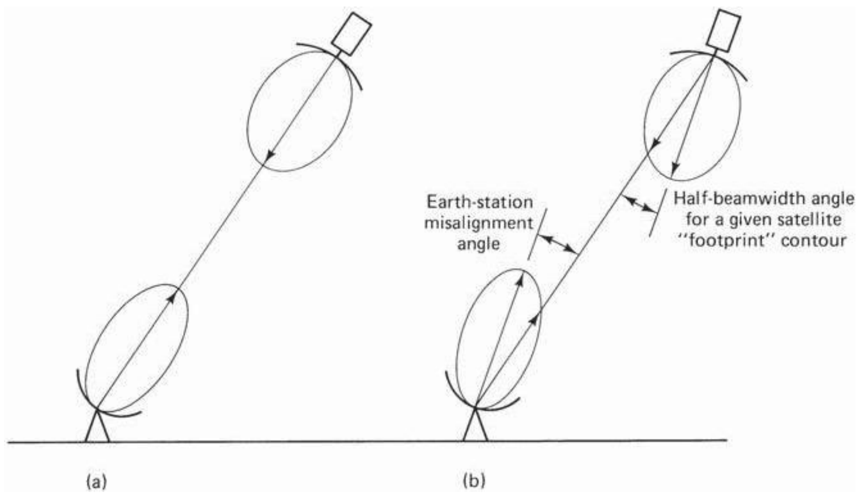


Figure 2.15 (*a*) Satellite and earth-station antennas aligned for maximum gain; (*b*) earth station situated on a given satellite "footprint," and earth-station antenna misaligned.

2.8 The Link-Power Budget Equation

Now that the losses for the link have been identified, the power at the receiver, which is the power output of the link, may be calculated simply as [EIRP] [LOSSES] [G_R], where the last quantity is the receiver antenna gain. Note carefully that decibel addition must be used.

The major source of loss in any ground-satellite link is the free-space spreading loss [FSL], the basic link-power budget equation taking into account this loss only. However, the other losses also must be taken into account, and these are simply added to [FSL]. The losses for clear-sky conditions are

[LOSSES] = [FSL] + [RFL] + [AML] + [AA] - [PL] equation for the received power is then

$$[P_R] = [\text{EIRP}] \ \text{x} \ [G_R] - [\text{LOSSES}]$$

where [PR] received power, dBW

[EIRP] - equivalent isotropic radiated power, dBW [FSL] free-space spreading loss, dB

[RFL] - receiver feeder loss, dB

[AML] - antenna misalignment loss, dB

[AA] - atmospheric absorption loss, dB [PL] polarization mismatch loss, dB

2.9 Amplifier noise temperature

Consider first the noise representation of the antenna and the *low noise amplifier* (LNA) shown in Fig. 2.15.

The available power gain of the amplifier is denoted as G, and the noise power output, as P_{no}.

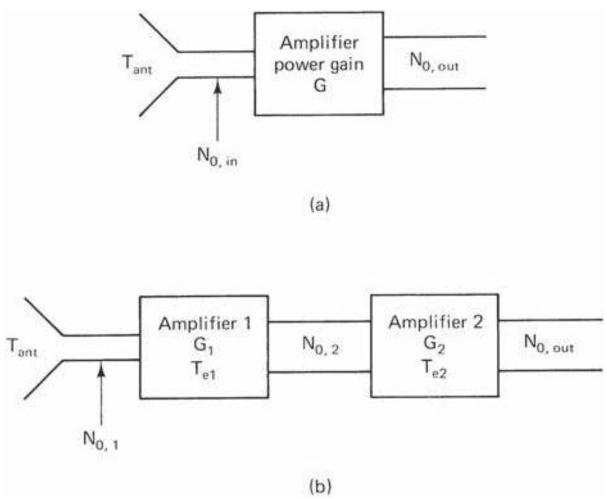


Figure 2.15 LNA Amplifier gain

For the moment we will work with the noise power per unit bandwidth, which is simply noise energy in joules as shown by Eq.

The input noise energy coming from the antenna is

$$N_{0,ant} = kT_{ant}$$

2.10 The Uplink

The uplink of a satellite circuit is the one in which the earth station is transmitting the signal and the satellite is receiving it specifically that the uplink is being considered.

$$\frac{C}{N} = [EIRP] - [LOSSES] \quad] + k]$$

In this Eq the values to be used are the earth station EIRP, the satellite receiver feeder losses, and satellite receiver *G*/*T*. The free-space loss and other losses which are frequency-dependent are calculated for the uplink frequency.

2.10.1 Input backoff

Number of carriers are present simultaneously in a TWTA, the operating point must be backed off to a linear portion of the transfer characteristic to reduce the effects of inter modulation distortion. Such multiple carrier operation occurs with *frequency- division multiple access* (FDMA), which is described in Chap. 14. The point to be made here is that *backoff* (BO) must be allowed for in the link- budget calculations.

Suppose that the saturation flux density for single-carrier operation is known. Input BO will be specified for multiple-carrier operation, referred to the single-carrier saturation level. The earth-station EIRP will have to be reduced by the specified BO, resulting in an uplink value of

$$[EIRP]_U = [EIRP_S]_U + [BO]_i$$

2.10.2 The earth station HPA

The earth station HPA has to supply the radiated power plus the transmit feeder losses, denoted here by TFL, or [TFL] dB. These include waveguide, filter, and coupler losses between the HPA output and the transmit antenna. Referring back to Eq. (12.3), the power output of

The earth station itself may have to transmit multiple carriers, and its output also will require back off, denoted by $[BO]_{HPA}$. The earth station HPA must be rated for a saturation power output given by

$$[P_{HPA,sat}] = [P_{HPA}] + [BO]_{HPA}$$

2.11 Downlink

The downlink of a satellite circuit is the one in which the satellite is transmitting the signal and the earth station is receiving it. Equation can be applied to the downlink, but subscript D will be used to denote specifically that the downlink is being considered. Thus Eq. becomes

$$\frac{C}{N} = [EIRP]] - [LOSSES] \quad] + [\quad]$$

In Eq. the values to be used are the satellite EIRP, the earth- station receiver feeder losses, and the earth-station receiver G/T. The free space and other losses are calculated for the downlink frequency. The resulting carrier-to-noise density ratio given by Eq. is that which appears at the detector of the earth station receiver.

2.11.1 Output back-off

Where input BO is employed as described in a corresponding output BO must be allowed for in the satellite EIRP. As the curve of Fig. 2.16 shows, output BO is not linearly related to input BO. A rule of thumb, frequently used, is to take the output BO as the point on the curve which is 5 dB below the extrapolated linear portion, as shown in Fig. 12.7. Since the linear portion gives a 1:1 change in decibels, the relationship between input and output BO is $[BO]_0$ $[BO]_i$ 5 dB. For example, with an input BO of $[BO]_i$ 11 dB, the corresponding output BO is $[BO]_0$

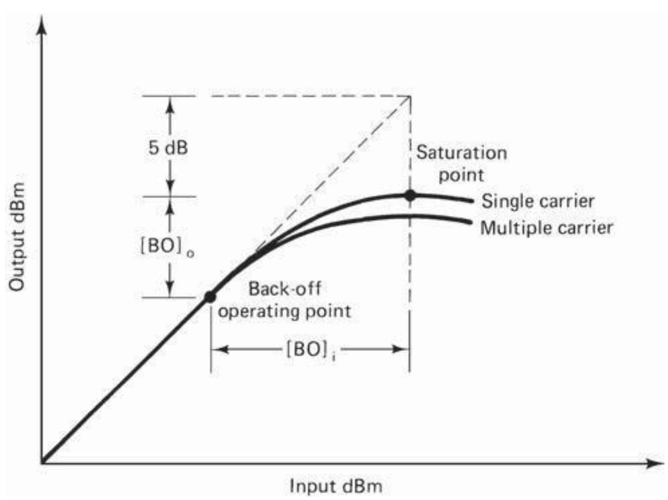


Figure 2.16 Input and output back- off relationship for the satellite traveling-wave-tube amplifier; $[BO]_i$ $[BO]_0$ 5 dB.

2.11.2 Effects of Rain

In the C band and, more especially, the Ku band, rainfall is the most significant cause of signal fading. Rainfall results in attenuation of radio waves by scattering and by absorption of energy from the wave.

Rain attenuation increases with increasing frequency and is worse in the Ku band compared with the C band.

This produces a depolarization of the wave; in effect, the wave becomes elliptically polarized. This is true for both linear and circular polar- izations, and the effect seems to be much worse for circular polarization (Freeman, 1981).

The C/N_0 ratio for the downlink alone, not counting the P_{NU} contri- bution, is P_R/P_{ND}, and the combined C/N_0 ratio at the ground receiver is

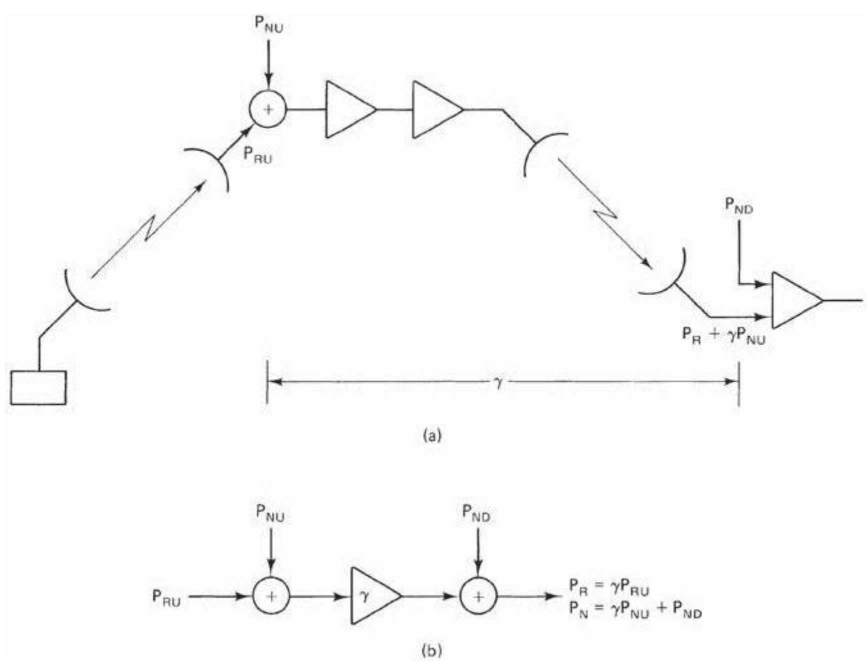


Figure 2.17 (*a*) Combined uplink and downlink; (*b*) power flow diagram

The reason for this reciprocal of the sum of the reciprocals method is that a single signal power is being transferred through the system, while the various noise powers, which are present are additive. Similar reasoning applies to the carrier-to-noise ratio, C/N.

2.12. inter modulation and interference

Intermodulation interference is the undesired combining of several signals in a nonlinear device, producing new, unwanted frequencies, which can cause interference in adjacent receivers located at repeater sites.

Not all interference is a result of intermodulation distortion. It can come from co-channel interference, atmospheric conditions as well as man-made noise generated by medical, welding and heating equipment.

Most intermodulation occurs in a transmitter's nonlinear power amplifier (PA). The next most common mixing point is in the front end of a receiver. Usually it occurs in the unprotected first mixer of older model radios or in some cases an overdriven RF front-end amp.

Intermodulation can also be produced in rusty or corroded tower joints, guy wires, turnbuckles and anchor rods or any nearby metallic object, which can act as a nonlinear "mixer/rectifier" device.

2.13. Propagation Characteristics and Frequency considerations

2.13.1 Introduction

A number of factors resulting from changes in the atmosphere have to be taken into account when designing a satellite communications system in order to avoid impairment of the wanted signal.

Generally, a margin in the required carrier-to-noise ratio is incorporated to accommodate such effects.

2.13.2 Radio Noise

Radio noise emitted by matter is used as a source of information in radio astronomy and in remote sensing. Noise of a thermal origin has a continuous spectrum, but several other radiation mechanisms cause the emission to have a spectral-line structure. Atoms and molecules are distinguished by their different spectral lines.

For other services such as satellite communications noise is a limiting factor for the receiving system; generally, it is inappropriate to use receiving systems with noise temperatures which are much less than those specified by the minimum external noise.

From about 30 MHz to about 1 GHz cosmic noise predominates over atmospheric noise except during local thunderstorms, but will generally be exceeded by man-made noise in populated areas.

In the bands of strong gaseous absorption, the noise temperature reaches maximum values of some 290 K. At times, precipitation will also increase the noise temperature at frequencies above 5 GHz.

Figure 6.1 gives an indication of sky noise at various elevation angles and frequencies.

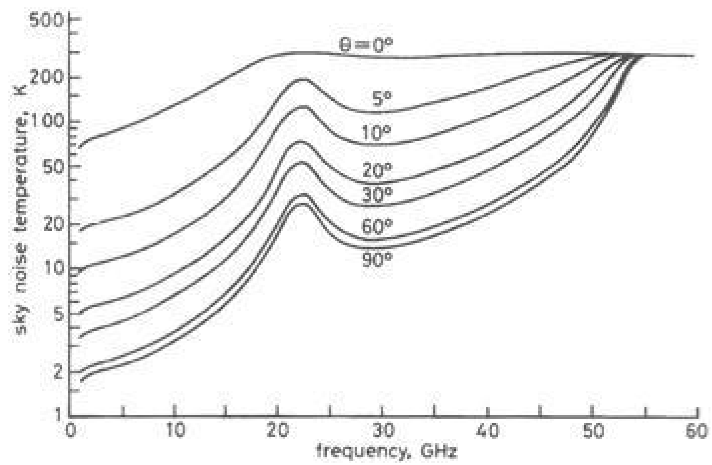


Figure 2.18 Sky-Noise Temperature for Clear Air

2.14. System reliability and design lifetime

2.14.1 System reliability

Satellites are designed to operate dependably throughout their operational life, usually a number of years.

This is achieved through stringent quality control and testing of parts and subsystems before they are used in the construction of the satellite.

Redundancy of key components is often built in so that if a particular part or subassembly fails, another can perform its functions.

In addition, hardware and software on the satellite are often designed so that ground controllers can reconfigure the satellite to work around a part that has failed.

2.14.2. Design lifetime

The Milstar constellation has demonstrated exceptional reliability and capability, providing vital protected communications to the warfighter," said Kevin Bilger, vice president and general manager, Global Communications Systems, Lockheed Martin Space Systems in Sunnyvale.

"Milstar's robust system offers our nation worldwide connectivity with flexible, dependable and highly secure satellite communications."

The five-satellite Milstar constellation has surpassed 63 years of combined successful operations, and provides a protected, global communication network for the joint forces of the U.S. military. In addition, it can transmit voice, data, and imagery, and offers video teleconferencing capabilities.

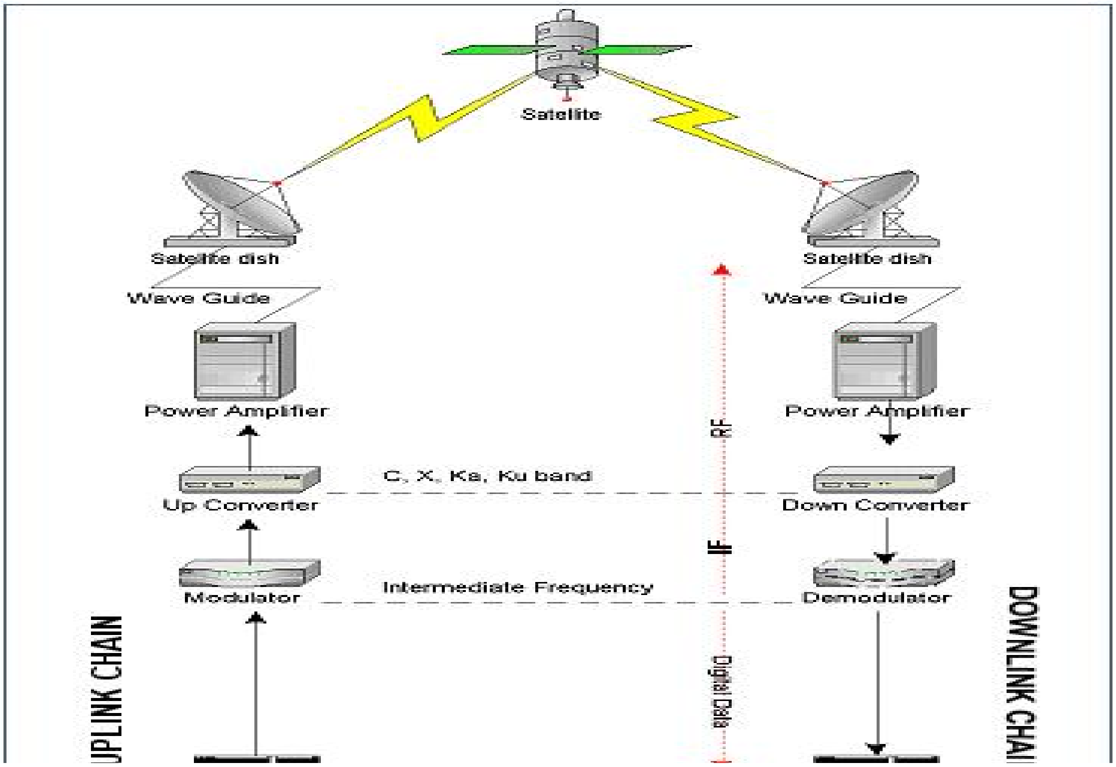
The system is the principal survivable, endurable communications structure that the President, the Secretary of Defense and the Commander, U.S. Strategic Command use to maintain positive command and control of the nation's strategic forces.

In addition to this 10-year milestone for Flight-5, each of the first two Milstar satellites have been on orbit for over 16 years – far exceeding their 10-year design life.

The next-generation Lockheed Martin-built Advanced EHF satellites, joining the Milstar constellation, provide five times faster data rates and twice as many connections, permitting transmission of strategic and tactical military communications, such as real-time video, battlefield maps and targeting data. Advanced EHF satellites are designed to be fully interoperable and backward compatible with Milstar.

Headquartered in Bethesda, Md., Lockheed Martin is a global security company that employs about 123,000 people worldwide and is principally engaged in the research, design, development, manufacture, integration and sustainment of advanced technology systems, products and services. The Corporation's net sales for 2011 were $46.5 billion.

APPLICATIONS



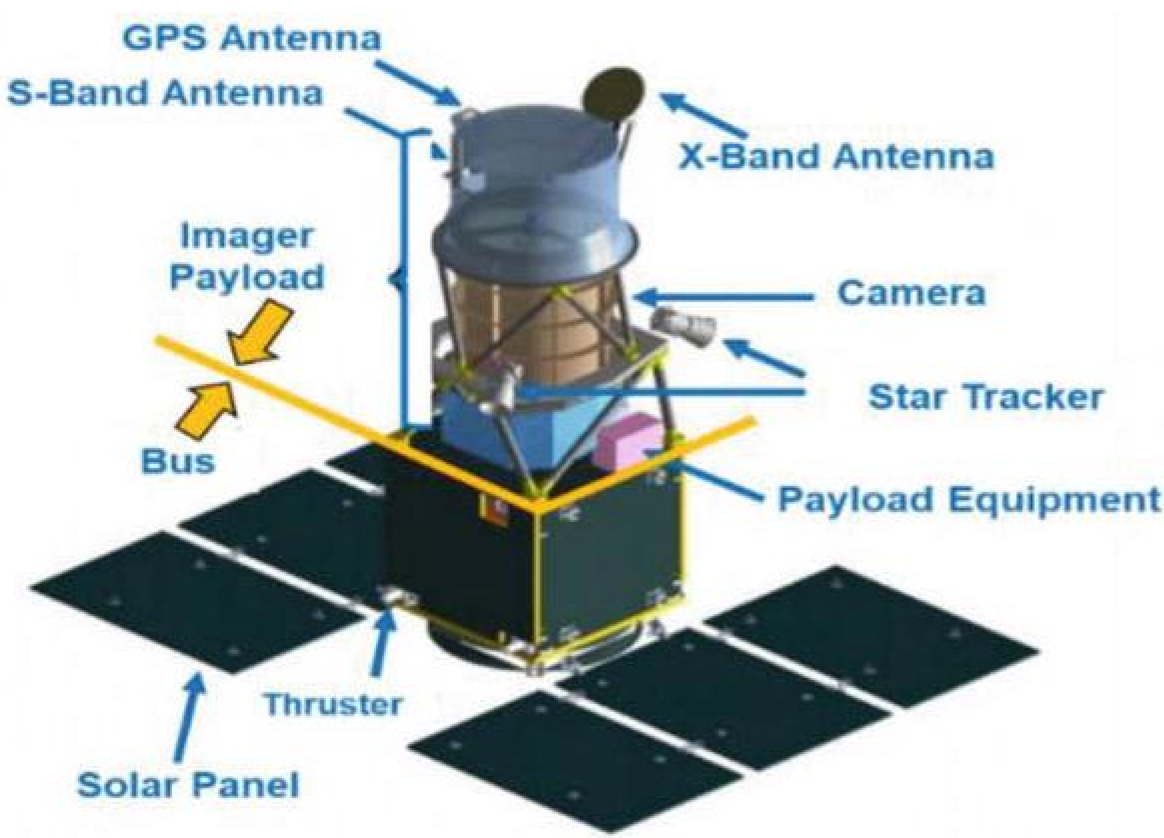


Figure typical satellite with bus and payload separation

POST TEST MCQ

1. The lowest frequency used in satellite communications is__________________ GHz.

a) 0.8

b) 3

c) 18

d) 30

2. For satellite transmission, analog signals may be converted into digital form with the help of ____________

a) modem

b) transponder

c) codec

d) compandor

3. Geosynchronous satellites are always launched in the equatorial plane because it is the only plane which provides

a) 24-hour orbit

b) stationary satellite

c) global communication

d) zero-gravity environs

4. The traffic-handling capacity of an Earth station on the uplink depends on _______

a) its EIRP

b) satellite antenna gain

c) noise associated with the satellite

d) all of the above

5. Phase modulation is commonly-used for data transmission mainly because

a) phase can be varied from + 180° to 180°

b) it is resistant to the effects of noise

c) demodulation is very easy

d) it gives highest data rates that can be transmitted over a given channel

6. Most of the communication satellites are stationed to the West of their service areas order to reduce their

a) eclipse period

b) loss of power

c) battery power provision

d) mass of station-keeping fuel

7. The echo heard by a telephone user on a satellite channel can be removed by using

a) a vocoder

b) a multiplexer

c) echo suppressor

d) digital techniques

8. A satellite link uses different frequencies for receiving and transmitting in order to

a) avoid interference from terrestrial microwave links

b) avoid interference between its powerful transmitted signal and weak in coming signal

c) minimise free-space losses

d) maximise antenna gain

9. System satellites orbit the Earth once in

a) 24 hours

b) 12 hours

c) 1 hour

d) 6 hours

10. A few minutes disturbance in space communications occurs twice a year during Sunblinding when __________ are in line.

a) Sun and satellite

b) Sun and Earth station

c) Satellite and Earth station

d) Sun, satellite and Earth station

11. In the context of error detection in satellite transmission, ARQ stands for

a) Automatic Repeat Request

b) Automatic Relay Request

c) Accelerated Recovery Request

d) Automatic Radiation Quenching

12. To cover all inhabited regions of the Earth, the number of geosynchronous communication satellites required

a) 5

b) 3

c) 10

d) 2

13. A modem is

a) a form of commutator

b) a device for digitizing speech

c) a circuit used for suppressing microwave interference

d) an electronic circuit which carries out modulation and demodulation of a carrier frequency

14. A typical signal strength received from a geosynchronous communication satellite is of the order of a few

a) milliwatts

b) kilowatts

c) picowatts

d) watts

15. A telephone user while talking to a person via a satellite has to wait for reply for about __________ millisecond.

a) 100

b) 270

c) 470

d) 540

CONCLUSION

- ❖ In this unit we describe the Spacecraft Technology.
- ❖ The Satellite Uplink and Downlink Analysis and Design were discussed.
- ❖ The Link Power Budget, C/N calculation and G/T ratio-Performance were elaborated.

REFERENCES

1. Wilbur L.Pritchard, Hendri G. Suyderhoud, Robert A. Nelson, "Satellite Communication Systems Engineering", Prentice Hall/Pearson, 2007.
2. N. Agarwal, "Design of Geosynchronous Space Craft", Prentice Hall, 1986.
3. Bruce R. Elbert, "Satellite Communication Applications", Hand Book, Artech House Bostan London,1997.
4. Tri T. Ha, "Digital Satellite Communication", II nd edition, 1990.
5. Robert G. Winch, "Telecommunication Transmission Systems", Mc Graw-Hill Book Co., 1983.
6. Brian Ackroyd, "World Satellite Communication and earth station Design", BSP professional Books, 1990.
7. G.B.Bleazard, "Introducing Satellite communications", NCC Publication, 1985.
8. M.Richharia, "Satellite Communication Systems-Design Principles", Macmillan 2003.
9. Emanuel Fthenakis, "Manual of Satellite Communications", Mc Graw Hill Book Co., 1984.

ASSIGNMENT

1. Explain the difference types of transmission losses in satellite communication with necessary expression. Write the link power budget equation.
2. Discuss the different types of noise and their significance in the design of a satellitelink with necessary expression.
3. Explain the following: input backoff, output backoff, earth station HPA andcombined uplink and downlink. C/N ratio.
4. Explain wideband receiver operation with neat diagram ORB.
5. Explain thermal control system. With a neat sketch,
6. Explain Telemetry, Tracking and command subsystem.

SATELLITE COMMUNICATION

UNIT III EARTH SEGMENT

Prepared by

Dr.M.A.ARCHANA

AP/ECE,SCSVMV

UNIT III EARTH SEGMENT

AIM & OBJECTIVE

- To understand the basics of earth segment.
- To understand the Indoor UNIT for analog (FM) TV.
- To analyze the Terrestrial Interface.
- To understand the Antenna Gain.

PRE TEST MCQ

1. To make antenna more directional, either its size must be increased or

a) the number of its feed horns must be increased

b) the frequency of its transmission must be increased

c) its effective isotropic radiated power (EIRP) must be increased

d) its footprint must be increased

2. The number of days when Earth's shadow falls on a geosynchronous satellite is

a) 88

b) 277

c) 5

d) 10

3. A helical antenna is used for satellite tracking because of ________

a) Circular polarization

b) Maneuverability

c) Beamwidth

d) Gain

3.1 Earth Station Technology

The earth segment of a satellite communications system consists of the transmit and receive earth stations. The simplest of these are the home *TV receive-only* (TVRO) systems, and the most complex are the terminal stations used for international communications networks. Also included in the earth segment are those stations which are on ships at sea, and commercial and military land and aeronautical mobile stations.

As mentioned in earth stations that are used for logistic sup- port of satellites, such as providing the *telemetry, tracking, and command* (TT&C) functions, are considered as part of the space segment.

3.1.1 Terrestrial Interface

Earth station is a vital element in any satellite communication network. The function of an earth station is to receive information from or transmit information to, the satellite network in the most cost-effective and reliable manner while retaining the desired signal quality. The design of earth station configuration depends upon many factors and its location. But it is fundamentally governed by its

Location which are listed below,

- In land
- On a ship at sea
- Onboard aircraft

The factors are

- Type of services
- Frequency bands used
- Function of the transmitter
- Function of the receiver
- Antenna characteristics

3.1.2 Transmitter and Receiver

Any earth station consists of four major subsystems

- Transmitter
- Receiver
- Antenna • Tracking equipment

Two other important subsystems are

- Terrestrial interface equipment
- Power supply

The earth station depends on the following parameters

- Transmitter power
- Choice of frequency
- Gain of antenna
- Antenna efficiency
- Antenna pointing accuracy
- Noise temperature

The functional elements of a basic digital earth station are shown in the below figure

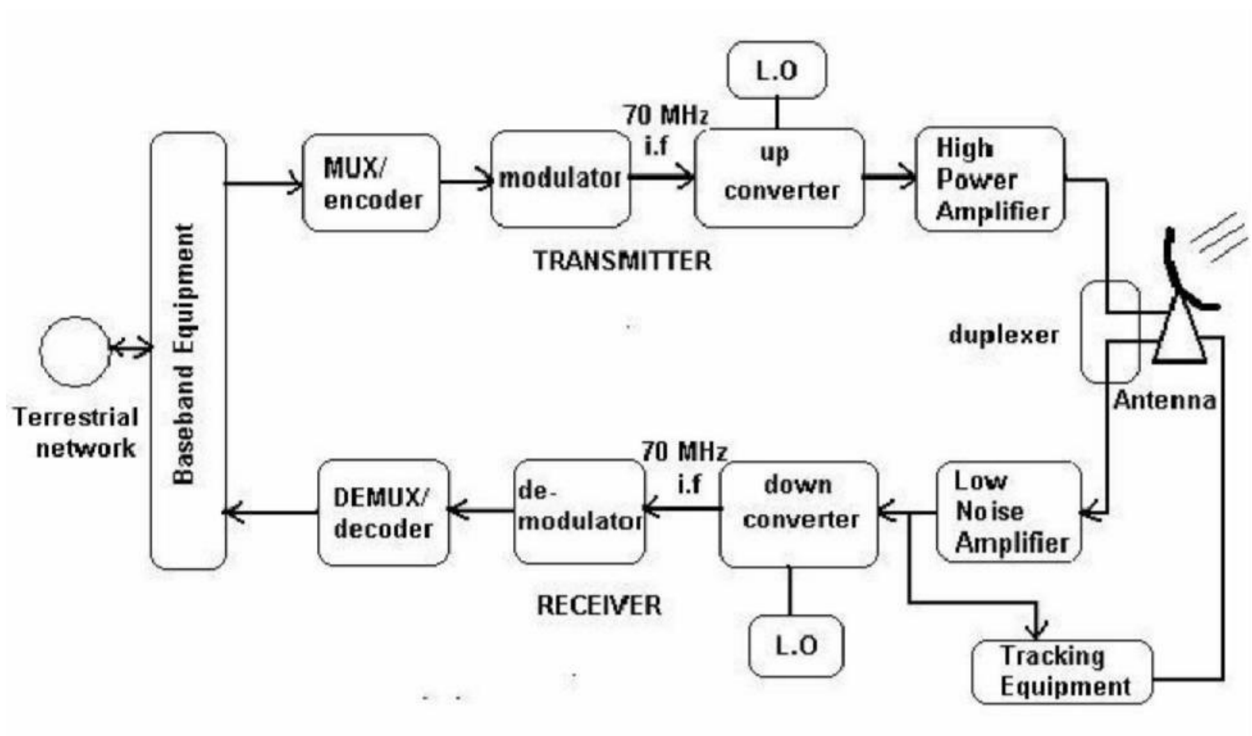


Figure 3.1 Transmitter- Receiver

Digital information in the form of binary digits from terrestrial networks enters earth station and is then processed (filtered, multiplexed, formatted etc.) by the base band equipment.

- The encoder performs error correction coding to reduce the error rate, by introducing extra digits into digital stream generated by the base band equipment. The extra digits carry information.

• In satellite communication, I.F carrier frequency is chosen at 70 MHz for communication using a 36 MHz transponder bandwidth and at 140 MHz for a transponder bandwidth of 54 or 72 MHz.

• On the receive side, the earth station antenna receives the low -level modulated R.F carrier in the downlink frequency spectrum.

• The low noise amplifier (LNA) is used to amplify the weak rec eived signals and improve the signal to Noise ratio (SNR). The error rate requirements can be met more easily.

• R.F is to be reconverted to I.F at 70 or 140 MHz because it is easier design a demodulation to work at these frequencies than 4 or 12 GHz.

• The demodulator estimate which of the possible symbols was transmitted based on observation of the received if carrier.

• The decoder performs a function opposite that of the encoder. Because the sequence of symbols recovered by the demodulator may cont ain errors, the decoder must use the uniqueness of the redundant digits introduced by the encoder to correct the errors and recover information-bearing digits.

• The information stream is fed to the base-band equipment for processing for delivery to the terrestrial network.

• The tracking equipments track the satellite and align the beam towards it to facilitate communication.

3.1.3. Earth Station Tracking System

Tracking is essential when the satellite drift, as seen by an earth station antenna is a significant fraction of an earth station's antenna beam width.

An earth station's tracking system is required to perform some of the functions such as

i)Satellite acquisition
ii)Automatic tracking
iii)Manual tracking
iv)Program tracking.

3.2 Antenna Systems

The antenna system consist of

- Feed System
- Antenna Reflector
- Mount
- Antenna tracking System

3.2.1 FEED SYSTEM

The feed along with the reflector is the radiating/receiving element of electromagnetic waves. The reciprocity property of the feed element makes the earth station antenna system suitable for transmission and reception of electromagnetic waves.

The way the waves coming in and going out is called feed configuration Earth Station feed systems most commonly used in satellite communication are:

i)Axi-Symmetric Configuration
ii)Asymmetric Configuration
i)Axi-Symmetric Configuration

In an axi-symmetric configuration the antenna axes are symmetrical with respect to the reflector ,which results in a relatively simple mechanical structure and antenna mount.

Primary Feed

In primary, feed is located at the focal point of the parabolic reflector. Many dishes use only a single bounce, with incoming waves reflecting off the dish surface to the focus in front of the dish, where the antenna is located. when the dish is used to transmit ,the transmitting antenna at the focus beams waves toward the dish, bouncing them off to space. This is the simplest arrangement.

Cassegrain

Many dishes have the waves make more than one bounce .This is generally called as folded systems. The advantage is that the whole dish and feed system is more compact. There are several folded configurations, but all have at least one secondary reflector also called a sub reflector, located out in front of the dish to redirect the waves.

A common dual reflector antenna called Cassegrain has a convex sub reflector positioned in front of the main dish, closer to the dish than the focus. This sub reflector bounces back the waves back toward a feed located on the main dish's center, sometimes behind a hole at the center of the main dish. Sometimes there are even more sub reflectors behind the dish to direct the waves to the fed for convenience or compactness.

Gregorian

This system has a concave secondary reflector located just beyond the primary focus. This also bounces the waves back toward the dish.

ii)Asymmetric Configuration

Offset or Off-axis feed

The performance of tan axi-symmetric configuration is affected by the blockage of the aperture by the feed and the sub reflector assembly. The result is a reduction in the antenna efficiency and an increase in the side lobe levels. The asymmetric configuration can remove this limitation..This is achieved by off - setting the mounting arrangement of the feed so that it does not obstruct the main beam. As a result ,the efficiency and side lobe level performance are improved.

3.2.2 ANTENNA REFLECTOR

Mostly parabolic reflectors are used as the main antenna for the earth stations because of the high gain available from the reflector and the ability of focusing a parallel beam into a point at the focus where the feed,i.e., the receiving/radiating element is located .For large antenna system more than one reflector surfaces may be used in as in the cassegrain antenna system.

Earth stations are also classified on the basis of services for example:

1.Two way TV ,Telephony and data

2. Two way TV

3.TV receive only and two way telephony and data

4.Two way data

From the classifications it is obvious that the technology of earth station will vary considerably on the performance and the service require ments of earth station

For mechanical design of parabolic reflector the following parameters are required to be considered:

- Size of the reflector
- Focal Length /diameter ratio
- RMS error of main and sub reflector
- Pointing and tracking accuracies
- Speed and acceleration
- Type of mount
- Coverage Requirement

Wind Speed

The size of the reflector depends on transmit and receive gain requirement and beamwidth of the antenna. Gain is directly proportional to the antenna diameter whereas the beamwidth is inversely proportional to the antenna diameter .for high inclination angle of the satellite ,the tracking of the earth station becomes necessary when the beamwidth is too narrow.

The gain of the antenna is given by

$$\text{Gain} = (\eta 4\Pi A_{eff}) / \lambda^2$$

Where Aeff is the aperture

Λ is wave length

H is efficiency of antenna system

For a parabolic antenna with circular aperture diameter D, the gain of the antenna is :

$$\text{Gain} = (\eta 4\Pi / \lambda^2)(\Pi D^2/4)$$

$$= \eta (\Pi D / \lambda)^2$$

The overall efficiency of the antenna is the net product of various factors such as

1. Cross Polarization
2. Spill over
3. Diffraction
4. Blockage
5. Surface accuracy
6. Phase error
7. Illumination

In the design of feed, the ratio of focal length F to the diameter of the reflector D of the antenna system control the maximum angle subtended by the reflector surface on the focal point. Larger the F/D ratio larger is the aperture illumination efficiency and lower the cross polarization.

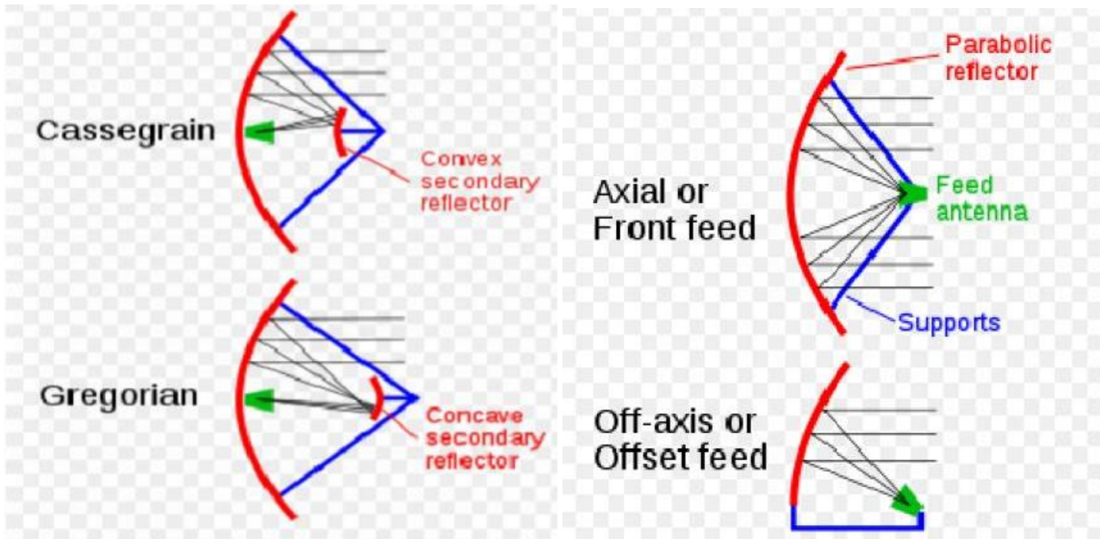


Figure 3.2 Antenna sub systems

3.2.3 ANTENNA MOUNT

Type of antenna mount is determined mainly by the coverage requirement and tracking requirements of the antenna systems. Different types of mounts used for earth station antenna are:

i) **The Azimuth –elevation mount**

This mount consists of a primary vertical axis. Rotation around this axis controls the azimuth angle. The horizontal axis is mounted over the primary axis, providing the elevation angle control.

ii) **The X-Y mount**

It consists of a horizontal primary axis (X-axis) and a secondary axis (Y-axis) and at right angles to it. Movement around these axes provides necessary steering.

3.2.4 ANTENNA TRACKING SYSTEM

Tracking is essential when the satellite drift, as seen by an earth station antenna is a significant fraction of an earth station's antenna beam width.

An earth station's tracking system is required to perform some of the functions such as

i)Satellite acquisition
ii)Automatic tracking
iii)Manual tracking
iv)Program tracking.

Recent Tracking Techniques

There have been some interesting recent developments in auto -track techniques which can potentially provide high accuracies at a low cost.

In one proposed technique the sequential lobing technique has been I implemented by using rapid electronic switching of a s single beam which effectively approximates simultaneous lobbing.

3.3 Receive-Only Home TV Systems

Planned broadcasting directly to home TV receivers takes place in the Ku (12-GHz) band. This service is known as *direct broadcast satellite* (DBS) service.

There is some variation in the frequency bands assigned to different geographic regions. In the Americas, for example, the down- link band is 12.2 to 12.7 GHz.

The comparatively large satellite receiving dishes [ranging in diame- ter from about 1.83 m (6 ft) to about 3-m (10 ft) in some locations], which may be seen in some "backyards" are used to receive downlink TV signals at C band (4 GHz).

Originally such downlink signals were never intended for home reception but for network relay to commercial TV outlets (VHF and UHF TV broadcast stations and cable TV "head-end" studios).

3.3.1 The Indoor unit

Equipment is now marketed for home reception of C-band signals, and some manufacturers provide dual C-band/Ku-band equipment. A single mesh type reflector may be used which focuses the signals into a dual feed- horn, which has two separate outputs, one for the C-band signals and onefor the Ku-band signals.

Much of television programming originates as *first generation signals*, also known as *master broadcast quality signals*.

These are transmitted via satellite in the C band to the network head- end stations, where they are retransmitted as compressed digital signals to cable and direct broadcast satellite providers.

- Another of the advantages, claimed for home C-band systems, is the larger number of satellites available for reception compared to what is available for direct broadcast satellite sys- terms.
- Although many of the C-band transmissions are scrambled, there are free channels that can be received, and what are termed “wild feeds.”
- These are also free, but unannounced programs, of which details can be found in advance from various publications and Internet sources.
- C-band users can also subscribe to pay TV channels, and another advantage claimed is that subscription services are cheaper than DBS or cable because of the multiple-source programming available.
- The most widely advertised receiving system for C-band system appears to be 4DTV manufactured by Motorola.

This enables reception of

- Free, analog signals and “wild feeds”
- Video Cipher ll plus subscription services
- Free Digi Cipher 2 services
- Subscription DigiCipher 2 services

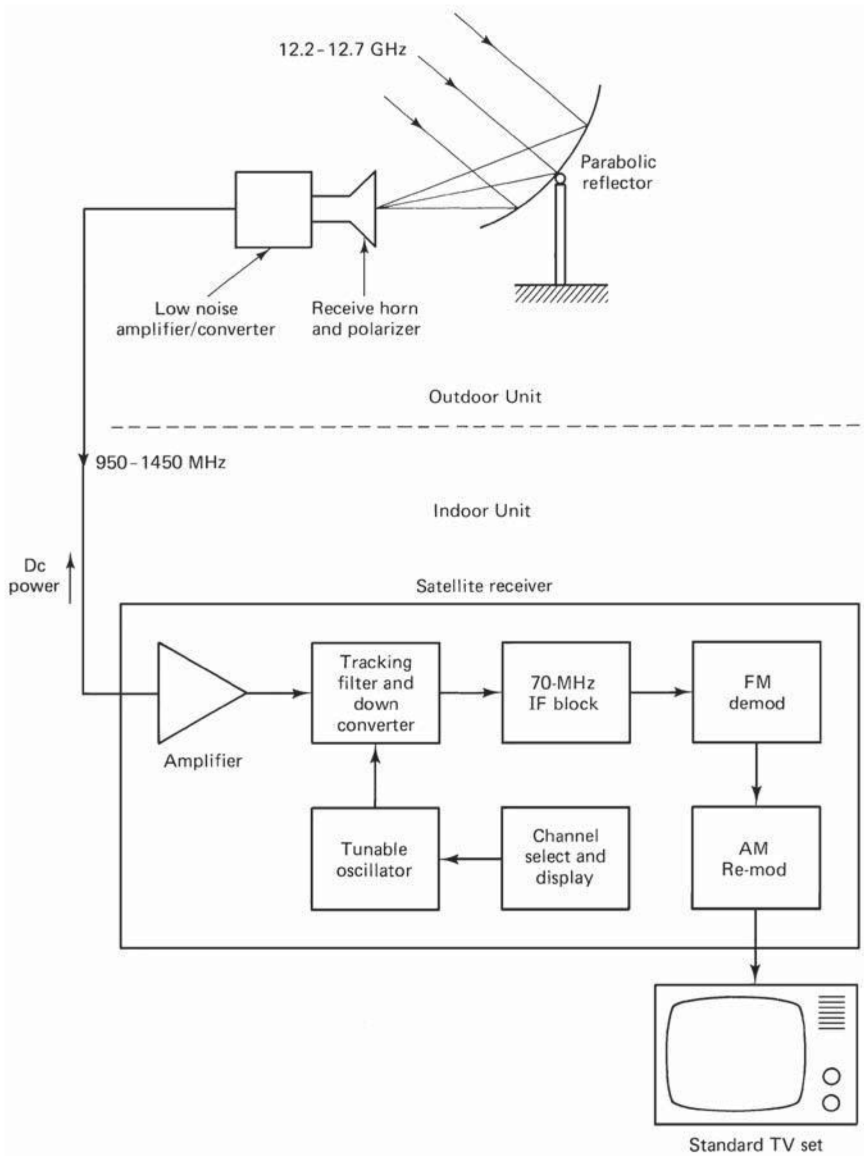


Figure 3.3 TVRO System block diagrams

3.3.2 The outdoor unit

This consists of a receiving antenna feeding directly into a low-noise amplifier/converter combination. A parabolic reflector is generally used, with the receiving horn mounted at the focus. A common design is to have the focus directly in front of the reflector, but for better interference rejection, an offset feed may be used as shown.

Comparing the gain of a 3-m dish at 4 GHz with a 1-m dish at 12 GHz, the ratio D/λ equals 40 in each case, so the gains will be about equal. Although the free-space losses are much higher at 12 GHz compared with 4 GHz.

The downlink frequency band of 12.2 to 12.7 GHz spans a range of 500 MHz, which accommodates 32 TV/FM channels, each of which is 24-MHz wide. Obviously, some overlap occurs between channels, but these are alternately polarized *left-hand circular* (LHC) and *right-hand circular* (RHC) or vertical/horizontal, to reduce interference to accept- able levels. This is referred to as *polarization interleaving*. A polarizer that may be switched to the desired polarization from the indoor con- trol unit is required at the receiving horn.

The receiving horn feeds into a *low-noise converter* (LNC) or possibly a combination unit consisting of a *low-noise amplifier* (LNA) followed by a converter.

The combination is referred to as an LNB, for *low-noise block*. The LNB provides gain for the broadband 12-GHz signal and then converts the signal to a lower frequency range so that a low-cost coaxial cable can be used as feeder to the indoor unit.

The signal fed to the indoor unit is normally a wideband signal cov- ering the range 950 to 1450 MHz. This is amplified and passed to a tracking filter which selects the desired channel, as shown in Fig.

As previously mentioned, polarization interleaving is used, and only half the 32 channels will be present at the input of the indoor unit for any one setting of the antenna polarizer. This eases the job of the tracking filter, since alternate channels are well separated in frequency.

The selected channel is again down converted, this time from the 950- to 1450-MHz range to a fixed intermediate frequency, usually 70 MHz although other values in the *very high frequency* (VHF) range are also used.

The 70-MHz amplifier amplifies the signal up to the levels required for demodulation. A major difference between DBS TV and conventional TV is that with DBS, frequency modulation is used, whereas with conventional TV, amplitude modulation in the form of *vestigial single side- band* (VSSB) is used.

The 70-MHz, FM *intermediate frequency* (IF) carrier therefore must be demodulated, and the baseband information used to generate a VSSB signal which is fed into one of the VHF/UHF channels of a standard TV set.

3.4 Master Antenna TV System

A *master antenna TV* (MATV) system is used to provide reception of DBS TV/FM channels to a small group of users, for example, to the tenants in an apartment building. It consists of a single outdoor unit (antenna and LNA/C) feeding a number of indoor units, as shown in Fig.

It is basically similar to the home system already described, but with each user having access to all the channels independently of the other users. The advantage is that only one outdoor unit is required, but as shown, separate LNA/Cs and feeder cables are required for each sense of polarization.

Compared with the single- user system, a larger antenna is also required (2- to 3-m diameter) in order to maintain a good signal-to-noise ratio at all the indoor units.

Where more than a few subscribers are involved, the distribution system used is similar to the *community antenna* (CATV) system described in the following section.

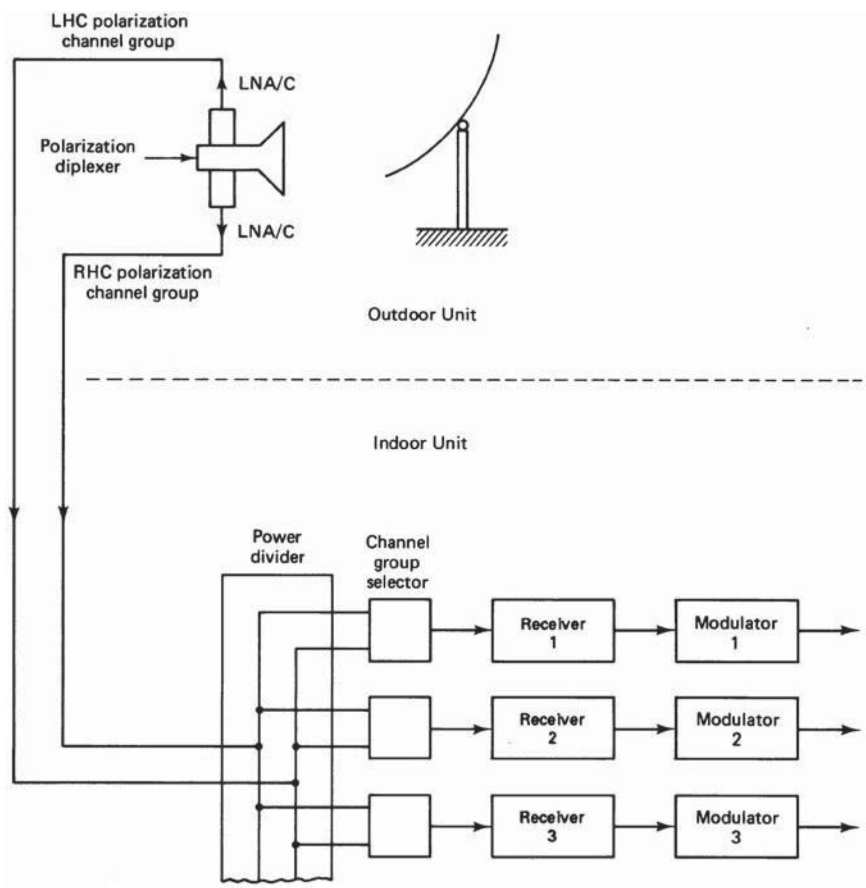


Figure 3.4 CATV System block diagrams

3.5 Community Antenna TV System

The CATV system employs a single outdoor unit, with separate feeds available for each sense of polarization, like the MATV system, so that all channels are made available simultaneously at the indoor receiver.

Instead of having a separate receiver for each user, all the carriers are demodulated in a common receiver-filter system, as shown in Fig. The channels are then combined into a standard multiplexed signal for transmission over cable to the subscribers.

In remote areas where a cable distribution system may not be installed, the signal can be rebroadcast from a low-power VHF TV transmitter.

Figure shows a remote TV station which employs an 8-m (26.2-ft) antenna for reception of the satellite TV signal in the C band.

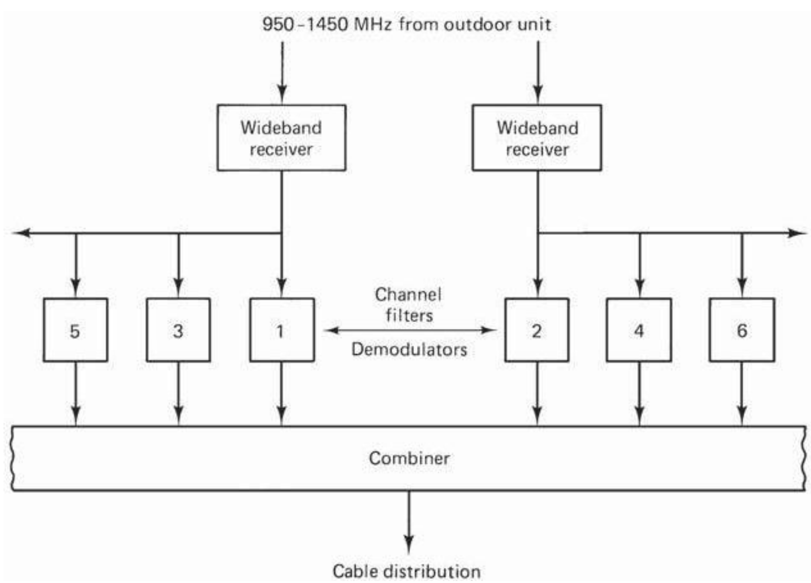


Figure 3.5 One possible arrangement for the indoor unit of a community antenna TV (CATV) system.

With the CATV system, local programming material also may be distributed to subscribers, an option which is not permitted in the MATV system.

3.6 Test Equipment Measurements on G/T, C/No, EIRP

Measurement of G/T of small antennas is easily and simply measured using the spectrum analyser method. For antennas with a diameter of less than 4.5 meters it is not normally necessary to point off from the satellite.

A step in frequency would be required into one of the satellite transponder guard bands.

However antennas with a G/T sufficiently large to enable the station to see the transponder noise floor either a step in frequency into one of the satellite transponder guard bands and/or in azimuth movement would be required.

The test signal can be provided from an SES WORLD SKIES beacon.

Procedure

(a) Set up the test equipment as shown below. Allow half an hour to warm up,
and then calibrate in accordance with the manufacturer’s procedures.

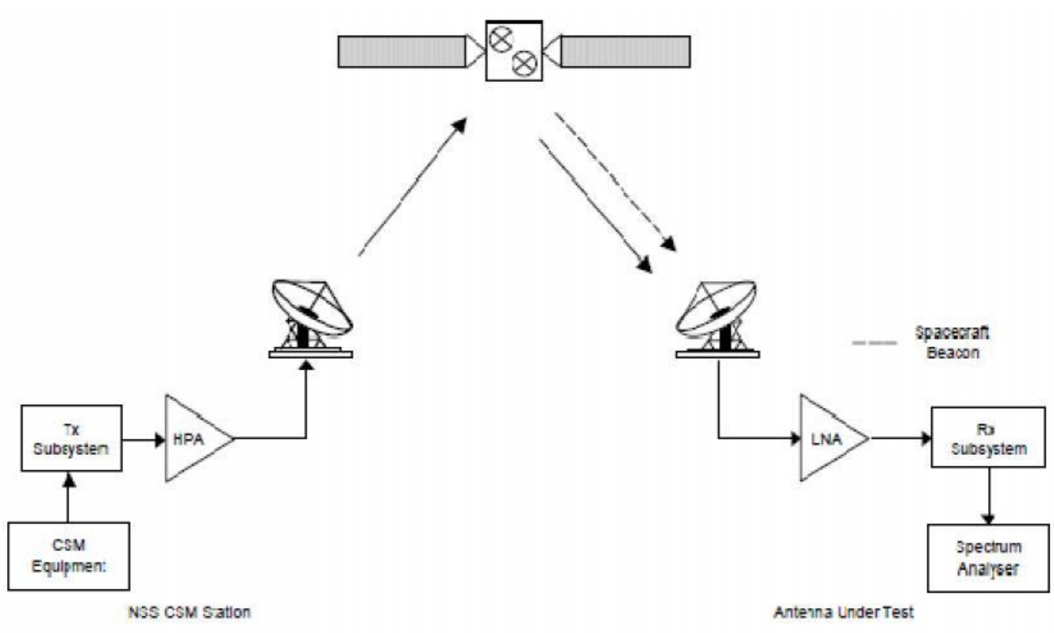


Figure 3.6 One possible arrangement for Measurement of G/T

(b) Adjust the centre frequency of your spectrum analyzer to receive the SES WORLD SKIES beacon (data to be provided on the satellite used for testing)

(c) Carefully peak the antenna pointing and adjust the polarizer by nulling the
cross polarized signal. You cannot adjust polarization when using the circularly polarized SES WORLD SKIES beacon.

(d) Configure the spectrum analyser as follows:

Centre Frequency: Adjust for beacon or test signal frequency (to be advised).
Use marker to peak and marker to centre functions.

- Frequency Span: 100 KHz
- Resolution Bandwidth: 1 KHz
- Video Bandwidth: 10 Hz (or sufficiently small to limit noise variance)
- Scale: 5 dB/div
- Sweep Time: Automatic
- Attenuator Adjust to ensure linear operation. Adjust to provide the "Noise floor delta" described in steps 7 and 8.

(e) To insure the best measurement accuracy during the following steps, adjust the spectrum analyser amplitude (reference level) so that the measured signal, carrier or noise, is approximately one division below the top line of the spectrum analyser display.

(f) Record the frequency and frequency offset of the test signal from the nominal frequency:
For example, assume the nominal test frequency is 11750 MHz but the spectrum analyser shows the peak at 11749 MHz. The frequency offset in this case is -1 MHz.

(g) Change the spectrum analyser centre frequency as specified by SES WORLD SKIES so that the measurement is performed in a transponder guard band so that only system noise power of the earth station and no satellite signals are received. Set the spectrum analyser frequency as follows:
Centre Frequency = Noise slot frequency provided by the PMOC

(h) Disconnect the input cable to the spectrum analyser and confirm that the noise floor drops by at least 15 dB but no more than 25dB. This confirms that the spectrum analyser's noise contribution has an insignificant effect on the measurement. An input attenuation value allowing a "Noise floor Delta" in excess of 25 dB may cause overloading of the spectrum analyser input. (i) Reconnect the input cable to the spectrum analyser.

(j) Activate the display line on the spectrum analyser.

(k) Carefully adjust the display line to the noise level shown on the spectrum analyser. Record the display line level.

(l) Adjust the spectrum analyser centre frequency to the test carrier frequency
recorded in step (e).

(m) Carefully adjust the display line to the peak level of the test carrier on the
spectrum analyser. Record the display line level.

(n) Determine the difference in reference levels between steps (l) and (j) which is the (C+N)/N.

(o) Change the (C+N)/N to C/N by the following conversion:

This step is not necessary if the (C+N)/N ratio is more than 20 dB because the resulting correction is less than 0.1 dB.

$$\left(\frac{C}{N}\right) = 10\log_{10}\left(10^{\frac{\left(\frac{C+N}{N}\right)}{10}} - 1\right) \quad \text{dB}$$

(p) Calculate the carrier to noise power density ratio (C/No) using:

$$\left(\frac{C}{No}\right) = \left(\frac{C}{N}\right) - 2.5 + 10\log_{10}(RBW \times SA_{corr}) \quad \text{dB}$$

The 2.5 dB figure corrects the noise power value measured by the log converters in the spectrum analyser to a true RMS power level, and the SA corr factor takes into account the actual resolution filter bandwidth. (q) Calculate the G/T using the following:

$$\left(\frac{G}{T}\right) - \left(\frac{C}{No}\right) - (EIRP_{SC} - A_{corr}) + (FSL + L_a) - 228.6 \qquad dB/K$$

where,
EIRPSC – Downlink EIRP measured by the PMOC
(dBW) Acorr – Aspect correction supplied by the PMOC
(dB)
FSL – Free Space Loss to the AUT supplied by the PMOC (dB)
La – Atmospheric attenuation supplied by the PMOC (dB)

(r) Repeat the measurement several times to check consistency of the result.

3.7 Antenna Gain

Antenna gain is usually **defined** as the ratio of the power produced by the **antenna** from a far-field source on the **antenna's** beam axis to the power produced by a hypothetical lossless isotropic **antenna**, which is equally sensitive to signals from all directions.

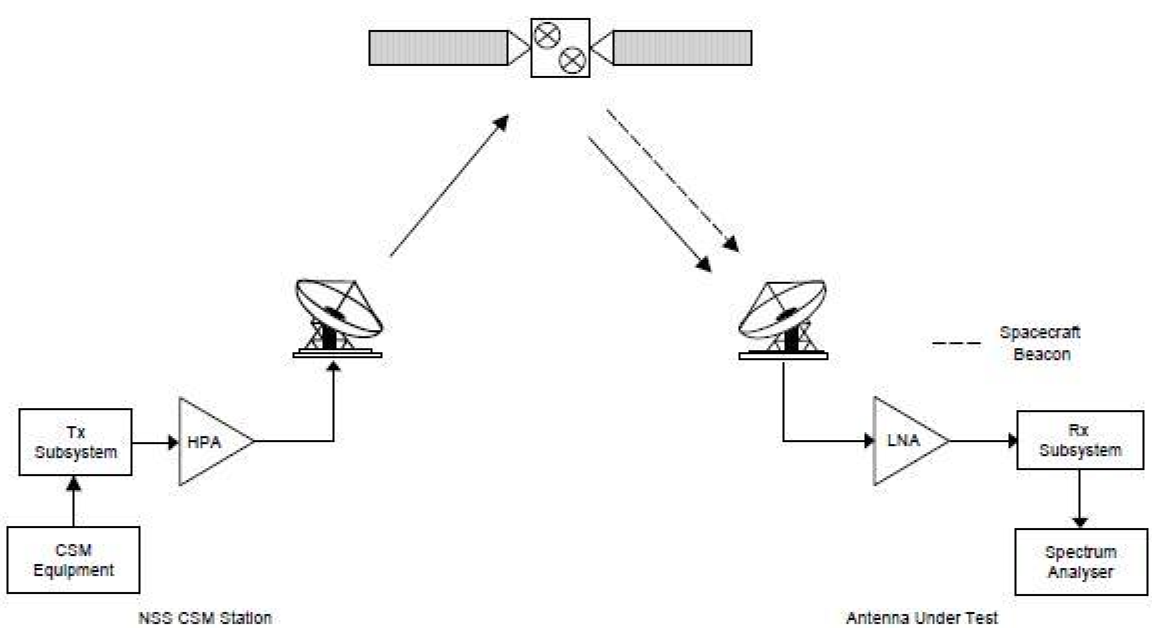


Figure 3.6 One possible arrangement for Measurement of Antenna Gain

Two direct methods of measuring the Rx gain can be used; integration of the Rx sidelobe pattern or by determination of the 3dB and 10dB beamwidths.

The use of pattern integration will produce the more accurate results but would require the AUT to have a tracking system. In both cases the test configurations for measuring Rx gain are identical, and are illustrated in Figure.

In order to measure the Rx gain using pattern integration the AUT measures the elevation and azimuth narrowband (±5° corrected) sidelobe patterns.

The AUT then calculates the directive gain of the antenna through integration of the sidelobe patterns. The Rx gain is then determine d by reducing the directive gain by the antenna inefficiencies.

In order to measure the Rx gain using the beamwidth method, the AUT measures the corrected azimuth and elevation 3dB/10dB beamwidths. From these results the Rx gain of the antenna can be directly calculated using the formula below.

$$G = 10\text{Log}_{10}\left[\frac{1}{2}\left(\frac{31000}{(Az_3)(El_3)} + \frac{91000}{(Az_{10})(El_{10})}\right)\right] - F_{Loss} - R_{Loss}$$

where:

G is the effective antenna gain (dBi)

Az3 is the corrected azimuth 3dB beamwidth

(°) El3 is the elevation 3dB beamwidth (°)

Az10 is the corrected azimuth 10dB beamwidth

(°) El10 is the elevation 10dB beamwidth (°)

FLoss is the insertion loss of the feed (dB)

RLoss is the reduction in antenna gain due to reflector inaccuracies, and is given by:

RLoss =4.922998677(Sdev f)2 dB

where: Sdev is the standard deviation of the actual reflector surface (inches)
f is the frequency (GHz)

APPLICATIONS

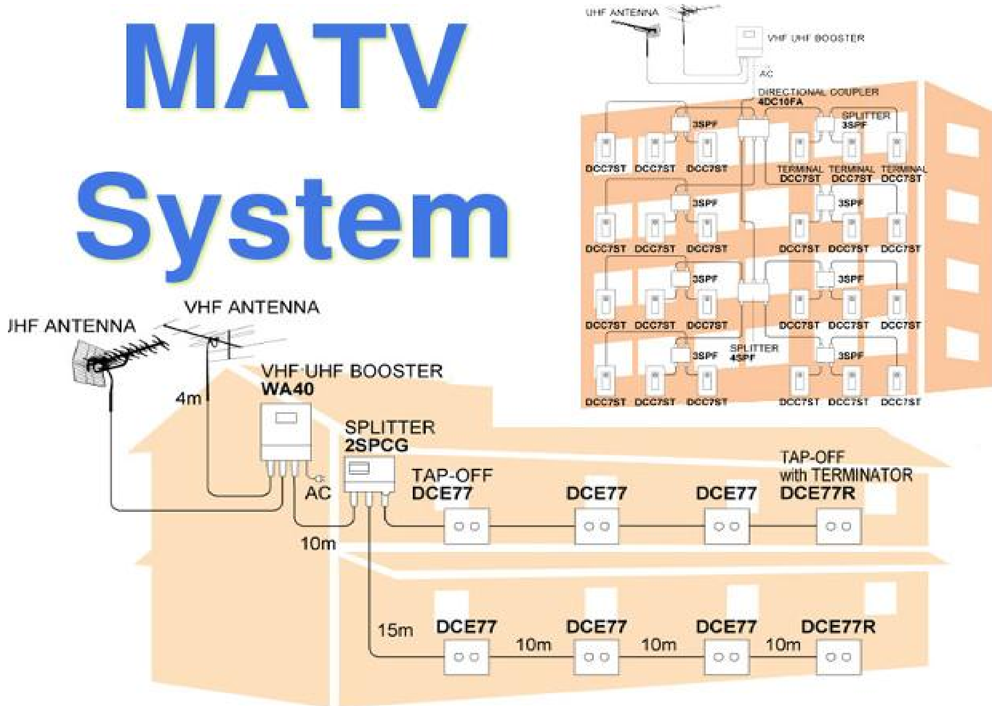


Figure an example of MATV system

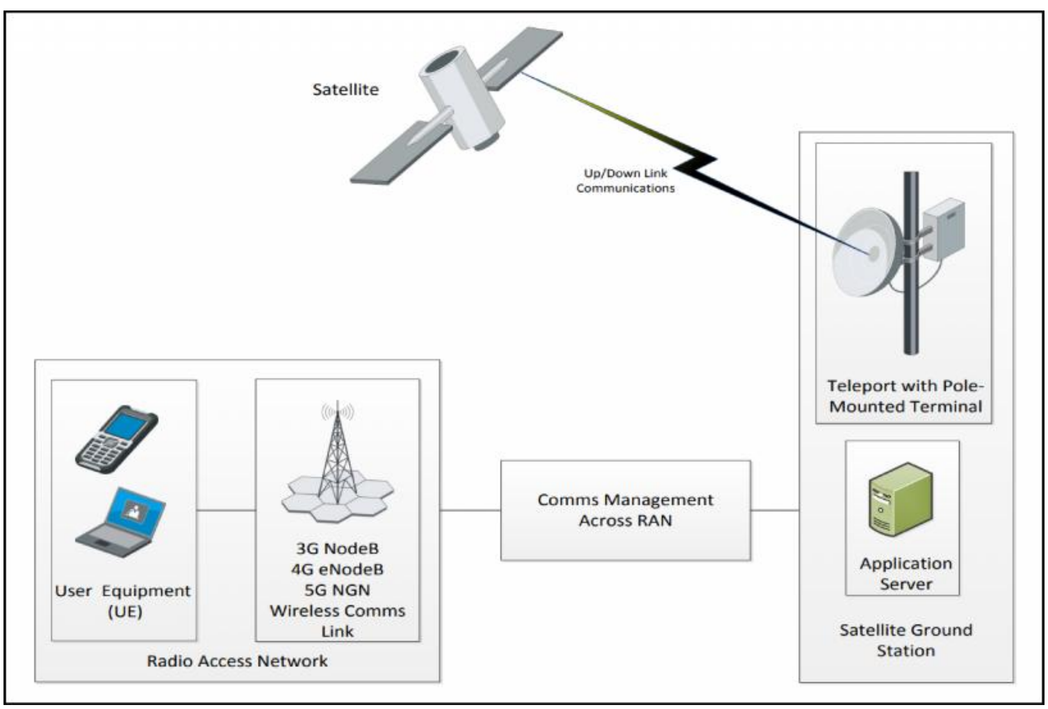


Figure an example of Satellite Earth Station

POST TEST MCQ

1. The earth segment of a satellite communications system consists of ________.

a) The earth segment of a satellite communications system consists of the transmit and receive earth stations

b) With active attitude control, there is no overall stabilizing torque present to resist the disturbance torques

c) Proper moment

d) None of these

2. In the Americas, for example, the down-link band is ________

a) 12.2 to 12.7 GHz

b) With active attitude control, there is no overall stabilizing torque present to resist the disturbance torques

c) The Ku (12-GHz) band

d) 64-GHz to 164 GHz

3. The major difference between the Ku-band and the C-band that satellites intended for DBS have much ________.

a) Low noise

b) The indoor unit

c) The home unit

d) higher equivalent isotropic radiated power(EIRP)

4. MATV stand as________

a) Multi amplitude TV

b) Maximum Angular TV

c) Master antenna TV

d) Multiplex All TV

5. The orbital spacing is ______ for the high-power satellites, so adjacent satellite interference is considered nonexistent.

a) 18°

b) 9°

c) 27°.

d) 45°

6. A satellite may carry ______ transponders

a) 32

b) 41

c) 24

d) 64

7. The frequencies for direct broadcast satellites vary from region to region throughout the world, although these are generally in the_____________

a) Ku band

b) Ka band

c) C band

d) W band

8. A master antenna TV (MATV) system is used to provide reception of DBS TV/FM channels to a small group of users, for example_____________

a) To the tenants in an apartment building

b) With attitude control, there is no overall stabilizing torque present to resist the disturbance torques

c) The Ku (12-GHz) band

d) None of these

9. A master antenna TV (MATV) system is used to provide reception of __________to a small group of users, for example to the tenants in an apartment building.

a) DBS TV/FM channels

b) FM channels

c) The Ku (12-GHz) band

d) None of these

10. CATV stands as ________

a) Community antenna TV

b) Carrier angular TV

c) Cost amplitude TV

d) Cost angular TV

11. The CATV system employs a single_________ with separate feeds available for each sense of polarization.

a) Outdoor unit

b) Indoor unit

c) TV unit

d) Input unit

12. With the CATV system, local programming material also may be distributed to subscribers, an option which is _________in the MATV system.

a) Not permitted

b) Permitted

c) Transmitted

d) None of these

13. Out of the following, find what is an polar antenna.

a) A single actuator is used which moves the antenna in a circular arc ie known as polar mount antenna.

b) An double actuator is used which moves the antenna in a circular arc ie known as polar mount antenna.

c) A single actuator is used which moves the antenna in a elliptical arc ie known as polar mount antenna.

d) None of above

14. Definition of a transponder is

a) In a communication satellite, the equipment which provides the connecting link between the satellite's transmit & receive antennas is referred to as the transponder.

b) In a communication satellite, the equipment which provides the power supply is referred to as the transponder

c) a & b

d) None of above

15. Definition of spot beam antenna is

a) A beam generated by a communication satellite antenna of sufficient size that the angular spread of sufficient size that the angular spread of the energy in the beam is very small with the result that a region that is only a few hundred km in diameter is illuminated on earth.

b) A beam generated by a communication satellite antenna of sufficient size that the angular spread of sufficient size that the angular spread of the energy in the beam is very slarge with the result that a region that is only a few hundred mm in diameter is illuminated on earth.

c) Either a or b.

d) None of above

16. Definition of an EIRP is

a) It is a measure of radiated or transmitted power of an antenna. It can be completed from the antenna gain & the power fed from the antenna output.

b) It is a measure of radiated or transmitted power of an antenna. It can be completed from the antenna gain & the power fed to the antenna input.

c) Either a or b

d) None of above

17. **Antenna losses is defined as**

a) It is add to noise received as radiation is in the sum of the equivalent noise temperature of all these sources.

b) It is add to noise received as radiation & the total antenna noise temperature is in the divider of the equivalent noise temperature of all these sources.

c) It is add to noise received as radiation & the total antenna noise temperature is in the sum of the equivalent noise temperature of all these sources.

d) None of above

18. Satellite launch sites are invariably located on Eastern seaboards to ensure that

a) launch takes place eastward

b) expenditure of propulsion fuel is reduced during plane changing

c) the satellite achieves circular orbit quickly

d) spent rocket motor and other launcher debris falls into the sea

19. Of the four INSAT-I satellites planned by India so for, only __________ has proved to be successful.

a) INSAT-IA

b) INSAT-IB

c) INSAT-IC

d) INSAT-ID

20. India's first domestic geostationary satellite 1NSAT-IA was launched on 10th April 1982 from

a) USSR

b) USA

c) UK

d) Ukraine

21. A transponder is a satellite equipment which

a) receives a signal from Earth station and amplifies

b) changes the frequency of the received signal

c) retransmits the received signal

d) does all of the above-mentioned functions

CONCLUSION

- In this unit we described the earth segment and space segment.
- The Outdoor UNIT – Indoor UNIT for analog (FM) TV were elaborated.
- The Equipment Measurements on G/T, C/N, EIRP was discussed.

REFERENCES

1. Wilbur L.Pritchard, Hendri G. Suyderhoud, Robert A. Nelson, “Satellite Communication Systems Engineering”, Prentice Hall/Pearson, 2007.
2. N. Agarwal, “Design of Geosynchronous Space Craft”, Prentice Hall, 1986.
3. Bruce R. Elbert, “Satellite Communication Applications”, Hand Book, Artech House Bostan London,1997.
4. Tri T. Ha, “Digital Satellite Communication”, II nd edition, 1990.
5. Robert G. Winch, “Telecommunication Transmission Systems”, Mc Graw-Hill Book Co., 1983.
6. Brian Ackroyd, “World Satellite Communication and earth station Design”, BSP professional Books, 1990.
7. G.B.Bleazard, “Introducing Satellite communications“, NCC Publication, 1985.
8. M.Richharia, “Satellite Communication Systems-Design Principles”, Macmillan 2003.
9. Emanuel Fthenakis, “Manual of Satellite Communications”, Mc Graw Hill Book Co., 1984.

ASSIGNMENT

1. What is Cassegrain antenna popular for large earth stations?
2. Why is G/T ratio a useful parameter to characterize earth stations?
3. Explain the major test equipments required at an earth station?
4. Write short notes on Community antenna TV system?
5. Write the short notes on TVRO system?
6. Explain about the feeder losses and antenna misalignment losses?
7. Derive link power budget equation.

SATELLITE COMMUNICATION

UNIT IV SATELLITE ACCESS

Prepared by

Dr.M.A.ARCHANA

AP/ECE,SCSVMV

UNIT IV SATELLITE ACCESS

AIM & OBJECTIVE

- To understand the basics of Modulation and Multiplexing
- To analyze the Assignment Methods.
- To understand the Spread Spectrum communication.

PRE TEST MCQ

1. In satellite communication modulation is used.

a) AM

b) FM

c) PWM

d) PAM

2. In TV broadcast via satellite the TV signal from the main broadcast station is routed to the earth station via

a) Low power transmitter

b) Microwave links

c) TV relay stations

d) Microwave repeater stations

3. The main advantage of satellite communication is

a) Low cost

b) Low distortion

c) High reliability

d) High band width

4.1 Modulation and Multiplexing: Voice, Data, Video

Communications satellites are used to carry telephone, video, and data signals, and can use both analog and digital modulation techniques.

Modulation

Modification of a carrier's parameters (amplitude, frequency, phase, or a combination of them) in dependence on the symbol to be sent.

Multiplexing

Task of multiplexing is to assign space, time, frequency, and code to each communication channel with a minimum of interference and a maximum of medium utilization Communication channel refers to an association of sender(s) and receiver(s) that want to exchange data One of several constellations of a carrier's parameters defined by the used modulation scheme.

4.1.1 Voice, Data, Video

The modulation and multiplexing techniques that were used at this time were analog, adapted from the technology developed for The change to digital voice signals made it easier for long-distance.

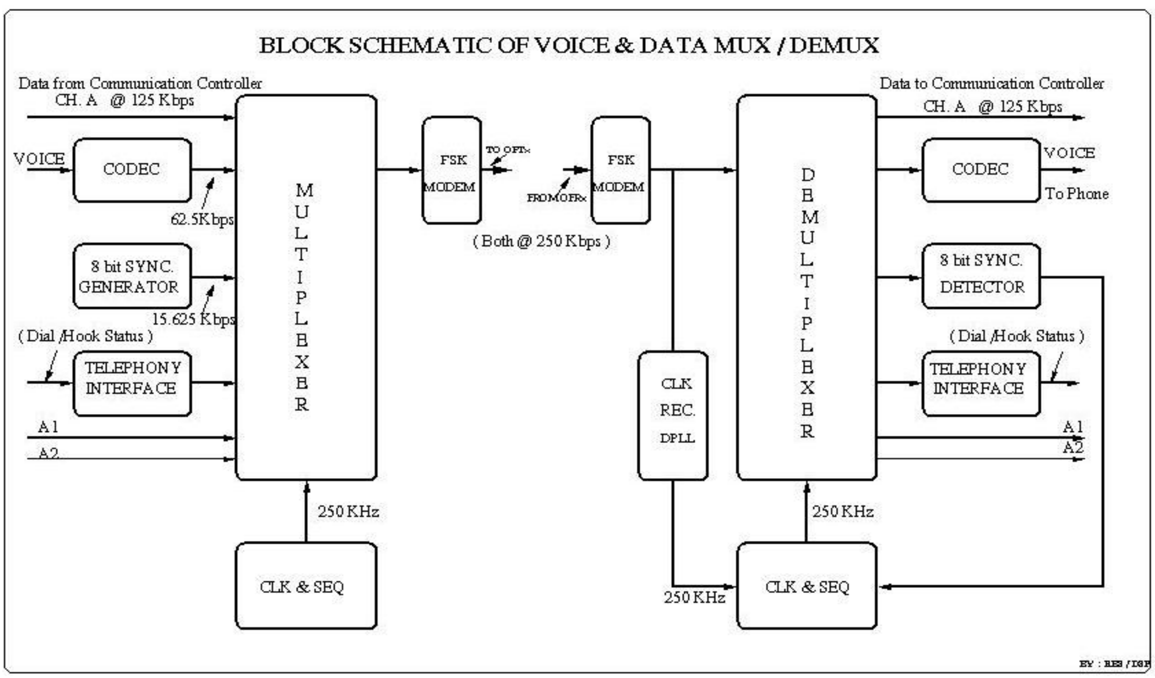


Figure 4.1 Modulation and Multiplexing: Voice/Data/Video

Communication carriers to mix digital data and telephone Fiber-optic Cable Transmission Standards System Bit rate (Mbps) 64 - kbps Voice channel capacity Stuffing bits and words are added to the satellite data stream as n eeded to fill empty bit and word spaces.

Primarily for video provided that a satellite link's overall carrier-to-noise but in to older receiving equipment at System and Satellite Specification Ku - band satellite parameters.

4.1.2 Modulation And Multiplexing

In analog television (TV) transmission by satellite, the baseband video signal and one or two audio subcarriers constitute a composite video signal.

Digital modulation is obviously the modulation of choice for transmitting digital data are digitized analog signals may conveniently share a channel with digital data, allowing a link to carry a varying mix of voice and data traffic.

Digital signals from different channels are interleaved for transmission through time division multiplexing TDM carry any type of traffic â€” the bent pipe transponder that can carry voice, video, or data as the marketplace demands.

Hybrid multiple access schemes can use time division multiplexing of baseband channels which are then modulate.

4.2 Analog – digital transmission system

4.2.1 Analog vs. Digital Transmission

Compare at two levels:

1. Data—continuous (audio) vs. discrete (text)
2. Signaling—continuously varying electromagnetic wave vs. sequence of voltage pulses.

Also Transmission—transmit without regard to signal content vs. being concerned with signal content. Difference in how attenuation is handled, but not focus on this. Seeing a shift towards digital transmission despite large analog base. Why?

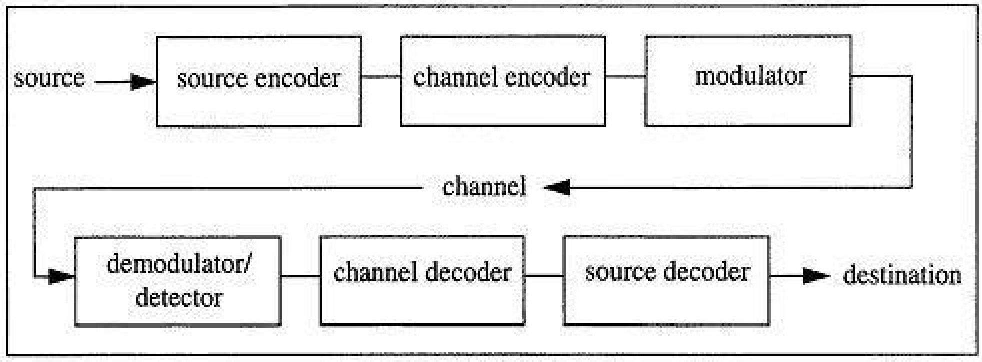


Figure 4.2 basic communication systems

- Improving digital technology
- Data integrity. Repeaters take out cumulative problems in transmission. Can thus transmit longer distances.
- Easier to multiplex large channel capacities with digital
- Easy to apply encryption to digital data
- Better integration if all signals are in one form. Can integrate voice, video and digital data.

4.2.2 Digital Data/Analog Signals

Must convert digital data to analog signal such device is a modem to translate between bit-serial and modulated carrier signals?

To send digital data using analog technology, the sender generates a carrier signal at some continuous tone (e.g. 1 -2 kHz in phone circuits) that looks like a sine wave. The following techniques are used to encode digital data into analog signals.

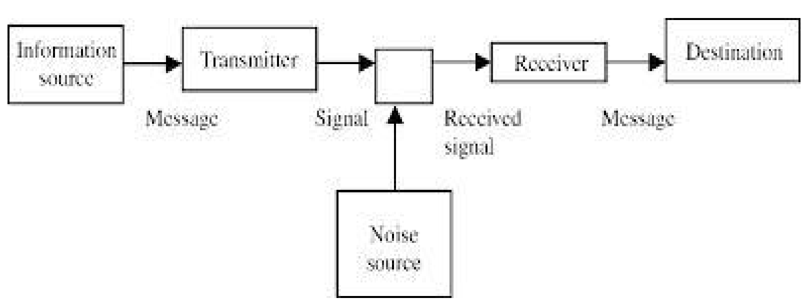


Figure 4.3 Digital /Analog Transmitter & receiver

Resulting bandwidth is centered on the carrier frequency.

- Amplitude-shift modulation (keying): vary the amplitude (e.g. voltage) of the signal. Used to transmit digital data over optical fiber.
- Frequency-shift modulation: two (or more tones) are used, which are near the carrier frequency. Used in a full-duplex modem (signals in both directions).
- Phase-shift modulation: systematically shift the carrier wave at uniformly spaced intervals.

For instance, the wave could be shifted by 45, 135, 225, 315 degree at each timing mark. In this case, each timing interval carries 2 bits of information.

Why not shift by 0, 90, 180, 270? Shifting zero degrees means no shift, and an extended set of no shifts leads to clock synchronization difficulties.

Frequency division multiplexing (FDM): Divide the frequency spectrum into smaller subchannels, giving each user exclusive use of a subchannel (e.g., radio and TV). One problem with FDM is that a user is given all of the frequency to use, and if the user has no data to send, bandwidth is wasted — it cannot be used by another user.

Time division multiplexing (TDM): Use time slicing to give each user the full bandwidth, but for only a fraction of a second at a time (analogous to time sharing in operating systems). Again, if the user doesn't have data to sent during his time slice, the bandwidth is not used (e.g., wasted).

Statistical multiplexing: Allocate bandwidth to arriving packets on demand. This leads to the most efficient use of channel bandwidth because it only carries useful data. That is, channel bandwidth is allocated to packets that are waiting for transmission, and a user generating no packets doesn't use any of the channel resources.

4.3. Digital Video Broadcasting (DVB)

- Digital Video Broadcasting (DVB) has become the synonym for digital television and for data broadcasting world-wide.

- DVB services have recently been introduced in Europe, in North- and
- South America, in Asia, Africa and Australia.

o This article aims at describing what DVB is all about and at introducing some of the technical background of a technology that makes possible the broadcasting.

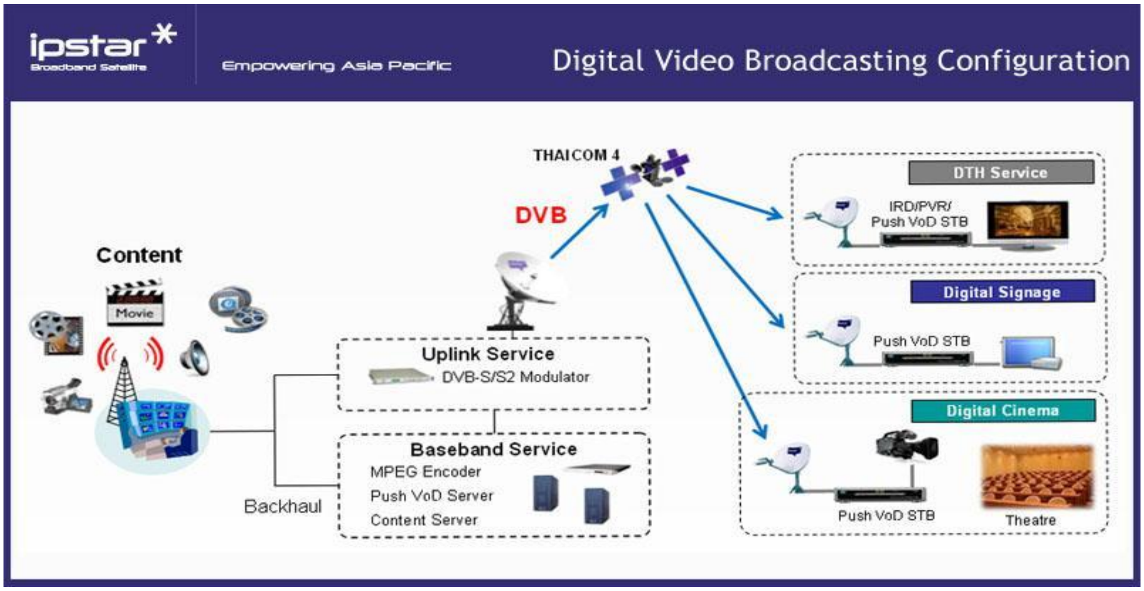


Figure 4.4 Digital Video Broadcasting systems

4.4 Multiple Access Techniques

❖ The transmission from the BS in the downlink can be heard by each and every mobile user in the cell, and is referred as *broadcasting*. Transmission from the mobile users in the uplink to the BS is many-to-one, and is referred to as multiple access.

❖ Multiple access schemes to allow many users to share simultaneously a finite amount of radio spectrum resources.

- Should not result in severe degradation in the performance of the system as compared to a single user scenario.
- Approaches can be broadly grouped into two categories: narrowband and wideband.

❖ Multiple Accessing Techniques : with possible conflict and conflict - free

- Random access
- Frequency division multiple access (FDMA)
- Time division multiple access (TDMA)
- Spread spectrum multiple access (SSMA) : an example is Code division multiple access (CDMA)
- Space division multiple access (SDMA)

Duplexing

For voice or data communications, must assure two way communication (duplexing, it is possible to talk and listen simultaneously). Duplexing may be done using frequency or time domain techniques.

- ❖ Forward (downlink) band provides traffic from the BS to the mobile
- ❖ Reverse (uplink) band provides traffic from the mobile to the BS.

4.4.1 Frequency division duplexing (FDD)

Provides two distinct bands of frequencies for every user, one for downlink and one for uplink.

A large interval between these frequency bands must be allowed so that interference is minimized.

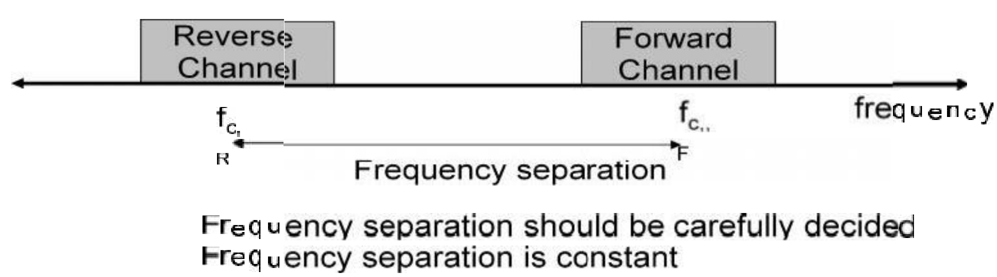


Figure 4.5 Frequency Separation

4.4.2. Time division duplexing (TDD)

- ❖ In TDD communications, both directions of transmission use one contiguous frequency allocation, but two separate time slots to provide both a forward and reverse link.

- Because transmission from mobile to BS and from BS to mobile alternates in time, this scheme is also known as “ping pong”.
- As a consequence of the use of the same frequency band, the communication quality in both directions is the same. This is different from FDD.

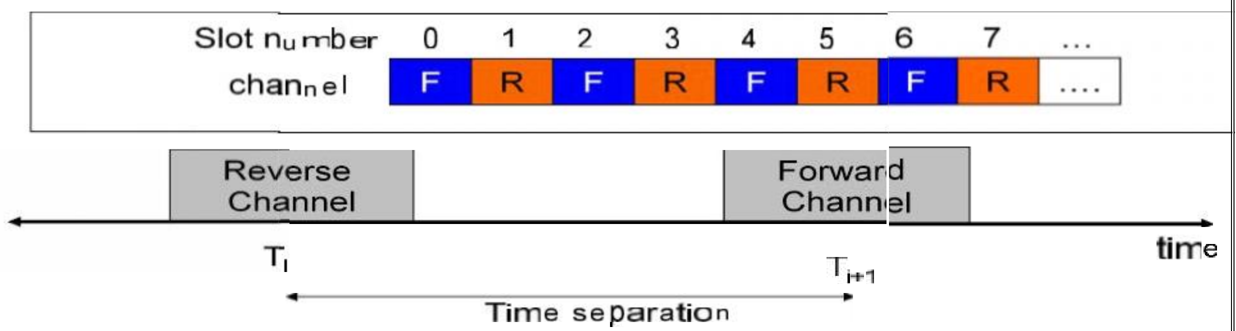


Figure 4.6 Time Slot

4.4.3 FDMA

- In FDMA, each user is allocated a unique frequency band or channel. During the period of the call, no other user can share the same frequency band.

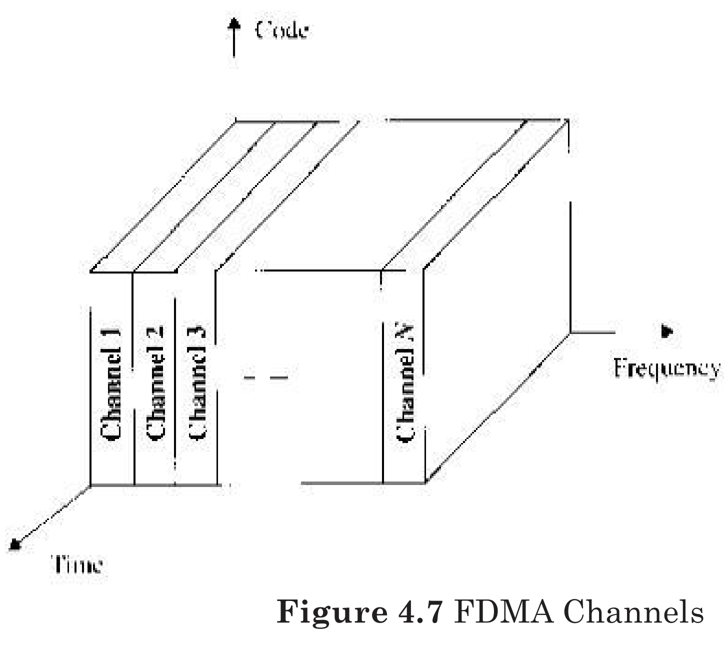


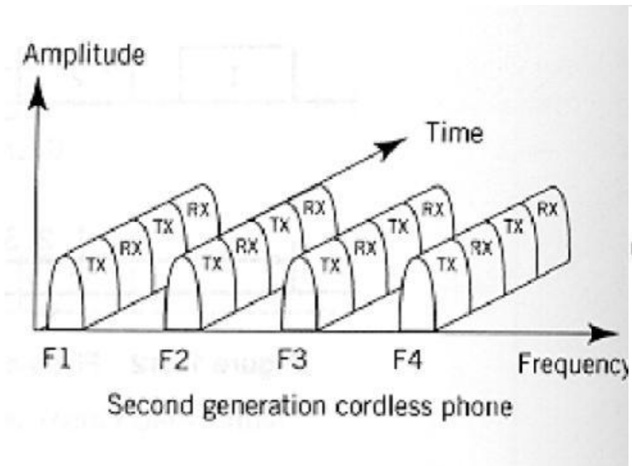
Figure 4.7 FDMA Channels

- All channels in a cell are available to all the mobiles. Channel assignment is carried out on a first-come first- served basis.

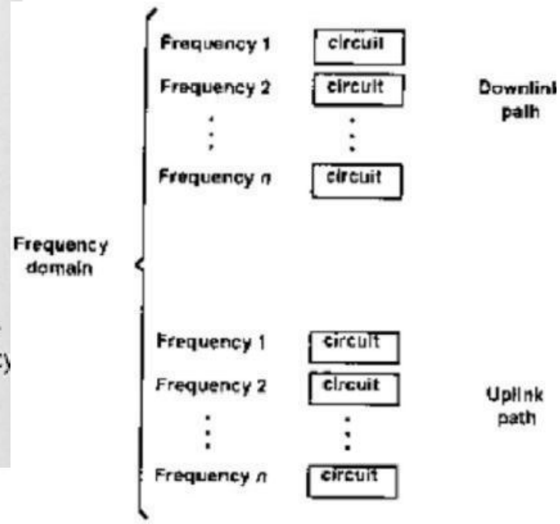
❖ The number of channels, given a frequency spectrum BT , depends on the modulation technique (hence Bw or Bc) and the guard bands between the channels $2Bguard$.

❖ These guard bands allow for imperfect filters and oscillators and can be used to minimize adjacent channel interference.

❖ FDMA is usually implemented in narrowband systems.



FDMA / TDD



FDMA / FDD

Figure 4.8 FDMA/FDD/TDD

Nonlinear effects in FDMA

- In a FD MA system, many channels share t he same antenna at the BS. The power amplifiers or the power combiners, when operated at or near saturation are non linear.
- The nonlinear ties generate inter-modulation frequencies.
- Undesirable harmonics generated outside the mobile radio band cause interference to adjacent services.
- Undesirable harmonics present inside the band ca use interference to other users in the mobile system.

4.4.4 TDMA

- TDMA systems divide the channel time into frames. Each frame is further partitioned into time slots. In each slot only one user is allowed to either transmit or receive.
- Unlike FDMA, only digital data and digital modulation must be used.
- Each user occupies a cyclically repeating time slot, so a channel may be thought of as a particular time slot of every frame, where *N* time slots comprise a frame.

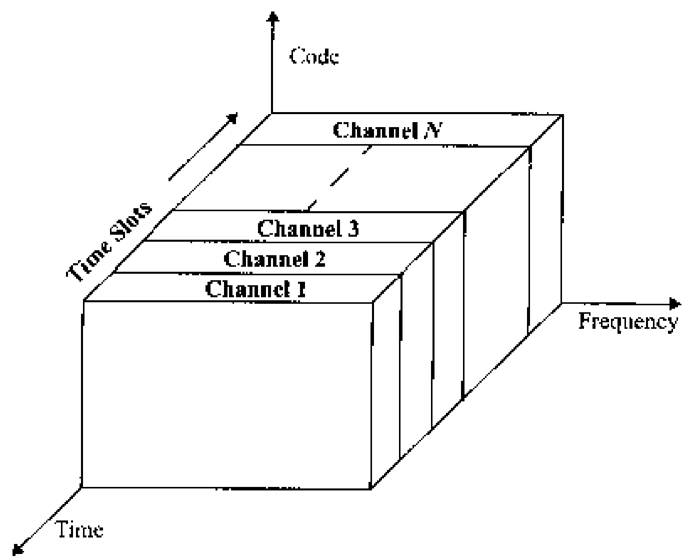


Figure 4.9 TDMA Channels

Features

- Multiple channels per carrier or RF channels.
- Burst transmission since channels are used on a timesharing basis.
  - Transmitter can be turned off during idle periods.
- Narrow or wide bandwidth – depends on factors such as modulation scheme, number of voice channels per carrier channel.
- High ISI – Higher transmission symbol rate, hence resulting in high ISI.
  - Adaptive equalizer required.

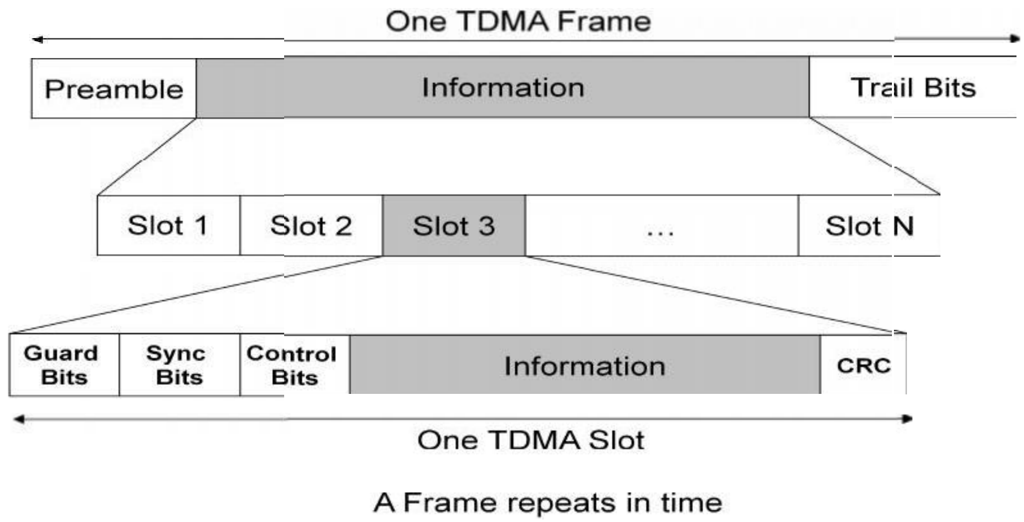


Figure 3.10 TDMA Channels time slot

- A guard time between the two time slots must be allowed in order to avoid
  - interference, especially in the uplink direction. All synchronize with BS to minimize interference.
- mobiles should
- Efficient power utilization : FDMA systems require a 3- to 6-dB power back
  - off in order to compensate for inter-modulation effects.
- Efficient handoff : TDMA systems can take advantage of the fact that the transmitter is switched off during idle time slots to improve the handoff procedure. An enhanced link control, such as that provided by mobile assisted handoff (MAHO) can be carried out by a subscriber by listening to
  - neighboring base station during the idle slot of the TDMA frame.
- Efficiency of TDMA
- Efficiency of TDMA is a measure of the percentage of bits per frame whi ch contain transmitted data. The transmitted data include source and channel coding bits.

- $$\eta_f = \frac{b_T - b_{OH}}{b_T} \cdot 100\%$$

- *bOH* includes all overhead bits such as preamble, guard bits, etc.

4.4.5 Code Division Multiple Access (CDMA)

- Spreading signal (code) consists of chips
  - Has Chip period and and hence, chip rate
  - Spreading signal use a pseudo-noise (PN) sequence (a pseudo-random sequence)
  - PN sequence is called a codeword
  - Each user has its own cordword
  - Codewords are orthogonal. (low autocorrelation)
  - Chip rate is order of magnitude larger than the symbol rate.
- The receiver correlator distinguishes the senders signal by examining the wideband signal with the same time-synchronized spreading code
- The sent signal is recovered by despreading process at the receiver.

CDMA Advantages:

- Low power spectral density.
  - Signal is spread over a larger frequency band
  - Other systems suffer less from the transmitter
- Interference limited operation
  - All frequency spectrum is used
- Privacy
  - The codeword is known only between the sender and receiver. Hence other users can not decode the messages that are in transit
- Reduction of multipath affects by using a larger spectrum

CDMA data

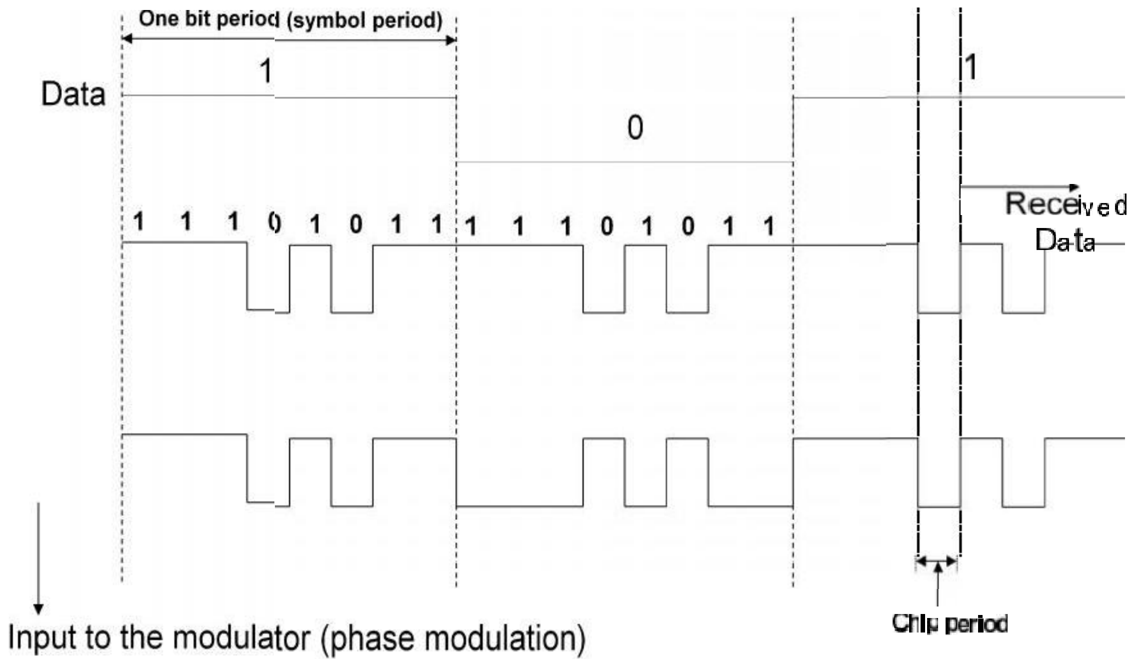


Figure 4.11 CDMA Channels transmission

DSSS Transmitter:

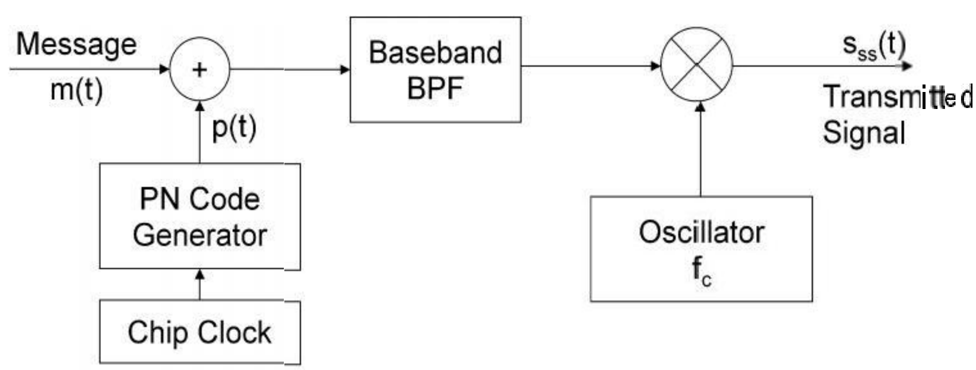
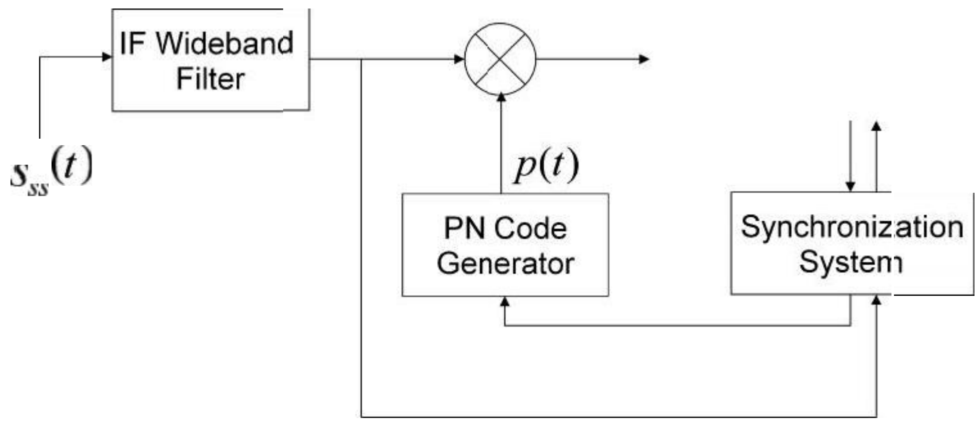


Figure 4.12 CDMA Transmitter

$$s_{ss}(t) = \sqrt{\frac{2E_s}{T_s}} m(t)\, p(t) \cos(2\pi f_c t + \theta)$$

DSSS Receiver



$$s_1(t) = \sqrt{\frac{2E_s}{T_s}} m(t) \cos(2\pi f_c t + \theta)$$

Figure 4.13 CDMA Receiver

- FDMA/CDMA
  - Available wideband spectrum is frequency divided into number narrowband radio channels. CDMA is employed inside each channel.
- DS/FHMA
  - The signals are spread using spreading codes (direct sequence signals
  - are obtained), but these signal are not transmitted over a constant
  - carrier frequency; they are transmitted over a frequency hopping carrier frequency.
- Time Division CDMA (TCDMA)

Each cell is using a different spreading code (CDMA employed between cells) that is conveyed to the mobiles in its range.

Inside each cell (inside a CDMA channel), TDMA is employed to multiplex multiple users.

- Time Division Frequency Hopping

  At each time slot, the user is hopped to a new frequency according to a pseudo-random hopping sequence.

  Employed in severe co-interference and multi-path environments.

Bluetooth and GSM are using this technique

- A large number of independently steered high-gain beams can be formed without any resulting degradation in SNR ratio.
- Beams can be assigned to individual users, thereby assuring that all links operate with maximum gain.
- Adaptive beam forming can be easily implemented to improve the system capacity by suppressing co channel interference.

Advantage of CDMA

- It is recognized that CDMA's capacity gains over TDMA
- FDMA are entirely due to Its tighter, dynamic control over the use of the power domain.
- Choosing a new non-orthogonal PN sequence a CDMA system does not encounter the difficulties of choosing a spare carrier frequency or time slot to carry a Traffic Channel
- Ensure that interference will not be too great if it begins to transmit -that there is still enough space left in the power domain.

Disadvantages of CDMA

- Satellite transponders are channelized too narrowly for roadband CDMA, which is the most attractive form of CDMA.
- Power control cannot be as tight as it is in a terrestrial system because of long round-trip delay.

4.5. Channel allocation schemes

In radio resource management for wireless and cellular network, channel allocation schemes are required to allocate bandwidth and communication channels to base stations, access points and terminal equipment.

The objective is to achieve maximum system spectral efficiency in bit/s/Hz/site by means of frequency reuse, but still assure a certain grade of service by avoiding co-channel interference and adjacent channel interference among nearby cells or networks that share the bandwidth. There are two types of strategies that are followed:-

- Fixed: FCA, fixed channel allocation: Manually assigned by the network operator
- Dynamic:
  - DCA, dynamic channel allocation,
  - DFS, dynamic frequency selection
  - Spread spectrum

4.5.1 FCA

In **Fixed Channel Allocation** or **Fixed Channel Assignment** (FCA) each cell is given a predetermined set of frequency channels.

FCA requires manual frequency planning, which is an arduous task in TDMA and FDMA based systems, since such systems are highly sensitive to co-channel interference from nearby cells that are reusing the same channel.

This results in traffic congestion and some calls being lost when traffic gets heavy in some cells, and idle capacity in other cells.

4.5.2. DCA and DFS

Dynamic Frequency Selection (DFS) may be applied in wireless networks with several adjacent non-centrally controlled access points.

A more efficient way of channel allocation would be **Dynamic Channel Allocation** or **Dynamic Channel Assignment** (DCA) in which voice channel are not allocated to cell permanently, instead for every call request base station request channel from MSC.

4.6 Spread spectrum

Spread spectrum can be considered as an alternative to complex DCA algorithms. Spread spectrum avoids cochannel interference between adjacent cells, since the probability that users in nearby cells use the same spreading code is insignificant.

Thus the frequency channel allocation problem is relaxed in cellular networks based on a combination of Spread spectrum and FDMA, for example IS95 and 3G systems.

In packet based data communication services, the communication is bursty and the traffic load rapidly changing. For high system spectrum efficiency, DCA should be performed on a packet-by-packet basis.

Examples of algorithms for packet-by-packet DCA are **Dynamic Packet Assignment** (DPA), Dynamic Single Frequency Networks (DSFN) and **Packet and resource plan scheduling** (PARPS).

4.6.1 Spread spectrum Techniques

1 In telecommunication and radio communication, spread-spectrum techniques are methods by which a signal (e.g. an electrical, electromagnetic, or acoustic signal) generated with a particular bandwidth is deliberately spread in the frequency domain, resulting in a signal with a wider bandwidth.

2 These techniques are used for a variety of reasons, including the establishment of secure communications, increasing resistance to natural interference, noise and jamming, to prevent detection, and to limit power flux density (e.g. in satellite downlinks).

3 Spread-spectrum telecommunications this is a technique in which a telecommunication signal is transmitted on a bandwidth considerably larger than the frequency content of the original information.

4 Spread-spectrum telecommunications is a signal structuring technique that employs direct sequence, frequency hopping, or a hybrid of these, which can be used for multiple access and/or multiple functions.

5 Frequency-hopping spread spectrum (FHSS), direct-sequence spread spectrum (DSSS), time-hopping spread spectrum (THSS), chirp spread spectrum (CSS).

6 Techniques known since the 1940s and used in military communication systems since the 1950s "spread" a radio signal over a wide frequency range several magnitudes higher than minimum requirement.

7 Resistance to jamming (interference). DS (direct sequence) is good at resisting continuous-time narrowband jamming, while FH (frequency hopping) is better at resisting pulse jamming.

8 Resistance to fading. The high bandwidth occupied by spread- spectrum signals offer some frequency diversity, i.e. it is unlikely that the

signal will encounter severe multipath fading over its whole bandwidth, and in other cases the signal can be detected using e.g. a Rake receiver.

9 Multiple access capability, known as code-division multiple access (CDMA) or code-division multiplexing (CDM). Multiple users can transmit simultaneously in the same frequency band as long as they use different spreading codes.

4.7 Compression – Encryption

At the broadcast center, the high-quality digital stream of video goes through an MPEG encoder, which converts the programming to MPEG-4 video of the correct size and format for the satellite receiver in your house.

Encoding works in conjunction with compression to analyze each vi deo frame and eliminate redundant or irrelevant data and extrapolate information from other frames. This process reduces the overall size of the file. Each frame can be encoded in one of three ways:

- As an **intraframe**, which contains the complete image data for that frame. This method provides the least ompression.
- As a **predicted** frame, which contains just enough information to tell the satellite receiver how to display the frame based on the most recently displayed intraframe or predicted frame.
- ❖ As a **bidirectional** frame, which displays information from the surrounding intraframe or predicted frames. Using data from the closest surrounding frames, the receiver **interpolates** the position and color of each pixel.

This process occasionally produces **artifacts** -- glitches in the video image. One artifact is **macroblocking,** in which the fluid picture temporarily dissolves into blocks. Macroblocking is often mistakenly called **pixilating,** a technically incorrect term which has been accepted as slang for this annoying artifact.

There really are pixels on your TV screen, but they're too small for your human eye to perceive them individually -- they're tiny squares of video data that make up the image you see. (For more information about pixels and perception, see How TV Works.)

The rate of compression depends on the nature of the programming. If the encoder is converting a newscast, it can use a lot more predicted frames because most of the scene stays the same from one frame to the next.

In more fast-paced programming, things change very quickly from one frame to the next, so the encoder has to create more intraframes. As a result, a newscast generally compresses to a smaller size than something like a car race.

4.7.1 Encryption and Transmission

After the video is compressed, the provider encrypts it to keep people from accessing it for free. Encryption scrambles the digital data in such a way that it can only be **decrypted** (converted back into usable data) if the receiver has the correct decryption algorithm and security keys.

Once the signal is compressed and encrypted, the broadcast center beams it directly to one of its satellites. The satellite picks up the signal with an onboard dish, amplifies the signal and uses another dish to beam the signal back to Earth, where viewers can pick it up.

In the next section, we'll see what happens when the signal reaches a viewer's house.

4.7.2 Video and Audio Compression

Video and Audio files are very large beasts. Unless we develop and maintain very high bandwidth networks (Gigabytes per second or more) we have to compress to data.

Relying on higher bandwidths is not a good option -- M25 Syndrome: Traffic needs ever increases and will adapt to swamp current limit whatever this is.

As we will compression becomes part of the representation or *coding* scheme which have become popular audio, image and video formats.

We will first study basic compression algorithms and then go on to study some actual coding formats.

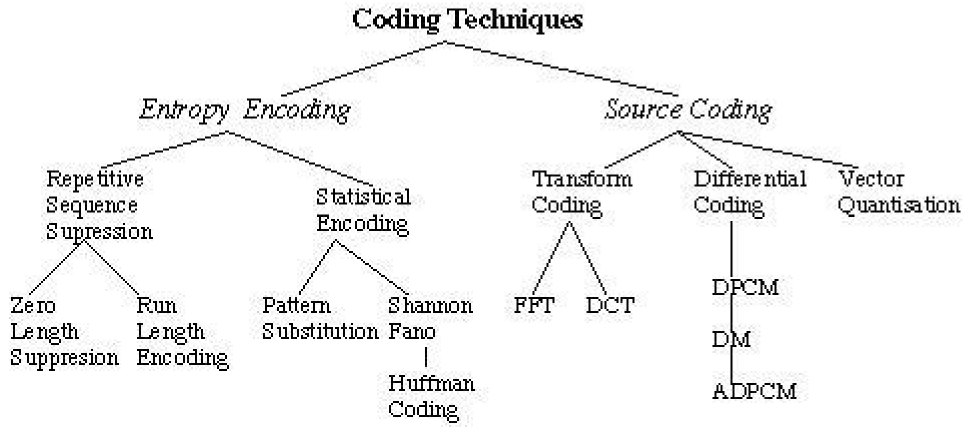


Figure 4.14 *coding* scheme

What is Compression?

Compression basically employs redundancy in the data:

- Temporal -- in 1D data, 1D signals, Audio etc.
- Spatial -- correlation between neighbouring pixels or data items
- Spectral -- correlation between colour or luminescence components. This uses the frequency domain to exploit relationships between frequency of change in data.
- psycho-visual -- exploit perceptual properties of the human visual system.

Compression can be categorized in two broad ways:

Lossless Compression

-- where data is compressed and can be reconstituted (uncompressed) without loss of detail or information. These are referred to as bit-preserving or reversible compression systems also.

Lossy Compression

-- where the aim is to obtain the best possible *fidelity* for a given bit-rate or minimizing the bit-rate to achieve a given fidelity measure. Video and audio compression techniques are most suited to this form of compression.

If an image is compressed it clearly needs to uncompressed (decoded) before it can viewed/listened to. Some processing of data may be possible in encoded form however. Lossless compression frequently involves some form of *entropy encoding* and are based in information theoretic techniques.

Lossy compression use source encoding techniques that may involve transform encoding, differential encoding or vector quantization.

4.7.3 MPEG Standards

All MPEG standards exist to promote system interoperability among your computer, television and handheld video and audio devices. They are:

- **MPEG-1:** the original standard for encoding and decoding streaming video and audio files.
- **MPEG-2:** the standard for digital television, this compresses files for transmission of high-quality video.
- **MPEG-4:** the standard for compressing high-definition video into smaller-scale files that stream to computers, cell phones and PDAs (personal digital assistants).
- **MPEG-21:** also referred to as the Multimedia Framework. The standard that interprets what digital content to provide to which individual user so that media plays flawlessly under any language, machine or user conditions.

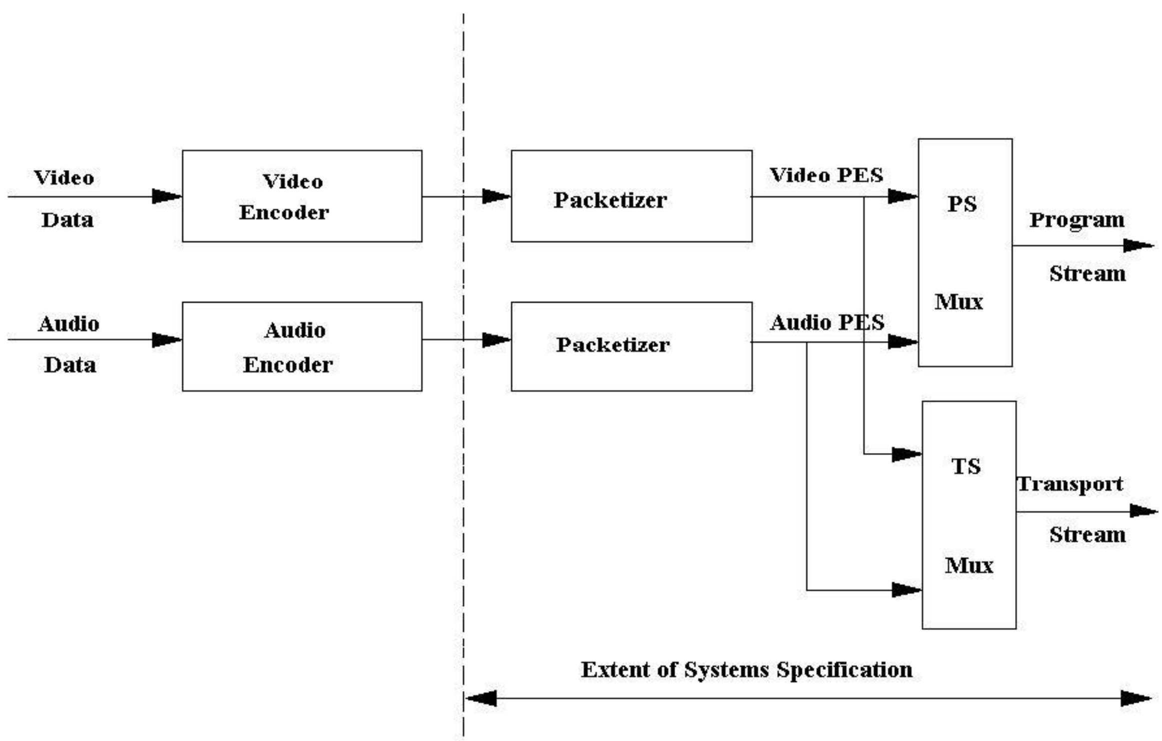


Figure 4.15 MPEG scheme

4.8 Encryption

It 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.**

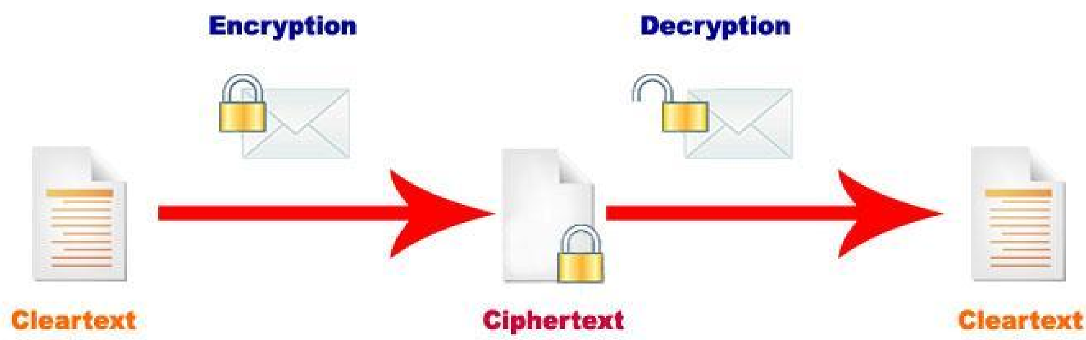


Figure 4.16 Encryption methods

4.8.1 Symmetric key encryption

In symmetric-key schemes, the encryption and decryption keys are the same. Thus communicating parties must have the same key before they can achieve secret communication.

In public-key encryption schemes, the encryption key is published for anyone to use and encrypt messages. However, only the receiving party has access to the decryption key that enables messages to be read.

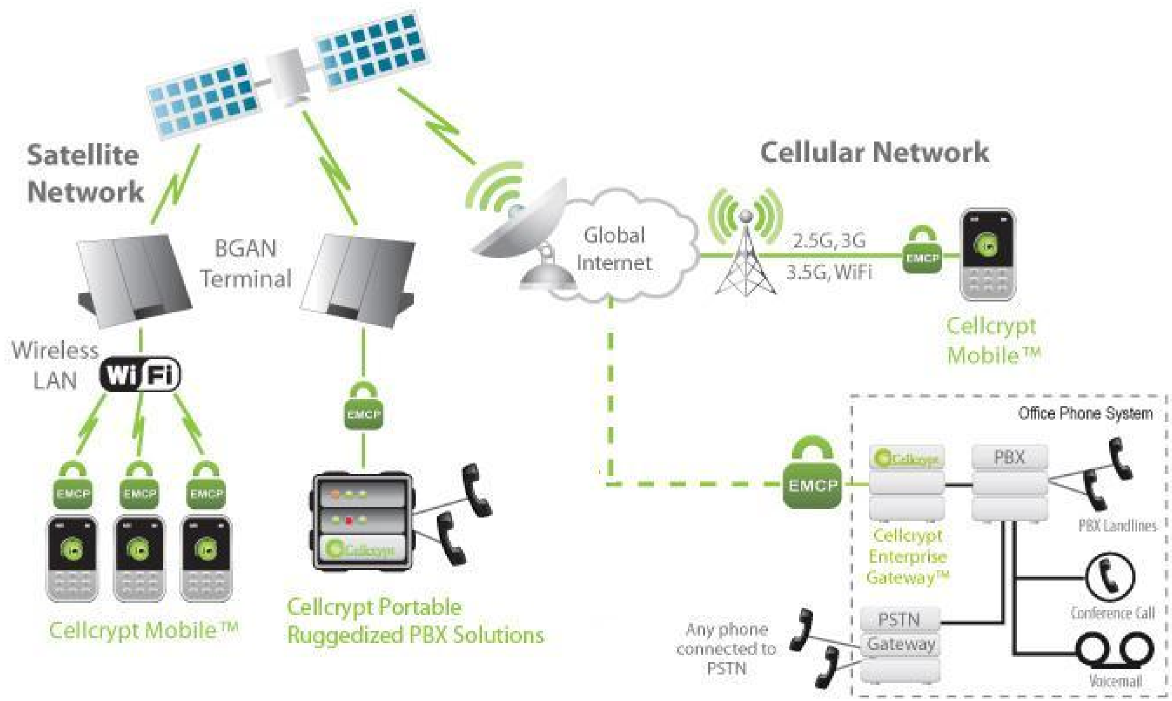


Figure 4.17 General block diagram Encryption methods

4.8.2 Decryption

It is the process of taking encoded or encrypted text or other data and converting it back into text that you or the computer are able to read and understand.

This term could be used to describe a method of un-encrypting the data manually or with un-encrypting the data using the proper codes or keys.

Data may be encrypted to make it difficult for someone to steal the information. Some companies also encrypt data for general protection of company data and trade secrets. If this data needs to be viewable, it may require decryption.

APPLICATIONS

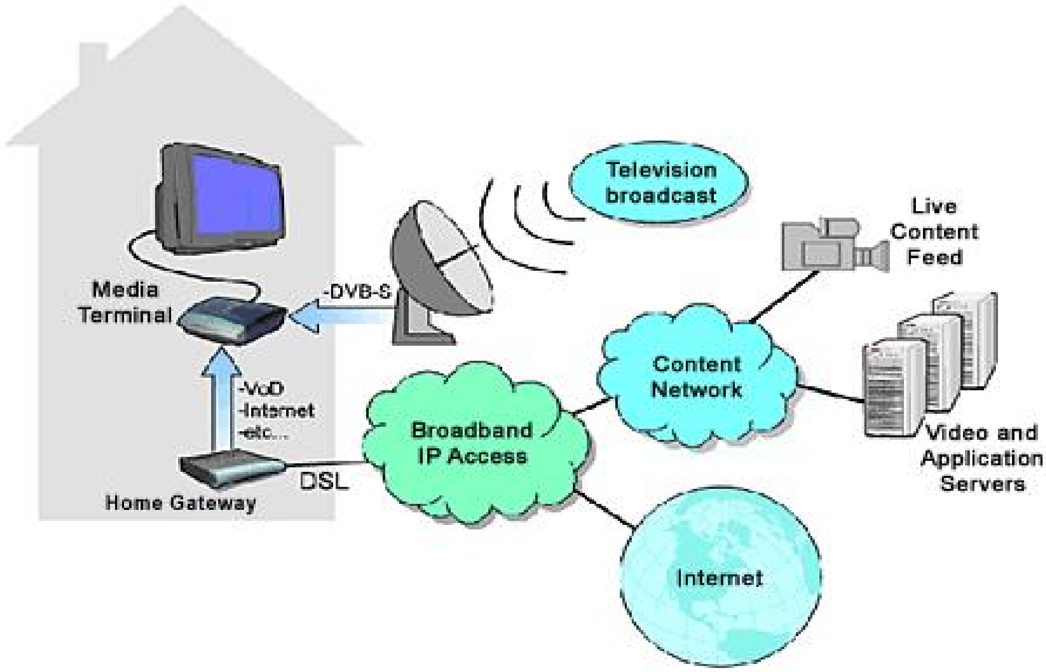


Figure an example of Digital video Broadcast

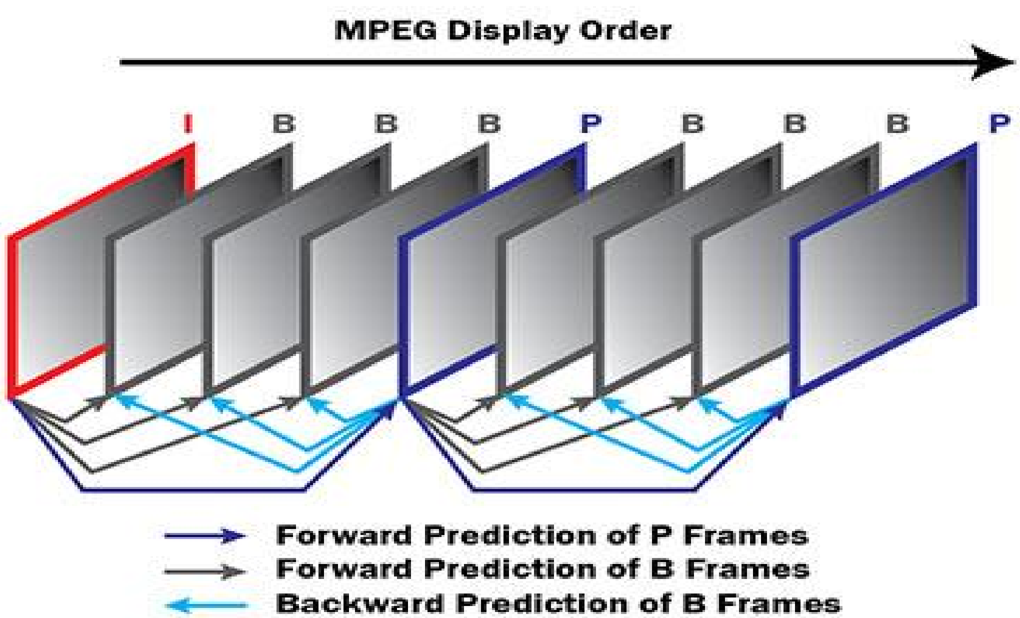


Figure an example of streaming video and compression

POST TEST MCQ

1. In a communication satellite, the telephone channels are assembled in

a) AM

b) FM

c) TDM

d) FDM

2. The 24-MHz bandwidth of a transponder is capable of carrying__________

a) One analog television channel

b) Two analog television channel

c) Four analog television channel

d) None of these

3. In DBS systems

a) MPEG-2 is used for video compression

b) MPEG-2 is used for video enhancing

c) MPEG-2 is used for audio compression

d) None of these

4. 4:4:4 Sampling means that____________

a) The sampling rates of Y, Cb, and Cr are equal

b) The sampling rates of Y, Cb, and Cr are unequal

c) a & b

d) None of these

5. A macro-block consists of ________

a) 16 * 16 pixels.

b) 8 * 8 pixels.

c) 2 * 2 pixels .

d) None of these.

6. MPEG stands as______.

a) Moving Pictures Expert Group

b) Most Pictures Expert Group

c) Moving Pictures Enhance Group

d) Motor Piston Expert Group

7. The bit rate for digital television depends very much on the___________.

a) Picture format

b) Information format

c) Voice format

d) None of these In DBS systems

8. MPEG-2 uses

a) 4:2:0 sampling

b) 2:0:0 sampling

c) 3:2:0 sampling

d) 0:0:9 sampling

9. The multiple access technique suitable only for digital transmission is

a) Packet Access

b) CDMA

c) FDMA

d) TDMA

10. MPEG-4 provides the major advantage is the ________in bit rate offered in satellite television.

a) Reduction

b)Increment

c) double

d) Triple

11. The most popular access method is________, which allows the use of comparatively low-power VSAT terminals

a) FDMA

b) TDMA

c) CDMA

d) TDD

12. The uplink Frequency of C-band?

a) 4GHz

b) 6GHz

c) 8GHz

d) 12GHz

13. Using the HDTV format having a pixel count per frame of _____and a refresh rate of 30 frames per second

a) 192 * 108

b) 1920 * 1080

c) 920 * 1000

d) 148*156

14. Identify which of the following is a type of CDMA.

a) Spread spectrum multiple access

b) pulse address multiple access

c) both a&b

d) None of above

15. VSAT stands as______.

a) Very small aperture terminal system

b) Vast small aperture terminal system

c) Virtual small aperture terminal system

d) Video small aperture terminal system

CONCLUSION

1. In this unit we discussed the Modulation and Multiplexing.
2. The multiple access: FDMA, TDMA, CDMA was discussed.
3. The compression – encryption were discussed.

REFERENCES

1. Wilbur L.Pritchard, Hendri G. Suyderhoud, Robert A. Nelson, "Satellite Communication Systems Engineering", Prentice Hall/Pearson, 2007.
2. N. Agarwal, "Design of Geosynchronous Space Craft", Prentice Hall, 1986.
3. Bruce R. Elbert, "Satellite Communication Applications", Hand Book, Artech House Bostan London,1997.
4. Tri T. Ha, "Digital Satellite Communication", II nd edition, 1990.
5. Robert G. Winch, "Telecommunication Transmission Systems", Mc Graw-Hill Book Co., 1983.
6. Brian Ackroyd, "World Satellite Communication and earth station Design", BSP professional Books, 1990.
7. G.B.Bleazard, "Introducing Satellite communications", NCC Publication, 1985.
8. M.Richharia, "Satellite Communication Systems-Design Principles", Macmillan 2003.
9. Emanuel Fthenakis, "Manual of Satellite Communications", Mc Graw Hill Book Co., 1984.

ASSIGNMENT

1. Distinguish single access and multiple accesses.
2. Compare the use of data compression and encryption in satellite communication.
3. Compare pre-assigned FDMA and demand-assigned FDMA.
4. Explain the following with respect to TDMA: Reference burst, preamble and post amble and carrier recovery.
5. Compare the uplink power requirements of FDMA and TDMA.
6. Explain principles of Code-Division multiple Access.

SATELLITE COMMUNICATION

UNIT V SATELLITE APPLICATIONS

Prepared by

Dr.M.A.ARCHANA

AP/ECE,SCSVMV

UNIT V SATELLITE APPLICATIONS

AIM & OBJECTIVE

- To understand the applications of satellites.
- To understand the Mobile satellite services.
- To understand the Video conferencing and Internet.

PRE -TEST MCQ

1. The familiar direct view cathode ray tube (CRT) used for analog TV is _____of displaying HDTV.

a) Not capable

b) Capable

c) Accept

d) None of these

2. DirecTV plans to use _________in its HDTV satellite broadcasts and all HDTV services in Europe are expected to use this rather than the MPEG-2.

a) H.264/AVC

b) H.24/AVC

c) H.64/AVC

d) H.456/AVC

3. The main factor governing performance of a DBS system will be the [Eb/N0] of the______

a) Uplink

b) Downlink

c) Up-downlink

d) None of these

THEORY

5.1 INTELSAT Series

INTELSAT stands for *International Telecommunications Satellite*. The organization was created in 1964 and currently has over 140 member countries and more than 40 investing entities (see http://www.intelsat.com/ for more details).

In July 2001 INTELSAT became a private company and in May 2002 the company began providing end-to-end solutions through a network of teleports, leased fiber, and *points of presence* (PoPs) around the globe.

Starting with the Early Bird satellite in 1965, a succes- sion of satellites has been launched at intervals of a few years. Figure 1.1 illustrates the evolution of some of the INTELSAT satellites. As the figure shows, the capacity, in terms of number of voice channels, increased dramatically with each succeeding launch, as well as the design lifetime.

These satellites are in *geostationary orbit,* meaning that they appear to be stationary in relation to the earth. At this point it may be noted that geosta- tionary satellites orbit in the earth's equatorial plane and their position is specified by their longitude.

For international traffic, INTELSAT covers three main regions—the *Atlantic Ocean Region* (AOR), the *Indian Ocean Region* (IOR), and the *Pacific Ocean Region* (POR) and what is termed *Intelsat America's Region*.

For the ocean regions the satellites are positioned in geostationary orbit above the particular ocean, where they provide a transoceanic telecommunications route. For example, INTELSAT satellite 905 is positioned at 335.5° east longitude.

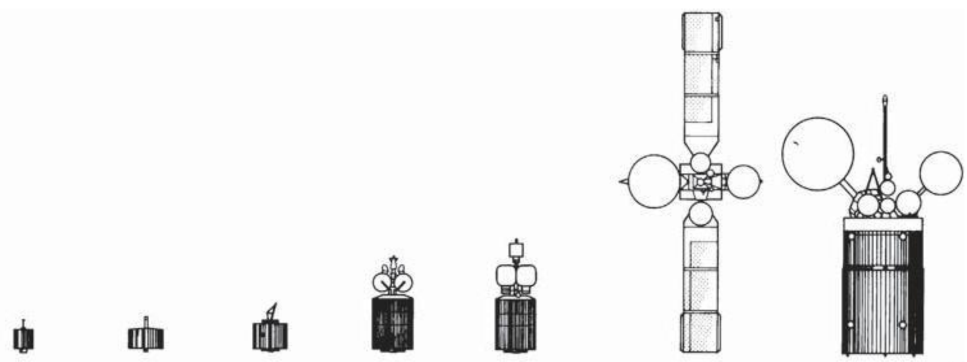
The INTELSAT VII-VII/A series was launched over a period from October 1993 to June 1996. The construction is similar to that for the V and VA/VB series, shown in Fig. in that the VII series has solar sails rather than a cylindrical body.

The VII series was planned for service in the POR and also for some of the less demanding services in the AOR. The antenna beam coverage is appropriate for that of the POR. Figure 1.3 shows the antenna beam footprints for the C -band hemispheric cover- age and zone coverage, as well as the spot beam coverage possible with the Ku-band antennas (Lilly, 1990; Sachdev et al., 1990).

When used in the AOR, the VII series satellite is inverted north for south (Lilly, 1990), minor adjustments then being needed only to optimize the antenna pat- terns for this region. The lifetime of these satellites ranges from 10 to 15 years depending on the launch vehicle.

Recent figures from the INTELSAT Web site give the capacity for the INTELSAT VII as 18,000 two-way telephone circuits and three TV channels; up to 90,000 two-way telephone circuits can be achieved with the use of “digital circuit multiplication.”

The INTELSAT VII/A has a capacity of 22,500 two-way telephone circuits and three TV channels; up to 112,500 two-way telephone circuits can be achieved with the use of digital circuit multiplication. As of May 1999, four satellites were in service over the AOR, one in the IOR, and two in the POR.



Designation: Intelsat	I	II	III	IV	IV A	V	V A/V B	VI
Year of first launch	1965	1966	1968	1971	1975	1980	1984/85	1986/87
Prime contractor	Hughes	Hughes	TRW	Hughes	Hughes	Ford Aerospace	Ford Aerospace	Hughes
Width (m)	0.7	1.4	1.4	2.4	2.4	2.0	2.0	3.6
Height (m)	0.6	0.7	1.0	5.3	6.8	6.4	6.4	6.4
Launch vehicles		Thor Delta		Atlas-Centaur		Atlas-Centaur and Ariane	Atlas-Centaur and Ariane	STS and Ariane
Spacecraft mass in transfer orbit (kg)	68	182	293	1385	1489	1946	2140	12,100/3720
Communications payload mass (kg)	13	36	56	185	190	235	280	800
End-of-life (EOL) power of equinox (W)	40	75	134	480	800	1270	1270	2200
Design lifetime (years)	1.5	3	5	7	7	7	7	10
Capacity (number of voice channels)	480	480	2400	8000	12,000	25,000	30,000	80,000
Bandwidth (MHz)	50	130	300	500	800	2137	2480	3520

Figure 5.1 INTELSAT Series

The INTELSAT VIII-VII/A series of satellites was launched over the period February 1997 to June 1998. Satellites in this series have similar capacity as the VII/A series, and the lifetime is 14 to 17 years.

It is standard practice to have a spare satellite in orbit on high-reliability routes (which can carry preemptible traffic) and to have a ground spare in case of launch failure.

Thus the cost for large international schemes can be high; for example, series IX, described later, represents a total investment of approximately $1 billion.

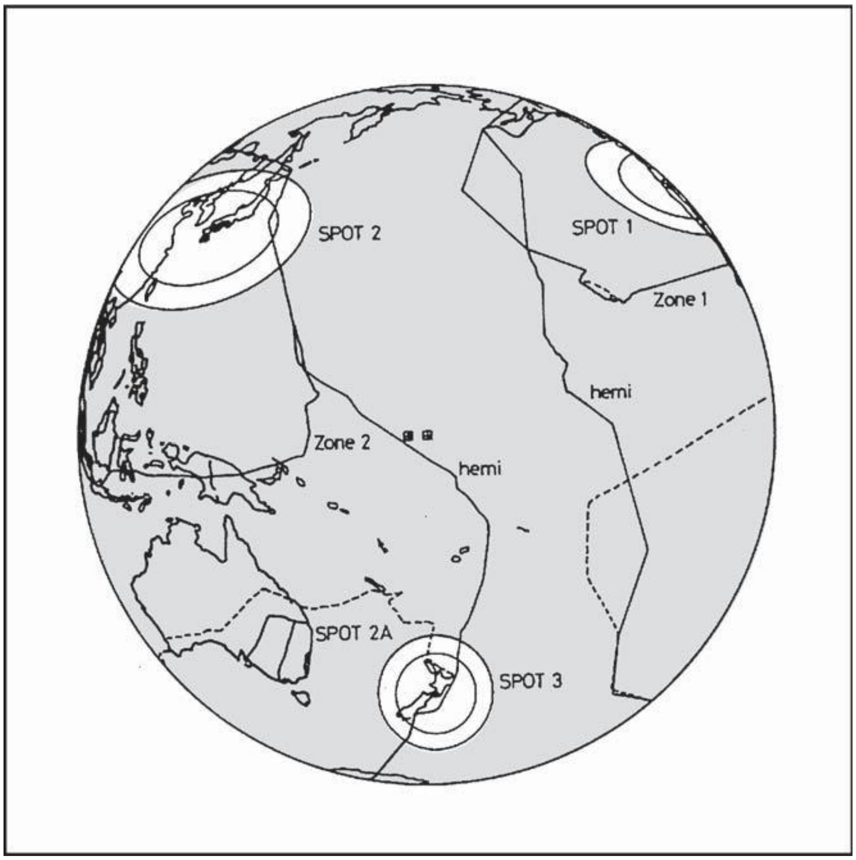


Figure 5.2 Region of glob

5.2 INSAT

INSAT or the ***Indian National Satellite System*** is a series of multipurpose geo-stationary satellites launched by ISRO to satisfy the telecommunications, broadcasting, meteorology, and search and rescue operations.

Commissioned in 1983, INSAT is the largest domestic communication system in the Asia Pacific Region. It is a joint venture of the Department of Space, Department of Telecommunications, India Meteorological Department, All India Radio and Doordarshan. The overall coordination and management of INSAT system rests with the Secretary-level INSAT Coordination Committee.

INSAT satellites provide transponders in various bands (C, S, Extended C and K_u) to serve the television and communication needs of India. Some of the satellites also have the Very High Resolution Radiometer (VHRR), CCD cameras for metrological imaging.

The satellites also incorporate transponder(s) for receiving distress alert signals for search and rescue missions in the South Asian and Indian Ocean Region, as ISRO is a member of the Cospas-Sarsat programme.

5.2.1 INSAT System

The Indian National Satellite (INSAT) System Was Commissioned With The Launch Of INSAT-1B In August 1983 (INSAT-1A, The First Satellite Was Launched In April 1982 But Could Not Fulfil The Mission).

INSAT System Ushered In A Revolution In India's Television And Radio Broadcasting, Telecommunications And Meteorological Sectors. It Enabled The Rapid Expansion Of TV And Modern Telecommunication Facilities To Even The Remote Areas And Off-Shore Islands.

5.2.2 Satellites In Service

Of The 24 Satellites Launched In The Course Of The INSAT Program, 10 Are Still In Operation.INSAT-2E

It Is The Last Of The Five Satellites In INSAT-2 Series{Prateek }. It Carries Seventeen C-Band And Lower Extended C-Band Transponders Providing Zonal And Global Coverage With An Effective Isotropic Radiated Power (EIRP) Of 36 Dbw.

It Also Carries A Very High Resolution Radiometer (VHRR) With Imaging Capacity In The Visible (0.55-0.75 µm), Thermal Infrared (10.5-12.5 µm) And Water Vapour (5.7-7.1 µm) Channels And Provides 2x2 Km, 8x8 Km And 8x8 Km Ground Resolution Respectively.

INSAT-3A

The Multipurpose Satellite, INSAT-3A, Was Launched By Ariane In April 2003. It Is Located At 93.5 Degree East Longitude. The Payloads On INSAT-3A Are As Follows:

12 Normal C-Band Transponders (9 Channels Provide Expanded Coverage From Middle East To South East Asia With An EIRP Of 38 Dbw, 3 Channels Provide India Coverage With An EIRP Of 36 Dbw And 6 Extended C -Band Transponders Provide India Coverage With An EIRP Of 36 Dbw).

A CCD Camera Provides 1x1 Km Ground Resolution, In The Visible (0.63-0.69 μm), Near Infrared (0.77-0.86 μm) And Shortwave Infrared (1.55-1.70 μm) Bands.

INSAT-3D

Launched In July 2013, INSAT-3D Is Positioned At 82 Degree East Longitude. INSAT-3D Payloads Include Imager, Sounder, Data Relay Transponder And Search & Rescue Transponder. All The Transponders Provide Coverage Over Large Part Of The Indian Ocean Region Covering India, Bangladesh, Bhutan,Maldives, Nepal, Seychelles, Sri Lanka And Tanzania For Rendering Distress Alert Services

INSAT-3E

Launched In September 2003, INSAT-3E Is Positioned At 55 Degree East Longitude And Carries 24 Normal C-Band Transponders Provide An Edge Of Coverage EIRP Of 37 Dbw Over India And 12 Extended C-Band Transponders Provide An Edge Of Coverage EIRP Of 38 Dbw Over India.

KALPANA-1

KALPANA-1 Is An Exclusive Meteorological Satellite Launched By PSLV In September 2002. It Carries Very High Resolution Radiometer And DRT Payloads To Provide Meteorological Services. It Is Located At 74 Degree East Longitude. Its First Name Was METSAT. It Was Later Renamed As KALPANA-1 To Commemorate Kalpana Chawla.

Edusat

Configured For Audio-Visual Medium Employing Digital Interactive Classroom Lessons And Multimedia Content, EDUSAT Was Launched By GSLV In September 2004. Its Transponders And Their Ground Coverage Are Specially Configured To Cater To The Educational Requirements.

GSAT-2

Launched By The Second Flight Of GSLV In May 2003, GSAT-2 Is Located At 48 Degree East Longitude And Carries Four Normal C-Band Transponders To Provide 36 Dbw EIRP With India Coverage, Two K_u Band Transponders With 42 Dbw EIRP Over India And An MSS Payload Similar To Those On INSAT-3B And INSAT-3C.

INSAT-4 Series:

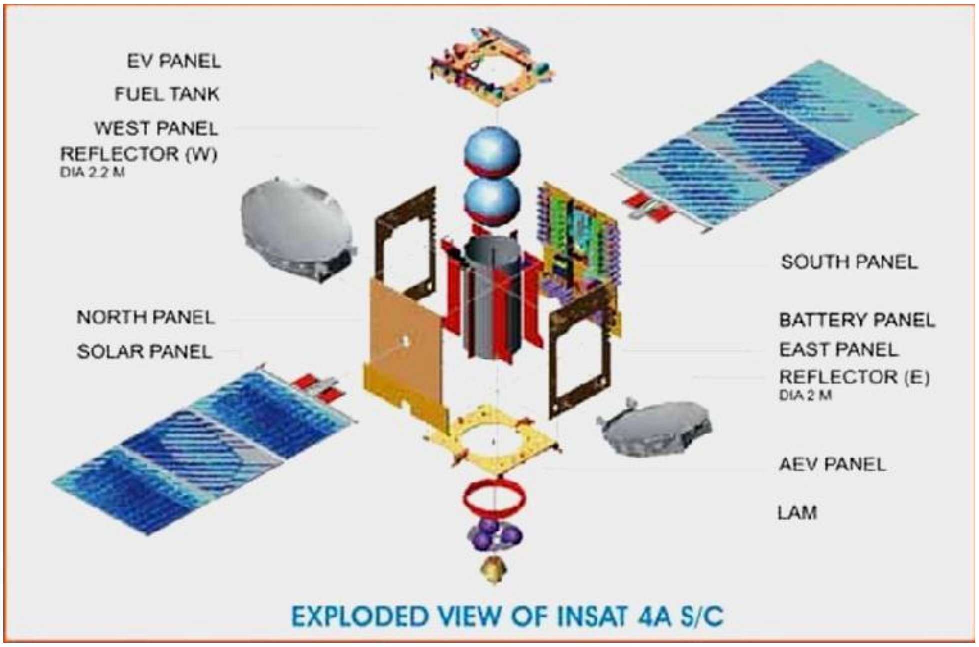


Figure 5.3 INSAT 4A

INSAT-4A is positioned at 83 degree East longitude along with INSAT-2E and INSAT-3B. It carries 12 K_u band 36 MHz bandwidth transponders employing 140 W TWTAs to provide an EIRP of 52 dBW at the edge of coverage polygon with footprint covering Indian main land and 12 C-band 36 MHz bandwidth transponders provide an EIRP of 39 dBW at the edge of coverage with expanded radiation patterns encompassing Indian geographical boun dary, area beyond India in southeast and northwest regions.[8] Tata Sky, a joint venture between the TATA Group and STAR uses INSAT-4A for distributing their DTH service.

- INSAT-4A
- INSAT-4B
- Glitch In INSAT 4B
- China-Stuxnet Connection
- INSAT-4CR
- GSAT-8 / INSAT-4G
- GSAT-12 /GSAT-10

5.3 VSAT

VSAT stands for *very small aperture terminal* system. This is the distinguishing feature of a VSAT system, the earth-station antennas being typically less than 2.4 m in diameter (Rana et al., 1990). The trend is toward even smaller dishes, not more than 1.5 m in diameter (Hughes et al., 1993).

In this sense, the small TVRO terminals for direct broadcast satellites could be labeled as VSATs, but the appellation is usually reserved for private networks, mostly providing two-way communications facilities.

Typical user groups include bank- ing and financial institutions, airline and hotel booking agencies, and large retail stores with geographically dispersed outlets.

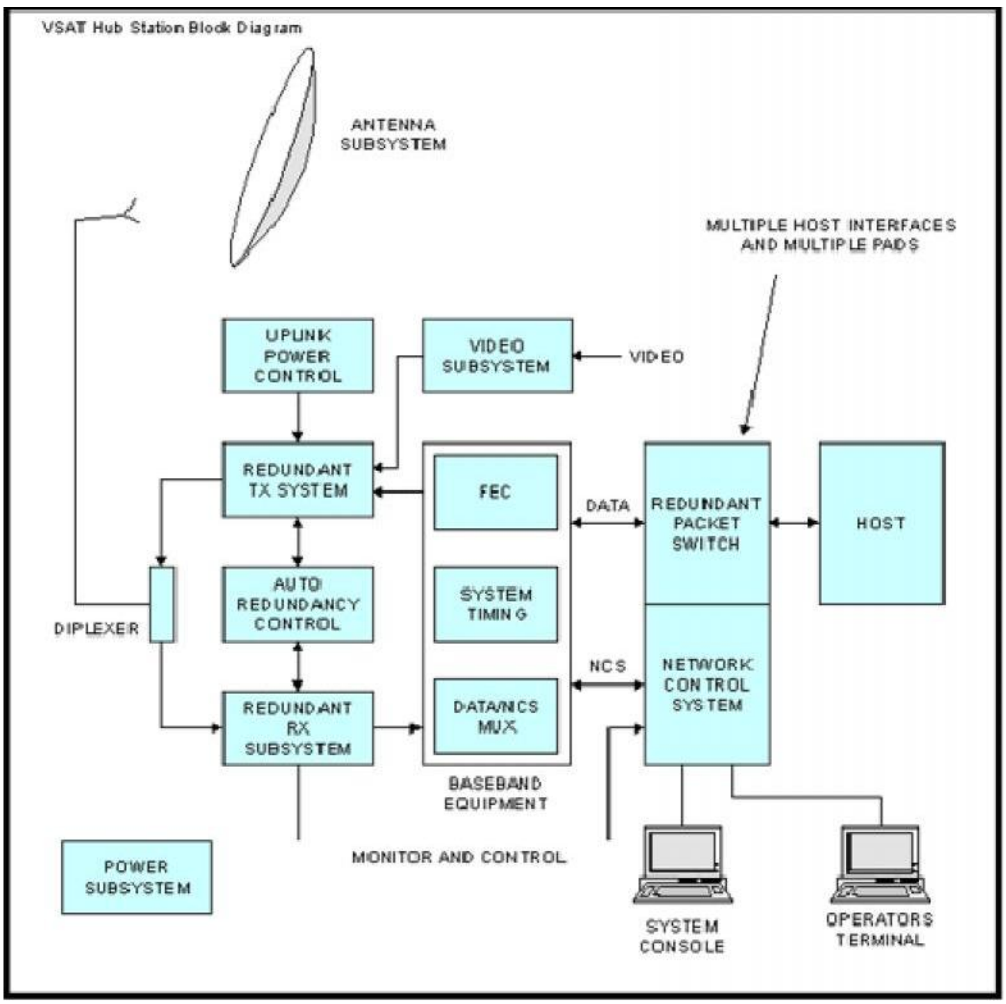


Figure 5.4 VSAT Block Diagrams

5.3.1 VSAT network

The basic structure of a VSAT network consists of a hub station which provides a broadcast facility to all the VSATs in the network and the VSATs themselves which access the satellite in some form of multiple- access mode.

The hub station is operated by the service provider, and it may be shared among a number of users, but of course, each user organ- ization has exclusive access to its own VSAT network.

Time division mul- tiplex is the normal downlink mode of transmission from hub to the VSATs, and the transmission can be broadcast for reception by all the VSATs in a network, or address coding can be used to direct messages to selected VSATs.

A form of *demand assigned multiple access* (DAMA) is employed in some systems in which channel capacity is assigned in response to the fluctuating demands of the VSATs in the network.

Most VSAT systems operate in the Ku band, although there are some C-band systems in existence (Rana et al., 1990).

5.3.2 Applications

- ✓ Supermarket shops (tills, ATM machines, stock sale updates and stock ordering).
- ✓ Chemist shops - Shoppers Drug Mart - Pharmaprix. Broadband direct to the home. e.g. Downloading MP3 audio to audio players.
- ✓ Broadband direct small business, office etc, sharing local use with many PCs.
- ✓ Internet access from on board ship Cruise ships with internet cafes, commercial shipping communications.

5.4 Mobile satellite services

5.4.1 GSM

5.4.1.1 Services and Architecture

If your work involves (or is likely to involve) some form of wireless public communications, you are likely to encounter the GSM standards. Initially developed to support a standardized approach to digital cellular communications in Europe, the "Global System for Mobile Communications" (GSM) protocols are rapidly being adopted to the next generation of wireless telecommunications systems.

In the US, its main competition appears to be the cellular TDMA systems based on the IS-54 standards. Since the GSM systems consist of a wide range of components, standards, and protocols.

The GSM and its companion standard DCS1800 (for the UK, where the 900 MHz frequencies are not available for GSM) have been developed over the last decade to allow cellular communications systems to move beyond the limitations posed by the older analog systems.

Analog system capacities are being stressed with more users that can be effectively supported by the available frequency allocations. Compatibility between types of systems had been limited, if non-existent.

By using digital encoding techniques, more users can share the same frequencies than had been available in the analog systems. As compared to the digital cellular systems in the US (CDMA [IS -95] and TDMA [IS-54]), the GSM market has had impressive success. Estimates of the numbers of telephones run from 7.5 million GSM phones to .5 million IS54 phones to .3 million for IS95.

GSM has gained in acceptance from its initial beginnings in Europe to other parts of the world including Australia, New Zealand, countries in the Middle East and the far east. Beyond its use in cellular frequencies (900 M Hz for GSM, 1800 MHz for DCS1800), portions of the GSM signaling protocols are finding their way into the newly developing PCS and LEO Satellite communications systems.

While the frequencies and link characteristics of these systems differ from the standard GSM air interface, all of these systems must deal with users roaming from one cell (or satellite beam) to another, and bridge services to public communication networks including the Public Switched Telephone Network (PSTN), and public data networks (PDN).

The GSM architecture includes several subsystems

The Mobile Station (MS) -- These digital telephones include vehicle, portable and hand-held terminals. A device called the Subscriber Identity Module (SIM) that is basically a smart -card provides custom information about users such as the services they've subscribed to and their identification in the network

The Base Station Sub-System (BSS) -- The BSS is the collection of devices that support the switching networks radio interface. Major components of the BSS include the Base Transceiver Station (BTS) that consists of the radio modems and antenna equipment.

In OSI terms, the BTS provides the physical interface to the MS where the BSC is responsible for the link layer services to the MS. Logically the transcoding equipment is in the BTS, however, an additional component.

The Network and Switching Sub-System (NSS) -- The NSS provides the switching between the GSM subsystem and external networks along with the databases used for additional subscriber and mobility management.

Major components in the NSS include the Mobile Servi ces Switching Center (MSC), Home and Visiting Location Registers (HLR, VLR). The HLR and VLR databases are interconnected through the telecomm standard Signaling System 7 (SS7) control network.

The Operation Sub-System (OSS) -- The OSS provides the support functions responsible for the management of network maintenance and services. Components of the OSS are responsible for network operation and maintenance, mobile equipment management, and subscription management and charging.

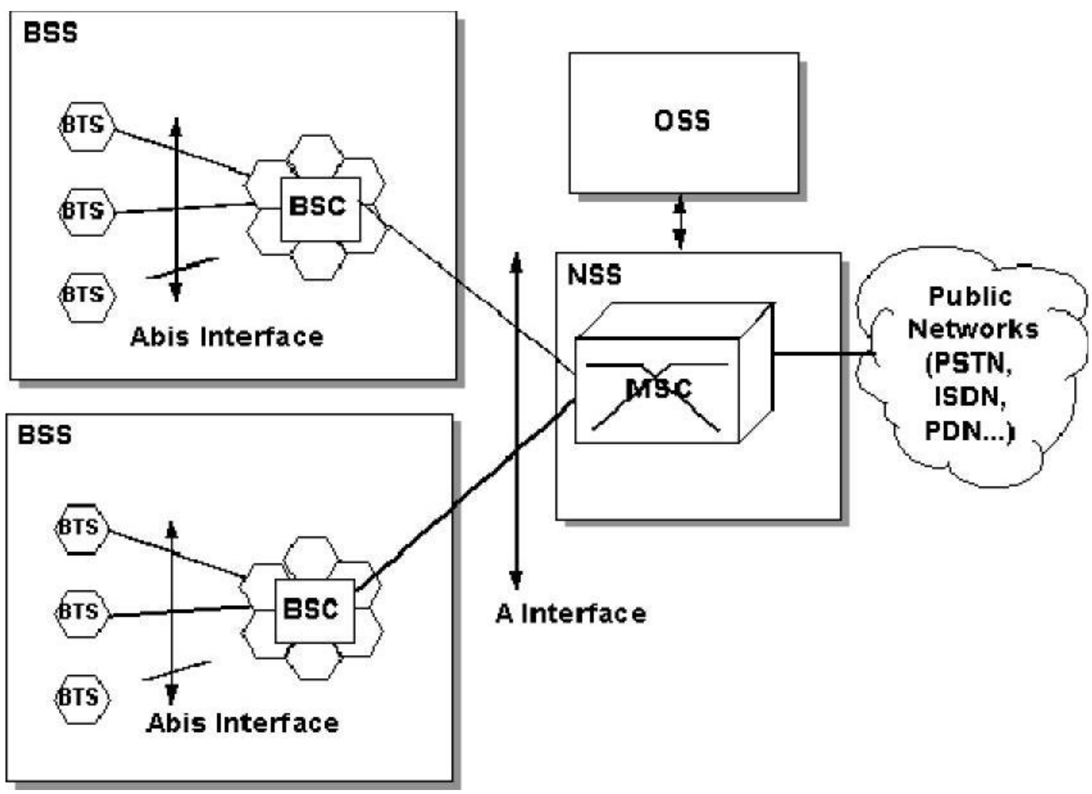


Figure 5.5 GSM Block Diagrams

Several channels are used in the air interface

- ✓ **FCCH** - the frequency correction channel - provides frequency synchronization information in a burst
- ✓ **SCH** - Synchronization Channel - shortly following the FCCH burst (8 bits later), provides a reference to all slots on a given frequency
- ✓ **PAGCH** - Paging and Access Grant Channel used for the transmission of paging information requesting the setup of a call to a MS.
- ✓ **RACH** - Random Access Channel - an inbound channel used by the MS to request connections from the ground network. Since this is used for the first access attempt by users of the network, a random access scheme is used to aid in avoiding collisions.
- ✓ **CBCH** - Cell Broadcast Channel - used for infrequent transmission of broadcasts by the ground network.
- ✓ **BCCH** - Broadcast Control Channel - provides access status information to the MS. The information provided on this channel is used by the MS to determine whether or not to request a transition to a new cell

- ✓ **FACCH** - Fast Associated Control Channel for the control of handovers
- ✓ **TCH/F** - Traffic Channel, Full Rate for speech at 13 kbps or data at 12, 6, or 3.6 kbps
- ✓ **TCH/H** - Traffic Channel, Half Rate for speech at 7 kbps, or data at 6 or 3.6 kbps

5.5 Mobility Management

One of the major features used in all classes of GSM networks (cellular, PCS and Satellite) is the ability to support roaming users. Through the control signaling network, the MSCs interact to locate and connect to users throughout the network.

"Location Registers" are included in the MSC databases to assist in the role of determining how, and whether connections are to be made to roaming users. Each user of a GSM MS is assigned a Home Location Register (HLR) that is used to contain the user's location and subscribed services.

Difficulties facing the operators can include

a. Remote/Rural Areas. To service remote areas, it is often economically unfeasible to provide backhaul facilities (BTS to BSC) via terrestrial lines (fiber/microwave).

b. Time to deploy. Terrestrial build-outs can take years to plan and implement.
c. Areas of 'minor' interest. These can include small isolated centers such as tourist resorts, islands, mines, oil exploration sites, hydro-electric facilities.
d. Temporary Coverage. Special events, even in urban areas, can overload the existing infrastructure.

5.5.1 GSM service security

GSM was designed with a moderate level of service security. GSM uses several cryptographic algorithms for security. The A5/1, A5/2, and A5/3 stream ciphers are used for ensuring over-the-air voice privacy.

GSM uses General Packet Radio Service (GPRS) for data transmissions like browsing the web. The most commonly deployed GPRS ciphers were publicly broken in 2011The researchers revealed flaws in the commonly used GEA/1.

5.5.2 Global Positioning System (GPS)

The Global Positioning System (GPS) is a satellite based navigation system that can be used to locate positions anywhere on earth. Designed and operated by the U.S. Department of Defense, it consists of satellites, control and monitor stations, and receivers. GPS receivers take information transmitted from the satellites and uses triangulation to calculate a user's exact location. GPS is used on incidents in a variety of ways, such as:

- ✓ To determine position locations; for example, you need to radio a helicopter pilot the coordinates of your position location so the pilot can pick you up.
- ✓ To navigate from one location to another; for example, you need to travel from a lookout to the fire perimeter.
- ✓ To create digitized maps; for example, you are assigned to plot the fire perimeter and hot spots.
- ✓ To determine distance between two points or how far you are from another location.

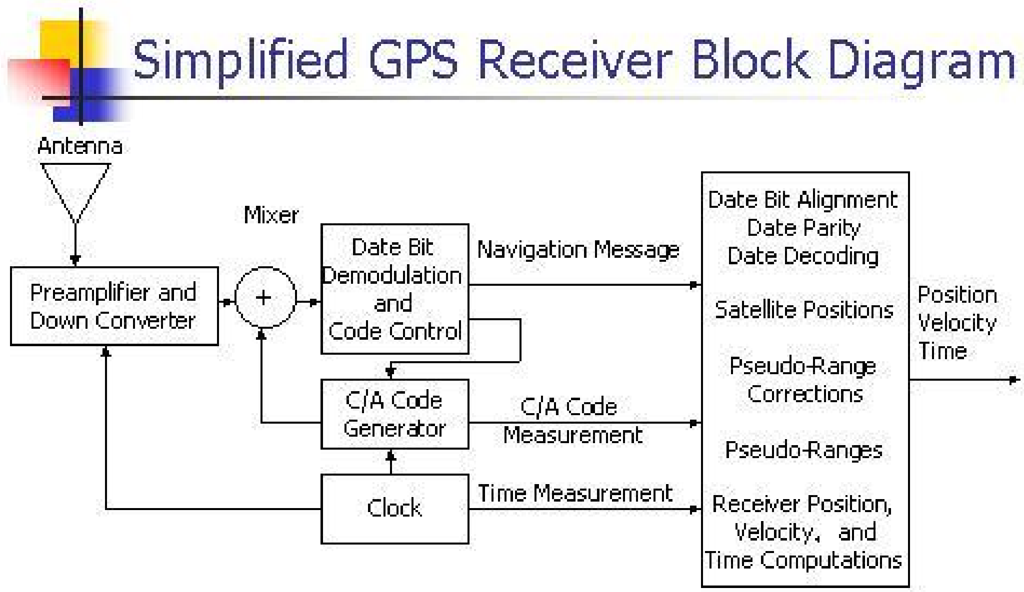


Figure 5.6 GPS Block Diagrams

The purpose of this chapter is to give a general overview of the Global Positioning System, not to teach proficiency in the use of a GPS receiver. To become proficient with a specific GPS receiver, study the owner's manual and practice using the receiver.

The chapter starts with a general introduction on how the global positioning system works. Then it discusses some basics on using a GPS receiver.

Three Segments of GPS:

Space Segment — Satellites orbiting the earth

The space segment consists of 29 satellites circling the earth every 12 hours at 12,000 miles in altitude. This high altitude allows the signals to cover a greater area. The satellites are arranged in their orbits so a GPS receiver on earth can receive a signal from at least four satellites at any given time. Each satellite contains several atomic clocks.

Control Segment — The control and monitoring stations

The control segment tracks the satellites and then provides them with corrected orbital and time information. The control segment consists of five unmanned monitor stations and one Master Control Station. The five unmanned stations monitor GPS satellite signals and then send that information to the Master Control Station where anomalies are corrected and sent back to the GPS satellites through ground antennas.

User Segment — The GPS receivers owned by civilians and military

The user segment consists of the users and their GPS receivers. The number of simultaneous users is limitless.

How GPS Determines a Position

The GPS receiver uses the following information to determine a position.

- ✓ Precise location of satellites

✓When a GPS receiver is first turned on, it downloads orbit information from all the satellites called an almanac. This process, the first time, can take as long as 12 minutes; but once this information is downloaded, it is stored in the receiver's memory for future use.

- ✓ Distance from each satellite

The GPS receiver calculates the distance from each satellite to the receiver by using the distance formula: distance = velocity x time. The receiver already knows the velocity, which is the speed of a radio wave or 186,000 miles per second (the speed of light).

- ✓ Triangulation to determine position

The receiver determines position by using triangulation. When it receives signals from at least three satellites the receiver should be able to calculate its approximate position (a 2D position). The receiver needs at least four or more satellites to calculate a more accurate 3D position.

Using a GPS Receiver

There are several different models and types of GPS receivers. Refer to the owner's manual for your GPS receiver and practice using it to become proficient.

- ✓ When working on an incident with a GPS receiver it is important to:
- ✓ Always have a compass and a map.
- ✓ Have a GPS download cable.
- ✓ Have extra batteries.
- ✓ Know memory capacity of the GPS receiver to prevent loss of data, decrease in accuracy of data,or other problems.
- ✓ Use an external antennae whenever possible, especially under tree canopy, in canyons, or while flying or driving.
- ✓ Set up GPS receiver according to incident or agency standard regulation; coordinate system.
- ✓ Take notes that describe what you are saving in the receiver.

5.6. INMARSAT

Inmarsat-Indian Maritime SATellite is still the sole IMO-mandated provider of satellite communications for the GMDSS.

- Availability for GMDSS is a minimum of 99.9%

Inmarsat has constantly and consistently exceeded this figure & Independently audited by IMSO and reported on to IMO.

Now Inmarsat commercial services use the same satellites and network &Inmarsat A closes at midnight on 31 December 2007 Agreed by IMO – MSC/Circ.1076. Successful closure programme almost concluded Overseen throughout by IMSO.

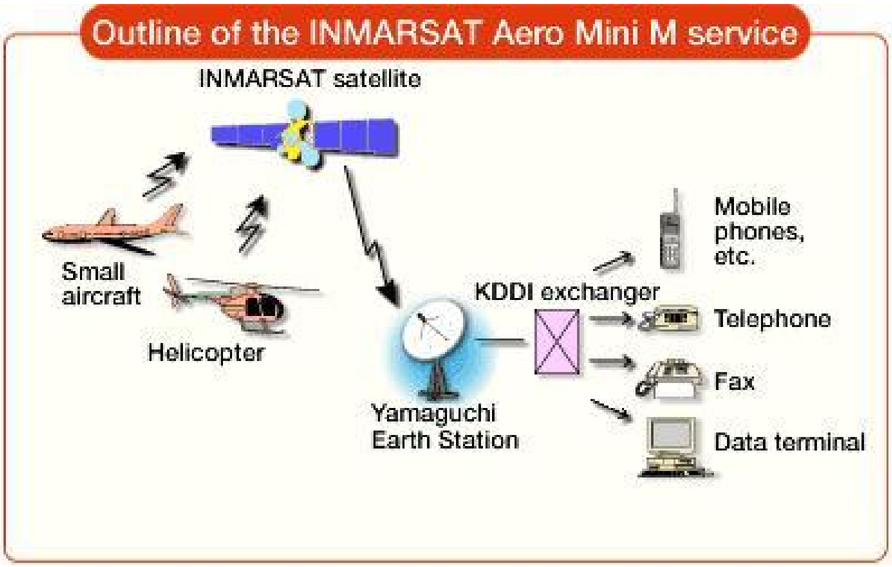


Figure 5.7 INMARSAT Satellite Service

GMDSS services continue to be provided by:

- Inmarsat B, Inmarsat C/mini-C and Inmarsat Fleet F77
- Potential for GMDSS on FleetBroadband being assessed

- The IMO Criteria for the Provision of Mobile Satellite Communications Systems in the Global Maritime Distress and Safety System (GMDSS)

- Amendments were proposed; potentially to make it simpler for other satellite systems to be approved

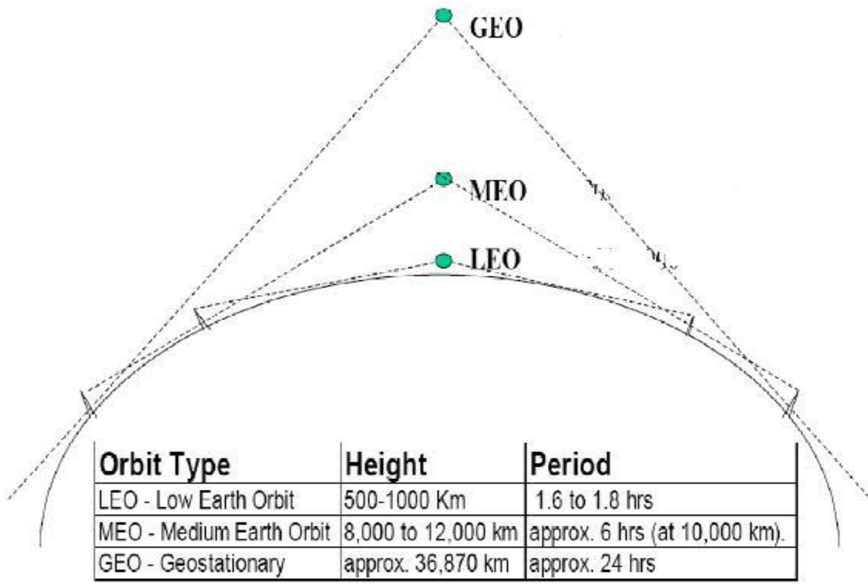
- The original requirements remain and were approved by MSC 83
  - No dilution of standards

- Minor amendments only; replacement Resolution expected to be approved by the IMO 25^{th} Assembly

- Inmarsat remains the sole, approved satcom provider for the GMDSS

5.7 LEO: Low Earth Orbit satellites have a small area of coverage. They are positioned in an orbit approximately 3000km from the surface of the earth

- They complete one orbit every 90 minutes
- The large majority of satellites are in low earth orbit
- The Iridium system utilizes LEO satellites (780km high)
- The satellite in LEO orbit is visible to a point on the earth for a very short time



Orbit Type	Height	Period
LEO - Low Earth Orbit	500-1000 Km	1.6 to 1.8 hrs
MEO - Medium Earth Orbit	8,000 to 12,000 km	approx. 6 hrs (at 10,000 km).
GEO - Geostationary	approx. 36,870 km	approx. 24 hrs

Figure 5.8 LEO, MEO & GEO range

5.8 MEO: *Medium Earth Orbit* satellites have orbital altitudes between 3,000 and 30,000 km.

- They are commonly used used in navigation systems such as GPS

5.9 GEO: *Geosynchronous (Geostationary) Earth Orbit* satellites are positioned over the equator. The orbital altitude is around 30,000-40,000 km

- There is only one geostationary orbit possible around the earth
  - Lying on the earth's equatorial plane.
  - The satellite orbiting at the same speed as the rotational speed of the earth on its axis.
  - They complete one orbit every 24 hours. This causes the satellite to appear stationary with respect to a point on the earth, allowing one satellite to provide continual coverage to a given area on the earth's surface
  - One GEO satellite can cover approximately 1/3 of the world's surface

They are commonly used in communication systems

- Advantages:
  - Simple ground station tracking.
  - Nearly constant range
  - Very small frequency shift

- Disadvantages:
  - Transmission delay of the order of 250 msec.
  - Large free space loss.
  - No polar coverage

- Satellite orbits in terms of the orbital height:
- According to distance from earth:
  - Geosynchronous Earth Orbit (GEO) ,
  - Medium Earth Orbit (MEO),
  - Low Earth Orbit (LEO)

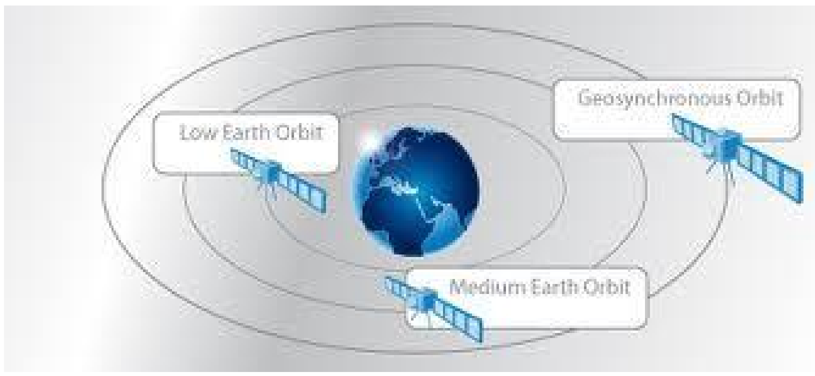


Figure 5.9 LEO, MEO & GEO Orbits

LEO / MEO / GEO / HEO (cont.)

	Name	Number	Panel	No./Panel	altitude	deg.
LEO	STARSYS	24	6	4	1300km	60
	ORBCOMM	24	4	6	785km	45
	GLOBALSTAR	48	8	6	1400km	52
	IRIDIUM	66	6	11	765km	86
MEO	**Name**	**Number**	**Panel**	**No./Panel**	**altitude**	**deg.**
	INMARSAT P	10	2	5	10300km	45
	ODYSEEY	12	3	4	10370km	55
	GPS	24	6	4	20200km	55
	GLONASS	24	3	8	19132km	64.8
HEO	**Name**	**Number**	**Panel**	**No./Panel**	**altitude**	**deg.**
	ELLIPSO	24	4	6	A:7800km P:520km	63.4
	MOLNIYA	4	1	4	A:39863km P:504km	63.4
	ARCHIMEDES	4	4	1	A:39447km P:926km	63.4

9

Figure 5.10 Diff b/w LEO, MEO & GEO Orbits

GEO: 35,786 km above the earth, MEO: 8,000-20,000 km above the earth & LEO: 500-2,000 km above the earth.

5.10 Satellite Navigational System: Benefits

- Enhanced Safety
- Increased Capacity
- Reduced Delays

Advantage

- Increased Flight Efficiencies
- Increased Schedule Predictability
- Environmentally Beneficial Procedures

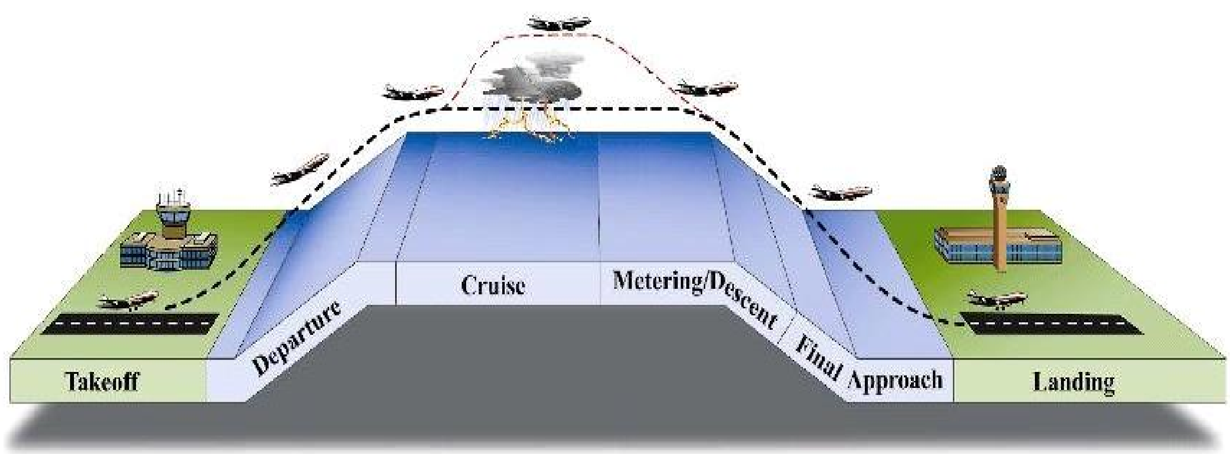


Figure 5.11 LEO, MEO & GEO Orbits

- Using ICAO GNSS Implementation Strategy and ICAO Standards and Recommended Practices
- GPS Aviation Use Approved for Over a Decade
  - Aircraft Based Augmentation Systems (ABAS) – (e.g. RAIM)
- Space Based Augmentation System (SBAS) since 2003
  - Wide Area Augmentation System (WAAS) augmenting GPS

Development of GNSS Ground Based Augmentation System (GBAS) Continues

- Local Area Augmentation System (LAAS)

GNSS is Cornerstone for National Airspace System

5.11 Direct Broadcast satellites (DBS)

Satellites provide *broadcast* transmissions in the fullest sense of the word, because antenna footprints can be made to cover large areas of the earth.

The idea of using satellites to provide direct transmissions into the home has been around for many years, and the services provided are known generally as *direct broadcast satellite* (DBS) services.

Broadcast services include audio, television, and Internet services.

5.11.1 Power Rating and Number of Transponders

From Table 1.4 it will be seen that satellites primarily intended for DBS have a higher [EIRP] than for the other categories, being in the range 51 to 60 dBW. At a *Regional Administrative Radio Council* (RARC) meeting in 1983, the value established for DBS was 57 dBW (Mead,2000). Transponders are rated by the power output of their high-power amplifiers.

Typically, a satellite may carry 32 transponders. If all 32 are in use, each will operate at the lower power rating of 120 W.

The available bandwidth (uplink and downlink) is seen to be 500 MHz. A total number of 32 transponder channels, each of bandwidth 24 MHz, can be accommodated.

The bandwidth is sometimes specified as 27 MHz, but this includes a 3-MHz guard band allowance. Therefore, when calculating bit-rate capacity, the 24 MHz value is used.

The total of 32 transponders requires the use of both *right- hand circular polarization* (RHCP) and *left-hand circular polarization* (LHCP) in order to permit frequency reuse, and guard bands are inserted between channels of a given polarization.

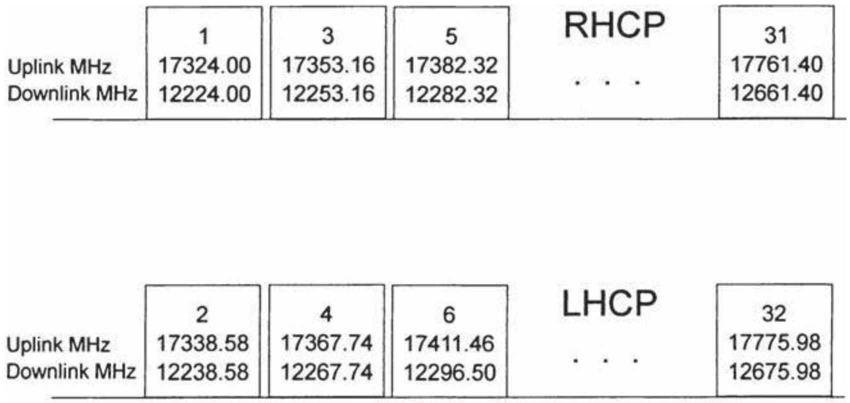


Figure 5.12 DBS Service

5.11.2 Bit Rates for Digital Television

The bit rate for digital television depends very much on the picture format. One way of estimating the uncompressed bit rate is to multiply the number of pixels in a frame by the number of frames per second, and multiply this by the number of bits used to encode each pixel.

5.11.3 MPEG Compression Standards

MPEG is a group within the *International Standards Organization and the International Electrochemical Commission* (ISO/IEC) that undertook the job of defining standards for the transmission and storage of moving pictures and sound.

The MPEG standards currently available are MPEG-1, MPEG-2, MPEG-4, and MPEG-7.

5.12 Direct to home Broadcast (DTH)

DTH stands for Direct-To-Home television. DTH is defined as the reception of satellite programmes with a personal dish in an individual home.

- ✓ DTH Broadcasting to home TV receivers take place in the ku band(12 GHz). This service is known as Direct To Home service.
- ✓ DTH services were first proposed in India in 1996.
- ✓ Finally in 2000, DTH was allowed.
- ✓ The new policy requires all operators to set up earth stations in India

within 12 months of getting a license. DTH licenses in India will cost $2.14 million and will be valid for 10 years.

Working principal of DTH is the satellite communication. Broadcaster modulates the received signal and transmit it to the satellite in KU Band and from satellite one can receive signal by dish and set top box.

5.12.1 DTH Block Diagram

- ✓ A DTH network consists of a broadcasting centre, satellites, encoders, multiplexers, modulators and DTH receivers
- ✓ The encoder converts the audio, video and data signals into the digital format and the multiplexer mixes these signals.

It is used to provide the DTH service in high populated area A Multi Switch is basically a box that contains signal splitters and A/B switches. A outputs of group of DTH LNBs are connected to the A and B inputs of the Multi Switch.

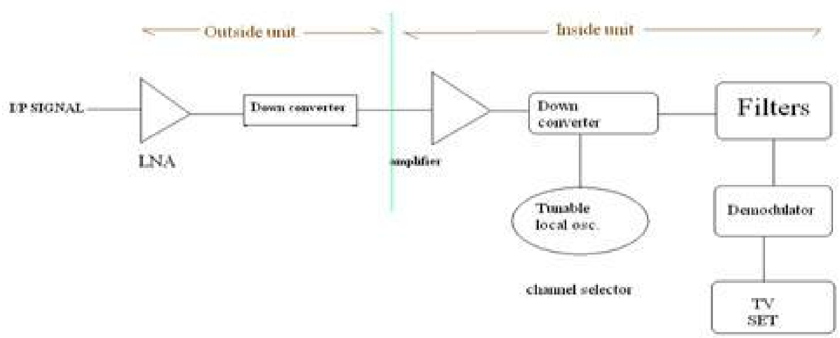


Figure 5.13 DTH Service

5.12.2 Advantage

- ✓ DTH also offers digital quality signals which do not degrade the picture or sound quality.
- ✓ It also offers interactive channels and program guides with customers having the choice to block out programming which they consider undesirable
- ✓ One of the great advantages of the cable industry has been the ability to provide local channels, but this handicap has been overcome by many

DTH providers using other local channels or local feeds.

- ✓ The other advantage of DTH is the availability of satellite broadcast in rural and semi-urban areas where cable is difficult to install.

5.13 Digital audio broadcast (DAB)

DAB Project is an industry-led consortium of over 300 companies

- ✓ The DAB Project was launched on 10th September, 1993
- ✓ In 1995 it was basically finished and became operational
- ✓ There are several sub-standards of the DAB standard
  - ❖ DAB-S (Satellite) – using QPSK – 40 Mb/s
  - ❖ DAB-T (Terrestrial) – using QAM – 50 Mb/s
  - ❖ DAB-C (Cable) – using OFDM – 24 Mb/s
- ✓ These three sub-standards basically differ only in the specifications to the physical representation, modulation, transmission and reception of the signal.
- ✓ The DAB stream consists of a series of fixed length packets which make up a Transport Stream (TS). The packets support ‘streams’ or ‘data sections’.
- ✓ Streams carry higher layer packets derived from an MPEG stream & Data sections are blocks of data carrying signaling and control data.
- ✓ DAB is actually a support mechanism for MPEG.& One MPEG stream needing higher instantaneous data can ‘steal’ capacity from another with spare capacity.

5.14 World space services

World Space (Nasdaq: WRSP) is the world's only global media and entertainment company positioned to offer a satellite radio experience to consumers in more than 130 countries with five billion people, driving 300 million cars. World Space delivers the latest tunes, trends and information from around the world and around the corner.

World Space subscribers benefit from a unique combination of local programming, original World Space content and content from leading brands around the globe, including the BBC, CNN, Virgin Radio, NDTV and RFI. World Space's satellites cover two-thirds of the globe with six beams.

Each beam is capable of delivering up to 80 channels of high quality digital audio and multimedia programming directly to World Space Satellite Radios anytime and virtually anywhere in its coverage area. World Space is a pioneer of satellite-based digital radio services (DARS) and was instrumental in the development of the technology infrastructure used today by XM Satellite Radio.

5.15 Business Television (BTV) - Adaptations for Education

Business television (BTV) is the production and distribution, via satellite, of video programs for closed user group audiences. It often has two -way audio interaction component made through a simple telephone line. It is being used by many industries including brokerage firms, pizza houses, car dealers and delivery services.

BTV is an increasingly popular method of information delivery for corporations and institutions. Private networks, account for about 70 percent of all BTV networks. It is estimated that by the mid-1990s BTV has the potential to grow to a $1.6 billion market in North America with more and more Fortune 1,000 companies getting involved. The increase in use of BTV has been dramatic.

Institution updates, news, training, meetings and other events can be broadcast live to multiple locations. The expertise of the best instructors can be delivered to thousands of people without requiring trainers to go to the site. Information can be disseminated to all employees at once, not just a few at a time. Delivery to the workplace at low cost provides the access to training that has been denied lower level employees. It may be the key to re-training America's work force.

Television has been used to deliver training and information within businesses for more than 40 years. Its recent growth began with the introduction of the video cassette in the early 1970s. Even though most programming is produced for video cassette distribution, business is using BTV to provide efficient delivery of specialized programs via satellite.

The advent of smaller receiving stations - called very small aperture terminals (VSATs) has made private communication networks much more economical to operate. BTV has a number of tangible benefits, such as reducing travel, immediate delivery of time-critical messages, and eliminating cassette duplication and distribution hassles.

The programming on BTV networks is extremely cost-effective compared to seminar fees and downtime for travel. It is an excellent way to get solid and current information very fast. Some people prefer to attend seminars and conferences where they can read, see, hear and ask questions in person. BTV provides yet another piece of the education menu and is another way to provide professional development.

A key advantage is that its format allows viewers to interact with presenters by telephone, enabling viewers to become a part of the program. The satellite effectively places people in the same room, so that sales personnel in the field can learn about new products at the same time.

Speed of transmission may well be the competitive edge which s ome firms need as they introduce new products and services. BTV enables employees in many locations to focus on common problems or issues that might develop into crises without quick communication and resolution.

BTV networks transmit information every business day on a broad range of topics, and provide instructional courses on various products, market trends, selling and motivation. Networks give subscribers the tools to apply the information they have to real world situations.

5.16 GRAMSAT

ISRO has come up with the concept of dedicated GRAMSAT satellites, keeping in mind the urgent need to eradicate illiteracy in the rural belt which is necessary for the all round development of the nation.

This Gramsat satellite is carrying six to eight high powered C -band transponders, which together with video compression techniques can disseminate regional and cultural specific audio-visual programmes of relevance in each of the regional languages through rebroadcast mode on an ordinary TV set.

The high power in C-band has enabled even remote area viewers outside the reach of the TV transmitters to receive programmers of their choice in a direct reception mode with a simple .dish antenna.

The salient features of GRAMSAT projects are:

i. Its communications networks are at the state level connecting the state capital to districts, blocks and enabling a reach to villages.

ii. It is also providing computer connectivity data broadcasting, TV - broadcasting facilities having applications like e- governance, development information, teleconferencing, helping disaster management.

iii. Providing rural-education broadcasting.

However, the Gramsat projects have an appropriate combination of following activities.

(i) Interactive training at district and block levels employing suit able configuration

(ii) Broadcasting services for rural development

(iii) Computer interconnectivity and data exchange services

(iv) Tele-health and tele-medicine services.

5.17 Specialized services

5.17.1Satellite-email services

The addition of Internet Access enables Astrium to act as an Internet Service Provider (ISP) capable of offering Inmarsat users a tailor -made Internet connection.

With Internet services added to our range of terrestrial networks, you will no longer need to subscribe to a third party for Internet access (available for Inmarsat A, B, M, mini-M, Fleet, GAN, Regional BGAN & SWIFT networks).

We treat Internet in the same way as the other terrestrial networks we provide, and thus offer unrestricted access to this service. There is no time - consuming log-on procedure, as users are not required to submit a user-ID or password.

Description of E-mail Service

Astrium's E-Mail service allows Inmarsat users to send and receive e-mail directly through the Internet without accessing a public telephone network.

Features and Benefits

- ✓ No need to configure an e-mail client to access a Astrium e-mail account
- ✓ Service optimized for use with low bandwidth Inmarsat terminals
- ✓ Filter e-mail by previewing the Inbox and deleting any unwanted e-mails prior to downloading
- ✓ No surcharge or monthly subscription fees
- ✓ Service billed according to standard airtime prices for Inmarsat service used

5.17.2 Video Conferencing (medium resolution)

Video conferencing technology can be used to provide the same full, two-way interactivity of satellite broadcast at much lower cost. For Multi-Site meetings, video conferencing uses bridging systems to connect each site to the others.

It is possible to configure a video conference bridge to show all sites at the same time on a projection screen or monitor. Or, as is more typical, a bridge can show just the site from which a person is speaking or making a presentation.

The technology that makes interactive video conferencing possible, compresses video and audio signals, thus creating an image quality lower than that of satellite broadcasts.

5.17.3. Satellite Internet access

Satellite Internet access is Internet access provided through communications satellites. Modern satellite Internet service is typically provided to users through geostationary satellites that can offer high data speeds, with newer satellites using Ka band to achieve downstream data speeds up to 50 Mbps.

Satellite Internet generally relies on three primary components: a satellite in geostationary orbit (sometimes referred to as a geosynchronous Earth orbit, or GEO), a number of ground stations known as gateways that relay Internet data to and from the satellite via radio waves (microwave), and a VSAT (very-small-aperture terminal) dish antenna with a transceiver, located at the subscriber's premises.

Other components of a satellite Internet system include a modem at the user end which links the user's network with the transceiver, and a centralized network operations center (NOC) for monitoring the entire system.

APPLICATIONS

Figure example of INSAT -3 satellites

Figure example of Weather forecasting satellite

Post Test MCQ

1. The INTELSAT VI satellite used ________.

a) The INTELSAT VI satellite used heaters to maintain propulsion thrusters and line temperatures

b) With active attitude control, there is no overall stabilizing torque present to resist the disturbance torques

c) Proper moment

d) None of these

2. Teledesic satellites are _________satellites.

a) GEO

b) MEO

c) LEO

d) none of the above

3. GPS satellites are ________ satellites.

a) GEO

b) MEO

c) LEO

d) none of the above

4. Low-Earth-orbit (LEO) satellites have _____ orbits

a) equatorial

b) polar

c) inclined

d) none of the above

5. MEO satellites are located at altitudes between km.

a) 3000 and 5000

b) 5000 and 10,000

c) 5000 and 15,000

d) 10 and 200

6. INTELSAT stands for

a) **International Telecommunications Satellite**

b) India Telecommunications Satellite

c) Inter Telecommunications Satellite

d) Italian Telemetry Satellite

7. What is application of satellite systems?

a) Weather forecasting

b) Terrestrial communication

c) point to point communication

d) Multipoint communication

8. Find out the service given by satellite systems.

a) Broadcasting satellite services

b) Signal transmission

c) Information transmission

d) None of the above

9. Universal time day is

a) UT day =1/24(hours+minutes/60+seconds/3600)

b) UT day =1/24(hours+minutes+seconds/3600)

c) UT day =1/24(hours+minutes/6+seconds/360)

d) UT day =1/12(hours+minutes/6+seconds/360)

10. **Identify the difference between the geodetic & geocentric latitudes.**

a) The latitudes reaches a maximum at a geocentric latitude of 45deg, when the geodetic latitude is 45.192deg.

b) The latitudes reaches a maximum at a geocentric latitude of 30deg, when the geodetic latitude is 45.192deg.

c) The latitudes reaches a maximum at a geocentric latitude of 45deg, when the geodetic latitude is 4.192deg.

(d) No difference.

11. The Orbital Communications Corporation (Orbcomm) system is a__________, which provides two-way message and data communications services and position determination.

a) MEO satellite system

b) GEO satellite system

c) LEO satellite system

d) None of these

12. The Orbital Communications Corporation (Orbcomm) system is a LEO satellite system, which provides __________message and data communications services and position determination.

a) Two-way

b) One-way

c) Half-way

d) None of these

13. The GPS system uses_____________, from satellites to users, so that the user does not require a transmitter, only a GPS receiver.

a) One-way transmissions

b) Two-way transmissions

c) Half-way transmissions

d) None of these

14. In the GPS system, a constellation of 24 satellites circles the earth in near-circular________.

a) GEO

(b) MEO

c) Inclined orbits

d) None of these

CONCLUSION

- ❖ In this unit we discussed satellite applications.
- ❖ The Direct Broadcast satellites (DBS) and Direct to Home (DTH) Broadcast were discussed.
- ❖ The Video conferencing, Internet were discussed.

REFERENCES

1. Wilbur L.Pritchard, Hendri G. Suyderhoud, Robert A. Nelson, "Satellite Communication Systems Engineering", Prentice Hall/Pearson, 2007.
2. N. Agarwal, "Design of Geosynchronous Space Craft", Prentice Hall, 1986.
3. Bruce R. Elbert, "Satellite Communication Applications", Hand Book, Artech House Bostan London,1997.
4. Tri T. Ha, "Digital Satellite Communication", IInd edition, 1990.
5. Emanuel Fthenakis, "Manual of Satellite Communications", Mc Graw Hill Book Co., 1984.
6. Robert G. Winch, "Telecommunication Transmission Systems", Mc Graw-Hill Book Co., 1983.
7. Brian Ackroyd, "World Satellite Communication and earth station Design", BSP professional Books, 1990.
8. G.B.Bleazard, "Introducing Satellite communications", NCC Publication, 1985.
9. M.Richharia, "Satellite Communication Systems-Design Principles", Macmillan 2003.

ASSIGNMENT

1. Explain with the neat diagram the indoor and outdoor units of DBS home receiver.
2. Discuss the satellite mobile services.
3. Explain the following satellite applications.
   i) GPS.
   ii) Satellite Navigational system.
4. What are the major short comings of present day VAST system?
5. Write note on Bit rate for digital television.
6. List out the MPEG compression standards.
7. Explain detail about the VSAT.
8. What is meant by DTH? What are the design issues to be considered for launching DTH systems?
9. What is the orbital spacing of satellites?
10. Write brief notes on the advantages and disadvantages of using satellite in LEOs ,MEOs and GEOs for mobile satellite communications.

LECTURE NOTES

ON

OPTICAL FIBER COMMUNICATION

Prepared by
Satya Prakash Rout

Assistant Professor, ECE

AIM & OBJECTIVES

- ❖ To learn the basic elements of optical fiber transmission link, fiber modes configurations and structures.
- ❖ To understand the different kind of losses, signal distortion, SM fibers.
- ❖ To learn the various optical sources, materials and fiber splicing.
- ❖ To learn the fiber optical receivers and noise performance in photo detector.
- ❖ To explore link budget, WDM, solitons and SONET/SDH network.

PRE TEST-MCQ TYPE

1. Which equations are best suited for the study of electromagnetic wave propagation?
a) Maxwell's equations
b) Allen-Cahn equations
c) Avrami equations
d) Boltzmann's equations

2. When λ is the optical wavelength in vacuum, k is given by k=2Π/λ. What does k stand for in the above equation?
a) Phase propagation constant
b) Dielectric constant
c) Boltzmann's constant
d) Free-space constant

3. When light is described as an electromagnetic wave, it consists of a periodically varying electric E and magnetic field H which are oriented at an angle?
a) 90 degree to each other
b) Less than 90 degree
c) Greater than 90 degree
d) 180 degree apart

4. Which is the most important velocity in the study of transmission characteristics of optical fiber?
a) Phase velocity
b) Group velocity
c) Normalized velocity
d) Average velocity

UNIT I INTRODUCTION TO OPTICAL FIBERS

Evolution of fiber optic system- Element of an Optical Fiber Transmission link- Total internal Reflection- Acceptance angle –Numerical aperture – Skew rays Ray Optics-Optical Fiber Modes and Configurations- Mode theory of Circular Wave guides- Overview of Modes-Key Modal concepts- Linearly Polarized Modes -Single Mode Fibers-Graded Index fiber structure

THEORY

Introduction

Fiber-optic communication is a method of transmitting information from one place to another by sending pulses of light through an optical fiber. The light forms an electromagnetic carrier wave that is modulated to carry information. Fiber is preferred over electrical cabling when high bandwidth, long distance, or immunity to electromagnetic interference are required. This type of communication can transmit voice, video, and telemetry through local area networks, computer networks, or across long distances.

Optical fiber is used by many telecommunications companies to transmit telephone signals, Internet communication, and cable television signals. Researchers at Bell Labs have reached internet speeds of over 100 peta bit ×kilometer per second using fiber-optic communication. The process of communicating using fiber-optics involves the following basic steps:

1. Creating the optical signal involving the use of a transmitter, usually from an electrical signal

2. Relaying the signal along the fiber, ensuring that the signal does not become too distorted or weak

3. Receiving the optical signal

4. Converting it into an electrical signal

Historical Development

First developed in the 1970s, fiber-optics have revolutionized the telecommunications industry and have played a major role in the advent of the Information Age. Because of its advantages over electrical transmission, optical fibers have largely replaced copper wire communications in core networks in the developed world.

In 1880 Alexander Graham Bell and his assistant Charles Sumner Tainter created a very early precursor to fiber-optic communications, the Photophone, at Bell's newly established Volta Laboratory in Washington, D.C. Bell considered it his most important invention. The device allowed for the transmission of sound on a beam of light. On June 3, 1880, Bell conducted the world's first wireless telephone transmission between two buildings, some 213 meters apart. Due to its use of an atmospheric transmission medium, the Photophone would not prove practical until advances in laser and optical fiber technologies permitted the secure transport of light. The Photophone's first practical use came in military communication systems many decades later.

In 1954 Harold Hopkins and Narinder Singh Kapany showed that rolled fiber glass allowed light to be transmitted. Initially it was considered that the light can traverse in only straight medium. Jun-ichi Nishizawa, a Japanese scientist at Tohoku University, proposed the use of optical fibers for communications in 1963. Nishizawa invented the PIN diode and the static induction transistor, both of which contributed to the development of optical fiber communications.

In 1966 Charles K. Kao and George Hockham at STC Laboratories (STL) showed that the losses of 1,000 dB/km in existing glass (compared to 5–10 dB/km in coaxial cable) were due to contaminants which could potentially be removed.

Optical fiber was successfully developed in 1970 by Corning Glass Works, with attenuation low enough for communication purposes (about 20 dB/km) and at the same time GaAs semiconductor lasers were developed that were compact and therefore suitable for transmitting light through fiber optic cables for long distances. In 1973, Optelecom, Inc., co-founded by the inventor of the laser, Gordon Gould, received a contract from APA for the first optical communication systems. Developed for Army Missile Command in Huntsville, Alabama, it was a laser on the ground and a spout of optical fiber played out by missile to transmit a modulated signal over five kilometers.

After a period of research starting from 1975, the first commercial fiber-optic communications system was developed which operated at a wavelength around 0.8 μm and used GaAs semiconductor lasers. This first-generation system operated at a bit rate of 45 Mbit/s with repeater spacing of up to 10 km. Soon on 22 April 1977, General Telephone and Electronics sent the first live telephone traffic through fiber optics at a 6 Mbit/s throughput in Long Beach, California.

In October 1973, Corning Glass signed a development contract with CSELT and Pirelli aimed to test fiber optics in an urban environment: in September 1977, the second cable in this test series, named COS-2, was experimentally deployed in two lines (9 km) in Turin, for the first time in a big city, at a speed of 140 Mbit/s.

The second generation of fiber-optic communication was developed for commercial use in the early 1980s, operated at 1.3 μm and used InGaAsP semiconductor lasers. These early systems were initially limited by multi mode fiber dispersion, and in 1981 the single-mode fiber was revealed to greatly improve system performance, however practical connectors capable of working with single mode fiber proved difficult to develop. Canadian service provider SaskTel had completed construction of what was then the world's longest commercial fiber optic network, which covered 3,268 km (2,031 mi) and linked 52 communities. By 1987, these systems were operating at bit rates of up to 1.7 Gb/s with repeater spacing up to 50 km (31 mi). The first transatlantic telephone cable to use optical fiber was TAT-8, based on Desurvire optimised laser amplification technology. It went into operation in 1988.

Third-generation fiber-optic systems operated at 1.55 μm and had losses of about 0.2 dB/km. This development was spurred by the discovery of Indium gallium arsenide and the development of the Indium Gallium Arsenide photodiode by Pearsall. Engineers overcame earlier difficulties with pulse- spreading at that wavelength using conventional InGaAsP semiconductor lasers. Scientists overcame this difficulty by using dispersion-shifted fibers designed to have minimal dispersion at 1.55 μm or by limiting the laser spectrum to a single longitudinal mode. These developments eventually allowed third-generation systems to operate commercially at 2.5 Gbit/s with repeater spacing in excess of 100 km (62 mi).

The fourth generation of fiber-optic communication systems used optical amplification to reduce the need for repeaters and wavelength-division multiplexing to increase data capacity. These two improvements caused a revolution that resulted in the doubling of system capacity every six months starting in 1992 until a bit rate of 10 Tb/s was reached by 2001. In 2006 a bit-rate of 14 Tbit/s was reached over a single 160 km (99 mi) line using optical amplifiers.

The focus of development for the fifth generation of fiber-optic communications is on extending the wavelength range over which a WDM system can operate. The conventional wavelength window, known as the C band, covers the wavelength range 1.53–1.57 μm, and dry fiber has a low-loss window promising an extension of that range to 1.30–1.65 μm. Other developments include the concept of "optical solutions", pulses that preserve their shape by counteracting the effects of dispersion with the nonlinear effects of the fiber by using pulses of a specific shape.

In the late 1990s through 2000, industry promoters, and research companies such as KMI, and RHK predicted massive increases in demand for communications bandwidth due to increased use of the Internet, and commercialization of various bandwidth-intensive consumer services, such as video on demand. Internet protocol data traffic was increasing exponentially, at a faster rate than integrated circuit complexity had increased under Moore's Law. From the bust of the dot-com bubble through 2006, however, the main trend in the industry has been consolidation of firms and off shoring of manufacturing to reduce costs.

Advantages of Fiber Optic Transmission

Optical fibers have largely replaced copper wire communications in core networks in the developed world, because of its advantages over electrical transmission. Here are the main advantages of fiber optic transmission.

Extremely High Bandwidth: No other cable-based data transmission medium offers the bandwidth that fiber does. The volume of data that fiber optic cables transmit per unit time is far great than copper cables.

Longer Distance: in fiber optic transmission, optical cables are capable of providing low power loss, which enables signals can be transmitted to a longer distance than copper cables.

Resistance to Electromagnetic Interference: in practical cable deployment, it’s inevitable to meet environments like power substations, heating, ventilating and other industrial sources of interference. However, fiber has a very low rate of bit error (10 EXP-13), as a result of fiber being so resistant to electromagnetic interference. Fiber optic transmission is virtually noise free.

Low Security Risk: the growth of the fiber optic communication market is mainly driven by increasing awareness about data security concerns and use of the alternative raw material. Data or signals are transmitted via light in fiber optic transmission. Therefore there is no way to detect the data being transmitted by "listening in" to the electromagnetic energy "leaking" through the cable, which ensures the absolute security of information.

Small Size: fiber optic cable has a very small diameter. For instance, the cable diameter of a single OM3 multimode fiber is about 2mm, which is smaller than that of coaxial copper cable. Small size saves mere space in fiber optic transmission.

Light Weight: fiber optic cables are made of glass or plastic, and they are thinner than copper cables. These make them lighter and easy to install.

Easy to Accommodate Increasing Bandwidth: with the use of fiber optic cable, new equipment can be added to existing cable infrastructure. Because optical cable can provide vastly expanded capacity over the originally laid cable and WDM (wavelength division multiplexing) technology, including CWDM and DWDM, enables fiber cables the ability to accommodate more bandwidth.

Disadvantages of Fiber Optic Transmission

Though fiber optic transmission brings lots of convenience, its disadvantages also cannot be ignored.

Fragility: usually optical fiber cables are made of glass, which lends to they are more fragile than electrical wires. In addition, glass can be affected by various chemicals including hydrogen gas (a problem in underwater cables), making them need more cares when deployed underground.

Difficult to Install: it’s not easy to splice fiber optic cable. And if you bend them too much, they will break. And fiber cable is highly susceptible to becoming cut or damaged during installation or construction activities. All these make it difficult to install.

Attenuation & Dispersion: as transmission distance getting longer, light will be attenuated and dispersed, which requires extra optical components like EDFA to be added.

Cost is Higher Than Copper Cable: despite the fact that fiber optic installation costs are dropping by as much as 60% a year, installing fiber optic cabling is still relatively higher than copper cables. Because copper cable installation does not need extra care like fiber cables. However, optical fiber is still moving into the local loop, and through technologies such as FTTx (fiber to the home, premises, etc.) and PONs (passive optical networks), enabling subscriber and end user broadband access.

Special Equipment Is Often Required: to ensure the quality of fiber optic transmission, some special equipment is needed. For example, equipment such as OTDR (optical time-domain reflectometry) is required and expensive, specialized optical test equipment such as optical probes and power meter are needed at most fiber endpoints to properly provide testing of optical fiber.

Applications of Optical Fiber Communications

Fiber optic cables find many uses in a wide variety of industries and applications. Some uses of fiber optic cables include:

Medical -Used as light guides, imaging tools and also as lasers for surgeries

Defense/Government-Used as hydrophones for seismic waves and SONAR , as wiring in aircraft, submarines and other vehicles and also for field networking

Data Storage- Used for data transmission

Telecommunications- Fiber is laid and used for transmitting and receiving purposes

Networking- Used to connect users and servers in a variety of network settings and help increase the speed and accuracy of data transmission

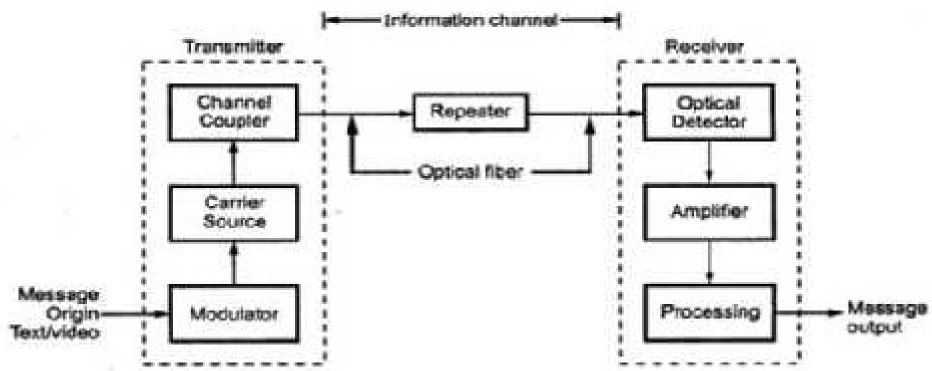
Industrial/Commercial- Used for imaging in hard to reach areas, as wiring where EMI is an issue, as sensory devices to make temperature, pressure and other measurements, and as wiring in automobiles and in industrial settings.

Broadcast/CATV-Broadcast/cable companies are using fiber optic cables for wiring CATV, HDTV, internet, video on- demand and other applications. Fiber optic cables are used for lighting and imaging and as sensors to measure and monitor a vast array of variables. Fiber optic cables are also used in research and development and testing across all the above mentioned industries

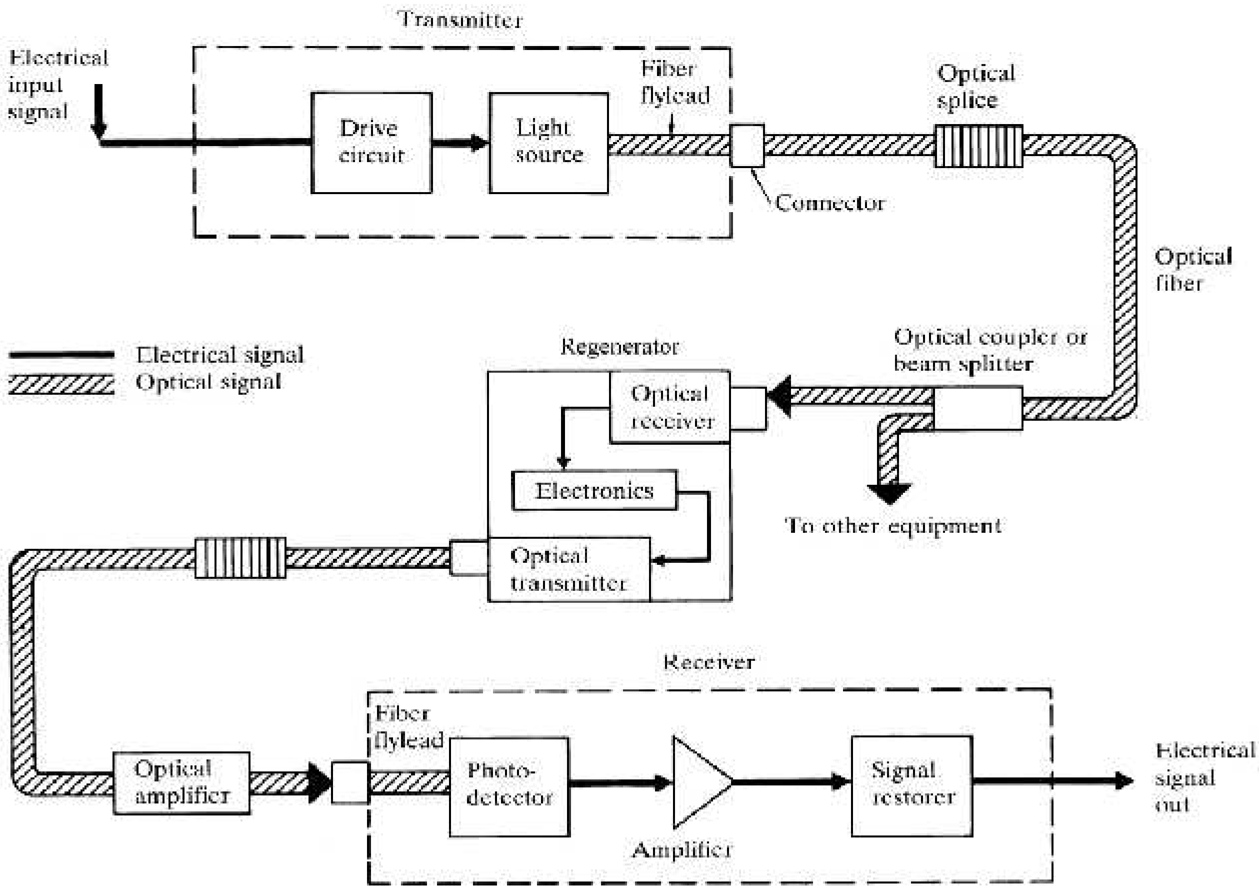
The optical fibers have many applications. Some of them are as follows

- ❖ Used in telephone systems
- ❖ Used in sub-marine cable networks
- ❖ Used in data link for computer networks, CATV Systems
- ❖ Used in CCTV surveillance cameras
- ❖ Used for connecting fire, police, and other emergency services.
- ❖ Used in hospitals, schools, and traffic management systems.
- ❖ They have many industrial uses and also used for in heavy duty constructions.

Block Diagram of Optical Fiber Communication System



Block Diagram of Optical Fiber Communication System



Message origin:

Generally message origin is from a transducer that converts a non-electrical message into an electrical signal. Common examples include microphones for converting sound waves into currents and video (TV) cameras for converting images into current. For data transfer between computers, the message is already in electrical form.

Modulator:

The modulator has two main functions.

1) It converts the electrical message into proper format.

2) It impresses this signal onto the wave generated by the carrier source.

Two distinct categories of modulation are used i.e. analog modulation and digital modulation.

Carrier source:

Carrier source generates the wave on which the information is transmitted. This wave is called the carrier. For fiber optic system, a laser diode (LD) or a light emitting diode (LED) is used. They can be called as optic oscillators, they provide stable, single frequency waves with sufficient power for long distance propagation.

Channel coupler:

Coupler feeds the power into information channel. For an atmospheric optic system, the channel coupler is a lens used for collimating the light emitted by the source and directing this light towards the receiver. The coupler must efficiently transfer the modulated light beam from the source to the optic fiber. The channel coupler design is an important part of fiber system because of possibility of high losses.

Information channel:

The information channel is the path between the transmitter and receiver. In fiber optic communications, a glass or plastic fiber is the channel. Desirable characteristics of the information channel include low attenuation and large light acceptance cone angle. Optical amplifiers boost the power levels of weak signals. Amplifiers are needed in very long links to provide sufficient power to the receiver. Repeaters can be used only for digital systems. They convert weak and distorted optical signals to electrical ones and then regenerate the original digital pulse trains for further transmission.

Another important property of the information channel is the propagation time of the waves travelling along it. A signal propagating along a fiber normally contains a range of fiber optic frequencies and divides its power along several ray paths. This results in a distortion of the propagation signal. In a digital system, this distortion appears as a spreading and deforming of the pulses. The spreading is so great that adjacent pulses begin to overlap and become unrecognizable as separate bits of information.

Optical detector:

The information begin transmitted is detected by detector. In the fiber system the optic wave is converted into an electric current by a photodetector. The current developed by the detector is proportional to the power in the incident optic wave. Detector output current contains the transmitted information. This detector output is then filtered to remove the constant bias and then amplified. The important properties of photodetectors are small size, economy, long life, low power consumption, high sensitivity to optic signals and fast response to quick variations in the optic power. Signal processing includes filtering, amplification. Proper filtering maximizes the ratio of signal to unwanted power. For a digital system decision circuit is an additional block. The bit error rate (BER) should be very small for quality communications.

Signal processing:

Signal processing includes filtering, amplification. Proper filtering maximizes the ratio of signal to unwanted power. For a digital syst5em decision circuit is an additional block. The bit error rate (BER) should be very small for quality communications.

Message output:

The electrical form of the message emerging from the signal processor is transformed into a sound wave or visual image. Sometimes these signals are directly usable when computers or other machines are connected through a fiber system.

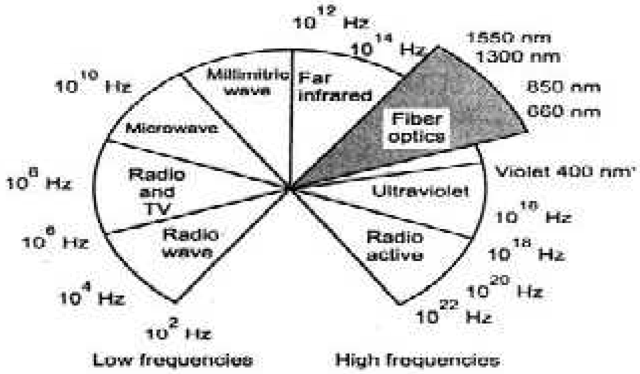
Electromagnetic Spectrum

The radio waves and light are electromagnetic waves. The rate at which they alternate in polarity is called their frequency (f) measured in hertz (Hz). The speed of electromagnetic wave (c) in free space is approximately 3 x 10^8 m/sec. The distance travelled during each cycle is called as wavelength (λ)

In fiber optics, it is more convenient to use the wavelength of light instead of the frequency with light frequencies; wavelength is often stated in microns or nanometers.

1 micron (μ) = 1 Micrometre (1 x 10^{-6}) ;1 nano (n) = 10^{-9} meter

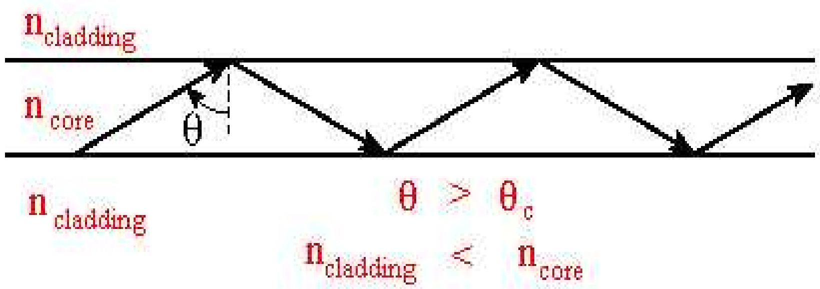
Fiber optics uses visible and infrared light. Infrared light covers a fairly wide range of wavelengths and is generally used for all fiber optic communications. Visible light is normally used for very short range transmission using a plastic fiber



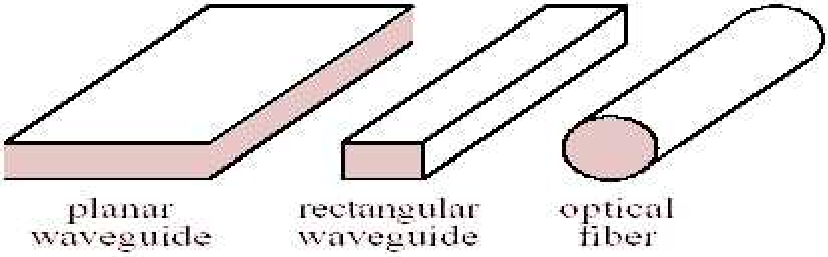
Electromagnetic Spectrum

Optical Fiber Waveguides

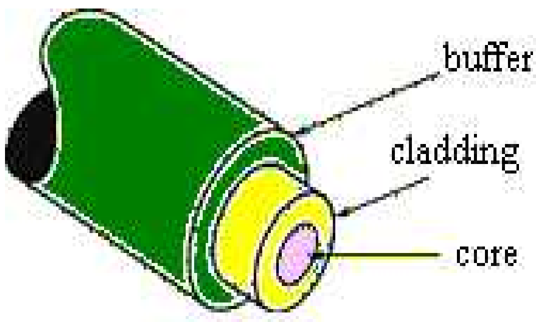
In free space light ravels as its maximum possible speed i.e. 3 x 10^8 m/s or 186 x 10^3 miles/sec. When light travels through a material it exhibits certain behavior explained by laws of reflection, refraction. An optical wave guide is a structure that "guides" a light wave by constraining it to travel along a certain desired path. If the transverse dimensions of the guide are much larger than the wavelength of the guided light, that explain how the optical waveguide works using geometrical optics and total internal reflection.



A wave guide traps light by surrounding a guiding region, called the core, made from a material with index of refraction n_{core}, with a material called the cladding, made from a material with index of refraction $n_{cladding} < n_{core}$. Light entering is trapped as long as $\sin\theta > n_{cladding}/n_{core}$.



Light can be guided by planar or rectangular wave guides, or by optical fibers. An optical fiber consists of three concentric elements, the core, the cladding and the outer coating, often called the buffer. The core is usually made of glass or plastic. The core is the light-carrying portion of the fiber. The cladding surrounds the core. The cladding is made of a material with a slightly lower index of refraction than the core. This difference in the indices causes total internal reflection to occur at the core-cladding boundary along the length of the fiber. Light is transmitted down the fiber and does not escape through the sides of the fiber.

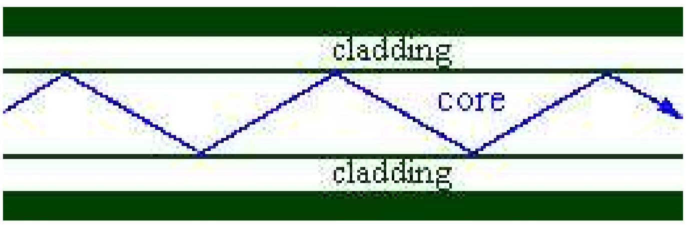


Fiber Optic Core: the inner light-carrying member with a high index of refraction.

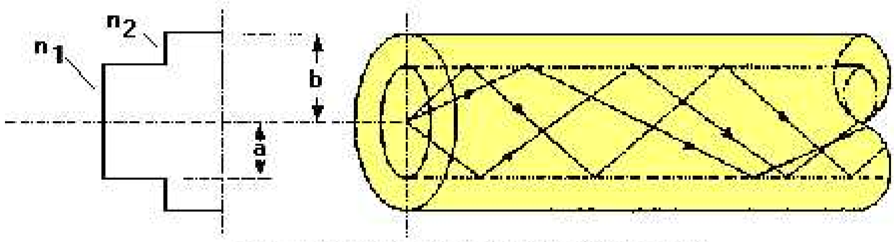
Cladding: the middle layer, which serves to confine the light to the core. It has a lower index of refraction.

Buffer: The outer layer, which serves as a "shock absorber" to protect the core and cladding from damage. The coating usually comprises one or more coats of a plastic material to protect the fiber from the physical environment.

o



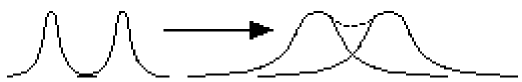
Light injected into the fiber optic core and striking the core-to-cladding interface at an angle greater than the critical angle is reflected back into the core. Since the angles of incidence and reflection are equal, the light ray continues to zigzag down the length of the fiber. The light is trapped within the core. Light striking the interface at less than the critical angle passes into the cladding and is lost.



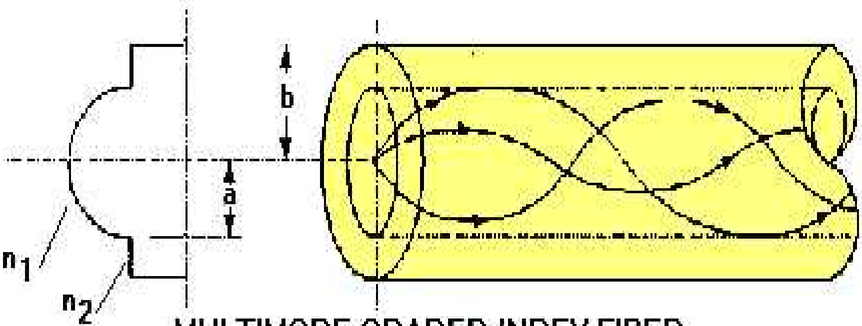
MULTIMODE STEP-INDEX FIBER

Fibers for which the refractive index of the core is a constant and the index changes abruptly at the core-cladding interface are called step-index fibers. Step-index fibers are available with core diameters of 100 mm to 1000 mm. They are well suited to applications requiring high-power densities, such as delivering laser power for medical and industrial applications.

Multimode step-index fibers trap light with many different entrance angles, each mode in a step-index multimode fiber is associated with a different entrance angle. Each mode therefore travels along a different path through the fiber. Different propagating modes have different velocities. As an optical pulse travels down a multimode fiber, the pulse begins to spread. Pulses that enter well separated from each other will eventually overlap each other. This limits the distance over which the fiber can transport data. Multimode step-index fibers are not well suited for data transport and communications.

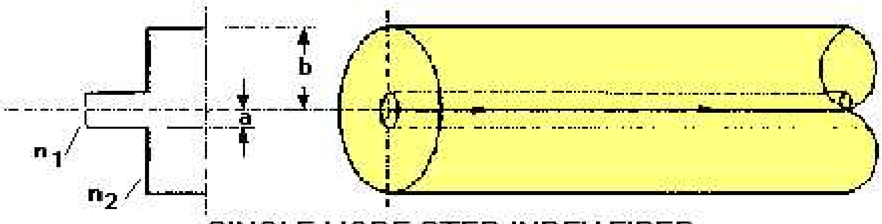


In a multimode graded-index fiber the core has an index of refraction that decreases as the radial distance from the center of the core increases. As a result, the light travels faster near the edge of the core than near the center. Different modes therefore travel in curved paths with nearly equal travel times. This greatly reduces the spreading of optical pulses.



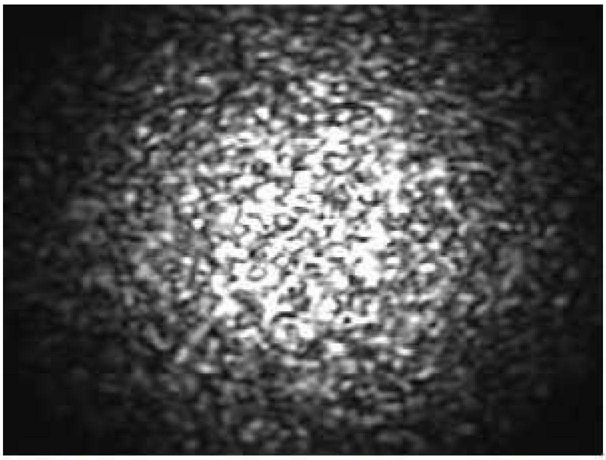
MULTIMODE GRADED-INDEX FIBER

A single mode fiber only allows light to propagate down its center and there are no longer different velocities for different modes. A single mode fiber is much thinner than a multimode fiber and can no longer be analyzed using geometrical optics. Typical core diameters are between 5 mm and 10 mm.

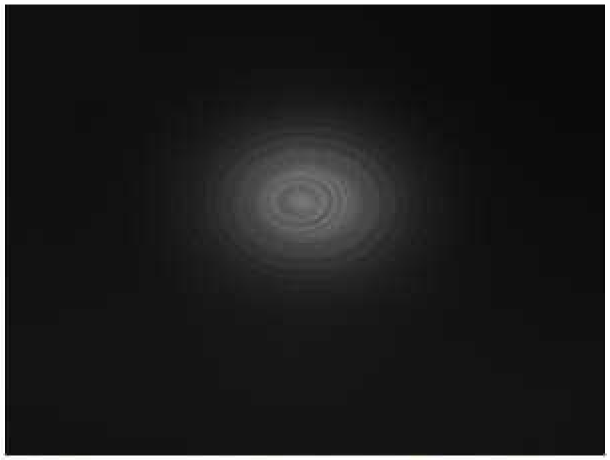


SINGLE MODE STEP-INDEX FIBER

When laser light is coupled into a fiber, the distribution of the light emerging from the other end reveals if the fiber is a multimode or single mode fiber.



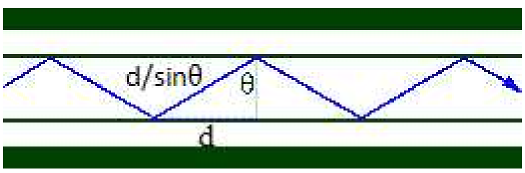
Light emerging from a multi-mode fiber



Light emerging from a single mode fiber

Optical fibers are used widely in the medical field for diagnoses and treatment. Optical fibers can be bundled into flexible strands, which can be inserted into blood vessels, lungs and other parts of the body. An Endoscope is a medical tool carrying two bundles of optic fibers inside one long tube. One bundle directs light at the tissue being tested, while the other bundle carries light reflected from the tissue, producing a detailed image. Endoscopes can be designed to look at regions of the human body, such as the knees, or other joints in the body

In a step-index fiber in the ray approximation, the ray propagating along the axis of the fiber has the shortest route, while the ray incident at the critical angle has the longest route. Determine the difference in travel time (in ns/km) for the modes defined by those two rays for a fiber with n_{core} = 1.5 and $n_{cladding}$ = 1.485.



Solution:

If a ray propagating along the axis of the fiber travels a distance d, then a ray incident at the critical angle θ_c travels a distance $L = d/\sin\theta_c$.

The respective travel times are $t_d = d n_{core}/c$ and $t_L = d n_{core}/(\sin\theta_c\, c)$.

$\sin\theta_C = n_{cladding}/n_{core}$.

θ_C = 81.9 deg.

For d = 1000 m , t_d = 5000 ns and t_L =5050.51 ns.

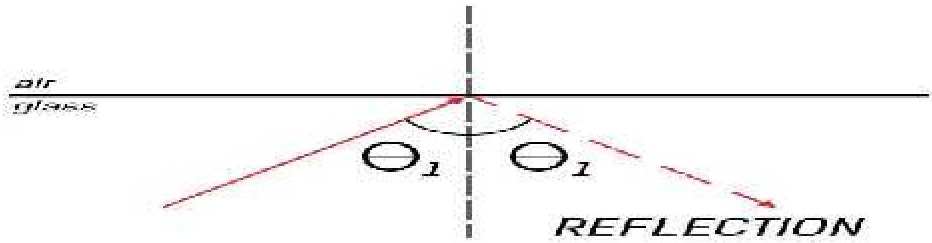
The difference in travel time is therefore 50.51 ns/km.

Ray theory

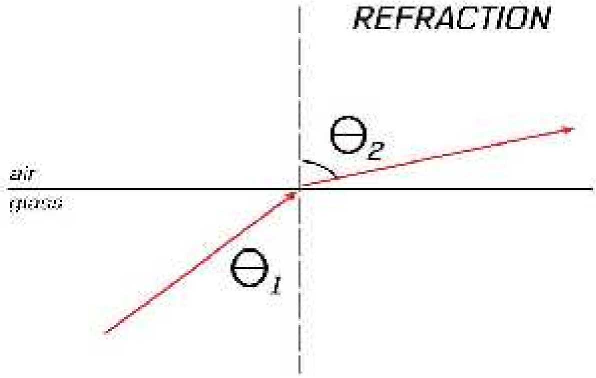
The phenomenon of splitting of white light into its constituents is known as dispersion. The concepts of reflection and refraction of light are based on a theory known as Ray theory or geometric optics, where light waves are considered as waves and represented with simple geometric lines or rays.

The basic laws of ray theory/geometric optics

- In a homogeneous medium, light rays are straight lines.
- Light may be absorbed or reflected.
- Reflected ray lies in the plane of incidence and angle of incidence will be equal to the angle of reflection.
- At the boundary between two media of different refractive indices, the refracted ray will lie in the plane of incidence. Snell's Law will give the relationship between the angles of incidence and refraction.



Reflection depends on the type of surface on which light is incident. An essential condition for reflection to occur with glossy surfaces is that the angle made by the incident ray of light with the normal at the point of contact should be equal to the angle of reflection with that normal. The images produced from this reflection have different properties according to the shape of the surface. For example, for a flat mirror, the image produced is upright, has the same size as that of the object and is equally distanced from the surface of the mirror as the real object. However, the properties of a parabolic mirror are different and so on.



Refraction is the bending of light in a particular medium due to the speed of light in that medium. The speed of light in any medium can be given by

$$v = \frac{c}{n}$$

$$\text{Refractive index n} = \frac{\text{Speed of light in air}}{\text{Speed of light in medium}} = \frac{c}{v}$$

The refractive index for vacuum and air is 1.0 for water it is 1.3 and for glass refractive index is 1.5. Here n is the refractive index of that medium. When a ray of light is incident at the interface of two media with different refractive indices, it will bend either towards or away from the normal depending on the refractive indices of the media. According to Snell's law, refraction can be represented as

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$$

n_1 = refractive index of first medium

n_2 θ_1 = angle of incidence, n_2= refractive index of second medium

θ_2 = angle of refraction

For $n_1 > n_2$, θ_2 is always greater than θ_1. Or to put it in different words, light moving from a medium of high refractive index (glass) to a medium of lower refractive index (air) will move away from the normal.

Total internal reflection

To consider the propagation of light within an optical fiber utilizing the ray theory model it is necessary to take account of the refractive index of the dielectric medium. Optical materials are characterized by their index of refraction, referred to as n. The refractive index of a medium is defined as the ratio of the velocity of light in a vacuum to the velocity of light in the medium.

When a beam of light passes from one material to another with a different index of refraction, the beam is bent (or refracted) at the interface.

$$n_I \sin I = n_R \sin R$$

where nI and nR are the indices of refraction of the materials through which the beam is refracted and I and R are the angles of incidence and refraction of the beam. If the angle of incidence is greater than the critical angle for the interface (typically about 82° for optical fibers), the light is reflected back into the incident medium without loss by a process known as total internal reflection .

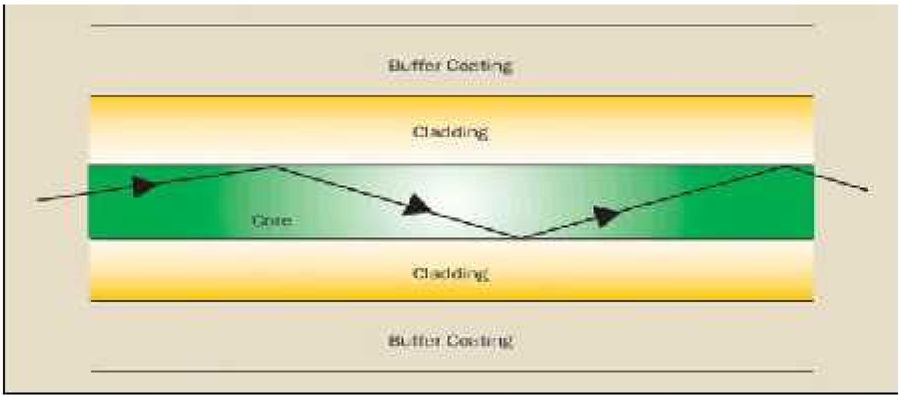


Figure Total Internal Reflection allows light to remain inside the core of the fiber

Refraction is described by Snell's law:

A ray of light travels more slowly in an optically dense medium than in one that is less dense, and the refractive index gives a measure of this effect. When a ray is incident on the interface between two dielectrics of differing refractive indices (e.g. glass–air), refraction occurs, as illustrated in Figure . It may be observed that the ray approaching the interface is propagating in a dielectric of refractive index *n* and is at an angle φ to the normal at the surface of the interface.

If the dielectric on the other side of the interface has a refractive index *n* which is less than *n*1, then the refraction is such that the ray path in this lower index medium is at an angle to the normal, where is greater than . The angles of incidence and refraction are related to each other and to the refractive indices of the dielectrics by Snell's law of refraction, which states that:

$$n_1 \sin \phi_1 = n_2 \sin \phi_2$$

Or

$$\frac{\sin \phi_1}{\sin \phi_2} = \frac{n_2}{n_1}$$

It may also be observed in Figure that a small amount of light is reflected back into the originating dielectric medium (partial internal reflection). As *n* is greater than *n*, the angle of refraction is always greater than the angle of incidence. Thus when the angle of refraction is 90° and the refracted ray emerges parallel to the interface between the dielectrics, the angle of incidence must be less than 90°.

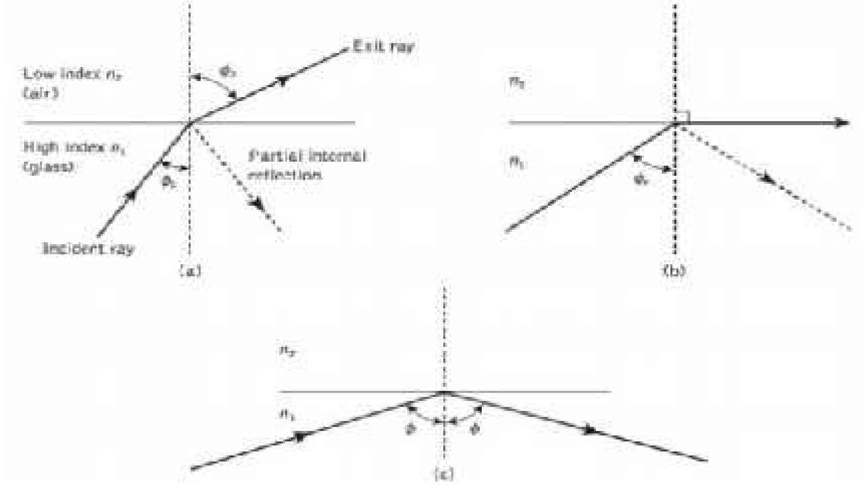


Figure Light rays incident on a high to low refractive index

This is the limiting case of refraction and the angle of incidence is now known as the critical angle φ_c, as shown in Figure. The value of the critical angle is given by

$$\sin \varphi_c = \frac{n_2}{n_1}$$

At angles of incidence greater than the critical angle the light is reflected back into the originating dielectric medium (total internal reflection) with high efficiency (around 99.9%). Hence, it may be observed in Figure that total internal reflection occurs at the inter- face between two dielectrics of differing refractive indices when light is incident on the dielectric of lower index from the dielectric of higher index, and the angle of incidence of the ray exceeds the critical value. This is the mechanism by which light at a sufficiently shallow angle (less than 90°) may be considered to propagate down an optical fiber with low loss.

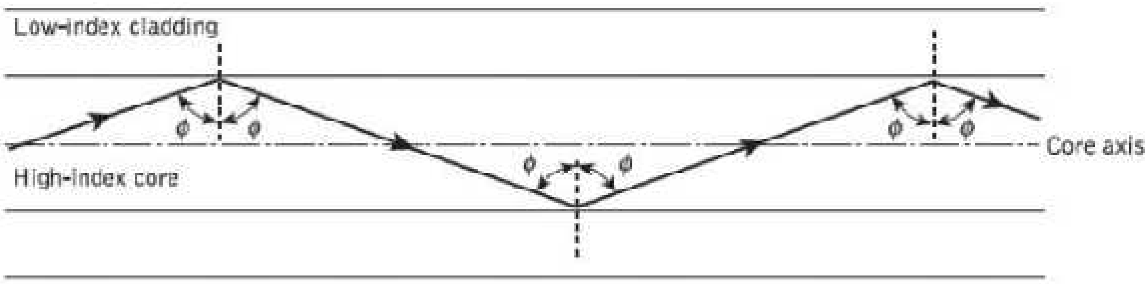


Figure Transmission of a light ray in a perfect optical fiber

The above figure illustrates the transmission of a light ray in an optical fiber via a series of total internal reflections at the interface of the silica core and the slightly lower refractive index silica cladding. The ray has an angle of incidence φ at the interface which is greater than the critical angle and is reflected at the same angle to the normal. The light ray shown in Figure is known as a meridional ray as it passes through the axis of the fiber core. This type of ray is the simplest to describe and is generally used when illustrating the fundamental transmission properties of optical fibers.

It must also be noted that the light transmission illustrated in Figure assumes a perfect fiber, and that any discontinuities or imperfections at the core–cladding interface would probably result in refraction rather than total internal reflection, with the subsequent loss of the light ray into the cladding.

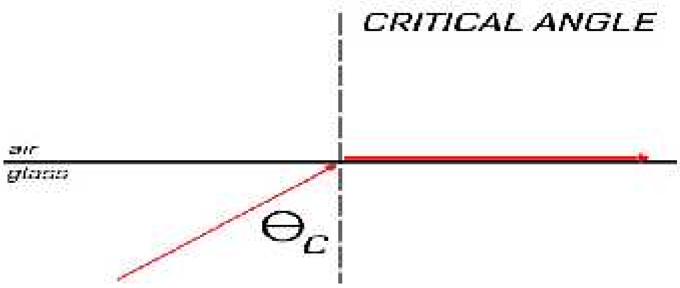
Critical Angle

When the angle of incidence is progressively increased, there will be progressive increase of refractive angle. At some condition the refractive angle becomes 90^{o} to the normal. When this happens the refracted light ray travels along the interface. The angle of incidence at the point at which the refractive angle becomes 90^{o} is called the critical angle. The critical angle is defined as the minimum angle of incidence at which the ray strikes the interface of two media and causes an angle of refraction equal to 90^{o}. Figure shows critical angle refraction. When the angle of refraction is 90 degree to the normal the refracted ray is parallel to the interface between the two media. Using Snell's law

$$\sin \phi_c = \frac{n_2}{n_1} \sin 90^o$$

$$\because \quad \sin 90^o = 1$$

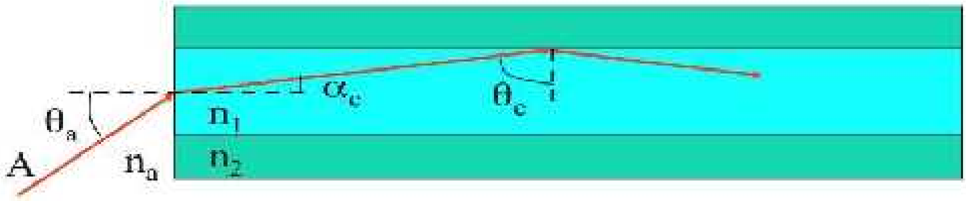
$$\text{Critical angle } \phi_c = \sin^{-1}\left(\frac{n_2}{n_1}\right)$$



It is important to know about this property because reflection is also possible even if the surfaces are not reflective. If the angle of incidence is greater than the critical angle for a given setting, the resulting type of reflection is called Total Internal Reflection, and it is the basis of Optical Fiber Communication.

Acceptance angle

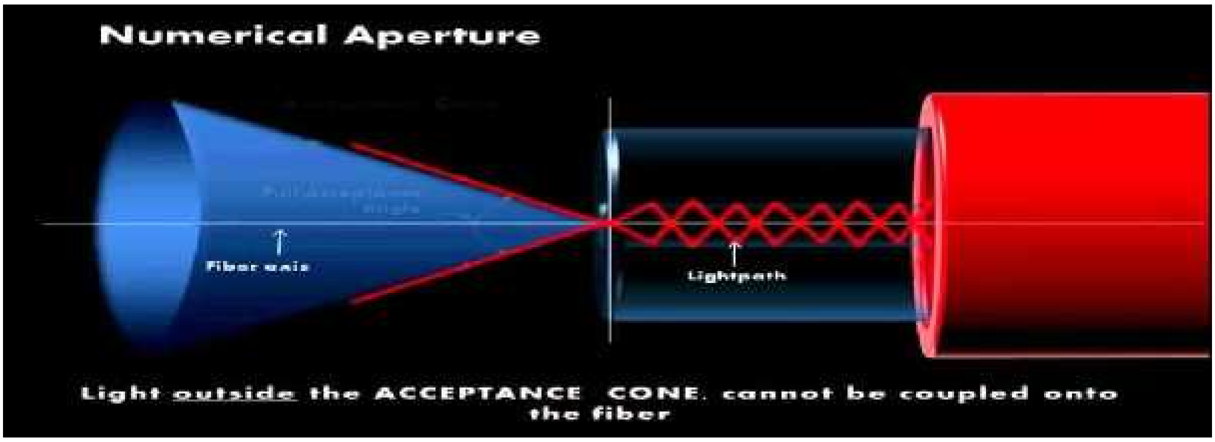
In an optical fiber, a light ray undergoes its first refraction at the air-core interface. The angle at which this refraction occurs is crucial because this particular angle will dictate whether the subsequent internal reflections will follow the principle of Total Internal Reflection. This angle, at which the light ray first encounters the core of an optical fiber is called Acceptance angle.

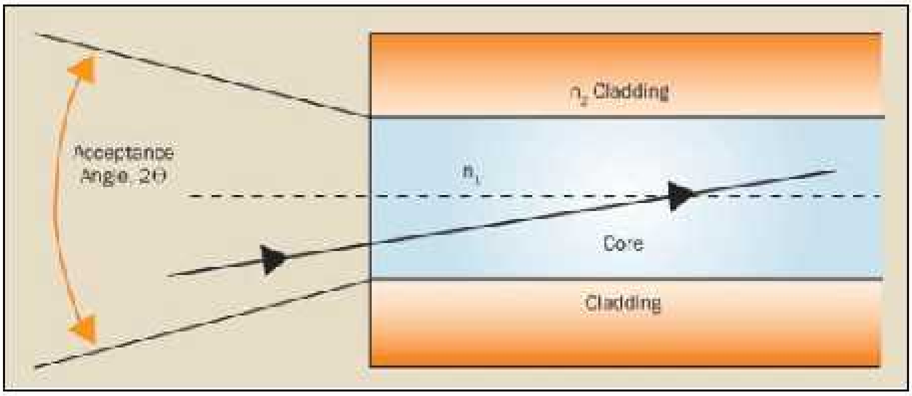


The objective is to have θ_c greater than the critical angle for this particular setting. As you can notice, θ_c depends on the orientation of the refracted ray at the input of the optical fiber. This in turn depends on θ_a, the acceptance angle. The acceptance angle can be calculated with the help of the formula below.

Numerical Aperture

Numerical Aperture is a characteristic of any optical system. For example, photo-detector, optical fiber, lenses etc. are all optical systems. Numerical aperture is the ability of the optical system to collect the entire light incident on it, in one area. The blue cone is known as the cone of acceptance. As you can see it is dependent on the Acceptance Angle of the optical fiber. Light waves within the acceptance cone can be collected in a small area which can then be sent into the optical fiber (Source).

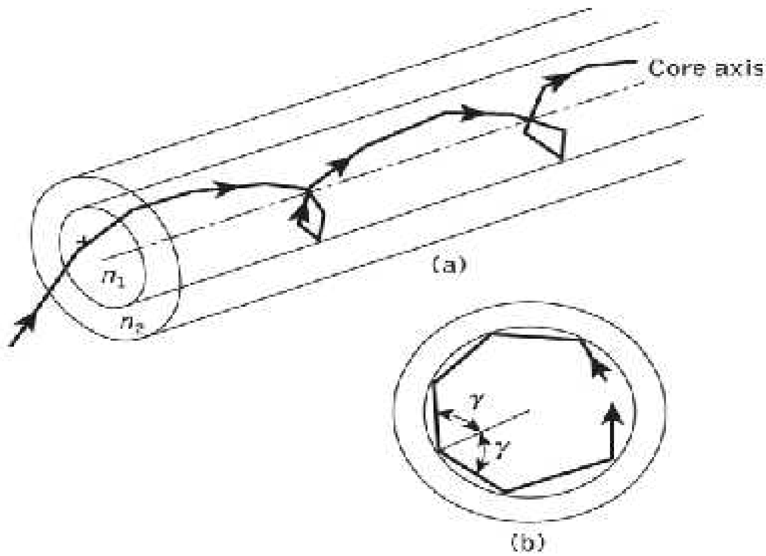




Numerical aperture (NA), shown in above Figure, is the measure of maximum angle at which light rays will enter and be conducted down the fiber. This is represented by the following equation:

$$NA = \sqrt{(n^2_{core} - n^2_{cladding})} = \sin\theta$$

Skew rays: In a multimode optical fiber, a bound ray that travels in a helical path along the fiber and thus (a) is not parallel to the fiber axis, (b) does not lie in a meridional plane, and (c) does not intersect the fiber axis is known as a Skew Ray.



Figure, view (a), provides an angled view and view (b) provides a front view.

1. Skew rays are rays that travel through an optical fiber without passing through its axis.

2. A possible path of propagation of skew rays is shown in figure.

3. Skew rays are those rays which follow helical path but they are not confined to a single plane. Skew rays are not confined to a particular plane so they cannot be tracked easily. Analyzing the meridional rays is sufficient for the purpose of result, rather than skew rays, because skew rays lead to greater power loss.

4. Skew rays propagate without passing through the center axis of the fiber. The acceptance angle for skew rays is larger than the acceptance angle of meridional rays.

5. Skew rays are often used in the calculation of light acceptance in an optical fiber. The addition of skew rays increases the amount of light capacity of a fiber. In large NA fibers, the increase may be significant.

6. The addition of skew rays also increases the amount of loss in a fiber. Skew rays tend to propagate near the edge of the fiber core. A large portion of the number of skew rays that are trapped in the fiber core are considered to be leaky rays.

7. Leaky rays are predicted to be totally reflected at the core-cladding boundary. However, these rays are partially refracted because of the curved nature of the fiber boundary. Mode theory is also used to describe this type of leaky ray loss.

Cylindrical fiber

1. Modes

When light is guided down a fiber (as microwaves are guided down a waveguide), phase shifts occur at every reflective boundary. There is a finite discrete number of paths down the optical fiber (known as modes) that produce constructive (in phase and therefore additive) phase shifts that reinforce the transmission. Because each mode occurs at a different angle to the fiber axis as the beam travels along the length, each one travels a different length through the fiber from the input to the output. Only one mode, the zero-order mode, travels the length of the fiber without reflections from the sidewalls. This is known as a single-mode fiber. The actual number of modes that can be propagated in a given optical fiber is determined by the wavelength of light and the diameter and index of refraction of the core of the fiber.

The exact solution of Maxwell's equations for a cylindrical homogeneous core dielectric waveguide* involves much algebra and yields a complex result. Although the presentation of this mathematics is beyond the scope of this text, it is useful to consider the resulting modal fields. In common with the planar guide TE (where $E_z = 0$) and TM (where $H_z = 0$) modes are obtained within the dielectric cylinder. The cylindrical waveguide, however, is bounded in two dimensions rather than one. Thus two integers, l and m, are necessary in order to specify the modes, in contrast to the single integer (m) required for the planar guide.

For the cylindrical waveguide, therefore refer to TE_{lm} and TM_{lm} modes. These modes correspond to meridional rays traveling within the fiber. However, hybrid modes where E_z and H_z are nonzero also occur within the cylindrical waveguide.

These modes, which result from skew ray propagation within the fiber, are designated HE_{lm} and EH_{lm} depending upon whether the components of H or E make the larger contribution to the transverse (to the fiber axis) field. Thus an exact description of the modal fields in a step index fiber proves somewhat complicated.

Fortunately, the analysis may be simplified when considering optical fibers for communication purposes. These fibers satisfy the weakly guiding approximation where the relative index difference Δ1. This corresponds to small grazing angles θ. In fact is usually less than 0.03 (3%) for optical communications fibers. For weakly guiding structures with dominant forward propagation, mode theory gives dominant transverse field components. Hence approximate solutions for the full set of HE, EH, TE and TM modes may be given by two linearly polarized components.

These linearly polarized (LP) modes are not exact modes of the fiber except for the fundamental (lowest order) mode. However, as in weakly guiding fibers is very small, then HE– EH mode pairs occur which have almost identical propagation constants. Such modes are said to be degenerate. The superposition of these degenerating modes characterized by a common propagation constant correspond to particular LP modes regardless of their HE, EH, TE or TM field configurations. This linear combination of degenerate modes obtained from the exact solution produces a useful simplification in the analysis of weakly guiding fibers.

The relationship between the traditional HE, EH, TE and TM mode designations and the LPlm mode designations is shown in Table. The mode subscripts l and m are related to the electric field intensity profile for a particular LP mode. There are in general 2l field maxima around the circumference of the fiber core and m field maxima along a radius vector. Furthermore, it may be observed from Table 1.1 that the notation for labeling the HE and EH modes has changed from that specified for the exact solution in the cylindrical waveguide mentioned previously.

Table 1.1 Correspondence between the lower order in linearly polarized modes and the traditional exact modes from which they are formed

Linearly polarized	*Exact*
LP_{01}	HE_{11}
LP_{11}	HE_{21}, TE_{01}, TM_{01}
LP_{21}	HE_{31}, EH_{11}
LP_{02}	HE_{12}
LP_{31}	HE_{41}, EH_{21}
LP_{12}	HE_{22}, TE_{02}, TM_{02}
LP_{lm}	HE_{2m}, TE_{0m}, TM_{0m}
LP_{lm} ($l \neq 0$ or 1)	$HE_{l+1,m}$, $EH_{l-1,m}$

2. Mode coupling

Thus, so far the propagation aspects of perfect dielectric waveguides were considered. However, waveguide perturbations such as deviations of the fiber axis from straightness, variations in the core diameter, irregularities at the core–cladding interface and refractive index variations may change the propagation characteristics of the fiber. These will have the effect of coupling energy traveling in one mode to another depending on the specific perturbation. Ray theory aids the understanding of this phenomenon, as shown in Figure which illustrates two types of perturbation. It may be observed that in both cases the ray no longer maintains the same angle with the axis. In electromagnetic wave theory this corresponds to a change in the propagating mode for the light. Thus individual modes do not normally propagate throughout the length of the fiber without large energy transfers to adjacent modes, even when the fiber is exceptionally good quality and is not strained or bent by its surroundings. This mode conversion is known as mode coupling or mixing. It is usually analyzed using coupled mode equations which can be obtained directly from Maxwell's equations.

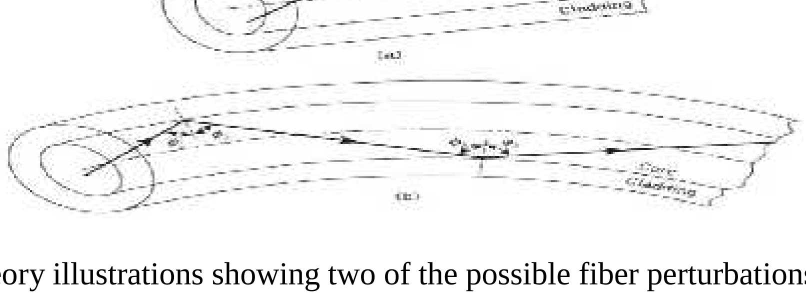


Figure Ray theory illustrations showing two of the possible fiber perturbations which give mode coupling: (a) irregularity at the core–cladding interface; (b) fiber bend

3. Step index fibers

The optical fiber considered in the preceding sections with a core of constant refractive index n_1 and a cladding of a slightly lower refractive index $n_2$is known as step index fiber. This is because the refractive index profile for this type of fiber makes a step change at the core–cladding interface, as indicated in Figure which illustrates the two major types of step index fiber.

Figure shows a multimode step index fiber with a core diameter of around 50µm or greater, which is large enough to allow the propagation of many modes within the fiber core. This is illustrated in Figure by the many different possible ray paths through the fiber. Figure shows a single-mode or monomode step index fiber which allows the propagation of only one transverse electromagnetic mode (typically HE11), and hence the core diameter must be of the order of 2 to 10µm. The propagation of a single mode is illustrated in Figure as corresponding to a single ray path only (usually shown as the axial ray) through the fiber. The single-mode step index fiber has the distinct advantage of low intermodal dispersion (broadening of transmitted light pulses), as only one mode is transmitted, whereas with multimode step index fiber considerable dispersion may occur due to the differing group velocities of the propagating modes. This in turn restricts the maximum bandwidth attainable with multimode step index fibers, especially when com- pared with single-mode fibers.

The refractive index profile may be defined as

$$n(r) = \begin{cases} n_1 & r < a \quad \text{(core)} \\ n_2 & r \geq a \quad \text{(cladding)} \end{cases}$$

in both cases

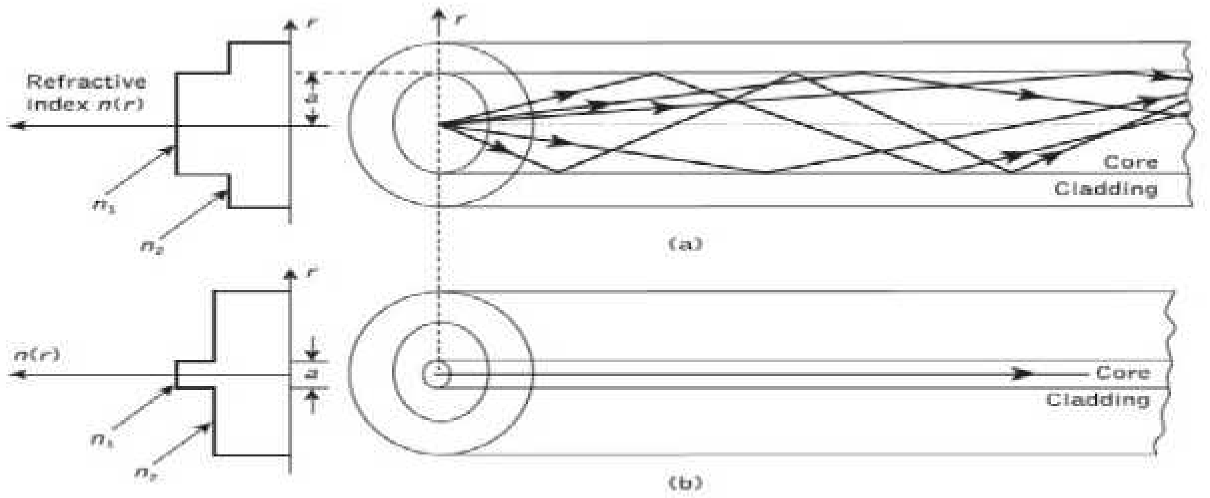


Figure Refractive index profile and ray transmission in step index a) multimode b) single mode

However, for lower bandwidth applications multimode fibers have several advantages over single- mode fibers. These are:

a) The use of spatially incoherent optical sources (e.g. most light-emitting diodes) which cannot be efficiently coupled to single-mode fibers.

b) Larger numerical apertures, as well as core diameters, facilitating easier coupling to optical sources

c) Lower tolerance requirements on fiber connectors

Multimode step index fibers allow the propagation of a finite number of guided modes along the channel. The number of guided modes is dependent upon the physical parameters (i.e. relative refractive index difference, core radius) of the fiber and the wavelengths of the transmitted light which are included in the normalized frequency *V* for the fiber.

Mode propagation does not entirely cease below cutoff. Modes may propagate as unguided or leaky modes which can travel considerable distances along the fiber. Nevertheless, it is the guided modes which are of paramount importance in optical fiber communications as these are confined to the fiber over its full length. The total number of guided modes or mode volume M_s for a step index fiber is related to the *V* value for the fiber by the approximate expression that allows an estimate of the number of guided modes propagating in a particular multimode step index fiber.

4. Graded index fibers

Graded index fibers do not have a constant refractive index in the core* but a decreasing core index $n(r)$ with radial distance from a maximum value of$n1$ at the axis to a constant value n_2 beyond the core radius *a* in the cladding. This index variation may be represented as:

$$n(r) = \begin{cases} n_1\left(1 - 2\Delta\left(\frac{r}{a}\right)^{\alpha}\right) & \text{when } r < a \ (core) \\ n_1(1-2\Delta)^{\frac{1}{2}} \approx n_2 & when\ r \geq a \ (\text{cladding}) \end{cases}$$

where is the relative refractive index difference and α is the profile parameter which gives the characteristic refractive index profile of the fiber core. Equation which is a convenient method of expressing the refractive index profile of the fiber core as a variation of α, allows representation of the step index profile when α = ∞, a parabolic profile when α = 2 and a triangular profile when α = 1. This range of refractive index profiles is illustrated in Figure. The graded index profiles which at present produce the best results for multimode optical propagation have a near parabolic refractive index profile core with ~~2. Fibers with such core index profiles are well established and consequently when the term ‘graded index’ is used without qualification it usually refers to a fiber with this profile.

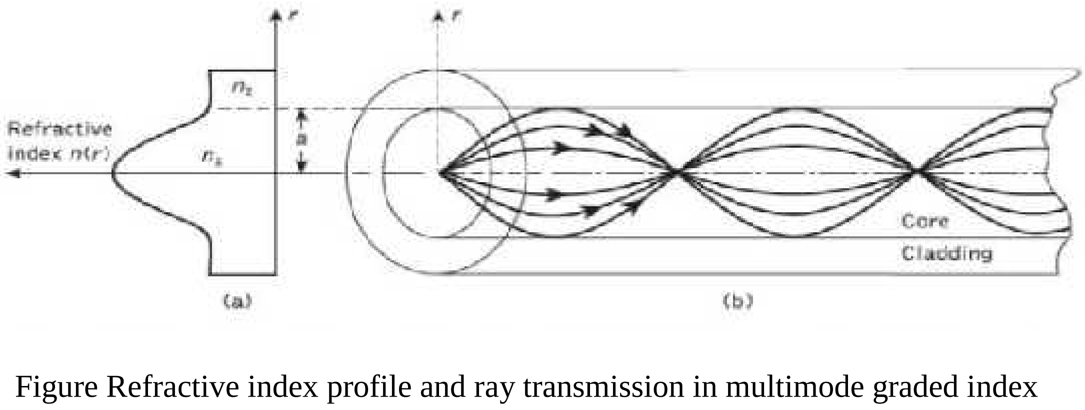


Figure Refractive index profile and ray transmission in multimode graded index

Where, r = Radial distance from fiber axis, a = Core radius, n_1= Refractive index of core, n_2 = Refractive index of cladding, α = Shape of index profile.

Profile parameter α determines the characteristic refractive index profile of fiber core. For this reason in this section, consider the waveguiding properties of graded index fiber with a parabolic refractive index profile core. A multimode graded index fiber with a parabolic index profile core is illustrated in Figure. It may be observed that the meridional rays shown appear to follow curved paths through the fiber core. Using the concepts of geometric optics, the gradual decrease in refractive index from the center of the core creates many refractions of the rays as they are effectively incident on a large number or high to low index interfaces. This mechanism is illustrated in Figure where a ray is shown to be gradually curved, with an ever- increasing angle of incidence, until the conditions for total internal reflection are met, and the ray travels back towards the core axis, again being continuously refracted.

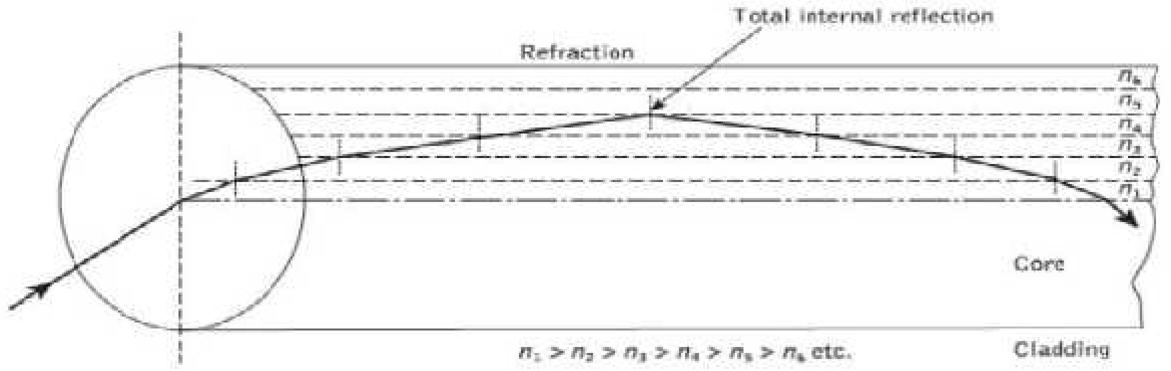


Figure An expanded ray diagram showing refraction

Multimode graded index fibers exhibit far less intermodal dispersion than multimode step index fibers due to their refractive index profile. Although many different modes are excited in the graded index fiber, the different group velocities of the modes tend to be normalized by the index grading. Again considering ray theory, the rays traveling close to the fiber axis have shorter paths when compared with rays which travel.

However, the near axial rays are transmitted through a region of higher refractive index and therefore travel with a lower velocity than the more extreme rays. This compensates for the shorter path lengths and reduces dispersion in the fiber. A similar situation exists for skew rays which follow longer helical paths, as illustrated in Figure. These travel for the most part in the lower index region at greater speeds, thus giving the same mechanism of mode transit time equalization. Hence, multi- mode graded index fibers with parabolic or near-parabolic index profile cores have trans- mission bandwidths which may be orders of magnitude greater than multimode step index fiber bandwidths.

Consequently, although they are not capable of the bandwidths attain- able with single- mode fibers, such multimode graded index fibers have the advantage of large core diameters (greater than 30 μm) coupled with bandwidths suitable for long- distance communication. The parameters defined for step index fibers (i.e. NA, Δ, V) may be applied to graded index fibers and give a comparison between the two fiber types.

However, it must be noted that for graded index fibers the situation is more complicated since the numerical aperture is a function of the radial distance from the fiber axis. Graded index fibers, therefore, accept less light than corresponding step index fibers with the same relative refractive index difference.

Single-mode fiber

The advantage of the propagation of a single mode within an optical fiber is that the signal dispersion caused by the delay differences between different modes in a multimode fiber may be avoided. Multimode step index fibers do not lend themselves to the propagation of a single mode due to the difficulties of maintaining single-mode operation within the fiber when mode conversion (i.e. coupling) to other guided modes takes place at both input mismatches and fiber imperfections. Hence, for the transmission of a single mode the fiber must be designed to allow propagation of only one mode, while all other modes are attenuated by leakage or absorption. Following the preceding discussion of multimode fibers, this may be achieved through choice of a suitable normalized frequency for the fiber. For single-mode operation, only the fundamental LP01 mode can exist. Hence the limit of single-mode operation depends on the lower limit of guided propagation for the LP11 mode. The cutoff normalized frequency for the LP11 mode in step index fibers occurs at V_c = 2.405. Thus single-mode propagation of the LP01 mode in step index fibers is possible over the range:

$$0 \leq V < 2.405$$

As there is no cutoff for the fundamental mode. It must be noted that there are in fact two modes with orthogonal polarization over this range, and the term single-mode applies to propagation of light of a particular polarization. Also, it is apparent that the normalized frequency for the fiber may be adjusted to within the range given in Equation by reduction of the core radius.

1. Cutoff wavelength

It may be noted that single-mode operation only occurs above a theoretical cutoff wavelength λ_c given by:

$$\lambda_c = \frac{2\pi a n_1}{V_c}(2\Delta)^{\frac{1}{2}}$$

Where V_c- Cut off normalized frequency.

Dividing above equation by

$$V = \frac{2\pi}{\lambda} a n_1 (2\Delta)^{\frac{1}{2}}$$

$$\frac{\lambda_c}{\lambda} = \frac{V}{V_c}$$

Thus for step index fiber where V_c=2.405, the cut-off wavelength is given by

$$\lambda_c = \frac{V\lambda}{2.405}$$

An effective cutoff wavelength has been defined by the ITU-T which is obtained from a 2 m length of fiber containing a single 14 cm radius loop. This definition was produced because the first higher order LP11 mode is strongly affected by fiber length and curvature near cutoff. Recommended cutoff wavelength values for primary coated fiber range from 1.1 to 1.28 μm for single-mode fiber designed for operation in the 1.3μm wavelength region in order to avoid modal noise and dispersion problems. Moreover, practical transmission systems are generally operated close to the effective cutoff wave- length in order to enhance the fundamental mode confinement, but sufficiently distant from cutoff so that no power is transmitted in the second-order LP11 mode.

2. Mode-field diameter and spot size

Many properties of the fundamental mode are determined by the radial extent of its electromagnetic field including losses at launching and jointing, micro bend losses, waveguide dispersion and the width of the radiation pattern. Therefore, the MFD is an important parameter for characterizing single-mode fiber properties which takes into account the wavelength-dependent field penetration into the fiber cladding. In this context it is a better measure of the functional properties of single- mode fiber than the core diameter. For step index and graded (near parabolic profile) single-mode fibers operating near the cutoff wavelength λ_c, the field is well approximated by a Gaussian distribution. In this case the MFD is generally taken as the distance between the opposite 1/e = 0.37 field amplitude points and the power 1/e2 = 0.135 points in relation to the corresponding values on the fiber axis. Another parameter which is directly related to the MFD of a single-mode fiber is the spot size (or mode-field radius) ω0. Hence MFD = $2\omega_0$, where ω_0 is the nominal half width of the input excitation.

The MFD can therefore be regarded as the single- mode analog of the fiber core diameter in multimode fibers. However, for many refractive index profiles and at typical operating wavelengths the MFD is slightly larger than the single-mode fiber core diameter. Often, for real fibers and those with arbitrary refractive index profiles, the radial field distribution is not strictly Gaussian and hence alternative techniques have been proposed. However, the problem of defining the MFD and spot size for non-Gaussian field distributions is difficult one and at least eight definitions exist**.**

3. Effective refractive index

The rate of change of phase of the fundamental LP01 mode propagating along a straight fiber is determined by the phase propagation constant. It is directly related to the wavelength of the LP01 mode λ_{01} by the factor 2π, since β gives the increase in phase angle per unit length. Hence:

$$\beta\lambda_{01} = 2\pi \quad \text{or} \quad \lambda_{01} = \frac{2\pi}{\beta}$$

Morever, it is convenient to define an effective refractive index for single mode fiber, sometimes referred to as a phase index or normalized phase change coefficient n_{eff} by the ratio of the propagation constant of the fundamental mode to that of the vaccum propagation constant.

$$n_{eff} = \frac{\beta}{k}$$

Hence, the wavelength of the fundamental mode is smaller than the vaccum wave by the factor 1/ n_{eff} ,where

$$\lambda_{01} = \frac{\lambda}{n_{eff}}$$

It should be noted that the fundamental mode propagates in a medium with a refractive index n(r) which is dependent on the distance r from the fiber axis. The effective refractive index can therefore be considered as an average over the refractive index of this medium. Within a normally clad fiber, not depressed-cladded fibers, at long wavelengths (i.e. small *V* values) the MFD is large compared to the core diameter and hence the electric field extends far into the cladding region. In this case the propagation constant β will be approximately equal to *n*2*k* (i.e. the cladding wave number) and the effective index will be similar to the refractive index of the cladding n_2. Physically, most of the power is transmitted in the cladding material. At short wavelengths, however, the field is concentrated in the core region and the propagation constant β approximates to the maximum wave number *nlk*. Following this discussion, and as indicated previously, then the propagation constant in single-mode fiber varies over the interval $n_2k < \beta < n_1k$. Hence, the effective refractive index will vary over the range $n_2 < n\mathrm{eff} < n_1$.

4. Group delay and mode delay factor

The transit time or group delay τ_g for a light pulse propagating along a unit length of fiber is the inverse of the group velocity, υ_g

$$\tau_g = \frac{1}{\upsilon_g} = \frac{d\beta}{d\omega} = \frac{1}{c}\frac{d\beta}{dk}$$

The group index of a uniform plane wave propagating in a homogenous medium has been identified as

$$N_g = \frac{c}{\upsilon_g}$$

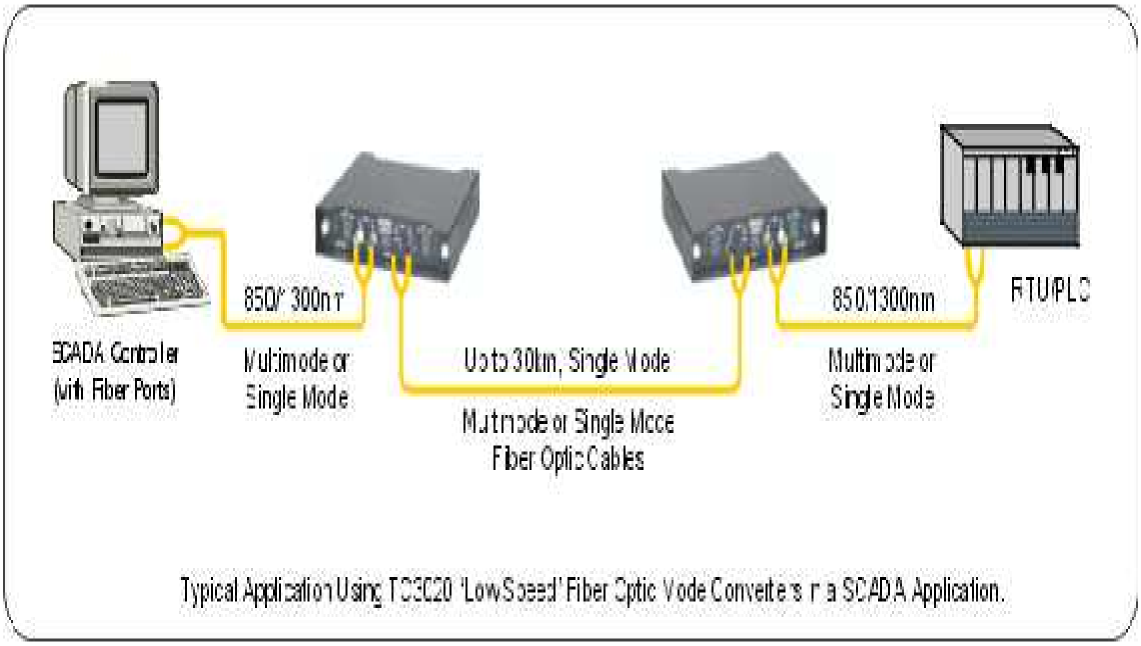
However, for a single mode fiber, it is usual to define an effective group index by

$$N_{ge} = \frac{c}{\upsilon_g}$$

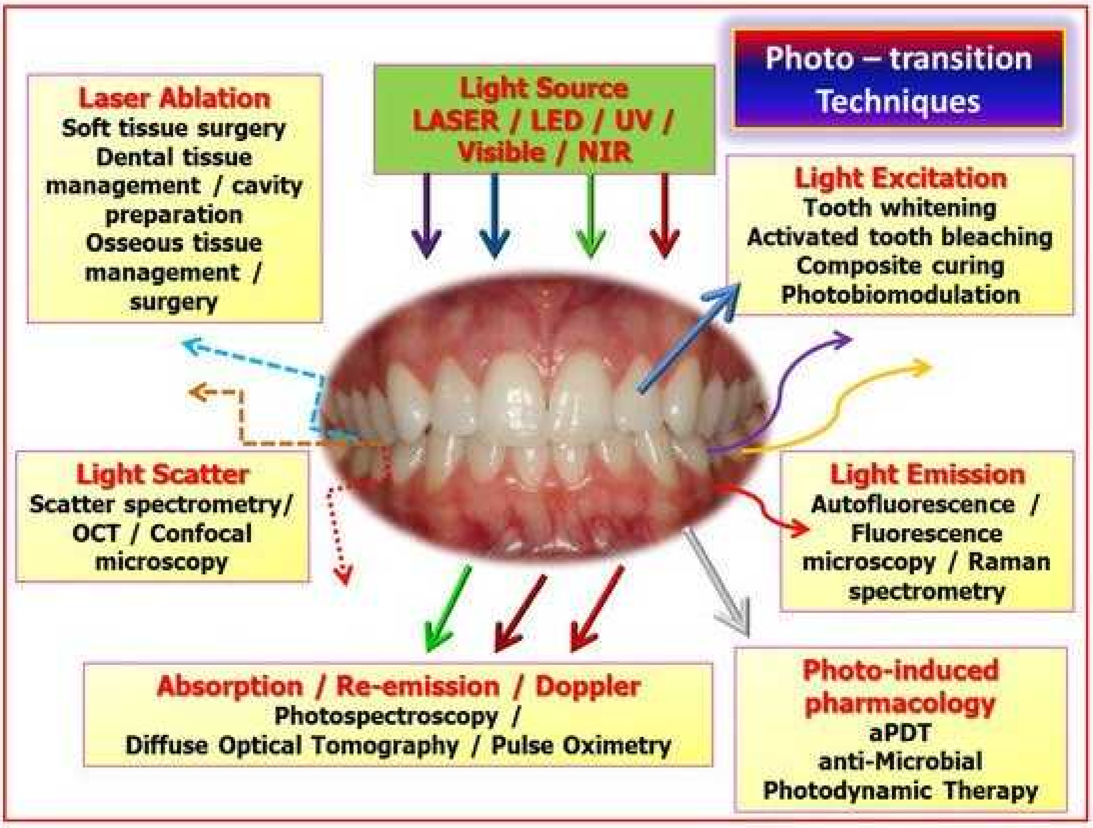
Hence, where υ_g is considered to be the group velocity of the fundamental fiber mode. Hence, the specific group delay of the fundamental fiber mode becomes:

$$\tau_g = \frac{N_{ge}}{c}$$

APPLICATIONS



Examples of Typical Application of Fiber Optic Mode in SCADA application



Examples of Application of optical fiber in Dentistry

POST TEST-MCQ TYPE

1. What is refraction?
a) Bending of light waves
b) Reflection of light waves
c) Diffusion of light waves
d) Scattering of light waves

2. The phenomenon which occurs when an incident wave strikes an interface at an angle greater than the critical angle with respect to the normal to the surface is called as
a) Refraction
b) Partial internal reflection
c) Total internal reflection
d) Limiting case of refraction

3. A monochromatic wave propagates along a waveguide in z direction. These points of constant phase travel in constant phase travel at a phase velocity Vp is given by?
a) Vp=ω/β
b) Vp=ω/c
c) Vp=C/N
d) Vp=mass/acceleration

4. Which law gives the relationship between refractive index of the dielectric?
a) Law of reflection
b) Law of refraction (Snell's Law)
c) Millman's Law
d) Huygen's Law

5. The light sources used in fibre optics communication are
a) LED's and Lasers
b) Phototransistors
c) Xenon lights
d) Incandescent

6. Which ray passes through the axis of the fiber core?
a) Reflected
b) Refracted
c) Meridional
d) Skew

7. Light incident on fibers of angles___the acceptance angle do not propagate into the fiber.
a) Less than
b) Greater than
c) Equal to
d) Less than and equal to

8. The ratio of speed of light in air to the speed of light in another medium is called as
a) Speed factor
b) Dielectric constant
c) Reflection index
d) Refraction index

9. When a ray of light enters one medium from another medium, which quality will not change?
a) Direction
b) Frequency
c) Speed
d) Wavelength

10. What is the numerical aperture of the fiber if the angle of acceptance is 16 degree?
a) 0.50
b) 0.36
c) 0.20
d) 0.27

11. For lower bandwidth applications
a) Single mode fiber is advantageous
b) Photonic crystal fibers are advantageous
c) Coaxial cables are advantageous
d) Multimode fiber is advantageous

12. Meridional rays in graded index fibers follow
a) Straight path along the axis
b) Curved path along the axis
c) Path where rays changes angles at core-cladding interface
d) Helical path

13. Skew rays follow a
a) Hyperbolic path along the axis
b) Parabolic path along the axis
c) Helical path
d) Path where rays changes angles at core-cladding interface

14. What is needed to predict the performance characteristics of single mode fibers?
a) The intermodal delay effect
b) Geometric distribution of light in a propagating mode
c) Fractional power flow in the cladding of fiber
d) Normalized frequency

15. Which equation is used to calculate MFD?
a) Maxwell's equations
b) Peterman equations
c) Allen Cahn equations
d) Boltzmann's equations

16. he difference between the modes' refractive indices is called as
a) Polarization
b) Cutoff
c) Fiber birefringence
d) Fiber splicing

17. How many propagation modes are present in single mode fibers?
a) One
b) Two
c) Three
d) Five

18. A device that reduces the intensity of light in optical fiber communications is
a) Compressor
b) Optical attenuator
c) Barometer
d) Reducer

19. The core of an optical fiber has a
a) Lower refracted index than air
b) Lower refractive index than the cladding
c) Higher refractive index than the cladding
d) Similar refractive index with the cladding

20. One of the following materials is sensitive to light. Identify it.
a) Photoresist
b) Photosensitive
c) Light Sensitive
d) Maser

21. If a mirror is used to reflect light, the reflected light angle is ____ as the incident angle
a) Smaller
b) Larger
c) The same
d) Independent

22. This is not a part of the optical spectrum. Identify it.
a) infrared
b) ultraviolet
c) visible color
d) x-rays

23. Which type of fiber has the highest modal dispersion.
a) Step-index multimode
b) Graded index multimode
c) Step-index single mode
d) Graded index mode

24. What is a specific path the light takes in an optical fiber corresponding to a certain angle and number of reflection?
a) Mode
b) Grade
c) Numerical Aperture
d) Dispersion

CONCLUSION

In this unit, an understanding of optical fiber communication link, structure, propagation and transmission properties of an optical fiber and the Estimate the losses and analyze the propagation characteristics of an optical signal in different types of fibers was done. The Introduction to Optical Fibers, its types and applications were discussed.

REFERENCES

1. Gerd Keiser, "Optical Fiber Communication" McGraw-Hill International,4th Edition. 2010.
2. John M. Senior, "Optical Fiber Communication", Second Edition, Pearson Education, 2007.
3. Ramaswami, Sivarajan and Sasaki "Optical Networks", Morgan Kaufmann, 2009.
4. J.Senior, Optical Communication, Principles and Practice, Prentice Hall of India, 3rd Edition, 2008.
5. J.Gower, "Optical Communication System", Prentice Hall of India, 2001.

ASSIGNMENT

1. Describe the Ray theory transmission.
2. Describe Single-mode fiber and its mode field diameter.
3. Describe in detail the Classification of fibers or Compare the structure and characteristics of step index and graded index fiber structures.
4. Derive the expression for linearly polarized modes in optical fibers and obtain the expression for normalized frequency.
5. A step index multimode fiber with a numerical aperture of 0.2 support approximately 1000 modes at 850 nm wavelength. What is the diameter of its core? How many modes does the fiber support at 850 nm and 1550 nm.
6. Draw the block diagram of optical fiber transmission link and explain.

AIM & OBJECTIVES

- ❖ To learn the basic elements of optical fiber transmission link, fiber modes configurations and structures.
- ❖ To understand the different kind of losses, signal distortion, SM fibers.
- ❖ To learn the various optical sources, materials and fiber splicing.
- ❖ To learn the fiber optical receivers and noise performance in photo detector.

PRE TEST-MCQ TYPE

1. What does ISI stand for in optical fiber communication?
a) Invisible size interference
b) Infrared size interference
c) Inter-symbol interference
d) Inter-shape interference

2. 3dB optical bandwidth is always ___________ the 3dB electrical bandwidth.
a) Smaller than
b) Larger than
c) Negligible than
d) Equal to

3. In waveguide dispersion, refractive index is independent of
a) Bit rate
b) Index difference
c) Velocity of medium
d) Wavelength

4. After Total Internal Reflection the Meridional ray
a) Makes an angle equal to acceptance angle with the axial ray
b) Makes an angle equal to critical angle with the axial ray
c) Travels parallel equal to critical angle with the axial ray
d) Makes an angle equal to critical angle with the axial ray

5. How many mechanisms are there which causes absorption?
a) One
b) Three
c) Two
d) Four

UNIT II SIGNAL DEGRADATION OPTICAL FIBERS

Attenuation - Absorption losses, Scattering losses, Bending Losses, Core and Cladding losses, Signal Distortion in Optical Waveguides-Information Capacity determination -Group Delay-Material Dispersion, Wave guide Dispersion, Signal distortion in SM fibers-Polarization Mode dispersion, Intermodal dispersion, Pulse Broadening in GI fibers-Mode Coupling -Design Optimization of SM fibers- RI profile and cut-off wavelength

THEORY

Introduction

One of the important property of optical fiber is signal attenuation. It is also known as fiber loss or signal loss. The signal attenuation of fiber determines the maximum distance between transmitter and receiver. The attenuation also determines the number of repeaters required, maintaining repeater is a costly affair. Another important property of optical fiber is distortion mechanism. As the signal pulse travels along the fiber length it becomes more broader. After sufficient length the broad pulses starts overlapping with adjacent pulses. This creates error in the receiver. Hence the distortion limits the information carrying capacity of fiber.

Attenuation

Attenuation is a measure of decay of signal strength or loss of light power that occurs as light pulses propagate through the length of the fiber. In optical fibers the attenuation is mainly caused by two physical factors absorption and scattering losses. Absorption is because of fiber material and scattering due to structural imperfection within the fiber. Nearly 90% of total attenuation is caused by Rayleigh scattering only. Microbending of optical fiber also contributes to the attenuation of signal.

The rate at which light is absorbed is dependent on the wavelength of the light and the characteristics of particular glass. Glass is a silicon compound, by adding different additional chemicals to the basic silicon dioxide the optical properties of the glass can be changed.

The Rayleigh scattering is wavelength dependent and reduces rapidly as the wavelength of the incident radiation increases. The attenuation of fiber is governed by the materials from which it is fabricated, the manufacturing process and the refractive index profile chosen. Attenuation loss is measured in dB/km.

Attenuation Units

As attenuation leads to a loss of power along the fiber, the output power is significantly less than the couples power. Let the couples optical power is p(0) i.e. at origin (z = 0). Then the power at distance z is given by,

$$P(z) = P(0)e^{-\alpha_p z}$$

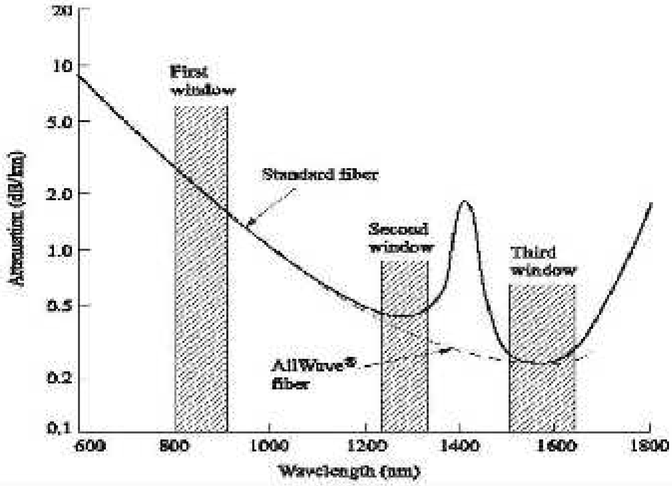
where, α_p is fiber attenuation constant (per km).

$$\alpha_p = \frac{1}{z} \ln \left[\frac{P(0)}{P(z)}\right]$$

$$\alpha_{dB/km} = 10 . \frac{1}{z} \log \left[\frac{P(0)}{P(z)}\right]$$

$$\alpha_{dB/km} = 4.343 \ \alpha_p \ \text{per km}$$

This parameter is known as fiber loss or fiber attenuation. Attenuation is also a function of wavelength. Optical fiber wavelength as a function of wavelength is shown in Figure.



Optical fiber attenuation as a function of wavelength yields nominal values of 0.5 dB/km at 1310 nm and 0.3 dB/km at 1550 nm for standard single mode fiber. Absorption by the water molecules causes the attenuation peak around 1400nm for standard fiber. The dashed curve is the attenuation for low water peak fiber.

Figure Optical fiber wavelength as a function of wavelength

Absorption

Absorption loss is related to the material composition and fabrication process of fiber. Absorption loss results in dissipation of some optical power as hear in the fiber cable. Although glass fibers are extremely pure, some impurities still remain as residue after purification. The amount of absorption by these impurities depends on their concentration and light wavelength.

Absorption in optical fiber is caused by these three mechanisms.

1. Absorption by atomic defects in the glass composition

2. Extrinsic absorption by impurity atoms in the glass material

3. Intrinsic absorption by the basic constituent atoms of the fiber material.

Absorption by Atomic Defects

Atomic defects are imperfections in the atomic structure of the fiber materials such as missing molecules, high density clusters of atom groups. These absorption losses are negligible compared with intrinsic and extrinsic losses.

The absorption effect is most significant when fiber is exposed to ionizing radiation in nuclear reactor, medical therapies, space missions etc. The radiation dames the internal structure of fiber. The damages are proportional to the intensity of ionizing particles. This results in increasing attenuation due to atomic defects and absorbing optical energy. The total dose a material receives is expressed in rad (Si), this is the unit for measuring radiation absorbed in bulk silicon. 1 rad (Si) = 0.01 J.kg. The higher the radiation intensity more the attenuation as shown in Figure

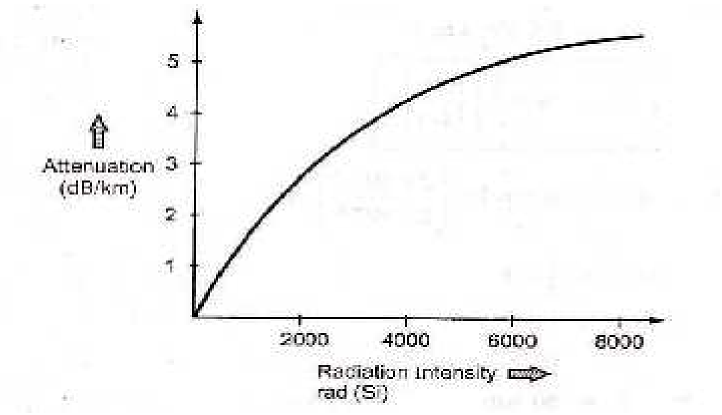


Figure ionizing radiation intensity vs fiber attenuation

Extrinsic Absorption

Extrinsic absorption occurs due to electronic transitions between the energy level and because of charge transitions from one ion to another. A major source of attenuation is from transition of metal impurity ions such as iron, chromium, cobalt and copper. These losses can be upto 1 to 10 dB/km. The effect of metallic impurities can be reduced by glass refining techniques. Another major extrinsic loss is caused by absorption due to OH (Hydroxil) ions impurities dissolved in glass. Vibrations occur at wavelengths between 2.7 and 4.2 µm. The absorption peaks occurs at 1400, 950 and 750 nm. These are first, second and third overtones respectively. Figure shows absorption spectrum for OH group in silica. Between these absorption peaks there are regions of low attenuation.

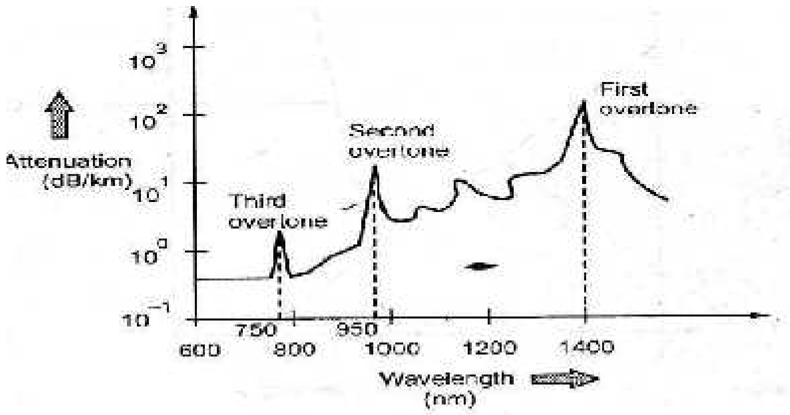


Figure Absorption spectra for OH groups

Intrinsic Absorption

Intrinsic absorption occurs when material is in absolutely pure state, no density variation and in homogenities. Thus intrinsic absorption sets the fundamental lower limit on absorption for any particular material. Intrinsic absorption results from electronic absorption bands in UV region and from atomic vibration bands in the near infrared region. The electronic absorption bands are associated with the band gaps of amorphous glass materials. Absorption occurs when a photon interacts with an electron in the valence band and excites it to a higher energy level. UV absorption decays exponentially with increasing wavelength (λ). In the IR (infrared) region above 1.2 µm the optical waveguide loss is determined by presence of the OH ions and inherent IR absorption of the constituent materials.

The inherent IR absorption is due to interaction between the vibrating band and the electromagnetic field of optical signal this results in transfer of energy from field to the band, thereby giving rise to absorption, this absorption is strong because of many bonds present in the fiber. The ultraviolet loss at any wavelength is expressed as,

$$\alpha_{uv} = \frac{154.2}{46.6\,x + 60} \times 10^{-2} \times e^{\left(\frac{4.63}{\lambda}\right)}$$

where, x is mole fraction of GeO_2. λ is operating wavelength. α_{uv} is in dB/km.

The loss in infrared (IR) region (above 1.2 μm) is given by expression

The expression is derived for GeO_2-SiO_2 glass fiber.

$$\alpha_{IR} = 7.81 \times 10^{11} \times e^{\left(\frac{-48.48}{\lambda}\right)}$$

Rayleigh Scattering Losses

Scattering losses exists in optical fibers because of microscopic variations in the material density and composition. As glass is composed by randomly connected network of molecules and several oxides (e.g. SiO2, GeO2 and P2O5), these are the major cause of compositional structure fluctuation. These two effects results to variation in refractive index and Rayleigh type scattering of light.

Rayleigh scattering of light is due to small localized changes in the refractive index of the core and cladding material. There are two causes during the manufacturing of fiber. The first is due to slight fluctuation in mixing of ingredients. The random changes because of this are impossible to eliminate completely. The other cause is slight change in density as the silica cools and solidifies. When light ray strikes such zones it gets scattered in all directions. The amount of scatter depends on the size of the discontinuity compared with the wavelength of the light so the shortest wavelength (highest frequency) suffers most scattering. The below figure shows graphically the relationship between wavelength and Rayleigh scattering loss.

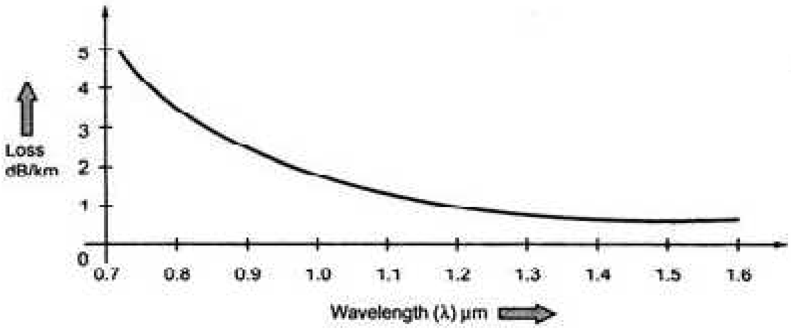


Figure Scattering loss

Scattering loss for single component glass is given by,

$$\propto_{scat} = \frac{8\pi^3}{3\lambda^4} (n^2 - 1)^2 k_B T_f \beta_T \text{ nepers}$$

where, n = Refractive index, B = Boltzmann's constant, β_T = Isothermal compressibility of material, T_f = Temperature at which density fluctuations are frozen into the glass as it solidifies (fictive temperature)

Another form of equation is

$$\propto_{scat} = \frac{8\pi^3}{3\lambda^4} n^8 p^2 k_B T_f \beta_T \text{ neper} \quad \propto_{scat} = \frac{8\pi^3}{3\lambda^4} (\delta_n^2)^2 \delta v$$

where, P = Photoelastic coefficient

where, δ_n^2 = Mean square refractive index fluctuation

δv = Volume of fiber

Multimode fibers have higher dopant concentrations and greater compositional fluctuations. The overall losses in these fibers are more as compared to single mode fibers.

Mie Scattering

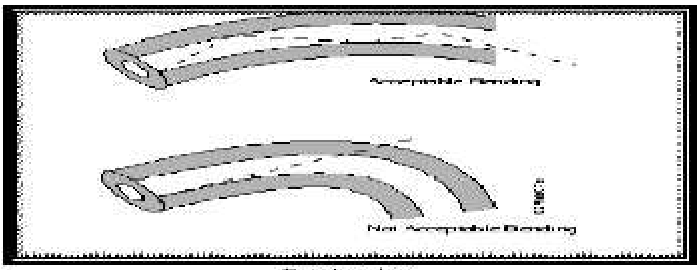
Linear scattering also occurs at in homogenities and these arise from imperfections in the fiber's geometry, irregularities in the refractive index and the presence of bubbles etc. caused during manufacture. Careful control of manufacturing process can reduce Mie scattering to insignificant levels.

Bending Loss

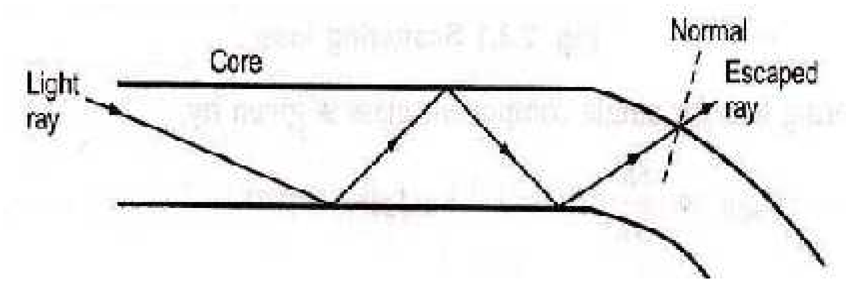
Radiative losses occur whenever an optical fiber undergoes a bend of finite radius of curvature. Fibers can be subjected to two types of bends:

a) Macroscopic bends (having radii that are large as compared with the fiber diameter)

b) Random microscopic bends of fiber axis Losses due to curvature and losses caused by an abrupt change in radius of curvature are referred to as 'bending losses.' The sharp bend of a fiber causes insignificant radiative losses and there is also possibility of mechanical failure.



Bending loss



As the core bends the normal will follow it and the ray will now find itself on the wrong side of critical angle and will escape. The sharp bends are therefore avoided. The radiation loss from a bent fiber depends on –Field strength of certain critical distance x_c from fiber axis where power is lost through radiation.

The radius of curvature R.

The higher order modes are less tightly bound to the fiber core, the higher order modes radiate out of fiber firstly. For multimode fiber, the effective number of modes that can be guided by curved fiber is

where, α is graded index profile.

Δ is core – cladding index difference. n_2 is refractive index of cladding, k is wave propagation constant $\left(\frac{2\pi}{\lambda}\right)$.

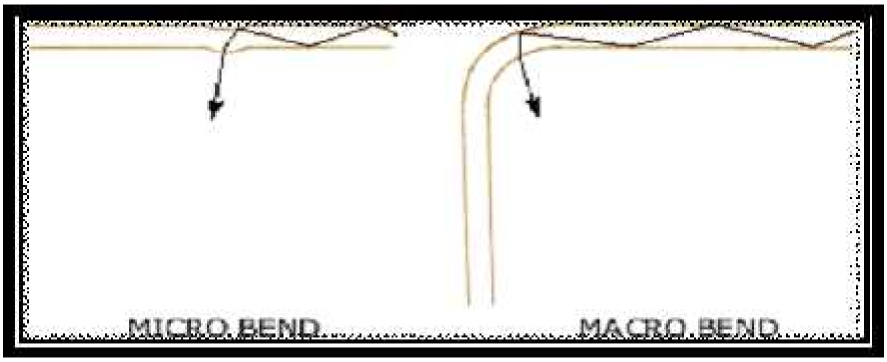
$$N_\infty = \frac{\alpha}{\alpha+2}(n_1\, k\, a)^2 \Delta$$

N_∞ is total number of modes in a straight fiber.

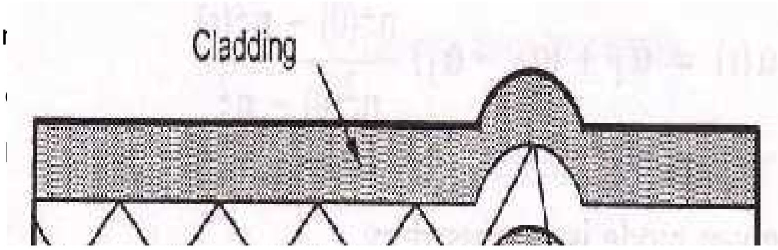
Micro bending Loss

Another form of radiation loss in optical waveguide results from mode coupling caused by random micro bends of the optical fiber. Micro bends are repetitive small scale fluctuations in the radius of curvature of the fiber axis. They are caused either by non uniformities in the manufacturing of the fiber or by non uniform lateral pressures created during the cabling of the fiber. An increase in attenuation results from micro bending because the fiber curvature causes repetitive coupling of energy between the guided modes and the leaky or non guided modes in the fiber.

Micro bending losses can be minimized by placing a compressible jacket over the fiber. When external forces are applied to this configuration, the jacket will be deformed but the fiber will tend to stay relatively straight. Microbending is a loss due to small bending or distortions. This small microbending is not visible. The losses due to this are temperature related, tensile related or crush related.



Macrobending



For slight bends, the loss is extremely small and is not observed. As the radius of curvature decreases, the loss increases exponentially until at a certain critical radius of curvature loss becomes observable. If the bend radius is made a bit smaller once this threshold point has been reached, the losses suddenly become extremely large. It is known that any bound core mode has an evanescent field tail in the cladding which decays exponentially as a function of distance from the core. Since this field tail moves along with the field in the core, part of the energy of a propagating mode travels in the fiber cladding. When a fiber is bent, the field tail on the far side of the centre of curvature must move faster to keep up with the field in the core, for the lowest order fiber mode.

At a certain critical distance x_c, from the centre of the fiber; the field tail would have to move faster than the speed of light to keep up with the core field. Since this is not possible the optical energy in the field tail beyond x_c radiates away.

The amount of optical radiation from a bent fiber depends on the field strength at x_c and on the radius of curvature R. Since higher order modes are bound less tightly to the fiber core than lower order modes, the higher order modes will radiate out of the fiber first. The change in spectral attenuation caused by macrobending is different to microbending. Usually there are no peaks and troughs because in a macrobending no light is coupled back into the core from the cladding as can happen in the case of microbends. The macrobending losses are cause by large scale bending of fiber. The losses are eliminated when the bends are straightened. The losses can be minimized by not exceeding the long term bend radii.

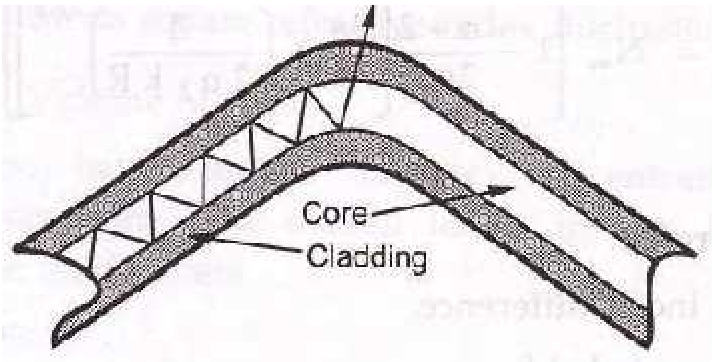


Figure macrobending loss

Core and Cladding Loss

Since the core and cladding have different indices of refraction hence they have different attenuation coefficients α_1 and α_2 respectively.

$$\propto(r) = \propto_1 + (\propto_2 - \propto_1)\frac{n^2(0) - n^2(r)}{n^2(0) - n_2^2}$$

For step index fiber, the loss for a mode order (v, m) is given by,

$$\propto_{vm} = \propto_1 \frac{P_{core}}{P} + \propto_2 \frac{P_{cladding}}{P} \qquad \ldots$$

For low-order modes, the expression reduced to

$$\propto_{vm} = \propto_1 + (\propto_2 + \propto_1)\frac{P_{cladding}}{P} \qquad \ldots$$

where, $\frac{P_{core}}{P}$ and $\frac{P_{cladding}}{P}$ are fractional powers.

For graded index fiber, loss at radial distance is expressed as,

The loss for a given mode is expressed by,

$\propto_{Graded\ Index} = \frac{\int_0^{\infty} \propto(r)\, P(r)\, r\, dr}{\int_0^{\infty} P(r)\, r\, dr}$ where, P(r) is power density of that model at radial distance r.

Signal Distortion in Optical Waveguide

The pulse gets distorted as it travels along the fiber lengths. Pulse spreading in fiber is referred as dispersion. Dispersion is caused by difference in the propagation times of light rays that takes different paths during the propagation. The light pulses travelling down the fiber encounter dispersion effect because of this the pulse spreads out in time domain. Dispersion limits the information bandwidth. The distortion effects can be analyzed by studying the group velocities in guided modes.

Information Capacity Determination

Dispersion and attenuation of pulse travelling along the fiber is shown in Figure. Figure shows, after travelling some distance, pulse starts broadening and overlap with the neighbouring pulses. At certain distance the pulses are not even distinguishable and error will occur at receiver. Therefore the information capacity is specified by bandwidth- distance product (MHz . km). For step index bandwidth distance product is 20 MHz .km and for graded index it is 2.5 MHz . km.

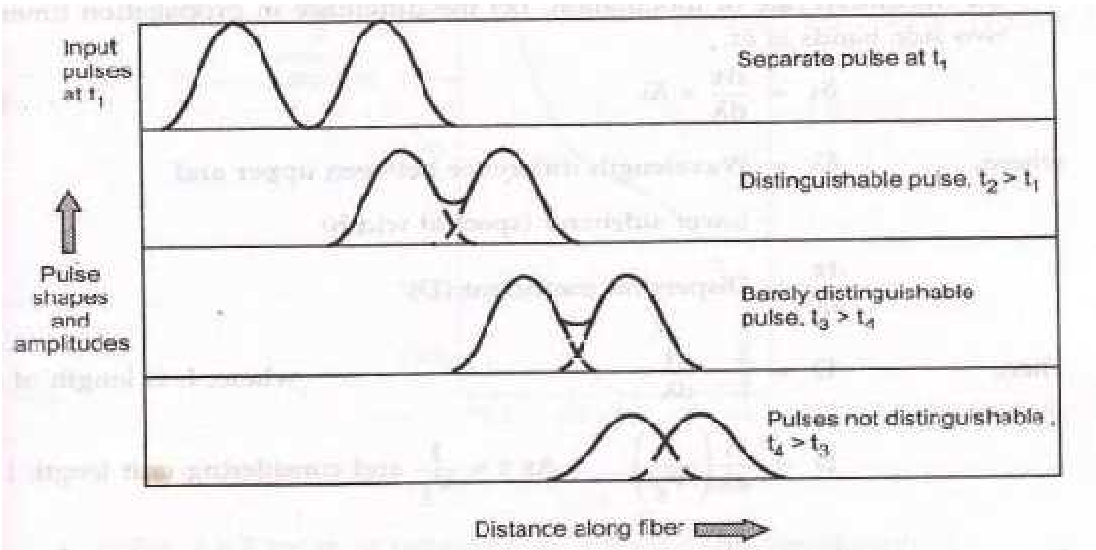
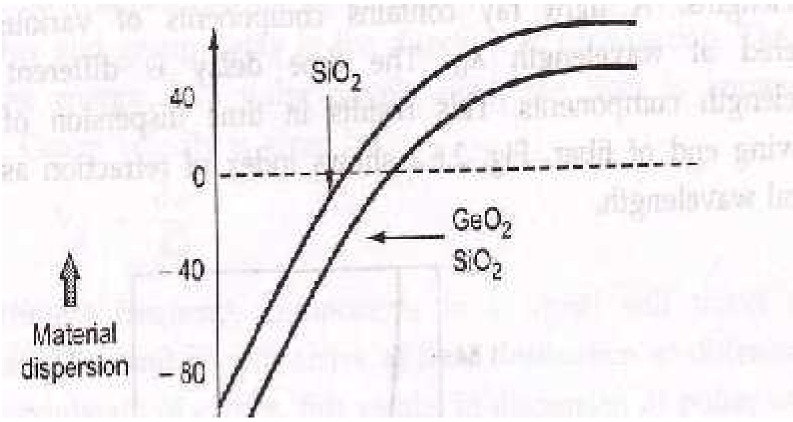


Figure Dispersion and Attenuation in fiber

Group Delay

Consider a fiber cable carrying optical signal equally with various modes and each mode contains all the spectral components in the wavelength band. All the spectral components travel independently and they observe different time delay and group delay in the direction of propagation. The velocity at which the energy in a pulse travels along the fiber is known as group velocity. Group velocity is given by,

$$V_g = \frac{\partial w}{\partial \beta}$$



Thus different frequency components in a signal will travel at different group velocities and so will arrive at their destination at different times, for digital modulation of carrier, this result in dispersion of pulse, which affects the maximum rate of modulation. Let the difference in propagation times for two side bands is δ_τ.

$$\delta\tau = \frac{d\tau}{d\lambda}\ x\ \delta\lambda$$

Where, $\delta\tau$ = Wavelength diff (spectral width)

$\frac{d\tau}{d\lambda}$ = Dispersion coefficient (D)

Then, $$D = \frac{1}{L}.\frac{d\tau}{d\lambda}$$

where, L is length of fiber.

$$D = \frac{d}{d\lambda}\left(\frac{1}{V_g}\right) \qquad \text{As } \tau = \frac{1}{V_g}$$

and considering unit length L = 1. Now $\frac{1}{V_g} = \frac{d\beta}{d\omega}$

$$\frac{1}{V_g} = \frac{d\lambda}{d\omega} x \frac{d\beta}{d\lambda}$$

$$\frac{1}{V_g} = \frac{-\lambda^2}{2\pi c} x \frac{d\beta}{d\lambda}$$

$$D = \frac{d}{d\lambda}\left(\frac{-\lambda^2}{2\pi c}.\frac{d\beta}{d\lambda}\right)$$

Dispersion is measured in picoseconds per nanometer per kilometer.

Material Dispersion

Material dispersion is also called as chromatic dispersion. Material dispersion exists due to change in index of refraction for different wavelengths. A light ray contains components of various wavelengths centered at wavelength λ_{10}. The time delay is different for different wavelength components. This results in time dispersion of pulse at the receiving end of fiber. Figure shows index of refraction as a function of optical wavelength. The material dispersion for unit length (L = 1) is given by

$$D_{mat} = \frac{-\lambda}{c} x \frac{d^2n}{d\lambda^2}$$

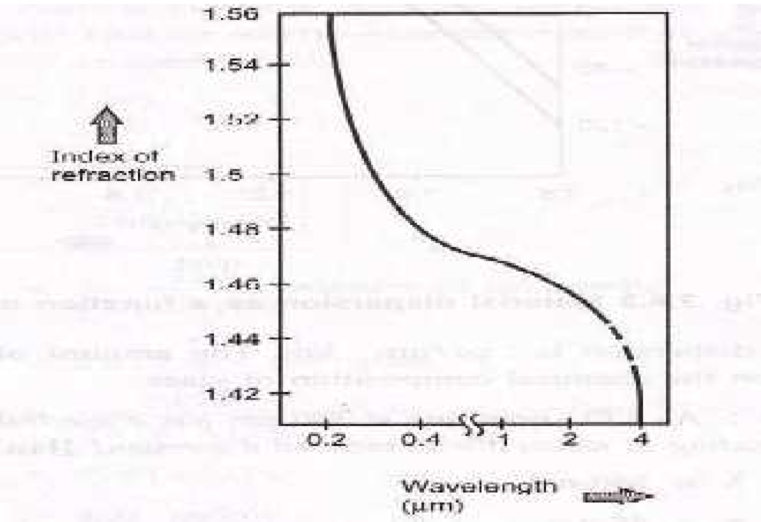


Figure Index of refraction as a function of wavelength

where, c = Light velocity, λ = Center wavelength

$\frac{d^2n}{d\lambda^2}$ = Second derivative of index of refraction w.r.t wavelength Negative sign shows that the upper sideband signal (lowest wavelength) arrives before the lower sideband (highest wavelength). The unit of dispersion is : ps/nm . km. The amount of material dispersion depends upon the chemical composition of glass.

Waveguide Dispersion

Waveguide dispersion is caused by the difference in the index of refraction between the core and cladding, resulting in a 'drag' effect between the core and cladding portions of the power. Waveguide dispersion is significant only in fibers carrying fewer than 5-10 modes. Since multimode optical fibers carry hundreds of modes, they will not have observable waveguide dispersion. The group delay (τ_{wg}) arising due to waveguide dispersion

$$\left(\tau_{wg}\right) = \frac{L}{c}\left[n_2 + n_2 \Delta \frac{d\,(kb)}{dk}\right]$$

Where, b = Normalized propagation constant k = 2π / λ (group velocity)

Normalized frequency V,

$$V = ka(n_1^2 - n_2^2)^{\frac{1}{2}}$$

$$\tau_{wg} = \frac{L}{c}\left[n_2 + n_2 \Delta \frac{d\,(V_b)}{dV}\right]$$

The second term $\frac{d(V_b)}{dV}$ is waveguide dispersion and is mode dependent term..

As frequency is a function of wavelength, the group velocity of the energy varies with frequency. The produces additional losses (waveguide dispersion) and the propagation constant (b) varies with wavelength, the causes of which are independent of material dispersion.

Chromatic Dispersion

The combination of material dispersion and waveguide dispersion is called chromatic dispersion. These losses primarily concern the spectral width of transmitter and choice of correct wavelength. A graph of effective refractive index against wavelength illustrates the effects of material, chromatic and waveguide dispersion.

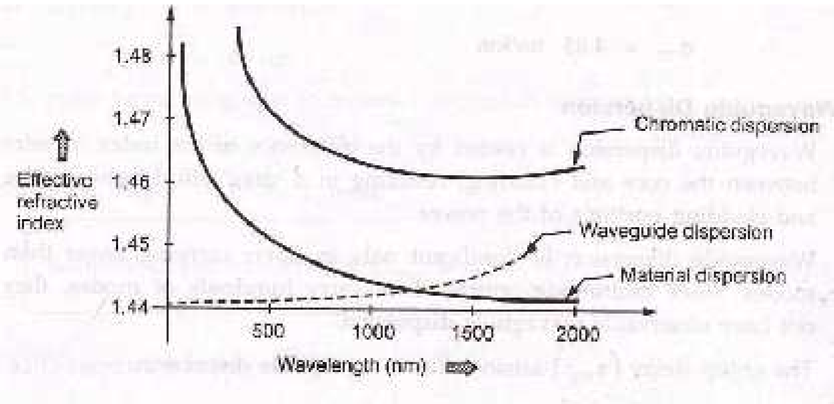


Figure Graph of refractive index against wavelength showing effects of chromatic, waveguide and material dispersion

Material dispersion and waveguide dispersion effects vary in vary in opposite senses as the wavelength increased, but at an optimum wavelength around 1300 nm, two effects almost cancel each other and chromatic dispersion is at minimum. Attenuation is therefore also at minimum and makes 1300 nm a highly attractive operating wavelength.

Modal Dispersion

As only a certain number of modes can propagate down the fiber, each of these modes carries the modulation signal and each one is incident on the boundary at a different angle, they will each have their own individual propagation times. The net effect is spreading of pulse, this form o dispersion is called modal dispersion. Modal dispersion takes place in multimode fibers. It is moderately present in graded index fibers and almost eliminated in single mode step index fibers.

Modal dispersion is given by,

$$\Delta t_{modal} = \frac{n_1 Z}{c}\left(\frac{\Delta}{1-\Delta}\right)$$

where Δt_{modal} = Dispersion

n1 = Core refractive index

Z = Total fiber length

c = Velocity of light in air

Fractional refractive index

$$\Delta t_{modal} = \frac{(NA^2)Z}{2n_1 \; c}$$

$$t_{r\,mod} = 0.44\,(\Delta t_{modal})\pi r^2$$

Polarization Mode Dispersion (PMD)

Different frequency component of a pulse acquires different polarization state (such as linear polarization and circular polarization). This result in pulse broadening is known as polarization mode dispersion (PMD). PMD is the limiting factor for optical communication system at high data rates. The effects of PMD must be compensated.

Pulse Broadening in GI Fibers

The core refractive index varies radially in case of graded index fibers, hence it supports multimode propagation with a low intermodal delay distortion and high data rate over long distance is possible. The higher order modes travelling in outer regions of the core, will travel faster than the lower order modes travelling in high refractive index region. If the index profile is carefully controlled, then the transit times of the individual modes will be identical, so eliminating modal dispersion. The r.m.s pulse broadening is given as:

$$\sigma = \left(\sigma^2_{intermodal} + \sigma^2_{intermodal}\right)^{1/2}$$

where, σintermodal – R.M.S pulse width due to intermodal delay distortion., $\sigma_{intermodal}$ – R.M.S pulse width resulting from pulse broadening within each mode.

The intermodal delay and pulse broadening are related by expression given by Personick.

$$\sigma_{intermodal} = \left(\langle\tau_g^2\rangle - \langle\tau_g\rangle^2\right)^{1/2}$$

Where τg is group delay.From this the expression for intermodal pulse broadening is given as:

$$\propto_{intermodal} = \frac{LN_1\Delta}{2c}\cdot\frac{\propto}{\propto+1}\left(\frac{\propto+2}{3\propto+2}\right)^{1/2} x$$

$$\left[c_1^2 + \frac{4c_1c_2(\propto+1)}{2\propto+1} + \frac{16\Delta^2c_2^2(\propto+1)^2}{(5\propto+2)(3\propto+2)}\right]^{1/2}$$

$$c_1 = \frac{\propto-2-E}{\propto+2} \; and \; c_2 = \frac{3\propto-2-2c}{2(\propto+2)}$$

The intramodal pulse broadening is given as :

$$\sigma_{intramodal}^{2} = \left(\frac{\sigma\lambda}{\lambda}\right)^{2} \left\langle \left(\lambda \frac{d\tau g}{d\lambda}\right)^{2} \right\rangle$$

Where σλ is spectral width of optical source.

Solving the expression gives :

$$\sigma_{intramodal}^{2} = \frac{L}{c} \cdot \frac{\sigma\lambda}{\lambda} \left[\left(-\lambda^{2} \frac{d^{2} n_{1}}{d\lambda^{2}} \right)^{2} - N_{1} c_{1} \Delta \right.$$

$$\left. \left(2\lambda^{2} \frac{d^{2} n_{1}}{d\lambda^{2}} \cdot \frac{\alpha}{\alpha+1} - N_{1} c_{1} \Delta \frac{4\alpha^{2}}{(\alpha+2)(3\alpha+2)} \right) \right]^{1/2}$$

Mode Coupling

After certain initial length, the pulse distortion increases less rapidly because of mode coupling. The energy from one mode is coupled to other modes because of Structural imperfections, Fiber diameter variations, Refractive index variations, Microbends in cable. Due to the mode coupling, average propagation delay become less and intermodal distortion reduces. Suppose certain initial coupling length = Lc, mode coupling length, over Lc = Z. Additional loss associated with mode coupling = h (dB/ km).Therefore the excess attenuation resulting from mode coupling = hZ. The improvement in pulse spreading by mode coupling is given as :

$$hZ \left(\frac{\sigma_c}{\sigma_0}\right) = C$$

Where, C is constant independent of all dimensional quantities and refractive indices. σ_c is pulse broadening under mode coupling. σ_0 is pulse broadening in absence of mode coupling. For long fiber length's the effect of mode coupling on pulse distortion is significant. For a graded index fiber, the effect of distance on pulse broading for various coupling losses are shown

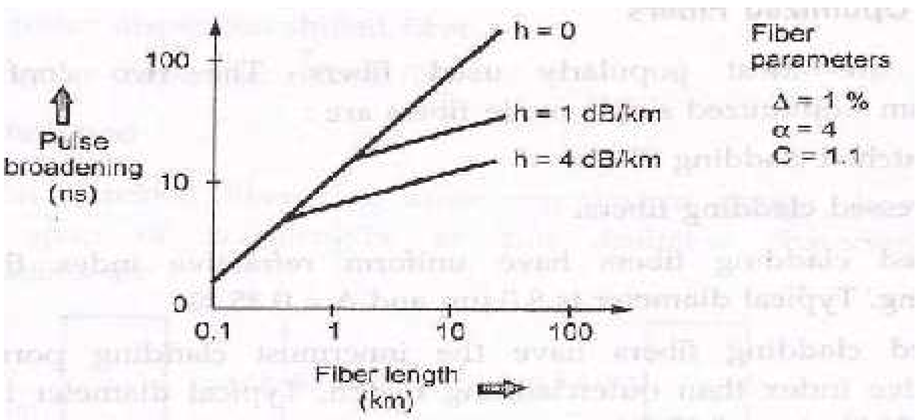


Figure Mode coupling effects on pulse broadening

Design Optimization

Features of single mode fibers are: Longer life, Low attenuation, Signal Transfer quality is good, Modal noise is absent, Largest BW-distance product. Basic design – optimization includes the following: Dispersion, Mode field, Diameter, bending loss, Refractive index profile, Cut-off wavelength.

Refractive Index Profile

Dispersion of single mode silica fiber is lowest at 1300 nm while its attenuation is minimum at 1550 nm. For archiving maximum transmission distance the dispersion null should be at the wavelength of minimum attenuation. The waveguide dispersion is easier to control than the material dispersion. Therefore a variety of core-cladding refractive.

1300nm – Optimized Fibers

These are most popularly used fibers. The two configurations of 1300 nm – optimized single mode fibers are

- Matched cladding fibers.
- Dressed cladding fibers.

Matched cladding fibers have uniform refractive index throughout its cladding. Typical diameter is 9.0 μm and Δ = 0.35 %. Dressed cladding fibers have the innermost cladding portion has low refractive index than outer cladding region. Typical diameter is 8.4 μm and Δ_1 = 0.25 %, Δ_2 = 0.12 %.

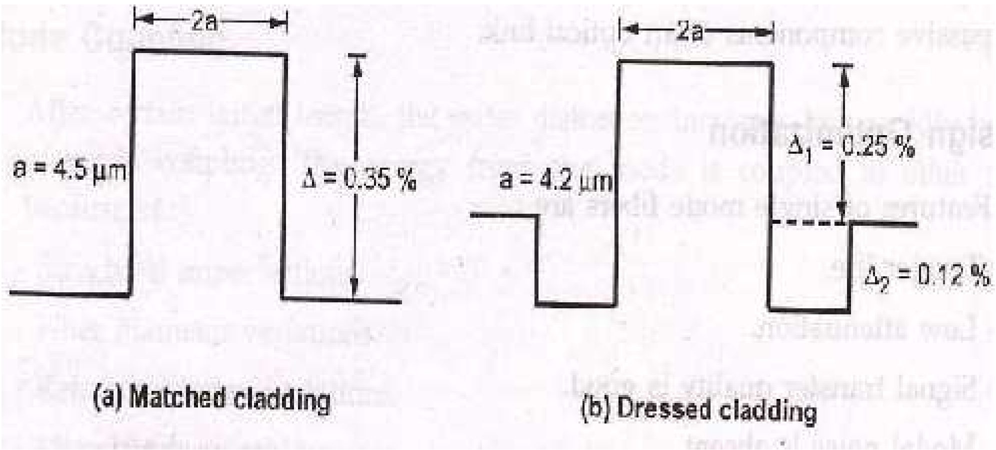


Figure 1300nm – optimized refractive index

Dispersion Shifted Fibers

The addition of wavelength and material dispersion can shift the zero dispersion point of longer wavelength. Two configurations of dispersion shifted fibers are

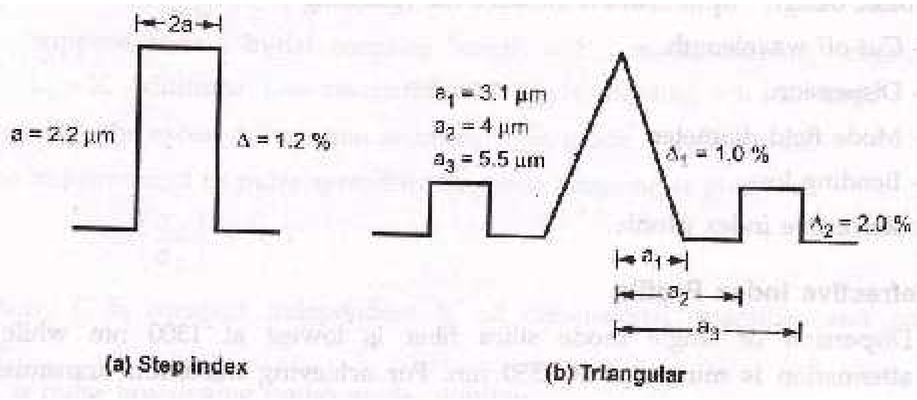


Figure Dispersion shifted fibers

Dispersion Flattened

Dispersion flattened fibers are more complex to design. It offers much broader span of wavelengths to suit desirable characteristics. Two configurations are:

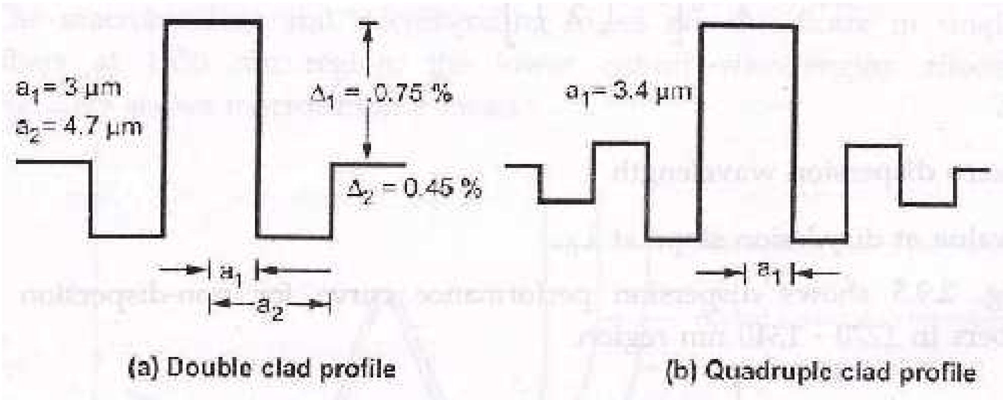


Figure Dispersion Flattened

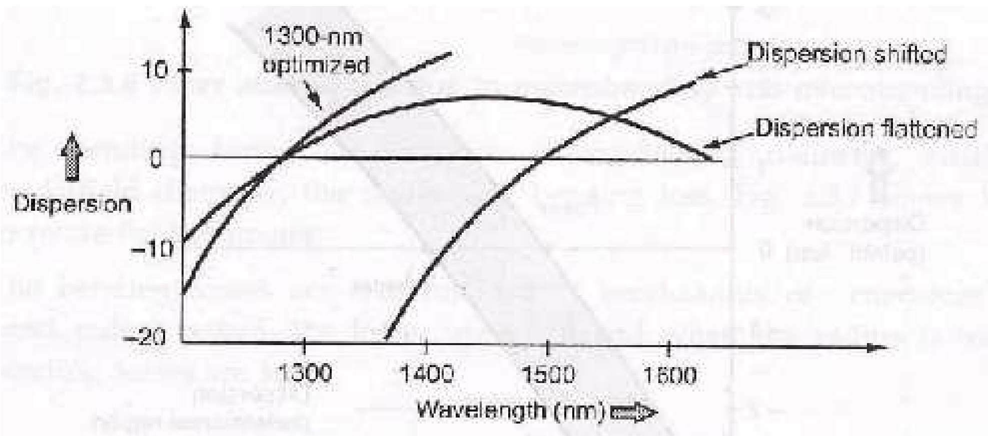


Figure Total resultant dispersion

Dispersion Calculations

The total dispersion consists of material and waveguide dispersions. The resultant intermodal dispersion is given as,

$$D(\lambda) = \frac{d\tau}{d\lambda}$$

Where, τ is group delay per unit length of fiber. The broadening σ of an optical pulse is given

$$\sigma = D(\lambda) L \sigma_\lambda$$

Where, σ_λ is half power spectral width of source.

As the dispersion varies with wavelength and fiber type, Different formulae are used to calculate dispersions for variety of fiber at different wavelength.

For Non-dispersion shifted fiber between 1270 nm to 1340 nm wavelength, the expression for dispersion is given as :

$$D(\lambda) = \frac{\lambda}{4} S_0 \left[1 - \left(\frac{\lambda_0}{\lambda} \right)^4 \right]$$

Where, λ_0 is zero dispersion wavelength. S_0 is value at dispersion slop at λ_0.

The below figure shows dispersion performance curve for non-dispersion shifted fibers in 1270 – 1340 nm region.

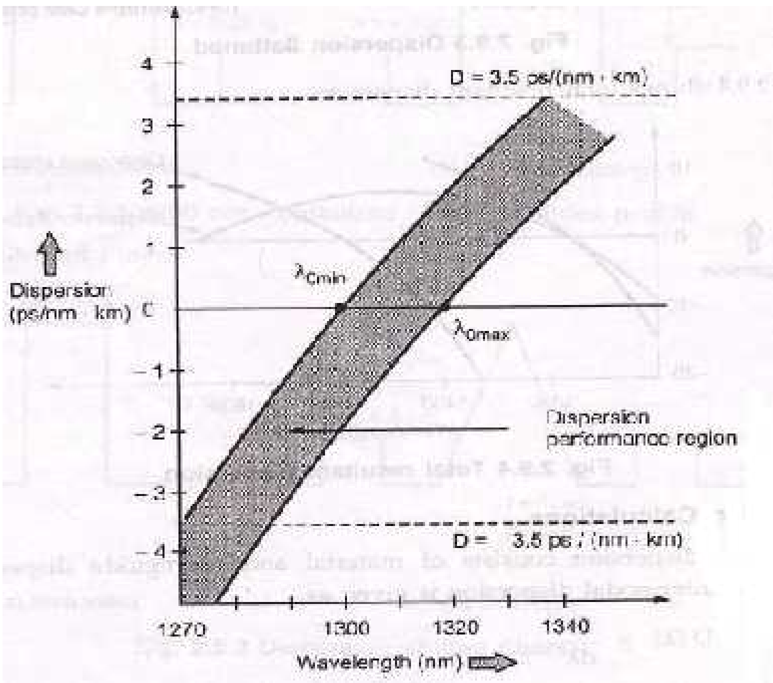


Figure Wavelength vs dispersion

Maximum dispersion specified as 3.5 ps/(nm . km) marked as dotted line in Figure

The cut-off frequency of an optical fiber

The cut-off frequency of an optical fiber is determined not only by the fiber itself (modal dispersion in case of multimode fibers and waveguide dispersion in case of single mode fibers) but also by the amount of material dispersion caused by the spectral width of transmitter.

Bending Loss Limitations

The macrobending and microbending losses are significant in single mode fibers at 1550 nm region, the lower cut-off wavelengths affects more. Figure shows macrobending losses.

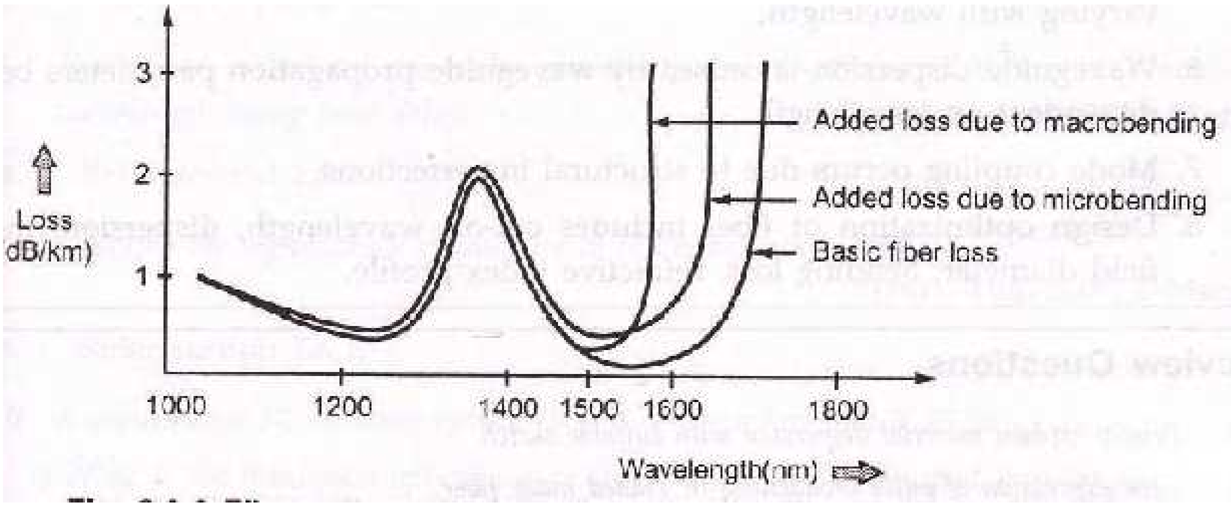
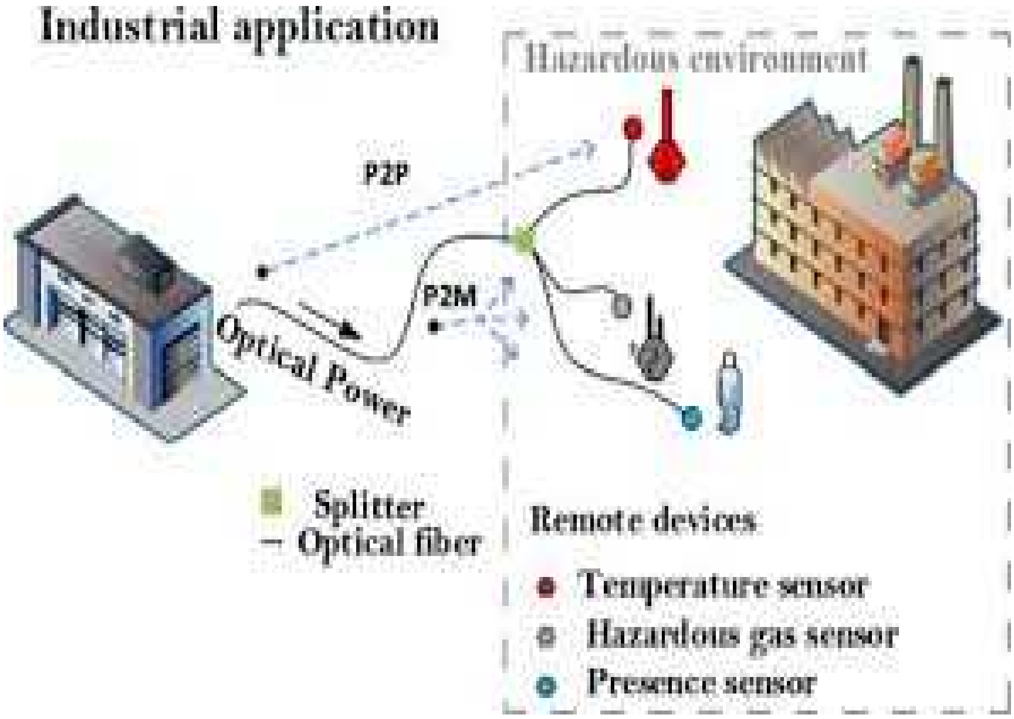


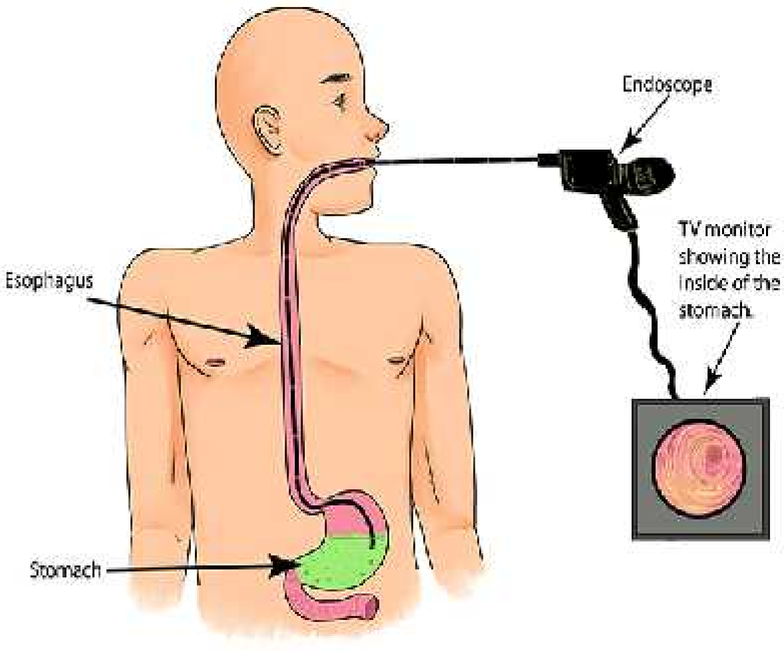
Figure Fiber attenuation due to microbending and macrobending

The bending losses are function of mode-filed diameter, smaller the mode-field diameter, the smaller the bending loss. Figure shows loss due to mode-field diameter. The bending losses are also function of bend-radius of curvature. If the bend radius is less, the losses are more and when the radius is more, the bending losses are less.

APPLICATIONS



Examples of Remote Optical Powering in Industrial Application



Examples of Medical use of Fiber Optics (Endoscopy)

POST TEST-MCQ TYPE

1. Which of the following statements best explain the concept of material absorption?

a) A loss mechanism related to the material composition and fabrication of fiber

b) A transmission loss for optical fibers

c) Results in attenuation of transmitted light

d) Causes of transfer of optical power

2. Polarization modal noise can **__________** the performance of communication system.

a) Degrade

b) Improve

c) Reduce

d) Attenuate

3. Absorption losses due to atomic defects mainly include

a) Radiation

b) Missing molecules, oxygen defects in glass

c) Impurities in fiber material

d) Interaction with other components of core

4. The effects of intrinsic absorption can be minimized by

a) Ionization

b) Radiation

c) Suitable choice of core and cladding components

d) Melting

5. Which of the following is not a metallic impurity found in glass in extrinsic absorption?

a) Fe^{2+}

b) Fe^{3+}

c) Cu

d) Si

6. A multimode fiber has refractive indices n1 = 1.15, n2 = 1.11 and an operating wavelength of 0.7µm. Find the radius of curvature?

a) 8.60µm

b) 9.30µm

c) 9.1µm

d) 10.2µm

7. A single mode fiber has refractive indices n1=1.50, n2 = 2.23, core diameter of 8µm, wavelength = 1.5µm cutoff wavelength = 1.214µm. Find the radius of curvature?

a) 12 mm

b) 20 mm

c) 34 mm

d) 36 mm

8. How the potential macro bending losses can be reduced in case of multimode fiber?

a) By designing fibers with large relative refractive index differences

b) By maintaining direction of propagation

c) By reducing the bend

d) By operating at larger wavelengths

9. Rayleigh scattering and Mie scattering are the types of
a) Linear scattering losses
b) Non-linear scattering losses
c) Fiber bends losses
d) Splicing losses

10. Dominant intrinsic loss mechanism in low absorption window between ultraviolet and infrared absorption tails is
a) Mie scattering
b) Rayleigh scattering
c) Stimulated Raman scattering
d) Stimulated Brillouin scattering

11. The scattering resulting from fiber imperfections like core-cladding RI differences, diameter fluctuations, strains, and bubbles is?
a) Rayleigh scattering
b) Mie scattering
c) Stimulated Brillouin scattering
d) Stimulated Raman scattering

12. Mie scattering has in-homogeneities mainly in
a) Forward direction
b) Backward direction
c) All direction
d) Core-cladding interface

13. Raman and Brillouin scattering are usually observed at
a) Low optical power densities
b) Medium optical power densities
c) High optical power densities
d) Threshold power densities

14. Stimulated Brillouin scattering is mainly a
a) Forward process
b) Backward process
c) Upward process
d) Downward process

15. Stimulated Raman scattering occur in
a) Forward direction
b) Backward direction
c) Upward direction
d) Forward and backward direction

16. Stimulated Raman scattering may have an optical power threshold of may be three orders of magnitude
a) Lower than Brillouin threshold
b) Higher than Brillouin threshold
c) Same as Brillouin threshold
d) Higher than Rayleigh threshold

17. What is dispersion in optical fiber communication?
a) Compression of light pulses
b) Broadening of transmitted light pulses along the channel
c) Overlapping of light pulses on compression
d) Absorption of light pulses

18. For no overlapping of light pulses down on an optical fiber link, the digital bit rate BT must be
a) Less than the reciprocal of broadened pulse duration
b) More than the reciprocal of broadened pulse duration
c) Same as that of than the reciprocal of broadened pulse duration
d) Negligible

19. What is pulse dispersion per unit length if for a graded index fiber, 0.1μs pulse broadening is seen over a distance of 13 km?
a) 6.12ns/km
b) 7.69ns/km
c) 10.29ns/km
d) 8.23ns/km

20. The optical source used in a fiber is an injection laser with a relative spectral width $\sigma\lambda/\lambda$ of 0.0011 at a wavelength of 0.70μm. Estimate the RMS spectral width.
a) 1.2 nm
b) 1.3 nm
c) 0.77 nm
d) 0.98 nm

21. Intermodal dispersion occurring in a large amount in multimode step index fiber results in
a) Propagation of the fiber
b) Propagating through the fiber
c) Pulse broadening at output
d) Attenuation of waves

22. Consider a single mode fiber having core refractive index n_1= 1.5. The fiber length is 12m. Find the time taken by the axial ray to travel along the fiber.
a) 1.00μsec
b) 0.06μsec
c) 0.90μsec
d) 0.30μsec

23. A 4 km optical link consists of multimode step index fiber with core refractive index of 1.3 and a relative refractive index difference of 1%. Find the delay difference between the slowest and fastest modes at the fiber output.
a) 0.173 μsec
b) 0.152 μsec
c) 0.96 μsec
d) 0.121 μsec

24. The modal noise can be reduced by
a) Decreasing width of signal longitudinal mode
b) Increasing coherence time
c) Decreasing number of longitudinal modes
d) Using fiber with large numerical aperture

25. Disturbance along the fiber such as vibrations, discontinuities, connectors, splices, source/detectors coupling result in
a) Modal noise
b) Inter-symbol interference
c) Infrared interference
d) Pulse broadening

26. Practical pulse broadening value for graded index fiber lies in the range of
a) 0.9 to 1.2 ns/km
b) 0.2 to 1 ns/km
c) 0.23 to 5 ns/km
d) 0.45 to 8 ns/km

27. Dispersion-shifted single mode fibers are created by
a) Increasing fiber core diameter and decreasing fractional index difference
b) Decreasing fiber core diameter and decreasing fractional index difference
c) Decreasing fiber core diameter and increasing fractional index difference
d) Increasing fiber core diameter and increasing fractional index difference

28. he fibers which relax the spectral requirements for optical sources and allow flexible wavelength division multiplying are known as
a) Dispersion-flattened single mode fiber
b) Dispersion-enhanced single mode fiber
c) Dispersion-compressed single mode fiber
d) Dispersion-standardized single mode fiber

29. The variant of non-zero-dispersion-shifted fiber is called as
a) Dispersion flattened fiber
b) Zero-dispersion fiber
c) Positive-dispersion fiber
d) Negative-dispersion fiber

30. The optical source used for detection of optical signal is
a) IR sensors
b) Photodiodes
c) Zener diodes
d) Transistors

31. An optical fiber behaves as a birefringence medium due to differences in
a) Effective R-I and core geometry
b) Core-cladding symmetry
c) Transmission/propagation time of waves
d) Refractive indices of glass and silica

CONCLUSION

In this unit, the different kind of losses, signal distortion, Single mode fibers were discussed. The estimation of the losses and analyzing the propagation characteristics of an optical signal in different types of fibers were dealt in this unit.

REFERENCES

1. Gerd Keiser, “Optical Fiber Communication” McGraw-Hill International,4th Edition. 2010.
2. John M. Senior, “Optical Fiber Communication”, Second Edition, Pearson Education, 2007.
3. Ramaswami, Sivarajan and Sasaki “Optical Networks”, Morgan Kaufmann, 2009.
4. J.Senior, Optical Communication, Principles and Practice,Prentice Hall of India, 3rd Edition, 2008.
5. J.Gower, “Optical Communication System”, Prentice Hall of India, 2001.

ASSIGNMENT

1. Explain the scattering and bending losses that occur in an optical fiber with relevant diagrams and expressions.
2. Write a brief note on design optimization of single mode fibers.
3. With diagram, explain intra and inter modal dispersion.
4. Discuss the attenuation encountered in optical fiber communication due to Bending, Scattering and Absorption.
5. What are the losses on signal attenuation mechanisms in a fiber? Explain in detail.
6. Discuss polarization mode dispersion and its limitations

AIM & OBJECTIVES

- ❖ To learn the basic elements of optical fiber transmission link, fiber modes configurations and structures.
- ❖ To understand the different kind of losses, signal distortion, SM fibers.
- ❖ To learn the various optical sources, materials and fiber splicing.
- ❖ To learn the fiber optical receivers and noise performance in photo detector.

PRE TEST-MCQ TYPE

1. How many types of sources of optical light are available?
a) One
b) Two
c) Three
d) Four

2. Which process gives the laser its special properties as an optical source?
a) Dispersion
b) Stimulated absorption
c) Spontaneous emission
d) Stimulated emission

3. A device which converts electrical energy in the form of a current into optical energy is called as
a) Optical source
b) Optical coupler
c) Optical isolator
d) Circulator

4. The majority of the carriers in a p-type semiconductor are ___________
a) Holes
b) Electrons
c) Photons
d) Neutrons

5. The hole concentration in extrinsic materials is _________ electron concentration.
a) much greater than
b) lesser than
c) equal to
d) negligible difference with

UNIT III FIBER OPTICAL SOURCES AND COUPLING

Direct and indirect Band gap Materials -LED structures -Light source materials -Quantum efficiency and LED power, Modulation of a LED, lasers Diodes-Modes and Threshold condition – Rate equations-External Quantum efficiency -Resonant frequencies -Laser Diodes, Temperature effects, Introduction to Quantum laser, Fiber amplifiers- Power Launching and coupling, Lencing schemes, Fiber -to- Fiber joints, Fiber splicing-Signal to Noise ratio , Detector response time.

THEORY

Introduction

Optical Sources:

The Optical transmitter converts electrical input signal into corresponding optical signal. optical signal is then launched into the fiber. Optical source is the major component in an optical transmitter .Popularly used optical transmitters are Light Emitting Diode (LED) and semiconductor

Laser Diodes (LD)

Characteristics of Light Source of Communication

To be useful in an optical link, a light source needs the following characteristics It must be possible to operate the device continuously at a variety of temperatures for many years. It must be possible to modulate the light output over a wide range of modulating frequencies. For fiber links, the wavelength of the output should coincide with one of transmission windows for the fiber type used. To couple large amount of power into an optical fiber, the emitting area should be small. To reduce material dispersion in an optical fiber link, the output spectrum should be narrow. The power requirement for its operation must be low. The light source must be compatible with the modern solid state devices. The optical output power must be directly modulated by varying the input current to the device. Better linearity of prevent harmonics and intermodulation distortion. High coupling efficiency. High optical output power. High reliability. Low weight and low cost.

Two types of light sources used in fiber optics are light emitting diodes (LEDs) and laser diodes (LDs).

Light Emitting Diodes (LEDs)

p-n Junction

Conventional p-n junction is called as homojunction as same semiconductor material is sued on both sides junction. The electron-hole recombination occurs in relatively layer = 10 μm. As the carriers are not confined to the immediate vicinity of junction, hence high current densities cannot be realized. The carrier confinement problem can be resolved by sandwiching a thin layer (= 0.1 μm) between p-type and n-type layers. The middle layer may or may not be doped. The carrier confinement occurs due to band gap discontinuity of the junction. Such a junction is called as heterojunction and the device is called double heterostructure.

In any optical communication system when the requirements is

1. Bit rate f 100-2—Mb/sec.

2. Optical power in tens of micro watts, LEDs are best suitable optical source.

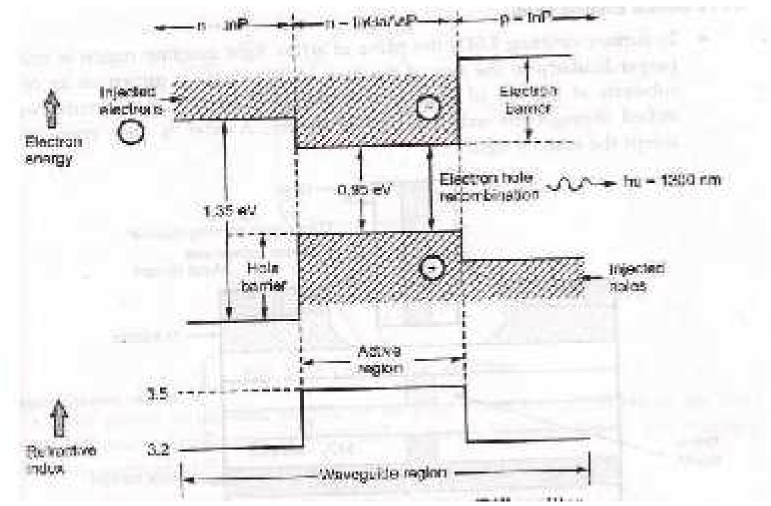
LED Structures

Heterojuncitons:

A heterojunction is an interface between two adjoining single crystal semiconductors with different band gap. Heterojuncitons are of two types, Isotype (n-n or p-p) or Antisotype (p-n).

Double Heterojuncitons (DH):

In order to achieve efficient confinement of emitted radiation double heterojunction are used in LED structure. A heterojunction is a junction formed by dissimilar semiconductors. Double heterojunction (DH) is formed by two different semiconductors on each side of active region. Figure shows double heterojunction (DH) light emitter.



The crosshatched regions represent the energy levels of free charge. Recombination occurs only in active In GaAsP layer. The two materials have different band gap energies and different refractive indices. The changes in band gap energies create potential barrier for both holes and electrons. The free charges can recombine only in narrow, well defined active layer side.

A double heterojunction (DH) structure will confine both hole and electrons to a narrow active layer. Under forward bias, there will be a large number of carriers injected into active region where they are efficiently confined. Carrier recombination occurs in small active region so leading to an efficient device. Another advantage DH structure is that the active region has a higher refractive index than the materials on either side, hence light emission occurs in an optical waveguide, which serves to narrow the output beam.

LED configurations

At present there are two main types of LED used in optical fiber links

- ❖ Surface emitting LED
- ❖ Edge emitting LED.

Both devices used a DH structure to constrain the carriers and the light to an active layer.

Surface Emitting LEDs

In surface emitting LEDs the plane of active light emitting region is oriented perpendicularly to the axis of the fiber. A DH diode is grown on an N-type substrate at the top of the diode as shown in Figure. A circular well is etched through the substrate of the device. A fiber is then connected to accept the emitted

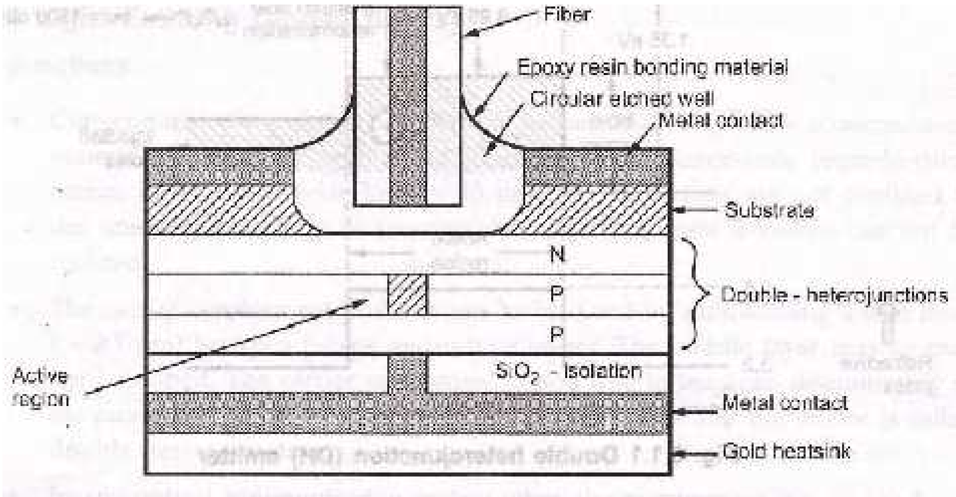


Figure Cross section through a typical surface emitting LED

At the back of device is a gold heat sink. The current flows through the p-type material and forms the small circular active region resulting in the intense beam of light.

Diameter of circular active area = 50 μm
Thickness of circular active area = 2.5 μm
Current density = 2000 A/cm2 half-power
Emission pattern = Isotropic, 120^o beamwidth.

The isotropic emission pattern from surface emitting LED is of Lambartian pattern. In Lambartian pattern, the emitting surface is uniformly bright, but its projected area diminishes as cos θ, where θ is the angle between the viewing direction and the normal to the surface as shown in Figure. The beam intensity is maximum along the normal.

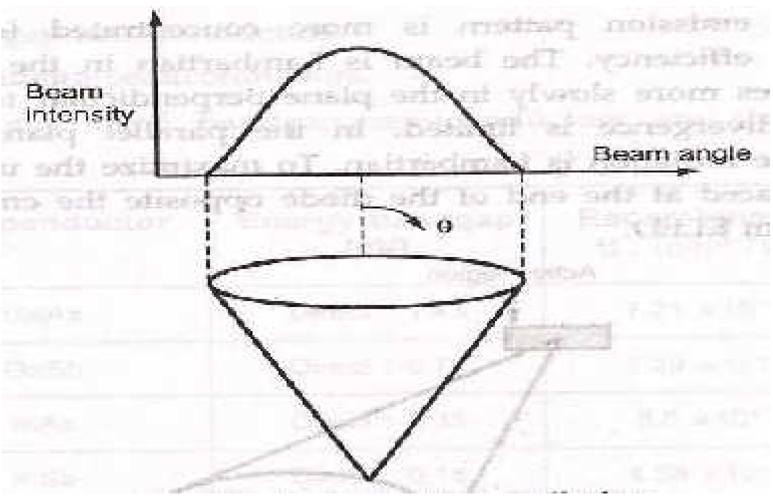


Figure Lambertian radiation

The power is reduced to 50% of its peak when θ = 60o, therefore the total half-power beamwidth is 120°. The radiation pattern decides the coupling efficiency of LED.

Edge Emitting LEDS (ELEDs)

In order to reduce the losses caused by absorption in the active layer and to make the beam more directional, the light is collected from the edge of the LED. Such a device is known as edge emitting LED or ELED. It consists of an active junction region which is the source of incoherent light and two guiding layers. The refractive index of guiding layers is lower than active region but higher than outer surrounding material. Thus a waveguide channel is form and optical radiation is directed into the fiber. Figure shows structure of LED

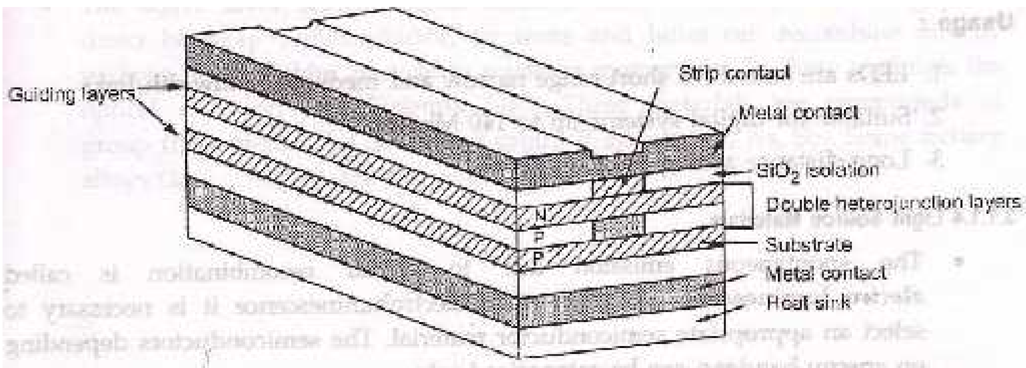


Figure Structure of Edge Emitting LED

Edge emitter's emission pattern is more concentrated (directional) providing improved coupling efficiency. The beam is Lambartian in the plane parallel to the junction but diverges more slowly in the plane perpendicular to the junction. In this plane, the beam divergence is limited. In the parallel plane, there is no beam confinement and the radiation is Lambartian. To maximize the useful output power, a reflector may be placed at the end of the diode opposite the emitting edge.

Features:

- Linear relationship between optical output and current.
- Spectral width is 25 to 400 nm for λ = 0.8 – 0.9 μm.

- Modulation bandwidth is much large.
- Not affected by catastrophic gradation mechanisms hence are more reliable.
- ELEDs have better coupling efficiency than surface emitter.
- ELEDs are temperature sensitive.

Usage:

1. LEDs are suited for short range narrow and medium bandwidth links.
2. Suitable for digital systems up to 140 Mb/sec.
3. Long distance analog links

Light Source Materials

The spontaneous emission due to carrier recombination is called electro luminescence. To encourage electroluminescence it is necessary to select as appropriate semiconductor material. The semiconductors depending on energy bandgap can be categorized into Direct bandgap semiconductors. Indirect bandgap semiconductors. Some commonly used bandgap semiconductors are shown in following table.

Direct bandgap semiconductors are most useful for this purpose. In direct bandgap semiconductors the electrons and holes on either side of bandgap have same value of crystal momentum. Hence direct recombination is possible. The recombination occurs within 10^{-8} to 10^{-10} sec. In indirect bandgap semiconductors, the maximum and minimum energies occur at different values of crystal momentum. The recombination in these semiconductors is quite slow i.e. 10^{-2} and 10^{-3} sec.

The active layer semiconductor material must have a direct bandgap. In direct bandgap semiconductor, electrons and holes can recombine directly without need of third particle to conserve momentum. In these materials the optical radiation is sufficiently high. These materials are compounds of group III elements (Al, Ga, In) and group V element (P, As, Sb). Some tertiary allos Ga1-x Alx As are also used.

Semiconductor	Energy bandgap (eV)	Recombination Br (cm3 sec)
GaAs	Direct : 1.43	7.21 x 10-10
GaAs	Direct : 0.73	2.39 x 10-10
InAs	Direct : 0.35	8.5 x 10-11
InSb	Direct : 0.18	4.58 x 10-11
Si	Indirect : 1.12	1.79 x 10-15
Ge	Indirect : 0.67	5.25 x 10-14
GaP	Indirect : 2.26	5.37 x 10-14

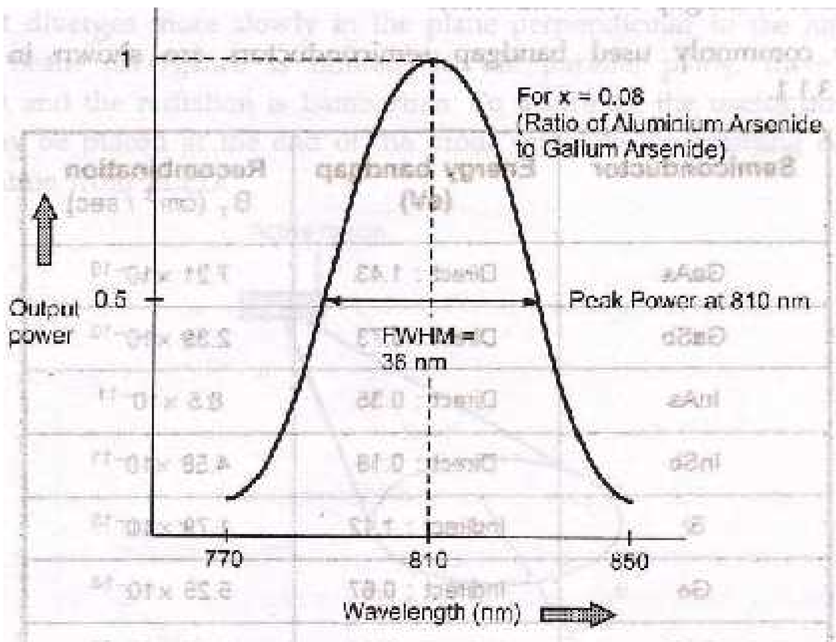


Figure Emission spectrum of Ga1-x AlxAs LED

The peak output power is obtained at 810 nm. The width of emission spectrum at half power (0.5) is referred as full width half maximum (FWHM) spectral width. For the given LED FWHM is 36 nm. The fundamental quantum mechanical relationship between gap energy E and frequency v is given as

$$E = hv$$

$$E = h\frac{c}{\lambda}$$

$$\lambda = \frac{hc}{E}$$

Where, energy (E) is in joules and wavelength (λ) is in meters. Expressing the gap energy (E_g) in electron volts and wavelength (λ) in micrometers for this application.

$$\lambda(\mu m) = \frac{1.24}{E_g(eV)}$$

Different materials and alloys have different band gap energies. The bandgap energy (E_g) can be controlled by two compositional parameters x and y, within direct bandgap region. The quaternary alloy In1-x Gax Asy P1-y is the principal material sued in such LEDs. Two expression relating Eg and x,y are

$$E_g = 1.424 + 1.266\,x + 0266\,x^2$$

$$E_g = 1.35 - 0.72\,y + 0.12\,y^2$$

Quantum Efficiency and Power

The internal quantum efficiency (η_{int}) is defined as the ratio of radiative recombination rate to the total recombination rate.

$$\eta_{int} = \frac{R_r}{R_r + R_{nr}}$$

Where, R_r is radiative recombination rate.

R_{nr} is non-radiative recombination rate.

$$\tau_r = \frac{n}{R_r}$$

If n are the excess carriers, then radiative life time, and non-radiative life time,

$$\tau_r = \frac{n}{R_{nr}}$$

The internal quantum efficiency is given and the recombination time of carriers in active region is τ

It is also known as bulk recombination life time.

$$\eta_{int} = \frac{1}{1 + \frac{R_{nr}}{R_r}}$$

$$\boxed{\frac{1}{\tau} = \frac{1}{\tau_r} + \frac{1}{\tau_{nr}}}$$

$$\eta_{int} = \frac{1}{1 + \frac{\tau_r}{\tau_{nr}}}$$

Therefore internal quantum efficiency is given as

$$\boxed{\eta_{int} = \frac{\tau}{\tau_r}}$$

If the current injected into the LED is I and q is electron charge then total number of recombinations per second is

$$R_r = R_{nr} = \frac{I}{q}$$

$$\eta_{int} = \frac{R_r}{I / q}$$

$$R_r = \eta_{int} \times \frac{I}{q}$$

Optical power generated internally in LED is given as

$$P_{int} = R_r . h\nu$$

$$P_{int} = \left(\eta_{int} \times \frac{I}{q}\right) . h\nu$$

$$P_{int} = \left(\eta_{int} \times \frac{I}{q}\right) . h\frac{c}{\lambda}$$

$$\boxed{P_{int} = \eta_{int} . \frac{hc\,I}{q\lambda}}$$

Not all internally generated photons will available from output of device. The external quantum efficiency is used to calculate the emitted power. The external quantum efficiency is defined as the ratio of photons emitted from LED to the number of photons generated internally. It is given by equation

$$\eta_{ext} = \frac{1}{n(n+1)^2}$$

The optical output power emitted from LED is given as

$$P = \eta_{ext} . P_{int}$$

$$P = \frac{1}{n\,(n+1)^2} \cdot P_{int}$$

Bulk Recombination Life time (τ) :

$$\eta_{int} = \frac{\tau}{\tau_r}$$

$$\frac{1}{\tau} = \frac{1}{\tau_r} + \frac{1}{\tau_{nr}}$$

Internal quantum efficiency (η_{int})

$$\eta_{int} = \frac{23.07}{30}$$

$$\eta_{int} = 0.769$$

Internal power level (P_{int}) :

$$P_{int} = \eta_{int} \cdot \frac{hc\,I}{q\lambda}$$

Advantages of LED

- Simple design.
- Ease of manufacture.
- Simple system integration.
- Low cost.
- High reliability.

Disadvantages of LED

- ❖ Refraction of light at semiconductor/air interface.
- ❖ The average life time of a radiative recombination is only a few nanoseconds, therefore
- ❖ Modulation BW is limited to only few hundred megahertz.
- ❖ Low coupling efficiency.
- ❖ Large chromatic dispersion.

Comparison of Surface and Edge Emitting LED

ED type	Max. modulation freq. (MHz)	Output power (mW)	Fiber coupled power
Surface emitting	60	< 4	< 0.2
Edge emitting	200	< 7	< 1.0

Injection Laser Diode (ILD)

The laser is a device which amplifies the light, hence the LASER is an acronym for light amplification by stimulated emission of radiation. The operation of the device may be described by the formation of an electromagnetic standing wave within a cavity (optical resonator) which provides an output of monochromatic highly coherent radiation.

Principle :

Material absorb light than emitting. Three different fundamental process occurs between the two energy states of an atom. 1) Absorption 2) Spontaneous emission 3) Stimulated emission. Laser action is the result of three process absorption of energy packets (photons) spontaneous emission, and stimulated emission.

(These processes are represented by the simple two-energy-level diagrams). Where E_1 is the lower state energy level. E_2 is the higher state energy level. Quantum theory states that any atom exists only in certain discrete energy state, absorption or emission of light causes them to make a transition from one state to another. The frequency of the absorbed or emitted radiation f is related to the difference in energy E between the two states. If E_1 is lower state energy level and E_2 is higher state energy level

E = $(E_2 – E_1)$ = h.f. Where, h = 6.626 x 10-34 J/s (Plank's constant).

An atom is initially in the lower energy state, when the photon with energy $(E_2 – E_1)$ is incident on the atom it will be excited into the higher energy state E_2 through the absorption of the photon

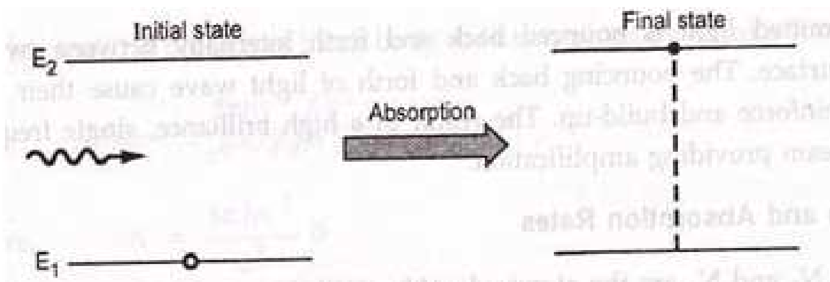


Figure Absorption

When the atom is initially in the higher energy state E_2, it can make a transition to the lower energy state E1 providing the emission of a photon at a frequency corresponding to E = h.f. The emission process can occur in two ways. By spontaneous emission in which the atom returns to the lower energy state in random manner by stimulated emission when a photon having equal energy to the difference between the two states $(E_2 – E_1)$ interacts with the atom causing it to the lower state with the creation of the second photon.

Spontaneous emission gives incoherent radiation while stimulated emission gives coherent radiation. Hence the light associated with emitted photon is of same frequency of incident photon, and in same phase with same polarization. It means that when an atom is stimulated to emit light energy by an incident wave, the liberated energy can add to the wave in constructive manner. The emitted light is bounced back and forth internally between two reflecting surface. The bouncing back and forth of light wave cause their intensity to reinforce and build-up. The result in a high brilliance, single frequency light beam providing amplification.

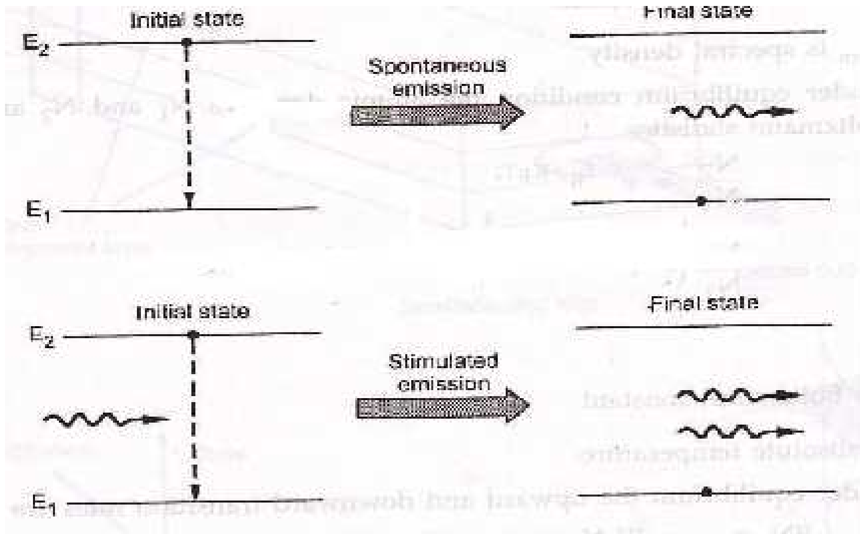


Figure Spontaneous and Stimulated emission

Emission and Absorption Rates

It N_1 and N_2 are the atomic densities in the ground and excited states.

Rate of spontaneous emission $R_{spon} = AN_2$

Rate of stimulated emission $R_{stim} = BN_2\, \rho em$ **R**

ate of absorption Rabs = B' N1 ρem

where, A, B and B' are constants. ρem is spectral density. Under equilibrium condition the atomic densities N_1 and N_2 are given by Boltzmann statistics.

$$\frac{N_2}{N_1} = eg^{(-E_B / K_B T)}$$

$$\frac{N_2}{N_1} = eg^{(-h_v / K_B T)}$$

where, KB is Boltzmann constant.

T is absolute temperature.

Under equilibrium the upward and downward transition rates are equal.

$AN_2 + BN_2\ \rho em = B'\ N1\ \rho em$

Spectral density ρem

Comparing spectral density of black body radiation given by Plank's formula, Therefore A and B are called Einstein's coefficient.

Fabry – Perot Resonator

Lasers are oscillators operating at frequency. The oscillator is formed by a resonant cavity providing a selective feedback. The cavity is normally a Fabry-Perot resonator i.e. two parallel plane mirrors separated by distance L,

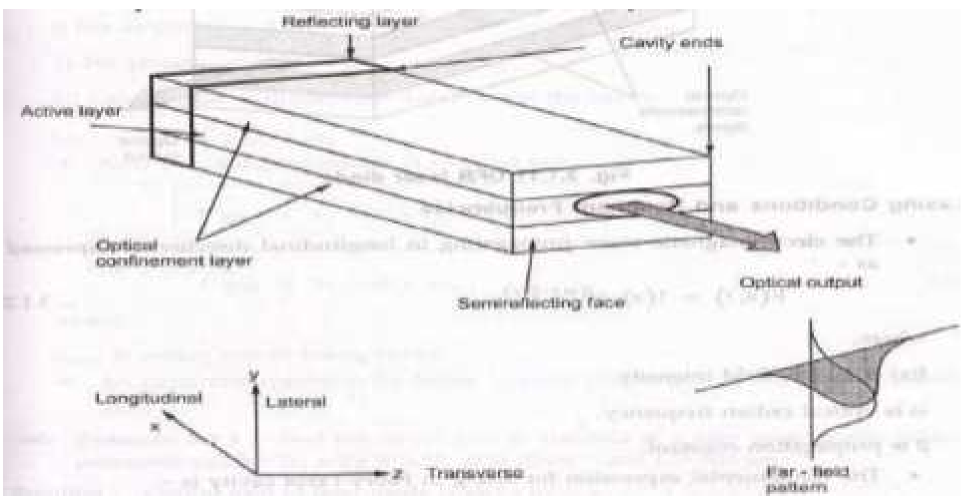


Figure Fabry –Perot Resonator

Light propagating along the axis of the interferometer is reflected by the mirrors back to the amplifying medium providing optical gain. The dimensions of cavity are 25-500 μm longitudinal 5-15 μm lateral and 0.1-0.2 μm transverse. Figure shows Fabry-Perot resonator cavity for a laser diode. The two Heterojuncitons provide carrier and optical confinement in a direction normal to the junction. The current at which lasing starts is the threshold current. Above this current the output power increases sharply.

Distributed Feedback (DFB) Laser

In DFB laser the lasing action is obtained by periodic variations of refractive index along the longitudinal dimension of the diode. Figure shows the structure of DFB laser diode

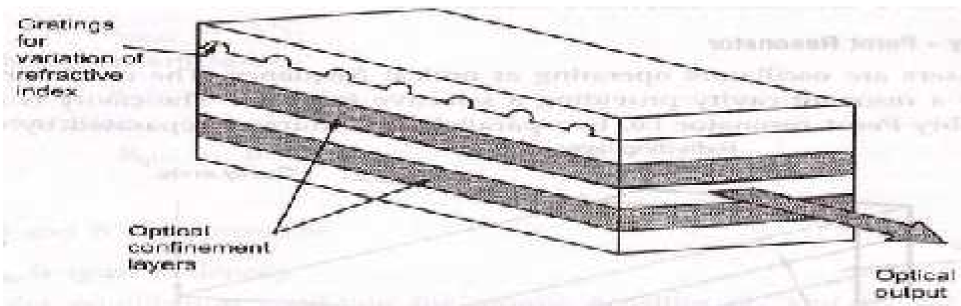


Figure DFB Laser Diode

Lasing conditions and resonant Frequencies

The electromagnetic wave propagating in longitudinal direction is expressed as – E(z, t) = I(z) ej(ω t-β z) where, I(z) is optical field intensity. is optical radian frequency. β is propagation constant. The fundamental expression for lasing in Fabry-Perot cavity is

$$I(z) = I(0)e^{[\{\Gamma g(h v)-\alpha(h v)\}z]}$$

where, is optical field confinement factor or the fraction of optical power in the active layer.

α is effective absorption coefficient of material.

g is gain coefficient.

h v is photon energy.

z is distance traverses along the lasing cavity. The condition of lasing threshold is given as

For amplitude : I (2L) = I (0)

For phase : $e^{-j2\beta L} = 1$

Optical gain at threshold = Total loss in the cavity.

i.e. $\Gamma\, g_{th} = \alpha_t$

Now the lasing expression is reduced to –

$$\Gamma\, g_{th} = a_t = \propto + \frac{1}{2L} \ln\left(\frac{1}{R_1 R_2}\right)$$

$$\Gamma g_{th} = \propto_t = \propto + \propto_{end}$$

where, alpha end is mirror loss in lasing cavity. An important condition for lasing to occur is that gain, $g \geq g_{th}$ i.e. threshold gain.

Power Current Characteristics

The output optic power versus forward input current characteristics is plotted in Figure for a typical laser diode. Below the threshold current (Ith) only spontaneous emission is emitted hence there is small increase in optic power with drive current. At threshold when lasing conditions are satisfied. The optical power increases sharply after the lasing threshold because of stimulated emission. The lasing threshold optical gain (g_{th}) is related by threshold current density (J^{th}) for stimulated emission by expression $g_{th} = \beta$ where, β is constant for device structure

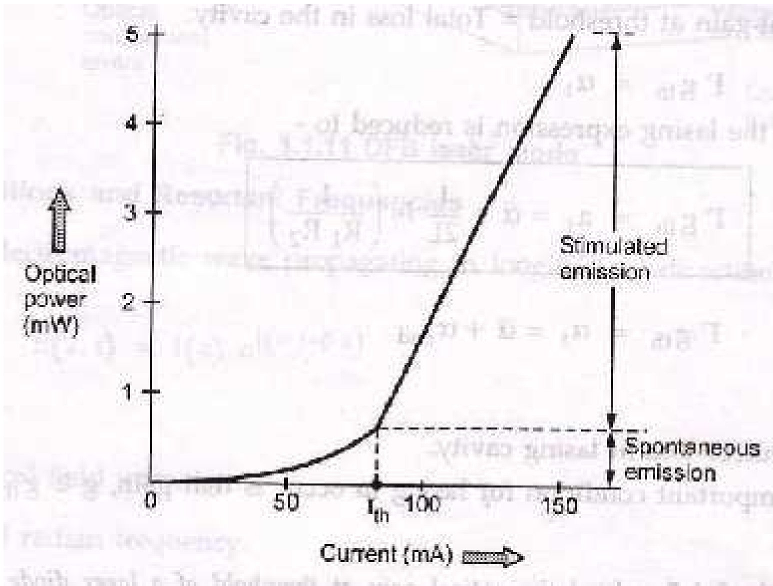


Figure Power Current Characteristics

Power current characteristics

External Quantum Efficiency

The external quantum efficiency is defined as the number of photons emitted per electron hole pair recombination above threshold point. The external quantum efficiency η_{ext} is given by

where, $$\eta_{ext} = \frac{\eta_i (g_{th} - \alpha)}{g_{th}}$$

η_i = Internal quantum efficiency (0.6-0.7). g_{th} = Threshold gain.

α = Absorption coefficient

Typical value of η_{ext} for standard semiconductor laser is ranging between 15-20 %.

Resonant Frequencies

At threshold lasing

$2\beta\, L = 2\pi\, m$

where, (propagation constant) m is an integer.

$$= 2L.\frac{n}{\lambda}$$

Since $c = v\lambda$

$$\lambda = \frac{c}{v}$$

Substituting λ in above

$$m = 2L\ \frac{nv}{c}$$

Gain in any laser is a function of frequency. For a Gaussian output the gain and frequency are related by expression

$$g(\lambda) = g(0)e^{\left[-\frac{(\lambda-\lambda_0)^2}{2\sigma^2}\right]}$$

where, g(0) is maximum gain. λ_0 is center wavelength in spectrum is spectral width of the gain. The frequency spacing between the two successive modes is

$$\Delta v = \frac{c}{2\,L\,n}$$

$$\Delta\lambda = \frac{\lambda^2}{2\,L\,n}$$

Advantages of Laser Diode

- Simple economic design.
- High optical power.
- Production of light can be precisely controlled.
- Can be used at high temperatures.
- Better modulation capability.
- High coupling efficiency.
- Low spectral width (3.5 nm)
- Ability to transmit optical output powers between 5 and 10 mW.
- Ability to maintain the intrinsic layer characteristics over long periods.

Disadvantages of Laser Diode

- At the end of fiber, a speckle pattern appears as two coherent light beams add or subtract their electric field depending upon their relative phases.
- Laser diode is extremely sensitive to overload currents and at high transmission rates, when laser is required to operate continuously the use of large drive current produces unfavourable thermal characteristics and necessitates the use of cooling and power stabilization.

Comparison of LED and Laser Diode

S. No.	Parameter	LED	LD (Laser Diode)
1	Principle of operation	Spontaneous emission.	Stimulated emission.
2	Output beam	Non – coherent	Coherent.
3	Spectral width	Board spectrum (20 nm – 100 nm)	Much narrower (1-5nm)
4	Data rate	Low.	Very high.
5	Transmission	Smaller.	Greater.
6	Temperature	Less sensitive.	More sensitive
7	Coupling efficiency	Very low.	High.
8	Compatible fibers	Multimode step index multimode	Single mode Sl
9	Circuit complexity	Simple	Complex
10	Life time	105 hours.	104 hours.
11	Cost	Low.	High.
12	Output power	Linearly proportional to	Proportional to current
13	Current required	Drive current 50 to 100 mA peak.	Threshold current 5 to
14	Applications	Moderate distance low data rate.	Long distance high data

Power Launching and coupling

Optical output of a luminescent source is usually measured by its radiance B at a given diode current.

Radiance: It is the optical power radiated into a unit solid angle per unit emitting surface area and is generally specified in terms of watts per square centimeter per steradian. Radiance = Power / per unit solid angle x per unit emitting surface area

Solid angle is defined by the projected area of a surface patch onto a unit sphere of a point. The angle that, seen from the center of a sphere, includes a given area on the surface of that sphere. The value of the solid angle is numerically equal to the size of that area divided by the square of the radius of the sphere.

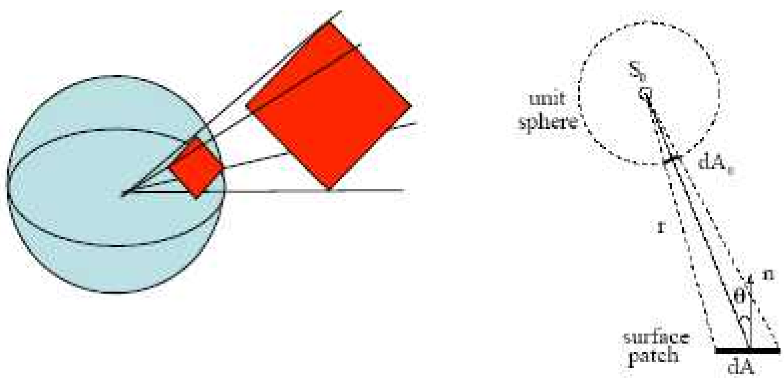


Figure Power launching

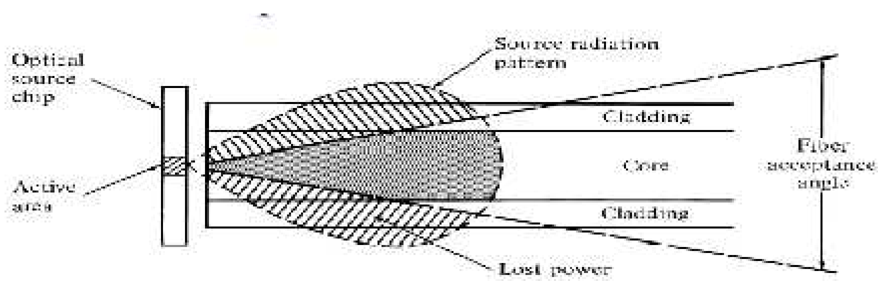


Figure Power coupled from source to fiber

Total power from LED to step index fiber

$$P_s = A_s \int_0^{2\pi}\int_0^{\pi/2} B(\theta,\phi)\sin\theta \, d\theta \, d\phi$$

$$P_s = \pi r_s^2 \, 2\pi B_0 \int_0^{\pi/2} \cos\theta \sin\theta \, d\theta = \pi^2 r_s^2 B_0$$

$$P_{\text{LED,step}} = \begin{cases} P_s(\text{NA})^2 & \text{if } r_s \le a \\ \left(\dfrac{a}{r_s}\right)^2 P_s(\text{NA})^2 & \text{if } r_s \ge a \end{cases}$$

Power coupled from the LED to the graded indexed fiber is given as

$$P_{LED,gin} = 2\pi^2 B_o \int_0^{r_s} \left[n^2(r) - n_2^2\right] r\,dr; \quad n(r) = \text{NA}(0)\sqrt{1-(r/a)^\alpha}$$

$$= 2P_s n_1^2 \Delta \left[1 - \frac{2}{\alpha+2}\left(\frac{r_s}{a}\right)^\alpha\right]$$

If the medium between source and fiber is different from the core material with refractive index n, the power coupled into the fiber will be reduced by the factor

$$R = \left(\frac{n_1 - n}{n_1 + n}\right)^2$$

Lensing Scheme for Coupling Improvement

Several Possible lensing schemes are:

1. Rounded end fiber

2. Nonimaging Microsphere (small glass sphere in contact with both the fiber and source)

3. Imaging sphere (a larger spherical lens used to image the source on the core area of the fiber end)

4. Cylindrical lens (generally formed from a short section of fiber)

5. Spherical surfaced LED and spherical ended fiber

6. Taper ended fiber.

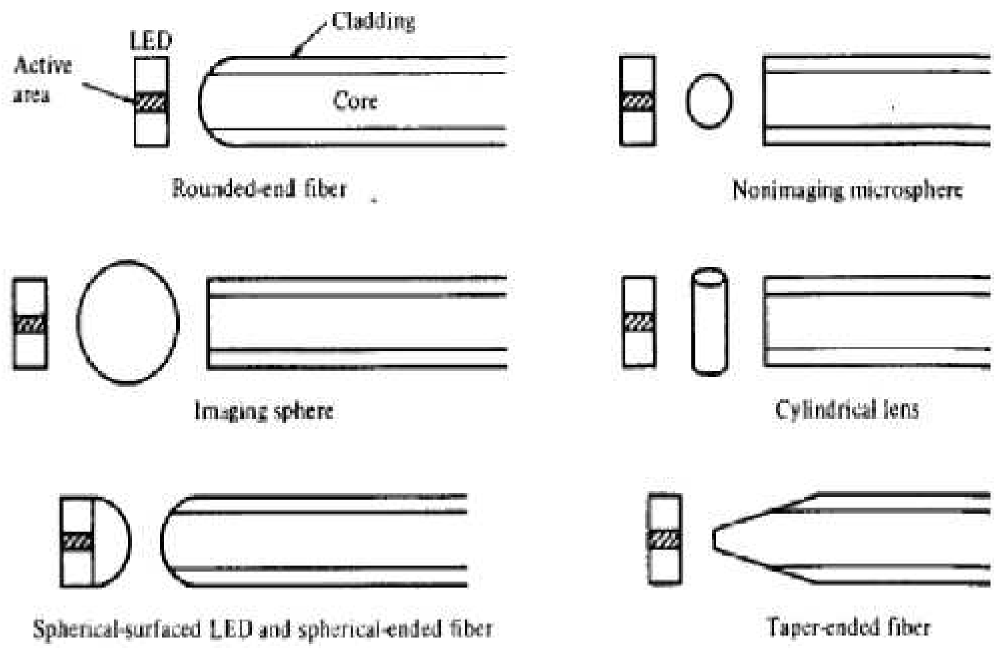


Figure Examples of possible lensing scheme used to improve optical source to fiber coupling efficiency

Optical fiber connectors

An optical fiber connector terminates the end of an optical fiber, and enables quicker connection and disconnection than splicing. The connectors mechanically couple and align the cores of fibers so light can pass. Better connectors lose very little light due to reflection or misalignment of the fibers. In all, about 100 different types of fiber optic connectors have been introduced to the market. An optical fiber connector is a flexible device that connects fiber cables requiring a quick connection and disconnection. Optical fibers terminate fiber-optic connections to fiber equipment or join two fiber connections without splicing. Hundreds of optical fiber connector types are available, but the key differentiator is defined by the mechanical coupling techniques and dimensions. Optical fiber connectors ensure stable connections, as they ensure the fiber ends are optically smooth and the end-to-end positions are properly aligned.

An optical fiber connector is also known as a fiber optic connector. 1980s. Most fiber connectors are spring loaded. The main components of an optical fiber connector are a ferrule, sub-assembly body, cable, stress relief boot and connector housing. The ferrule is mostly made of hardened material like stainless steel and tungsten carbide, and it ensures the alignment during connector mating. The connector body holds the ferrule and the coupling device serves the purpose of male-female configuration

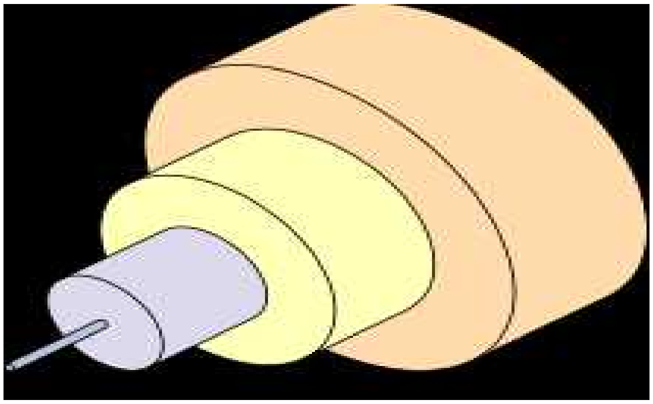
The fiber types for fiber optic connectors are categorized into simplex, duplex and multiple fiber connectors. A simplex connector has one fiber terminated in the connector, whereas duplex has two fibers terminated in the connector. Multiple fiber connectors can have two or more fibers terminated in the connector. Optical fiber connectors are dissimilar to other electronic connectors in that they do not have a jack and plug design. Instead they make use of the fiber mating sleeve for connection purposes.

Common optical fiber connectors include biconic, D4, ESCON, FC, FDDI, LC and SC.

- Biconic connectors use precision tapered ends to have low insertion loss.
- D4 connectors have a keyed body for easy intermateability.
- ESCON connectors are commonly used to connect from a wall outlet to a device.

- FC connector (fixed connection connector) is used for single-mode fibers and high-speed communication links.
- FDDI connector is a duplex connector which makes use of a fixed shroud.
- LC connector (local connection connector) has the benefit of small-form-factor optical transmitter/receiver assemblies and is largely used in private and public networks.
- SC connector (subscriber connector) is used in simplex and multiple applications and is best suited for high-density applications.

In fiber-optic communication, a single-mode optical fiber (SMF) is an optical fiber designed to carry only a single mode of light - the transverse mode. Modes are the possible solutions of the Helmholtz equation for waves, which is obtained by combining Maxwell's equations and the boundary conditions. These modes define the way the wave travels through space, i.e. how the wave is distributed in space. Waves can have the same mode but have different frequencies. This is the case in single-mode fibers, where it can have waves with different frequencies, but of the same mode, which means that they are distributed in space in the same way, and that gives us a single ray of light. Although the ray travels parallel to the length of the fiber, it is often called transverse mode since its electromagnetic oscillations occur perpendicular (transverse) to the length of the fiber.



The structure of a typical single-mode fiber.

1. Core 8 -9 µm diameter

2.Cladding 125 µm diameter

3.Buffer 250 µm diameter

4. Jacket 900 µm diameter

Like multi-mode optical fibers, single-mode fibers do exhibit modal dispersion resulting from multiple spatial modes but with narrower modal dispersion. Single-mode fibers are therefore better at retaining the fidelity of each light pulse over longer distances than multi-mode fibers. For these reasons, single-mode fibers can have a higher bandwidth than multi-mode fibers. Equipment for single-mode fiber is more expensive than equipment for multi-mode optical fiber, but the single-mode fiber itself is usually cheaper in bulk.

Cross section of a single-mode optical fiber patch cord end, taken with a Fiberscope. The round circle is the cladding, 125 microns in diameter. Debris is visible as a streak on the cross- section, and glows due to the illumination.

A typical single-mode optical fiber has a core diameter between 8 and 10.5 µm and a cladding diameter of 125 µm. There are a number of special types of single-mode optical fiber which have been chemically or physically altered to give special properties, such as dispersion-shifted fiber and nonzero dispersion-shifted fiber. Data rates are limited by polarization mode dispersion and chromatic dispersion. As of 2005, data rates of up to 10 gigabits per second were possible at distances of over 80 km (50 mi) with commercially available transceivers (Xenpak). By using optical amplifiers and dispersion-compensating devices, state-of-the-art DWDM optical systems can span thousands of kilometers at 10 Gbit/s, and several hundred kilometers at 40 Gbit/s

The lowest-order bounds mode is ascertained for the wavelength of interest by solving Maxwell's equations for the boundary conditions imposed by the fiber, which are determined by the core diameter and the refractive indices of the core and cladding. The solution of Maxwell's equations for the lowest order bound mode will permit a pair of orthogonally polarized fields in the fiber, and this is the usual case in a communication fiber.

In step-index guides, single-mode operation occurs when the normalized frequency, *V*, is less than or equal to 2.405. For power-law profiles, single-mode operation occurs for a normalized frequency, *V*, less than approximately where *g* is the profile parameter.

,

Cross section of a single-mode optical fiber patch cord end, taken with a Fiberscope. The round circle is the cladding, 125 microns in diameter. Debris is visible as a streak on the cross- section, and glows due to the illumination.

In practice, the orthogonal polarizations may not be associated with degenerate modes.OS1 and OS2 are standard single-mode optical fiber used with wavelengths 1310 nm and 1550 nm (size 9/125 µm) with a maximum attenuation of 1 dB/km (OS1) and 0.4 dB/km (OS2). OS1 is defined in ISO/IEC 11801 and OS2 is defined in ISO/IEC 24702.

Optical fiber connectors

Optical fiber connectors are used to join optical fibers where a connect/disconnect capability is required. The basic connector unit is a connector assembly. A connector assembly consists of an adapter and two connector plugs. Due to the sophisticated polishing and tuning procedures that may be incorporated into optical connector manufacturing, connectors are generally assembled onto optical fiber in a supplier's manufacturing facility. However, the assembly and polishing operations involved can be performed in the field, for example to make cross-connect jumpers to size.

Optical fiber connectors are used in telephone company central offices, at installations on customer premises, and in outside plant applications. Their uses include:

- Making the connection between equipment and the telephone plant in the central office
- Connecting fibers to remote and outside plant electronics such as Optical Network Units (ONUs) and Digital Loop Carrier (DLC) systems
- Optical cross connects in the central office
- Patching panels in the outside plant to provide architectural flexibility and to interconnect fibers belonging to different service providers

- Connecting couplers, splitters, and Wavelength Division Multiplexers (WDMs) to optical fibers
- Connecting optical test equipment to fibers for testing and maintenance.

Outside plant applications may involve locating connectors underground in subsurface enclosures that may be subject to flooding, on outdoor walls, or on utility poles. The closures that enclose them may be hermetic, or may be "free-breathing." Hermetic closures will prevent the connectors within being subjected to temperature swings unless they are breached. Free-breathing enclosures will subject them to temperature and humidity swings, and possibly to condensation and biological action from airborne bacteria, insects, etc. Connectors in the underground plant may be subjected to groundwater immersion if the closures containing them are breached or improperly assembled.

The latest industry requirements for optical fiber connectors are in Telcordia GR-326, Generic Requirements for Single mode Optical Connectors and Jumper Assemblies. A multi-fiber optical connector is designed to simultaneously join multiple optical fibers together, with each optical fiber being joined to only one other optical fiber. The last part of the definition is included so as not to confuse multi-fiber connectors with a branching component, such as a coupler.

The latter joins one optical fiber to two or more other optical fibers. Multi-fiber optical connectors are designed to be used wherever quick and/or repetitive connects and disconnects of a group of fibers are needed. Applications include telecommunications companies' Central Offices (COs), installations on customer premises, and Outside Plant (OSP) applications. The multi-fiber optical connector can be used in the creation of a low-cost switch for use in fiber optical testing. Another application is in cables delivered to a user with pre- terminated multi-fiber jumpers. This would reduce the need for field splicing, which could greatly reduce the number of hours necessary for placing an optical fiber cable in a telecommunications network. This, in turn, would result in savings for the installer of such cable.

The return loss R_L is a measure of the portion of light that is reflected back to the source at the junction. It is expressed in decibel. The higher the RL value in decibels, the lower is the reflections. Typical R_L values lie between 35 and 50 dB for PC, 60 to 90 dB for APC and 20 to 40 dB for multimode fibres. In the early days of fibre-optic plug-in connectors, the abutting end faces were polished to an angle of 90° to the fibre axis, while current standards require PC (Physical Contact) polishing or APC (Angled Physical Contact) polishing. The term HRL (High Return Loss) is frequently used, but it has the same meaning as APC.

In PC polishing, the ferrule is polished to a convex end to ensure that the fibre cores touch at their highest point. This reduces the occurrence of reflections at the junction. A further improvement in return loss is achieved by using the APC polishing technique. Here, the convex end surfaces of the ferrules are polished to an angle (8°) relative to the fibre axis. SC connectors are also sold with a 9° angle. They possess I_L and R_L values identical to 8° versions, and for this reason they have not established themselves worldwide.

Return Loss

In optics (particularly in fiber optics) a loss that takes place at discontinuities of refractive index, especially at an air-glass interface such as a fiber end face. At those interfaces, a fraction of the optical signal is reflected back toward the source. This reflection phenomenon is also called "Fresnel reflection loss," or simply "Fresnel loss."

Fiber optic transmission systems use lasers to transmit signals over optical fiber, and a high optical return loss (ORL) can cause the laser to stop transmitting correctly. The measurement of ORL is becoming more important in the characterization of optical networks as the use of wavelength-division multiplexing increases. These systems use lasers that have a lower tolerance for ORL, and introduce elements into the network that are located in close proximity to the laser.

Definition of Return Loss

In technical terms, RL is the ratio of the light reflected back from a device under test, P_{out}, to the light launched into that device, P_{in}, usually expressed as a negative number in dB.

$$R_L = 10 \log_{10}(P_{out}/P_{in})$$

where P_{out} is the reflected power and P_{in} is the incident, or input, power.
Sources of loss include reflections and scattering along the fiber network. A typical R_L value for an Angled Physical Contact (APC) connector is about -55dB, while the R_L from an open flat polish to air is typically about -14dB. High R_L is a large concern in high bit rate digital or analog single mode systems and is also an indication of a potential failure point, or compromise, in any optical network.

Fiber alignment and joint loss

A major consideration with all types of fiber–fiber connection is the optical loss encountered at the interface. Even when the two jointed fiber ends are smooth and perpendicular to the fiber axes, and the two fiber axes are perfectly aligned, a small proportion of the light may be reflected back into the transmitting fiber causing attenuation at the joint. This phenomenon, known as Fresnel reflection,

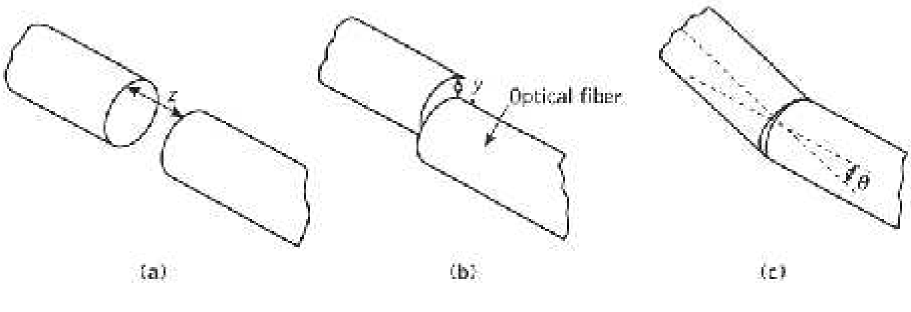


Figure the three possible types of misalignment which may occur when jointing compatible optical fibers: (a) longitudinal misalignment; (b) lateral misalignment; (c) angular misalignment

It is apparent that Fresnel reflection may give a significant loss at a fiber joint even when all other aspects of the connection are ideal. However, the effect of Fresnel reflection at a fiber–fiber connection can be reduced to a very low level through the use of an index-matching fluid in the gap between the jointed fibers. When the index-matching fluid has the same refractive index as the fiber core, losses due to Fresnel reflection are in theory eradicated.

Unfortunately, Fresnel reflection is only one possible source of optical loss at a fiber joint. A potentially greater source of loss at a fiber–fiber connection is caused by misalignment of the two jointed fibers. In order to appreciate the development and relative success of various connection techniques it is useful to discuss fiber alignment in greater detail.

Any deviations in the geometrical and optical parameters of the two optical fibers which are jointed will affect the optical attenuation (insertion loss) through the connection. It is not possible within any particular connection technique to allow for all these variations. Hence, there are inherent connection problems when jointing fibers with, for instance

- ✓ different core and/or cladding diameters;
- ✓ different numerical apertures and/or relative refractive index differences;
- ✓ different refractive index profiles;
- ✓ fiber faults (core ellipticity, core concentricity, etc.).

The losses caused by the above factors together with those of Fresnel reflection are usually referred to as intrinsic joint losses. The best results are therefore achieved with compatible (same)

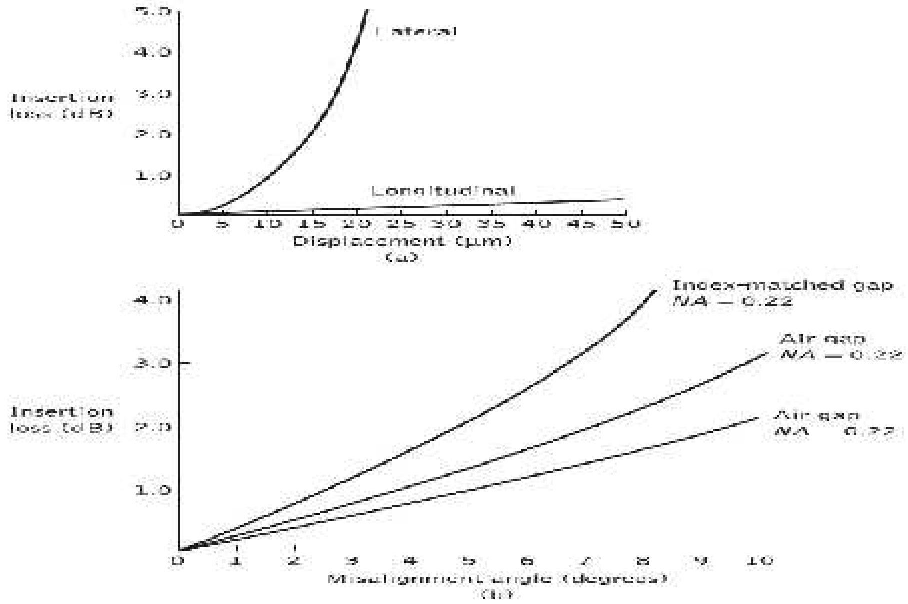


Figure Insertion loss characteristics for jointed optical fibers with various types of misalignment: (a) insertion loss due to lateral and longitudinal misalignment for a graded index fiber of 50 µm core diameter.

In this case there is still the problem of the quality of the fiber alignment provided by the jointing mechanism. Examples of possible misalignment between coupled compatible optical fibers are illustrated in Figure. It is apparent that misalignment may occur in three dimensions: the separation between the fibers (longitudinal misalignment), the offset perpendicular to the fiber core axes (lateral/radial/ axial misalignment) and the angle between the core axes (angular misalignment). Optical losses resulting from these three types of misalignment depend upon the fiber type, core diameter and the distribution of the optical power between the propagating modes.

Examples of the measured optical losses due to the various types of misalignment are shown in Figure and shows the attenuation characteristic for both longitudinal and lateral misalignment of a graded index fiber of 50 µm core diameter.

It may be observed that the lateral misalignment gives significantly greater losses per unit displacement than the longitudinal misalignment. For instance, in this case a lateral displacement of 10 µm gives about 1 dB insertion loss whereas a similar longitudinal displacement gives an insertion loss of around 0.1 dB. Figure shows the attenuation characteristic for the angular misalignment of two multimode step index fibers with numerical apertures of 0.22 and 0.3. An insertion loss of around 1 dB is obtained with angular misalignment of 4° and 5° for the NA =0.22 and NA =0.3 fibers respectively It may also be observed in figure that the effect of an index-matching fluid in the fiber gap causes increased losses with angular misalignment. Therefore, it is clear that relatively small levels of lateral and/or angular misalignment can cause significant attenuation at a fiber joint. This is especially the case for fibers of small core diameter (less than 150 µm) which are currently employed for most telecommunication purposes.

Multimode fiber joints

Theoretical and experimental studies of fiber misalignment in optical fiber connections allow approximate determination of the losses encountered with the various misalignments of different fiber types. Here some of the expressions used to calculate losses due to lateral and angular misalignment of optical fiber joints. Longitudinal misalignment is not discussed in detail as it tends to be the least important effect and may be largely avoided in fiber connection. Both groups of workers claim good agreement with experimental results, which is perhaps understandable when considering the number of variables involved in the measurement. Also, all groups predict higher losses for fibers with larger numerical apertures, which is consistent with intuitive considerations (i.e. the larger the numerical aperture, the greater the spread of the output light and the higher the optical loss at a longitudinally misaligned joint).

Theoretical expressions for the determination of lateral and angular misalignment losses are by no means definitive, although in all cases they claim reasonable agreement with experimental results. However, experimental results from different sources tend to vary (especially for angular misalignment losses) due to difficulties of measurement. It is therefore not implied that the expressions given in the text are necessarily the most accurate, as at present the choice appears somewhat arbitrary. Lateral misalignment reduces the overlap region between the two fiber cores. Assuming uniform excitation of all the optical modes in a multimode step index fiber, the overlapped area between both fiber cores approximately gives the lateral coupling efficiency µlat. Hence, the lateral coupling efficiency for two similar step index fibers may be written as

$$\eta_{lat} \simeq \frac{16(n_1/n)^2}{[1+(n_1/n)]^4}\frac{1}{\pi}\left\{2\cos^{-1}\left(\frac{y}{2a}\right)-\left(\frac{y}{a}\right)\left[1-\left(\frac{y}{2a}\right)^2\right]^{\frac{1}{2}}\right\}$$

where n_1 is the core refractive index, n is the refractive index of the medium between the fibers, y is the lateral offset of the fiber core axes, and a is the fiber core radius. The lateral misalignment loss in decibels may be determined using:

$$\text{Loss}_{lat} = -10\log_{10}\eta_{lat}\ \text{dB}$$

The predicted losses obtained using the formulas given are generally slightly higher than the measured values due to the assumption that all modes are equally excited. This assumption is only correct for certain cases of optical fiber transmission. Also, certain authors assume index matching and hence no Fresnel reflection, which makes the first term in equation equal to unity (as $n_1/n = 1$). This may be valid if the two fiber ends are assumed to be in close contact (i.e. no air gap in between) and gives lower predicted losses. Nevertheless, bearing in mind these possible inconsistencies, useful estimates for the attenuation due to lateral misalignment of multimode step index fibers may be obtained. Lateral misalignment loss in multimode graded index fibers assuming a uniform distribution of optical power throughout all guided modes was calculated by Gloge. He estimated that the lateral misalignment loss was dependent on the refractive index gradient α for small lateral offset and may be obtained from:

$$L_t = \frac{2}{\pi}\left(\frac{y}{a}\right)\left(\frac{\alpha+2}{\alpha+1}\right) \quad \text{for } 0 \le y \le 0.2a$$

Where the lateral coupling efficiency was given by:

$$\eta_{lat} = 1 - L_t$$

Hence it may be utilized to obtain the lateral misalignment loss in decibels. With a parabolic refractive index profile where α = 2, gives: A further estimate including the leaky modes gave a revised expression for the lateral misalignment loss given in Equation of 0.75(y/a). This analysis was also extended to step index fibers (where α = ∞) and gave lateral misalignment losses of 0.64(y/a) and 0.5(y/a) for the cases of guided modes only and both guided plus leaky modes respectively.

Factors causing fiber–fiber intrinsic losses were listed in previous Section; the major ones comprising a mismatch in the fiber core diameters, a mismatch in the fiber numerical apertures and differing fiber refractive index profiles are illustrated in Figure. Connections between multimode fibers with certain of these parameters being different can be quite common, particularly when a pigtailed optical source is used, the fiber pigtail of which has different characteristics from the main transmission fiber. Moreover, as indicated previously, diameter variations can occur with the same fiber type. Assuming all the modes are equally excited in a multimode step or graded index fiber, and that the numerical apertures and index profiles are the same, then the loss resulting from a mismatch of core diameters is given by:

$$\text{Loss}_{CD} = \begin{cases} -10\log_{10}\left(\frac{a_2}{a_1}\right)^2 \text{ (dB)} & a_2 < a_1 \\ 0 \text{ (dB)} & a_2 \ge a_1 \end{cases}$$

where a1 and a2 are the core radii of the transmitting and receiving fibers respectively. It may be observed from Equation that no loss is incurred if the receiving fiber has a larger core diameter than the transmitting one. In addition, only a relatively small loss (0.09 dB) is obtained when the receiving fiber core diameter is 1% smaller than that of the transmitting fiber. When the transmitting fiber has a higher numerical aperture than the receiving fiber, then some of the emitted light rays will fall outside the acceptance angle of the receiving fiber and they will therefore not be coupled through the joint.

Again assuming a uniform modal power distribution, and fibers with equivalent refractive index profiles and core diameters, then the loss caused by a mismatch of numerical apertures

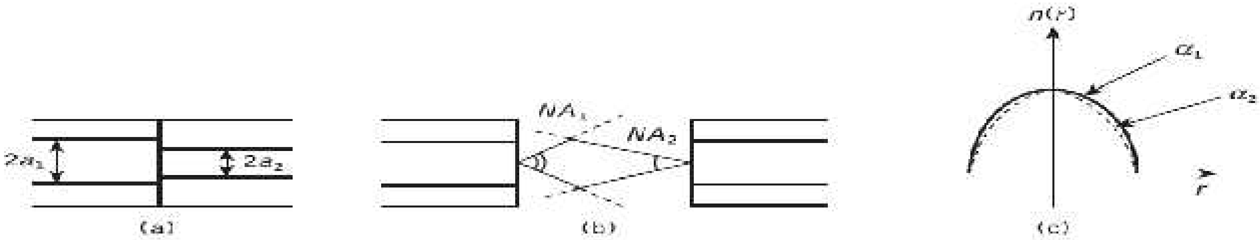


Figure Some intrinsic coupling losses at fiber joints: (a) core diameter mismatch; (b) numerical aperture mismatch; (c) refractive index profile difference

$$\text{Loss}_{NA} = \begin{cases} -10 \log_{10}\left(\dfrac{NA_2}{NA_1}\right)^2 \text{ (dB)} & NA_2 < NA_1 \\ 0 \text{ (dB)} & NA_2 \geq NA_1 \end{cases}$$

$$\text{Loss}_{RI} = \begin{cases} -10 \log_{10}\dfrac{\alpha_2(\alpha_1+2)}{\alpha_1(\alpha_2+2)} \text{ (dB)} & \alpha_2 < \alpha_1 \\ 0 \text{ (dB)} & \alpha_2 \geq \alpha_1 \end{cases}$$

$$\text{Loss}_{int} = \begin{cases} -10 \log_{10}\dfrac{(a_2 NA_2)^2(\alpha_1+2)\alpha_2}{(a_1 NA_1)^2(\alpha_2+2)\alpha_1} \text{ (dB)} & a_2 > a_1,\ NA_2 > NA_1,\ \alpha_2 > \alpha_1 \\ 0 \text{ (dB)} & a_2 \leq a_1,\ NA_2 \leq NA_1,\ \alpha_2 \leq \alpha_1 \end{cases}$$

Single-mode fiber joints

Misalignment losses at connections in single-mode fibers have been theoretically considered by Marcuse and Gambling. The theoretical analysis which was instigated by Marcuse is based upon the Gaussian or near-Gaussian shape of the modes propagating in single-mode fibers regardless of the fiber type (i.e. step index or graded index). Further development of this theory by Gambling *et al.* gave simplified formulas for both the lateral and angular misalignment losses at joints in single mode fibers. In the absence of angular misalignment Gambling *et al.* calculated that the loss T_l due to lateral offset y was given by:

$$T_l = 2.17\left(\frac{y}{\omega}\right)^2 \text{dB}$$

Where ω is the normalized spot size of the fundamental mode. However, the normalized spot size for the LP01 mode (which corresponds to the HE mode) may be obtained from the empirical formula:

$$\omega = a\,\frac{(0.65 + 1.62\,V^{-\frac{3}{2}} + 2.88\,V^{-6})}{2^{\frac{1}{2}}}$$

where ω is the spot size in μm, a is the fiber core radius and V is the normalized frequency for the fiber. Alternatively, the insertion loss T_a caused by an angular misalignment θ (in radians) at a joint in a single-mode fiber may be given by

$$T_a = 2.17\left(\frac{\theta\omega\, n_1 V}{aNA}\right)^2 \text{dB}$$

where n_1 is the fiber core refractive index and *NA* is the numerical aperture of the fiber. It must be noted that the formulas given in assume that the spot sizes of the modes in the two coupled fibers are the same. Gambling *et al.* also derived a somewhat complicated formula which gave a good approximation for the combined losses due to both lateral and angular misalignment at a fiber joint. However, they indicate that for small total losses (less than 0.75 dB) a reasonable approximation is obtained by simply combining Assuming that no losses are present due to the extrinsic factors, the intrinsic coupling loss is given by where ω_{01} and ω_{02} are the spot sizes of the transmitting and receiving fibers respectively. Equation therefore enables the additional coupling loss resulting from mode-field diameter mismatch between two single-mode fibers to be calculated.

Fiber splices

A permanent joint formed between two individual optical fibers in the field or factory is known as a fiber splice. Fiber splicing is frequently used to establish long-haul optical fiber links where smaller fiber lengths need to be joined, and there is no requirement for repeated connection and disconnection. Splices may be divided into two broad categories depending upon the splicing technique utilized. These are fusion splicing or welding and mechanical splicing.

Fusion splicing is accomplished by applying localized heating (e.g. by a flame or an electric arc) at the interface between two butted, prealigned fiber ends causing them to soften and fuse. Mechanical splicing, in which the fibers are held in alignment by some mechanical means, may be achieved by various methods including the use of tubes around the fiber ends (tube splices) or V-grooves into which the butted fibers are placed (groove splices). All these techniques seek to optimize the splice performance (i.e. reduce the insertion loss at the joint) through both fiber end preparation and alignment of the two joint fibers.

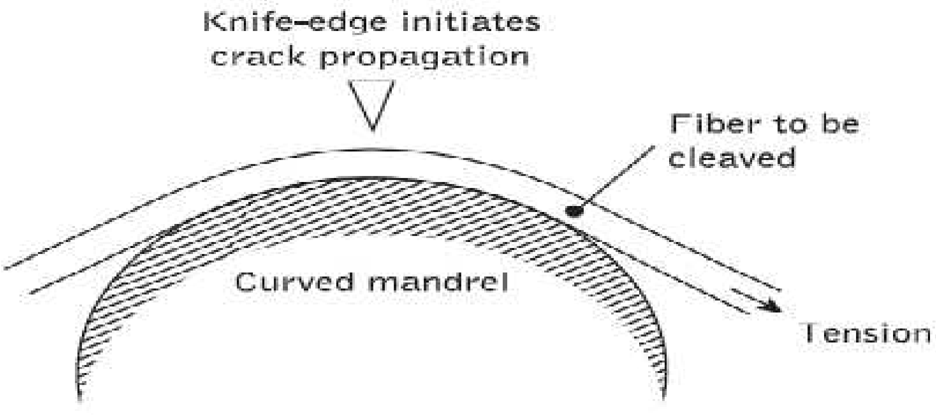


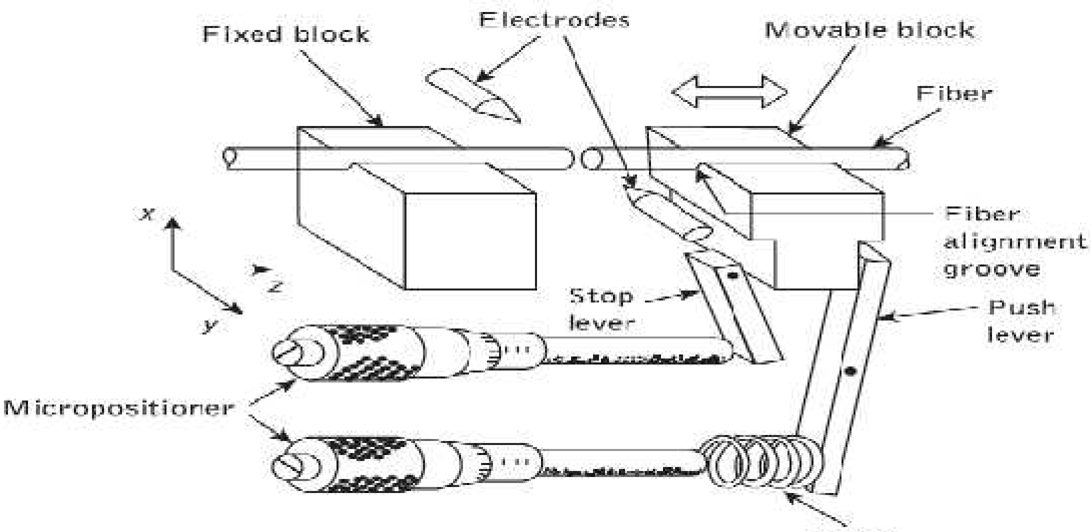
Figure Optical fiber end preparation: the principle of scribe and break cutting

Typical average splice insertion losses for multimode fibers are in the range 0.1 to 0.2 dB which is generally a better performance than that exhibited by demountable connections. It may be noted that the insertion losses of fiber splices are generally much less than the possible Fresnel reflection loss at a butted fiber–fiber joint. This is because there is no large step change in refractive index with the fusion splice as it forms a continuous fiber connection, and some method of index matching (e.g. a fluid) tends to be utilized with mechanical splices.

A requirement with fibers intended for splicing is that they have smooth and square end faces. In general this end preparation may be achieved using a suitable tool which cleaves the fiber as illustrated in Figure. This process is often referred to as scribe and break or score and break as it involves the scoring of the fiber surface under tension with a cutting tool (e.g. sapphire, diamond, tungsten carbide blade). The surface scoring creates failure as the fiber is tensioned and a clean, reasonably square fiber end can be produced. Figure illustrates this process with the fiber tensioned around a curved mandrel. However, straight pull, scribe and break tools are also utilized, which arguably give better results.

Fusion splices

The fusion splicing of single fibers involves the heating of the two prepared fiber ends to their fusing point with the application of sufficient axial pressure between the two optical fibers. It is therefore essential that the stripped (of cabling and buffer coating) fiber ends are adequately positioned and aligned in order to achieve good continuity of the transmission medium at the junction point. Hence the fibers are usually positioned and clamped with the aid of an inspection microscope. Flame heating sources such as microplasma torches (argon and hydrogen) and oxhydric microburners (oxygen, hydrogen and alcohol vapor) have been utilized with some success. However, the most widely used heating source is an electric arc. This technique offers advantages of consistent, easily controlled heat with adaptability for use under field conditions. which involves the rounding of the fiber ends with a low-energy discharge before pressing the fibers together and fusing with a stronger arc. This technique, known as prefusion, removes the requirement for fiber end preparation which has a distinct advantage in the field environment. It has been utilized with multimode fibers giving average splice losses of 0.09 db.



(a)

Fusion splicing of single-mode fibers with typical core diameters between 5 and 10 μm presents problems of more critical fiber alignment (i.e. lateral offsets of less than 1 μm are required for low loss joints). However, splice insertion losses below 0.3 dB may be achieved due to a self-alignment phenomenon which partially compensates for any lateral offset.

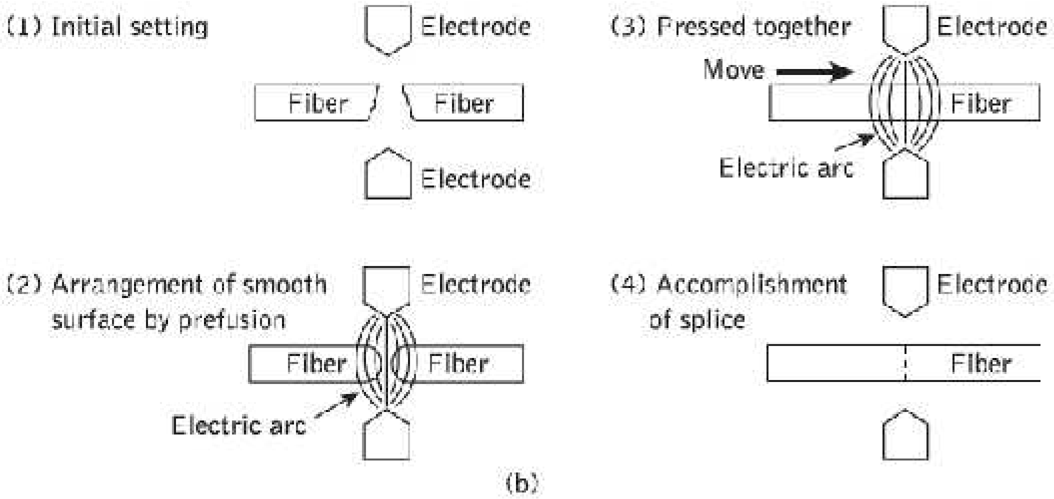


Figure Electric arc fusion splicing: (a) an example of fusion splicing apparatus; (b) schematic illustration of the prefusion method for accurately splicing optical fibers

Self-alignment, illustrated in Figure, is caused by surface tension effects between the two fiber ends during fusing. An early field trial of single-mode fiber fusion splicing over a 31.6 km link gave mean splice insertion losses of 0.18 and 0.12 dB at wavelengths of 1.3 and 1.55 μm respectively. Mean splice losses of only 0.06 dB have also been obtained with a fully automatic single-mode fiber fusion splicing machine weaken the fiber in the vicinity of the splice. It has been found that even with careful handling, the tensile strength of the fused fiber may be as low as 30% of that of the uncoated fiber before fusion.

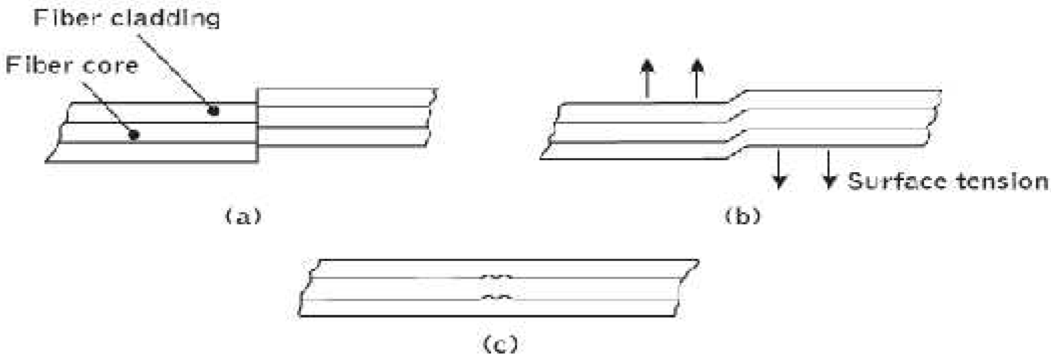


Figure Self-alignment phenomenon which takes place during fusion splicing: (a) before fusion; (b) during fusion; (c) after fusion

The fiber fracture generally occurs in the heat affected zone adjacent to the fused joint. The reduced tensile strength is attributed to the combined effects of surface damage caused by handling, surface defect growth during heating and induced residential stresses due to changes in chemical composition. It is therefore necessary that the completed splice is packaged so as to reduce tensile loading upon the fiber in the vicinity of the splice.

Mechanical splices

A number of mechanical techniques for splicing individual optical fibers have been developed. A common method involves the use of an accurately produced rigid alignment tube into which the prepared fiber ends are permanently bonded. This snug tube splice is illustrated in Figure and may utilize a glass or ceramic capillary with an inner diameter just large enough to accept the optical fibers. Transparent adhesive (e.g. epoxy resin) is injected through a transverse bore in the capillary to give mechanical sealing and index matching of the splice. Average insertion losses as low as 0.1 dB have been obtained with multimode graded index and single-mode fibers using ceramic capillaries. However, in general, snug tube splices exhibit problems with capillary tolerance requirements. Hence as a commercial product they may exhibit losses of up to 0.5 dB.

Mechanical splicing technique which avoids the critical tolerance requirements of the snug tube splice is shown in Figure. This loose tube splice uses an oversized square-section metal tube which easily accepts the prepared fiber ends. Transparent adhesive is first inserted into the tube followed by the fibers. The splice is self-aligning when the fibers are curved in the same plane, forcing the fiber ends simultaneously into the same corner of the tube, as indicated in Figure. Mean splice insertion losses of 0.073 dB have been achieved using multimode graded index fibers with the loose tube approach.

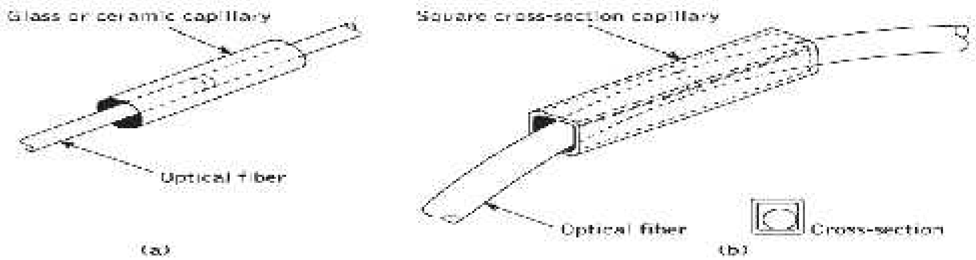


Figure Techniques for tube splicing of optical fibers: (a) snug tube splice; (b) loose tube splice utilizing square cross-section capillary

Other common mechanical splicing techniques involve the use of grooves to secure the fibers to be jointed. A simple method utilizes a V-groove into which the two prepared fiber ends are pressed. The V-groove splice which is illustrated in Figure gives alignment of the prepared fiber ends through insertion in the groove. The splice is made permanent by securing the fibers in the V-groove with epoxy resin. Jigs for producing Vgroove splices have proved quite successful, giving joint insertion losses of around 0.1 dB.

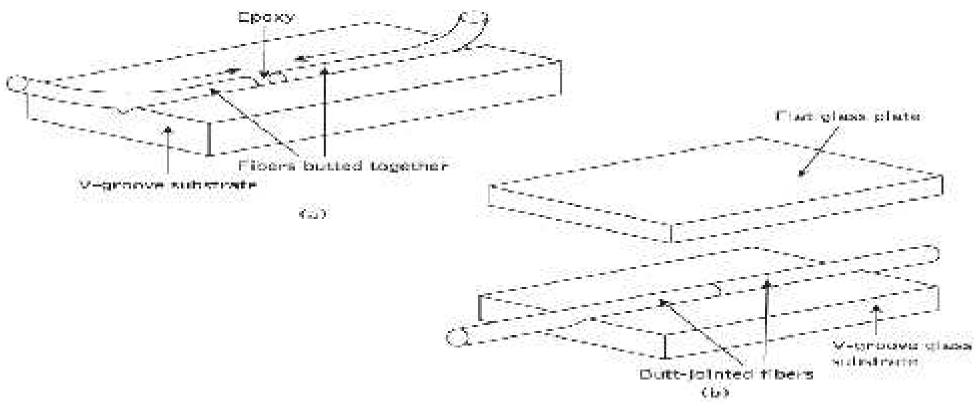


Figure V-groove splices

V-groove splices formed by sandwiching the butted fiber ends between a V-groove glass substrate a flat glass retainer plate, as shown in Figure, have also proved very successful in the laboratory. Splice insertion losses of less than 0.01 dB when coupling single-mode fibers have been reported using this technique. However, reservations are expressed regarding the field implementation of these splices with respect to manufactured fiber geometry, and housing of the splice in order to avoid additional losses due to local fiber bending.

A further variant on the V-groove technique is the elastic tube or elastomeric splice shown. The device comprises two elastomeric internal parts, one of which contains a V-groove. An outer sleeve holds the two elastic parts in compression to ensure alignment of the fibers in the V-groove, and fibers with different diameters tend to be centered and hence may be successfully spliced. Although originally intended for multimode fiber connection, the device has become a widely used commercial product which is employed with single-mode fibers, albeit often as a temporary splice for laboratory investigations. The splice loss for the elastic tube device was originally reported as 0.12 dB or less but is generally specified as around 0.25 dB for the commercial product. In addition, index-matching gel is normally employed within the device to improve its performance.

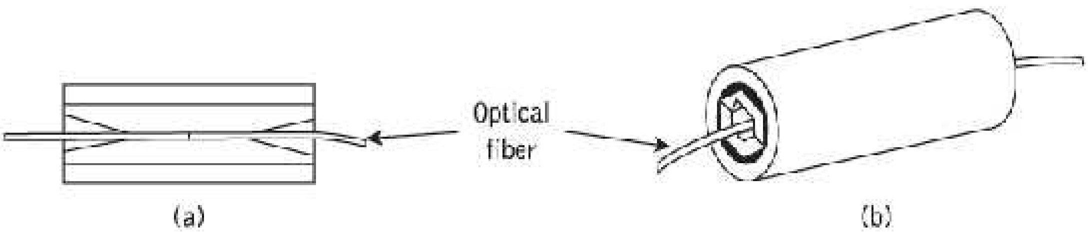


Figure The elastomeric splice: (a) cross-section; (b) assembly

A slightly more complex groove splice known as the Springroove splice utilized a bracket containing two cylindrical pins which serve as an alignment guide for the two prepared fiber ends. The cylindrical pin diameter was chosen to allow the fibers to protrude above the cylinders, as shown in Figure. An elastic element (a spring) was used to press the fibers into a groove and maintain the fiber end alignment, as illustrated in Figure. The complete assembly was secured using a drop of epoxy resin. Mean splice insertion losses of 0.05 dB were obtained using multimode graded index fibers with the Springroove splice. This device found practical use in Italy.

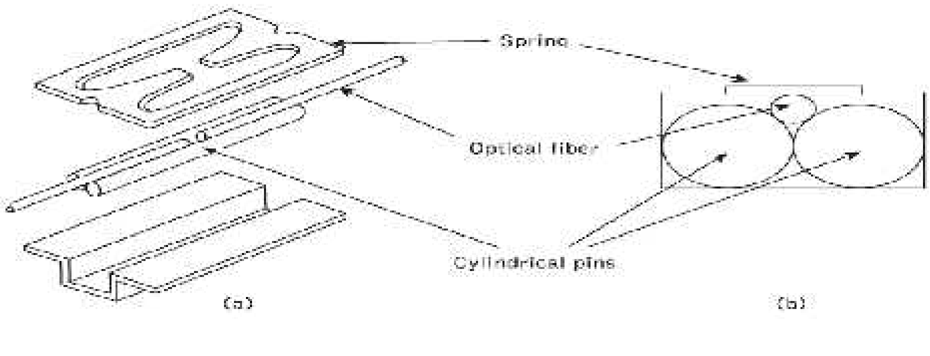
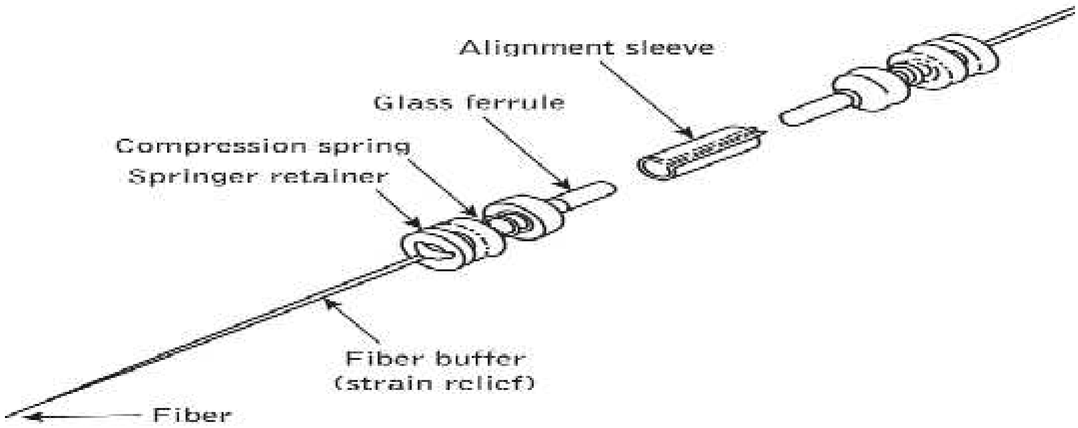


Figure the Springroove splice: (a) expanded overview of the splice; (b) schematic cross-section of the splice

An example of a secondary aligned mechanical splice for multimode fiber is shown in Figure. This device uses precision glass capillary tubes called ferrules as the secondary elements with an alignment sleeve of metal or plastic into which the glass tubed fibers are inserted. Normal assembly of the splice using 50 µm core diameter fiber yields an average loss of around 0.2 dB.



Fiber connectors

Demountable fiber connectors are more difficult to achieve than optical fiber splices. This is because they must maintain similar tolerance requirements to splices in order to couple light between fibers efficiently, but they must accomplish it in a removable fashion. Also, the connector design must allow for repeated connection and disconnection without problems of fiber alignment, which may lead to degradation in the performance of the transmission line at the joint. Hence to operate satisfactorily the demountable connector must provide reproducible accurate alignment of the optical fibers.

In order to maintain an optimum performance the connection must also protect the fiber ends from damage which may occur due to handling (connection and disconnection), must be insensitive to environmental factors (e.g. moisture and dust) and must cope with tensile load on the cable. Additionally, the connector should ideally be a low-cost component which can be fitted with relative ease. Hence optical fiber connectors may be considered in three major areas, which are:

- the fiber termination, which protects and locates the fiber ends;
- the fiber end alignment to provide optimum optical coupling;
- the outer shell, which maintains the connection and the fiber alignment, protects the fiber ends from the environment and provides adequate strength at the joint.

Cylindrical ferrule connectors

The basic ferrule connector (sometimes referred to as a concentric sleeve connector), which is perhaps the simplest optical fiber connector design. The two fibers to be connected are permanently bonded (with epoxy resin) in metal plugs known as ferrules which have an accurately drilled central hole in their end faces where the stripped (of buffer coating) fiber is located. Within the connector the two ferrules are placed in an alignment sleeve which, using accurately machined components, allows the fiber ends to be butt jointed. The ferrules are held in place via a retaining mechanism as shown in figure.

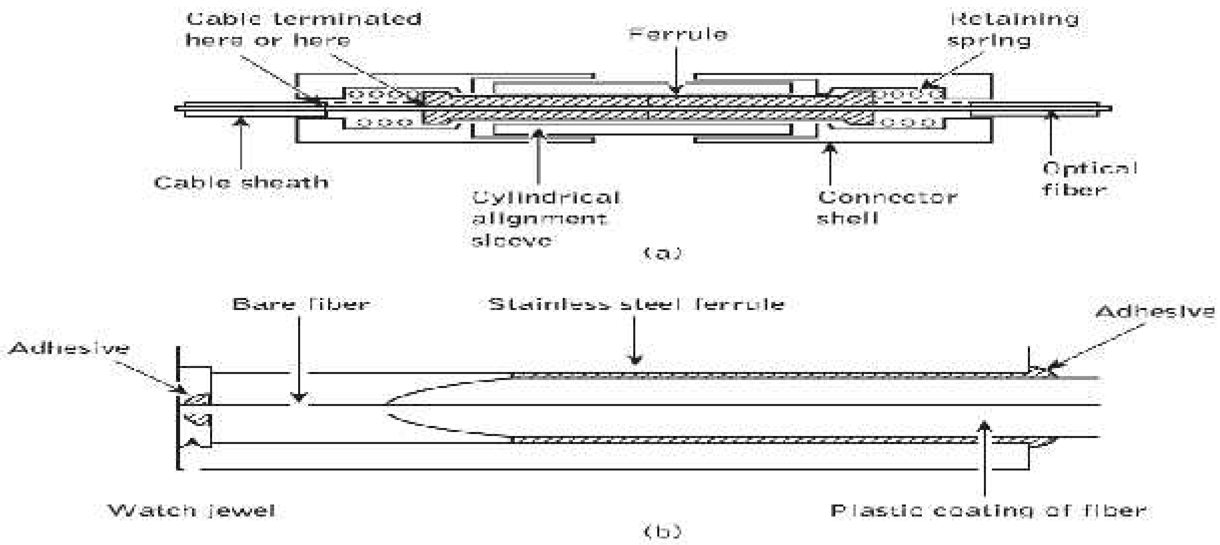


Figure Ferrule connectors: (a) structure of a basic ferrule connector; (b) structure of a watch jewel connector ferrule

It is essential with this type of connector that the fiber end faces are smooth and square (i.e. perpendicular to the fiber axis). This may be achieved with varying success by:

- cleaving the fiber before insertion into the ferrule;
- inserting and bonding before cleaving the fiber close to the ferrule end face;
- using either (a) or (b) and polishing the fiber end face until it is flush with the end of the ferrule.

Polishing the fiber end face after insertion and bonding provides the best results but it tends to be time consuming and inconvenient, especially in the field. The fiber alignment accuracy of the basic ferrule connector is largely dependent upon the ferrule hole into which the fiber is inserted. Hence, some ferrule connectors have incorporated a watch jewel in the ferrule end face (jeweled ferrule connector), as illustrated in Figure. In this case the fiber is centered with respect to the ferrule through the watch jewel hole. The use of the watch jewel allows the close diameter and tolerance requirements of the ferrule end face hole to be obtained more easily than simply through drilling of the metallic ferrule end face alone. Nevertheless, typical concentricity errors between the fiber core and the outside diameter of the jeweled ferrule are in the range 2 to 6 μm giving insertion

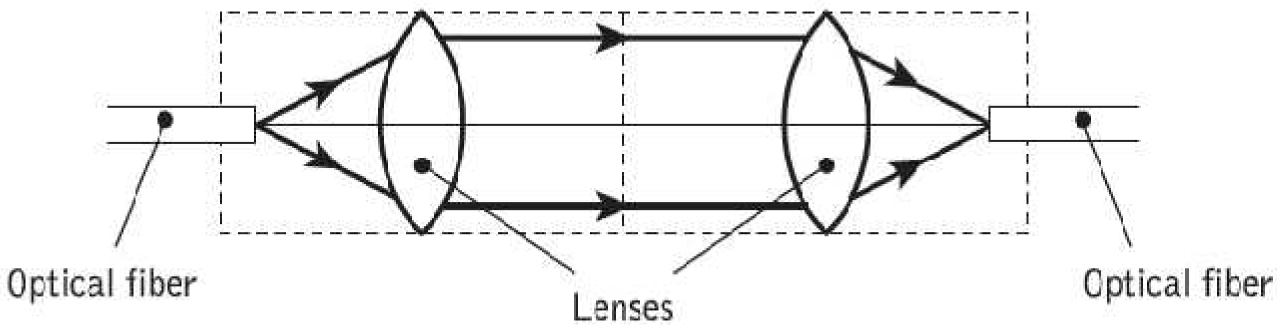


Figure Schematic illustration of an expanded beam connector showing the principle of operation

Also, the longitudinal separation between the two mated halves of the connector ceases to be critical. However, this is achieved at the expense of more stringent angular alignment. Nevertheless, expanded beam connectors are useful for multifiber connection and edge connection for printed circuit boards where lateral and longitudinal alignment are frequently difficult to achieve.

Two examples of lens-coupled expanded beam connectors are illustrated in Figure The connector shown in Figure utilized spherical microlenses for beam expansion and reduction. It exhibited average losses of 1 dB which were reduced to 0.7 dB with the application of an antireflection coating on the lenses and the use of graded index fiber of 50 µm core diameter. A similar configuration has been used for single-mode fiber connection in which the lenses have a 2.5 mm diameter. Again with antireflection-coated lenses, average losses around 0.7 dB were obtained using single-mode fibers of 8 µm core diameter. Furthermore, successful single-mode fiber connection has been achieved with a much smaller (250 µm diameter) sapphire ball lens expanded beam design.

In this case losses in the range 0.4 to 0.7 dB were demonstrated over 1000 connections. Figure shows an expanded beam connector which employs a molded spherical lens. The fiber is positioned approximately at the focal length of the lens in order to obtain a collimated beam and hence minimize lens-to-lens longitudinal misalignment effects. A lens alignment sleeve is used to minimize the effects of angular misalignment which, together with a ferrule, grommet, spring and external housing, provides the complete connector structure. The repeatability of this relatively straightforward lens design was found to be good, incurring losses of around 0.7 dB.

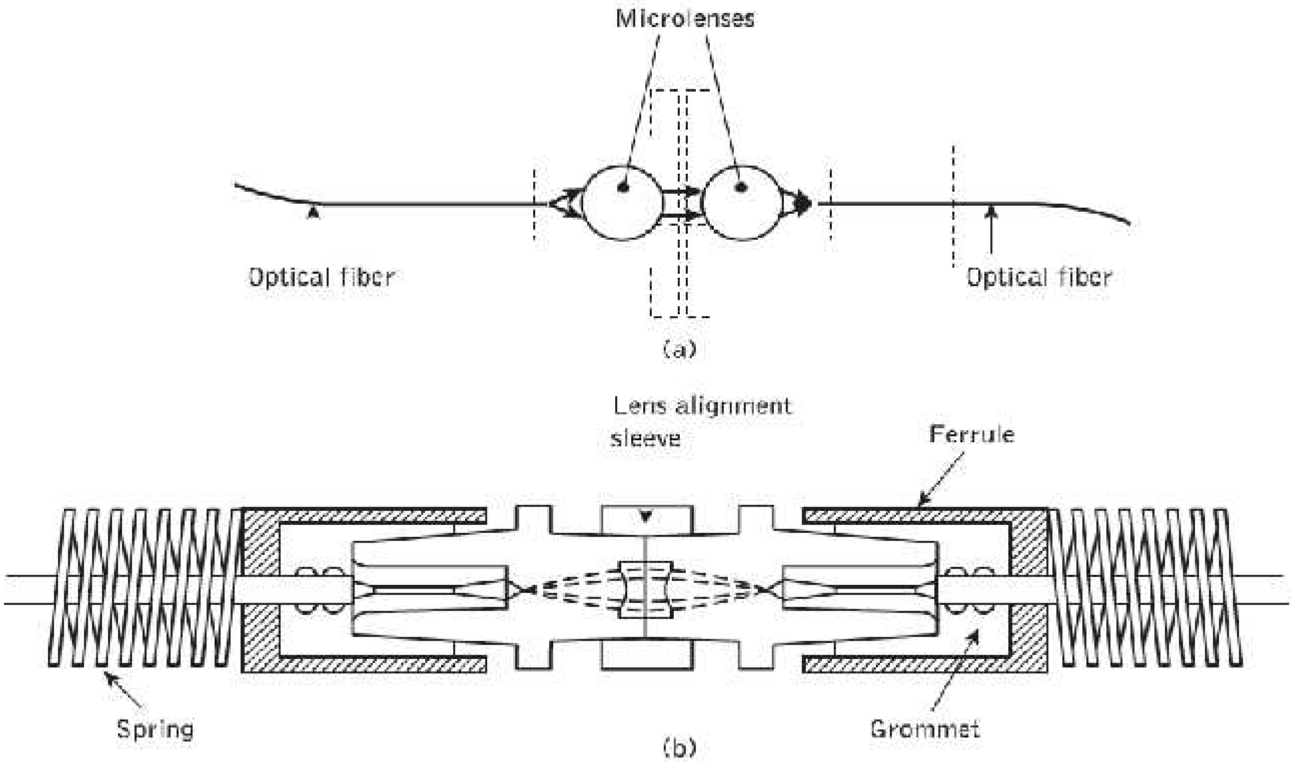


Figure Lens-coupled expanded beam connectors: (a) schematic diagram of a connector with two microlenses making a 1:1 image of the emitting fiber upon the receiving one; (b) molded plastic lens connector assembly

Fiber couplers

An optical fiber coupler is a device that distributes light from a main fiber into one or more branch fibers. The latter case is more normal and such devices are known as multiport fiber couplers. Requirements are increasing for the use of these devices to divide or combine optical signals for application within optical fiber information distribution systems including data buses, LANs, computer networks and telecommunication access networks.

Optical fiber couplers are often passive devices in which the power transfer takes place either: (a) through the fiber core cross-section by butt jointing the fibers or by using some form of imaging optics between the fibers (core interaction type); or (b) through the fiber surface and normal to its axis by converting the guided core modes to both cladding and refracted modes which then enable the power-sharing mechanism (surface interaction type).

Multiport optical fiber couplers can also be subdivided into the following three main groups, as illustrated in Figure

1. Three- and four-port* couplers, which are used for signal splitting, distribution and combining.

2. Star couplers, which are generally used for distributing a single input signal to multiple outputs.

3. Wavelength division multiplexing (WDM) devices, which are a specialized form of coupler designed to permit a number of different peak wavelength optical signals to be transmitted in parallel on a single fiber.

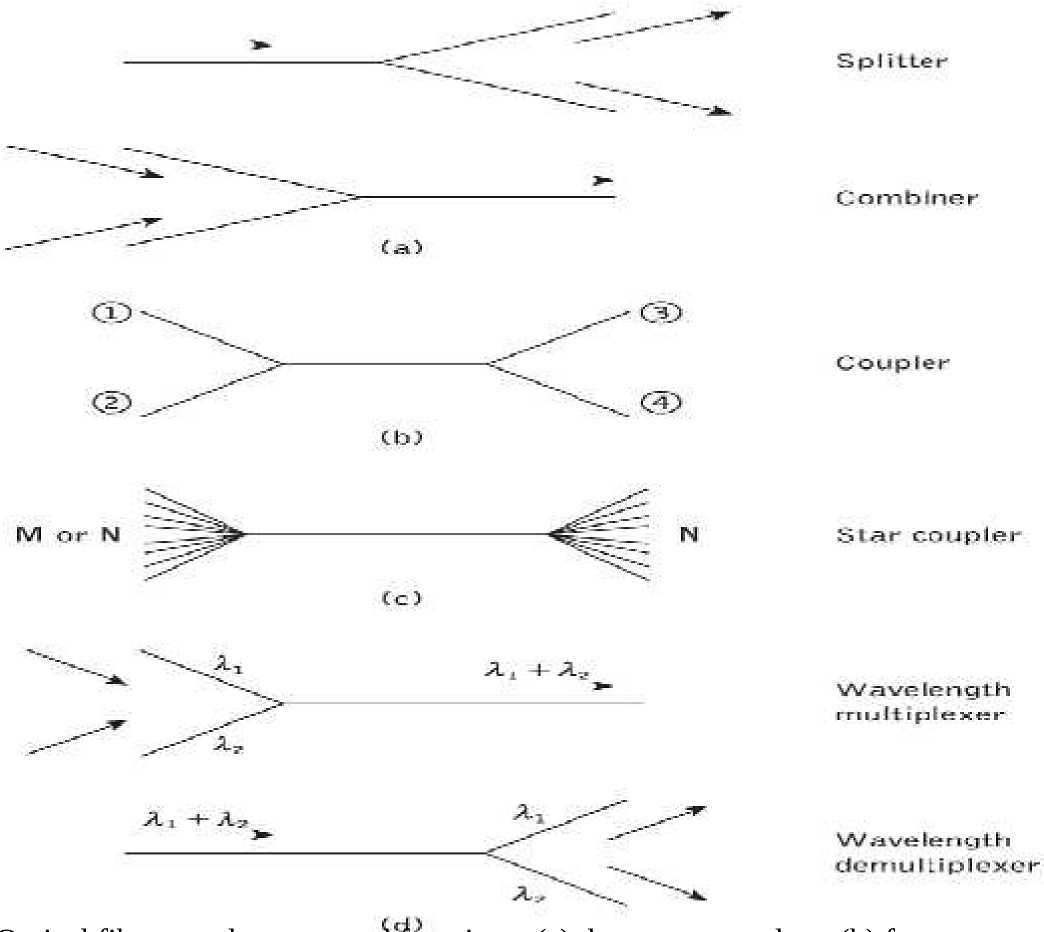
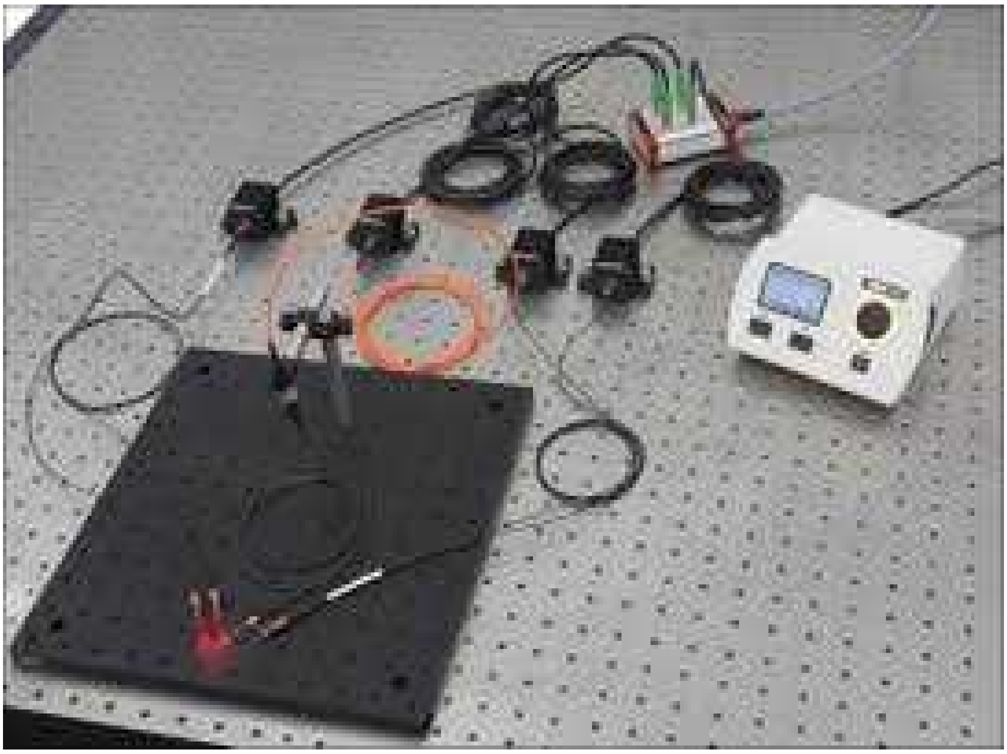


Figure Optical fiber coupler types and functions: (a) three-port couplers; (b) four-port coupler; (c) star coupler; (d) wavelength division multiplexing and demultiplexing couplers

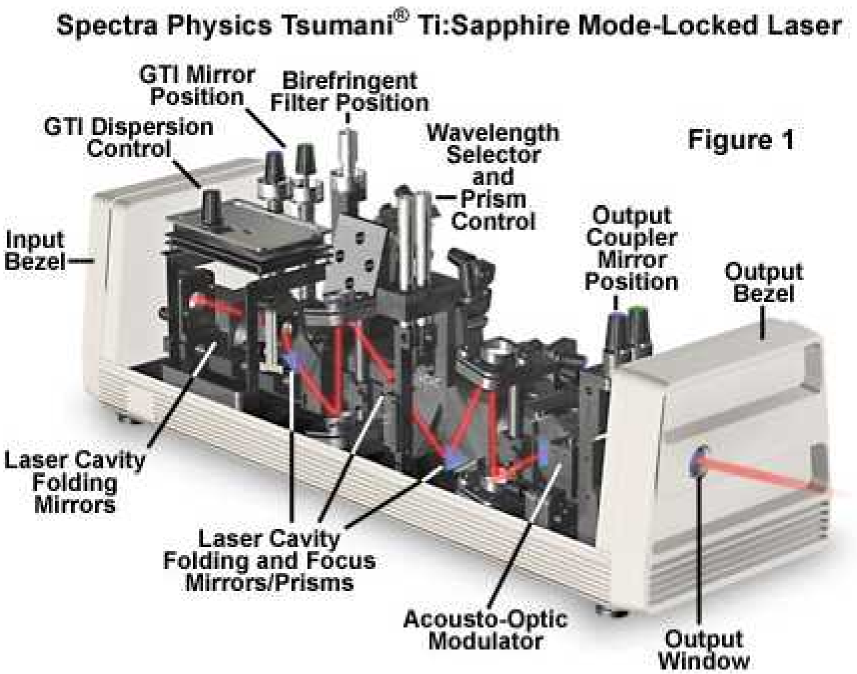
In this context WDM couplers either combine the different wavelength optical signal onto the fiber (i.e. multiplex) or separate the different wavelength optical signals output from the fiber (i.e. demultiplex). Ideal fiber couplers should distribute light among the branch fibers with no scattering loss† or the generation of noise, and they should function with complete insensitivity to factors including the distribution of light between the fiber modes, as well as the state of polarization of the light. Unfortunately, in practice passive fiber couplers do not display all of the above properties and hence the characteristics of the devices affect the performance of optical fiber networks.

This technique, which can provide a bidirectional coupling capability, is well suited for use with multimode step index fibers but may incur higher excess losses than other methods as all the input light cannot be coupled into the output fibers. Another coupling technique is to incorporate a beam splitter element between the fibers. The semitransparent mirror method provides an ingenious way to accomplish such a fiber coupler, as shown in Figure. A partially reflecting surface can be applied directly to the fiber end face cut at an angle of 45° to form a thin-film beam splitter.

APPLICATIONS



Examples of Fiber coupled LEDs for Optogenetics



Examples of LASER system in Optical Microscopy

POST TEST-MCQ TYPE

1. The radiation emission process (emission of a proton at frequency) can occur in how many ways.
a) Two
b) Three
c) Four
d) One

2. The lower energy level contains more atoms than upper level under the conditions of
a) Isothermal packaging
b) Population inversion
c) Thermal equilibrium
d) Pumping

3. Which of the following in the laser occurs when photon colliding with an excited atom causes the stimulated emission of a second photon.
a) Light amplification
b) Attenuation
c) Dispersion
d) Population inversion

4. Which of the following factors does not cause divergence of the collimated beam from a GRIN-rod lens?
a) Lens cut length
b) Size of fiber core
c) Refractive index profile
d) Chromatic aberration
5. Which is used when the optical emission results from the application of electric field.
a) Radiation
b) Efficiency
c) Electro-luminescence
d) Magnetron oscillator

6. Population inversion is obtained at a p-n junction by
a) Heavy doping of p-type material
b) Heavy doping of n-type material
c) Light doping of p-type material
d) Heavy doping of both p-type and n-type material

7. The absence of ________________ in LEDs limits the internal quantum efficiency.
a) Proper semiconductor
b) Adequate power supply
c) Optical amplification through stimulated emission
d) Optical amplification through spontaneous emission

8. What is the use of interposed optics in expanded beam connectors?

a) To achieve lateral alignment less critical than a butt-joined fiber connector

b) To make a fiber loss free

c) To make a fiber dispersive

d) For index-matching

9. Determine the internal quantum efficiency generated within a device when it has a radiative recombination lifetime of 80 ns and total carrier recombination lifetime of 40 ns.

a) 20 %

b) 80 %

c) 30 %

d) 40 %

10. The Lambertian intensity distribution __________ the external power efficiency by some percent.

a) Reduces

b) Does not affects

c) Increases

d) Have a negligible effect

11. Determine coupling efficiency into the fiber when GaAs LED is in close proximity to fiber core having numerical aperture of 0.3.

a) 0.9

b) 0.3

c) 0.6

d) 0.12

12. The amount of radiance in planer type of LED structures is

a) Low

b) High

c) Zero

d) Negligible

13. In surface emitter LEDs, more advantage can be obtained by using

a) BH structures

b) QC structures

c) DH structures

d) Gain-guided structure

14. In a multimode fiber, much of light coupled in the fiber from an LED is

a) Increased

b) Reduced

c) Lost

d) Unaffected

15. The InGaAsP is emitting LEDs are realized in terms of restricted are
a) Length strip geometry
b) Radiance
c) Current spreading
d) Coupled optical power

16. The internal quantum efficiency of LEDs decreasing ________________ with _________________ temperature.
a) Exponentially, decreasing
b) Exponentially, increasing
c) Linearly, increasing
d) Linearly, decreasing

17. The optical 3 dB point occurs when currents ratio is equal to
a) 8/3
b) 2/2
c) 1/2
d) 3/4

18. Laser modes are generally separated by few
a) Tenths of micrometer
b) Tenths of nanometer
c) Tenths of Pico-meter
d) Tenths of millimeter

19. The spectral width of emission from the single mode device is
a) Smaller than broadened transition line-width
b) Larger than broadened transition line-width
c) Equal the broadened transition line-width
d) Cannot be determined

20. Gain guided laser structure are
a) Chemical laser
b) Gas laser
c) DH injection laser
d) Quantum well laser

21. In a DH laser, the sides of cavity are formed by
a) Cutting the edges of device
b) Roughening the edges of device
c) Softening the edges of device
d) Covering the sides with ceramics

22. In Buried hetero-junction (BH) lasers, the optical field is confined within
a) Transverse direction
b) Lateral direction
c) Outside the strip
d) Both transverse and lateral direction

23. Quantum well lasers are also known as
a) BH lasers
b) DH lasers
c) Chemical lasers
d) Gain-guided lasers

24. Better confinement of optical mode is obtained in
a) Multi Quantum well lasers
b) Single Quantum well lasers
c) Gain guided lasers
d) BH lasers

25. The phenomenon resulting in the electrons to jump from one state to another each time emitting of photon is known as
a) Inter-valence band absorption
b) Mode hopping
c) Quantum cascading
d) Quantum confinement

26. Which lasers are presently the major laser source for optical fiber communications?
a) Semiconductor
b) Non-Semiconductor
c) Injection
d) Solid-state

27. It is a resonant cavity formed by two parallel reflecting mirrors separated by a mirror separated by a medium such as air or gas is?
a) Optical cavity
b) Wheatstone’s bridge
c) Oscillator
d) Fabry-perot resonator

28. Which of the following co-dopant is NOT employed by neodymium and erbium doped silica fiber lasers?
a) Phosphorus pent oxide
b) Germania
c) Nitrogen
d) Alumina

29. Dopants levels in glass fiber lasers are generally
a) Low
b) High
c) Same as that of GRIN rod lens laser
d) Same as that of semiconductor laser

30. The lasing output of the basic Fabry-perot cavity fiber is restricted to between
a) 1 and 2 nm
b) 5 and 10 nm
c) 3 and 6 nm
d) 15 and 30 nm

31. In Fabry-perot laser, the lower threshold is obtained by
a) Increasing the refractive index
b) Decreasing the refractive index
c) Reducing the slope efficiency
d) Increasing the slope efficiency

32. $Y_3A_{l5}\ O_{12}$ is a molecular formula for
a) Ytterbium aluminate
b) Yttrium oxide
c) Ytterbium oxy-aluminate
d) Yttrium-aluminum garnet

33**.** A measure of amount of optical fiber emitted from source that can be coupled into a fiber is termed as
a) Radiance
b) Angular power distribution
c) Coupling efficiency
d) Power-launching

34. How many types of misalignments occur when joining compatible fiber?
a) One
b) Two
c) Five
d) Three

35. Losses caused by factors such as core-cladding diameter, numerical aperture, relative refractive index differences, different refractive index profiles, fiber faults are known as
a) Intrinsic joint losses
b) Extrinsic losses
c) Insertion losses
d) Coupling losses

36. A permanent joint formed between two different optical fibers in the field is known as a
a) Fiber splice
b) Fiber connector
c) Fiber attenuator
d) Fiber dispersion

37. How many types of fiber splices are available?
a) One
b) Two
c) Three
d) Four

38. What is the main requirement with the fibers that are intended for splicing?
a) Smooth and oval end faces
b) Smooth and square end faces
c) Rough edge faces
d) Large core diameter

39. In score and break process, which of the following is NOT used as a cutting tool?
a) Diamond
b) Sapphire
c) Tungsten carbide
d) Copper

40. The heating of the two prepared fiber ends to their fusing point with the application of required axial pressure between the two optical fibers is called as
a) Mechanical splicing
b) Fusion splicing
c) Melting
d) Diffusion

41. Which of the following is not used as a flame heating source in fusion splicing?
a) Microprocessor torches
b) Ox hydric burners
c) Electric arc
d) Gas burner

42. Which is caused by surface tension effects between the two fiber ends during fusing.
a) Pre-fusion
b) Diffusion
c) Self-alignment
d) Splicing

43. What are formed by sandwiching the butted fiber ends between a V-groove glass substrate and a flat glass retainer plate.

a) Springroove splices

b) V-groove splices

c) Elastic splices

d) Fusion splices

44. What is the use of an index-matching material in the connector between the two jointed fibers?

a) To decrease the light transmission through the connection

b) To increase the light transmission through the connection

c) To induce losses in the fiber

d) To make a fiber dispersive

45. How many categories of fiber connectors exist?

a) One

b) Three

c) Two

d) Four

46. What is the use of watch jewel in cylindrical ferrule connector?

a) To obtain the diameter and tolerance requirements of the ferrule

b) For polishing purposes

c) Cleaving the fiber

d) To disperse a fiber

47. In connectors, the fiber ends are separated by some gap. This gap ranges from

a) 0.040 to 0.045 mm

b) 0.025 to 0.10 mm

c) 0.12 to 0.16 mm

d) 0.030 to 0.2mm

CONCLUSION

In this unit, the various optical sources, materials and fiber splicing were learnt. The principles of different optical sources and power launching-coupling methods were described in detailed.

REFERENCES

1. Gerd Keiser, "Optical Fiber Communication" McGraw-Hill International,4th Edition. 2010.
2. John M. Senior, "Optical Fiber Communication", Second Edition, Pearson Education, 2007.
3. Ramaswami, Sivarajan and Sasaki "Optical Networks", Morgan Kaufmann, 2009.
4. J.Senior, Optical Communication, Principles and Practice,Prentice Hall of India, 3rd Edition, 2008.
5. J.Gower, "Optical Communication System", Prentice Hall of India, 2001.

ASSIGNMENT

1. Draw and explain the LED structures based Double Heterostucture configuration
2. Discuss the principle of operation of LASER diodes.
3. What are the effects of temperature on the performance of a LASER diode?
4. Explain the lensing schemes used to improve optical source - to- fiber coupling efficiency.
5. What are the advantages of Quantum well LASERS?
6. Explain briefly the three key processes involved in the laser action.
7. Describe for a fabry-perot resonator laser diode, modes and threshold conditions. Obtain its rate equations for steady state output.
8. Describe various fiber splicing techniques with their diagrams.
9. Describe the various types of fiber connectors and couplers.
10. Explain mechanical splices with neat diagrams.

AIM & OBJECTIVES

- ❖ To learn the basic elements of optical fiber transmission link, fiber modes configurations and structures.
- ❖ To understand the different kind of losses, signal distortion, SM fibers.
- ❖ To learn the various optical sources, materials and fiber splicing.
- ❖ To learn the fiber optical receivers and noise performance in photo detector.
- ❖ To explore link budget, WDM, solitons and SONET/SDH network.

PRE TEST-MCQ TYPE

1. How many design considerations are considered while determining the receiver performance?
a) Three
b) Two
c) One
d) Four

2. Which circuits extends the dynamic range of the receiver?
a) Monolithic
b) Trans-impedance
c) Automatic Error Control (AEC)
d) Automatic Gain Control (AGC)

3. The sensitivity of the low-impedance configuration is _____________
a) Good
b) Poor
c) Great
d) Same as that of high-impedance configuration

4. What is generally used to determine the receiver performance characteristics?
a) Noise
b) Resistor
c) Dynamic range & sensitivity characteristics
d) Impedance

UNIT IV FIBER OPTIC RECEIVER AND MEASUREMENTS

Fundamental receiver operation, Pre-amplifiers, Error sources – Receiver Configuration – Probability of Error – Quantum limit, Fiber Attenuation measurements- Dispersion measurements-Fiber Refractive index profile measurements – Fiber cut- off Wave length Measurements – Fiber Numerical Aperture Measurements – Fiber diameter measurements

THEORY
Introduction

Fundamental Receiver Operation

The receiver must first detect weak, distorted signal and then make decisions on what type of data was sent based on amplified version of the distorted signal. To understand the function of the receiver, we first examine what happens to the signal as it is sent through the optical data link which is shown in the following figure

- ❖ Digital Signal Transmission
- ❖ Error Sources

• Digital Receiver Performance

- ❖ Probability of Error
- ❖ Receiver Sensitivity
- ❖ The Quantum Limit

• Coherent Detection

• Analog Receiver

Optical Receiver Operation

Digital Signal Transmission

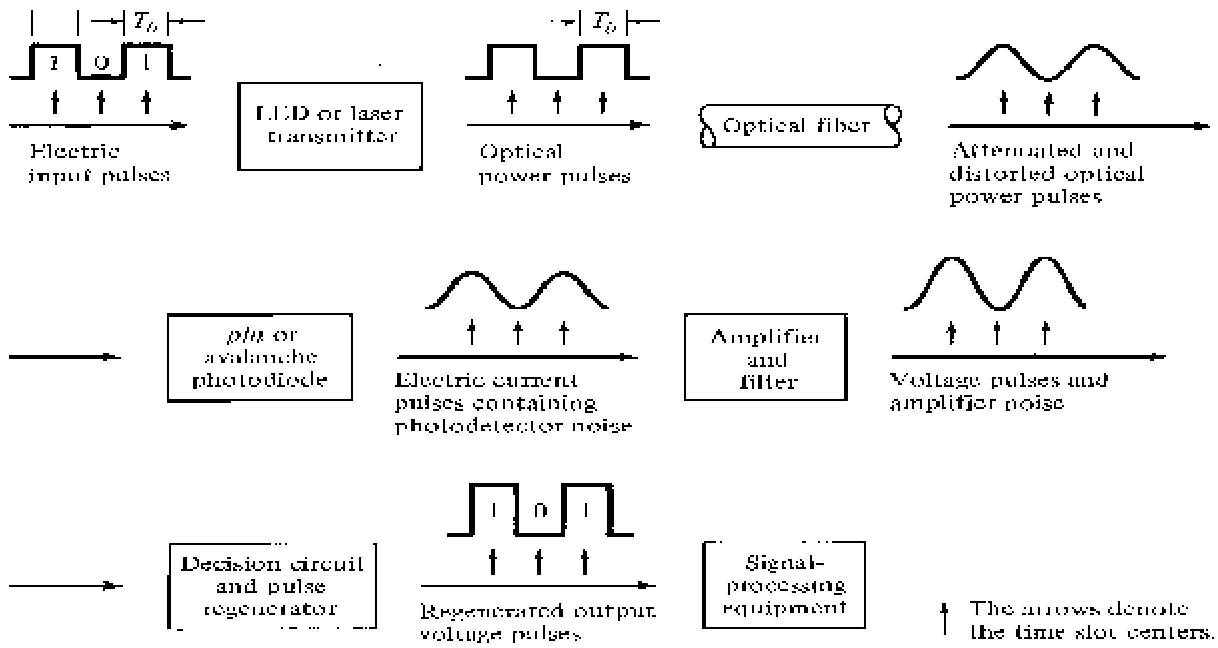


Figure Signal path through an optical data link.

- A typical digital fiber transmission link is shown in Figure. The transmitted signal is a two- level binary data stream consisting of either a ‘0’ or a ‘1’ in a bit period T_b.
- The simplest technique for sending binary data is amplitude-shift keying, wherein a voltage level is switched between on or off values.
- The resultant signal wave thus consists of a voltage pulse of amplitude V when a binary 1 occurs and a zero-voltage-level space when a binary 0 occurs.
- An electric current i(t) can be used to modulate directly an optical source to produce an optical output power P(t).
- In the optical signal emerging from the transmitter, a ‘1’ is represented by a light pulse of duration Tb, whereas a ‘0’ is the absence of any light.
- The optical signal that gets coupled from the light source to the fiber becomes attenuated and distorted as it propagates along the fiber waveguide. Upon reaching the receiver, either a PIN or an APD converts the optical signal back to an electrical format.
- A decision circuit compares the amplified signal in each time slot with a threshold level.
- If the received signal level is greater than the threshold level, a ‘1’ is said to have been received.
- If the voltage is below the threshold level, a ‘0’ is assumed to have been received.

Error Sources

- Errors in the detection mechanism can arise from various noises and disturbances associates with the signal detection system.
- The two most common samples of the spontaneous fluctuations are shot noise and thermal noise.
- Shot noise arises in electronic devices because of the discrete nature of current flow in the device
- Thermal noise arises from the random motion of electrons in a conductor.
- The random arrival rate of signal photons produces a quantum (or shot) noise at the photodetector.
- This noise depends on the signal level.
- This noise is of particular importance for PIN receivers that have large optical input levels and for APD receivers.
- When using an APD, an additional shot noise arises from the statistical nature

of the multiplication process. This noise level increases with increasing avalanche gain *M*.

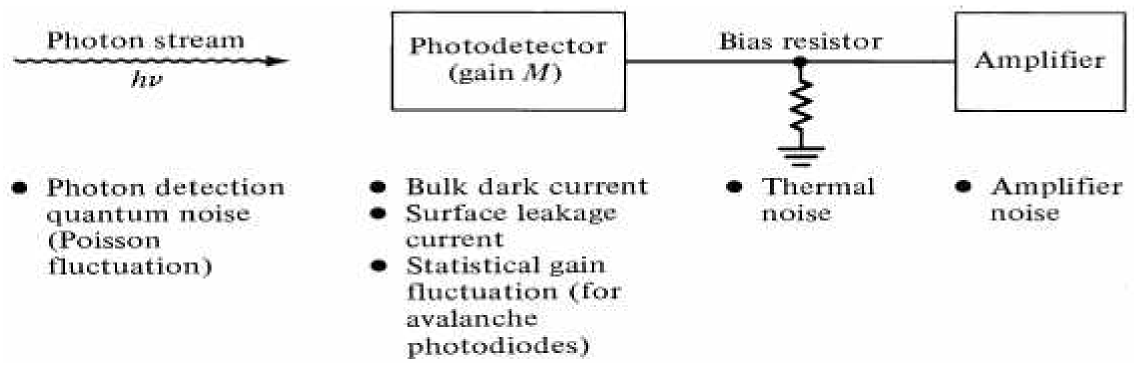


Figure Noise sources and disturbances in the optical pulse detection mechanism.

- ❖ Thermal noises arising from the detector load resistor and from the amplifier electronics tend to dominate in applications with low SNR when a PIN photodiode is used.
- ❖ When an APD is used in low-optical-signal level applications, the optimum avalanche gain is determined by a design tradeoff between the thermal noise and the gain-dependent quantum noise.
- ❖ The primary photocurrent generated by the photodiode is a time-varying Poisson process.
- ❖ If the detector is illuminated by an optical signal *P*(*t*), then the average number of electron-hole pairs generated in a time t is

$$\overline{N} = \frac{\eta}{h\nu}\int_0^{\tau} P(t)\,dt = \frac{\eta E}{h\nu}$$

Where h is the detector quantum efficiency, *h*n is the photon energy, and *E* is the energy received in a time interval.

- ❖ The actual number of electron-hole pairs *n* that are generated fluctuates from the average according to the Poisson distribution

where *P*r(*n*) is the probability that *n* electrons are emitted in an interval t. For a detector with a mean avalanche gain *M* and an ionization rate ratio *k*, the excess noise factor *F*(*M*) for electron injection is

$$F(M) = kM + \left(2 - \frac{1}{M}\right)(1 - k)$$

Or

$$F(M) \simeq M^{x}$$

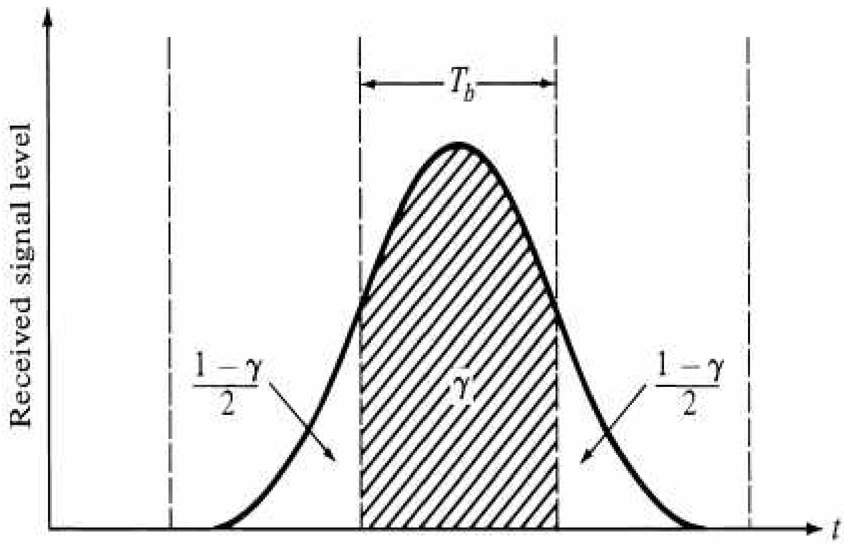


Figure Pulse spreading in an optical signal that leads to ISI.

Where the factor x ranges between 0 and 1.0 depending on the photodiode material. A further error source is attributed to inter symbol interference (ISI), which results from pulse spreading in the optical fiber. The fraction of energy remaining in the appropriate time slot is designated by g, so that 1-g is the fraction of energy that has spread into adjacent time slots.

Receiver Configuration

A typical optical receiver is shown in Figure. The three basic stages of the receiver are a photodetector, an amplifier, and an equalizer. The photo-detector can be either an APD with a mean gain *M* or a PIN for which M=1. The photodiode has a quantum efficiency h and a capacitance C_d. The detector bias resistor has a resistance R_b which generates a thermal noise current *i*b(*t*).

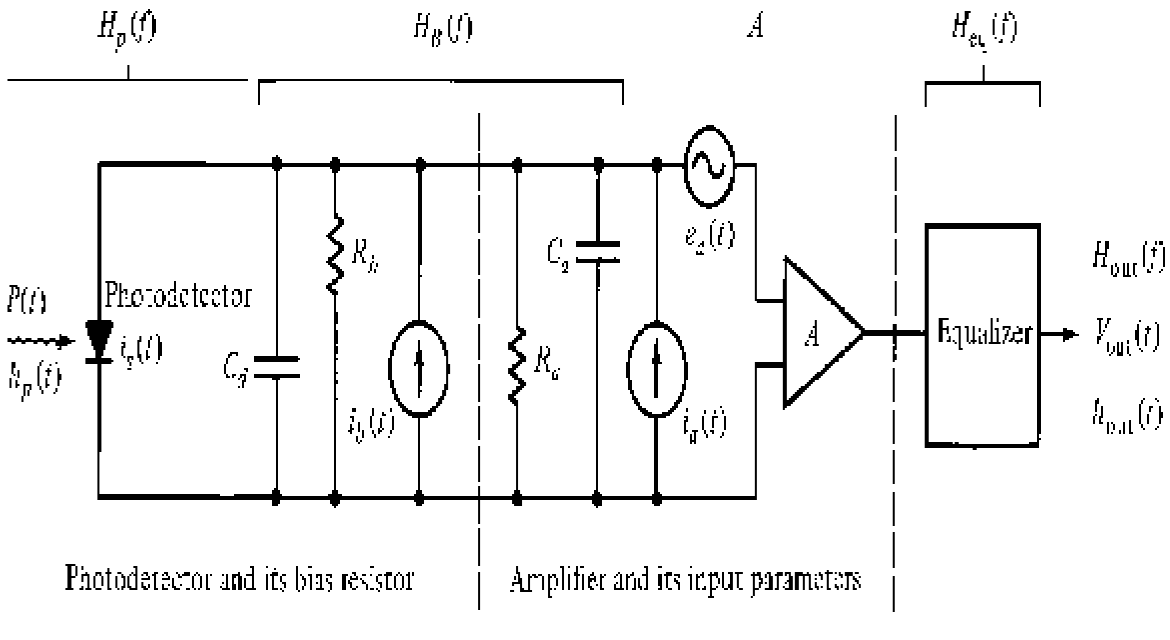


Figure Schematic diagram of a typical optical receiver.

Amplifier Noise Sources:

The input noise current source ia(t) arises from the thermal noise of the amplifier input resistance Ra; The noise voltage source ea(t) represents the thermal noise of the amplifier channel. The noise sources are assumed to be Gaussian in statistics, flat in spectrum (which characterizes white noise), and uncorrelated (statistically independent).

The Linear Equalizer:

The equalizer is normally a linear frequency shaping filter that is used to mitigate the effects of signal distortion and inter symbol interference (ISI). The equalizer accepts the combined frequency response of the transmitter, the fiber, and the receiver, and transforms it into a signal response suitable for the following signal processing electronics.

The binary digital pulse train incident on the photo-detector can be described by

$$P(t) = \sum_{n=-\infty}^{\infty} b_n h_p(t - nT_b)$$

Here, P(t) is the received optical power, Tb is the bit period, b_n is an amplitude parameter representing the nth message digit, and hp(t) is the received pulse shape.

Let the nonnegative photodiode input pulse $h_p(t)$ be normalized to have unit area

$$\int_{-\infty}^{\infty} h_p(t)\,dt = 1$$

Then b_n represents the energy in the n^{th} pulse.

The mean output current from the photodiode at time *t* resulting from the pulse train given previously is

$$\langle i(t) \rangle = \frac{\eta q}{h\nu} MP(t) = R_0 \sum_{n=-\infty}^{\infty} b_n h_p(t - nT_b)$$

Where *R*o = h*q*/*h*n is the photodiode responsivity. The above current is then amplified and filtered to produce a mean voltage at the output of the equalizer.

Probability of error:

The digital receiver performance can be evaluated by measuring the probability of error and quantum limit. In practice, several standards ways are available to measuring the rate of error occurrences in a digital data stream.

$$P_r(n) = N^n \frac{e^{-N}}{n!}$$

The Quantum Limit:

Consider an ideal photodetector which has unity quantum efficiency and which produces no dark current that is no electron hole pairs are generated in the absence of an optical pulse.

With this condition it is possible to find the minimum received optical power required for a specific bit error rate performance in a digital system. This minimum received power level is known as the quantum limit, since all system parameters are assumed ideal and the performance is only limited by the photodetection statistics.

Assume that an optical pulse of energy E falls on the photodetector in a time interval τ. This can only be interpreted by the receiver as a 0 pulse if no electron hole pairs are generated with the pulse present. The probability that n=0 electrons are emitted in a time interval τ is

$$P_r(0) = e^{-\bar{N}}$$

Thus for a given error probability $P_r(0)$, we can find the minimum energy E required at a specific wavelength λ.

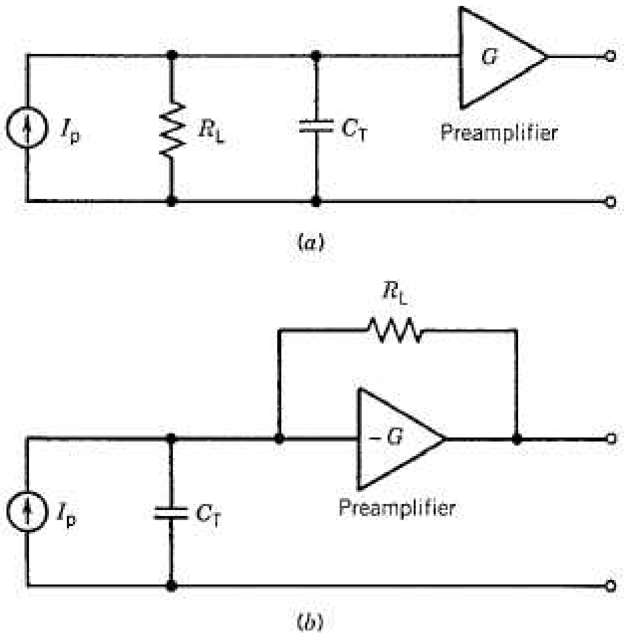
Preamplifier:

Front End

The front end of a receiver consists of a photodiode followed by a preamplifier. The optical signal is coupled onto the photodiode by using a coupling scheme similar to that used for optical transmitters but coupling is often used in practice. The photodiode converts the optical bit stream into an electrical time-varying signal. The role of the preamplifier is to amplify the electrical signal for further processing. The design of the front end requires a trade-off between speed and sensitivity. Since the input voltage to the preamplifier can be increased by using a large load resistor *R L*, a high-impedance front end is often used. Furthermore, as discussed, a large *RL* reduces the thermal noise and improves the receiver sensitivity.

The main drawback of high-impedance front end is its low bandwidth given by $\Delta f = (2\pi R_L C_T)^{-1}$, where *Rs _RL* is assumed and $C_T = C_p + C_A$ is the total capacitance, which includes the contributions from the photodiode (C_p) and the transistor used for amplification (C_A). The receiver bandwidth is limited by its slowest component. A high-impedance front end cannot be used if Δf is considerably less than the bit rate. An equalizer is sometimes used to increase the bandwidth. The equalizer acts as a filter that attenuates low-frequency components of the signal more than the high-frequency components, thereby effectively increasing the front-end bandwidth. If the receiver sensitivity is not of concern, one can simply decrease *RL* to increase the bandwidth, resulting in a low-impedance front end. Transimpedance front ends provide a configuration that has high sensitivity together with a large bandwidth. Its dynamic range is also improved compared with high-impedance front ends. As seen in Figure the load resistor is connected as a feedback resistor around an inverting amplifier.

Even though R_L is large, the negative feedback reduces the effective input impedance by a factor of G, where G is the amplifier gain. The bandwidth is thus enhanced by a factor of G compared with high impedance front ends. Transimpedance front ends are often used in optical receivers because of their improved characteristics. A major design issue is related to the stability of the feedback loop.



The receiver amplifiers are the front end preamplifiers. Preamplifier bandwidth must be greater than or equal to signal bandwidth. It must reduce all source of noise. It must have high receiver sensitivity. There are three basic preamplifier structures

1. Low –Impedance preamplifier

2. High- impedance preamplifier

3. Trans-impedance preamplifier

The preamplifier should have low noise level, high bandwidth, high dynamic range, high sensitivity and high gain.

Fiber attenuation measurements

Fiber attenuation measurement techniques have been developed in order to determine the total fiber attenuation of the relative contributions to this total from both absorption losses and scattering losses. The overall fiber attenuation is of greatest interest to the system designer, but the relative magnitude of the different loss mechanisms is important in the development and fabrication of low-loss fibers. Measurement techniques to obtain the total fiber attenuation give either the spectral loss characteristic or the loss at a single wavelength (spot measurement).

Total fiber attenuation

A commonly used technique for determining the total fiber attenuation per unit length is the cut- back or differential method. Figure shows a schematic diagram of the typical experimental setup for measurement of the spectral loss to obtain the overall attenuation spectrum for the fiber. It consists of a 'white' light source, usually a tungsten halogen or xenon are lamp. The focused light is mechanically chopped at a low frequency of a few hundred hertz. This enables the lock-in amplifier at the receiver to perform phase-sensitive detection.

The chopped light is then fed through a monochromator which utilizes a prism or diffraction grating arrangement to select the required wavelength at which the attenuation is to be measured. Hence the light is filtered before being focused onto the fiber by means of a microscope objective lens. A beam splitter may be incorporated before the fiber to provide light for viewing optics and a reference signal used to compensate for output power fluctuations. When the measurement is performed on multimode fibers it is very dependent on the optical launch conditions. Therefore unless the launch optics are arranged to give the steady-state mode distribution at the fiber input, or a dummy fiber is used, then a mode scrambling device is attached to the fiber within the first meter.

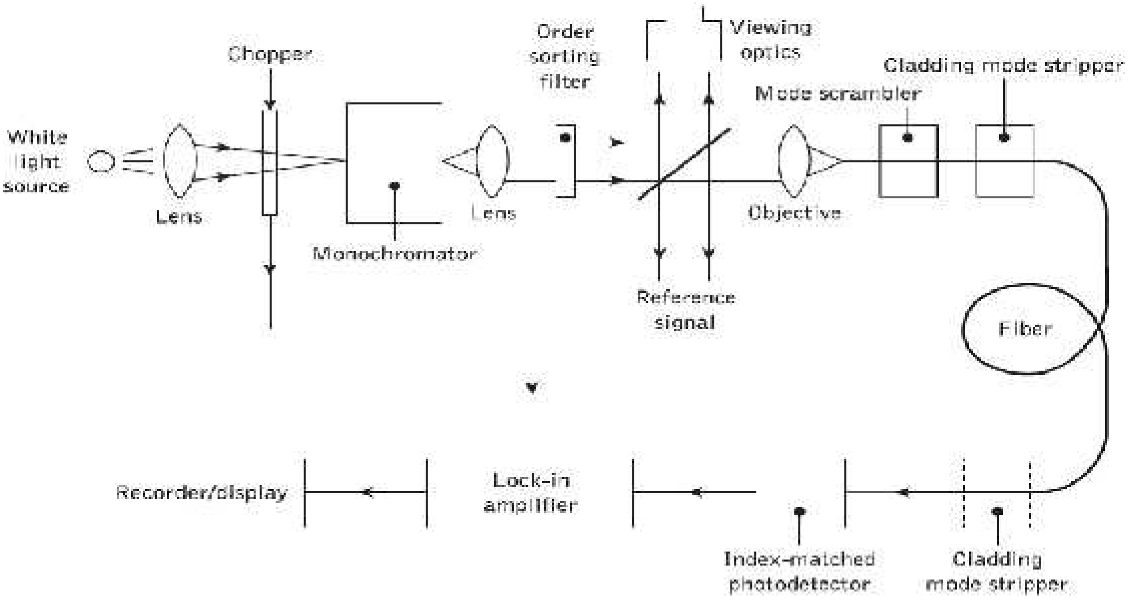


Figure A typical experimental arrangement for the measurement of spectral loss in optical fibers using the cut-back technique

The fiber is also usually put through a cladding mode stripper, which may consist of an S-shaped groove cut in the Teflon and filled with glycerine. This device removes light launched into the fiber cladding through radiation into the index-matched (or slightly higher refractive index) glycerine. A mode stripper can also be included at the fiber output end to remove any optical power which is scattered from the core into the cladding down the fiber length. This tends to be pronounced when the fiber cladding consists of a low-refractive-index silicone resin.

The optical power at the receiving end of the fiber is detected using a *p–i–n* or avalanche photodiode. In order to obtain reproducible results the photodetector surface is usually index matched to the fiber output end face using epoxy resin or an index-matching gell. Finally, the electrical output from the photodetector is fed to a lock-in amplifier, the output of which is recorded.

The cut-back method* involves taking a set of optical output power measurements over the required spectrum using a long length of fiber (usually at least a kilometer). This fiber is generally uncabled having only a primary protective coating. Increased losses due to cabling do not tend to change the shape of the attenuation spectrum as they are entirely radiative, and for multimode fibers are almost wavelength independent. The fiber is then cut back to a point 2m from the input end and, maintaining the same launch conditions, another set of power output measurements is taken.

$$\alpha_{dB} = \frac{10}{L_1 - L_2} \log_{10} \frac{P_{02}}{P_{01}}$$

L_1 and L_2 are the original and cut-back fiber lengths respectively, and P_{01} and P_{02} are the corresponding output optical powers at a specific wavelength from the original and cut-back fiber lengths. Hence when L_1 and L_2 are measured in kilometers, α_{dB} has units of dB km−1.

$$\alpha_{dB} = \frac{10}{L_1 - L_2} \log_{10} \frac{V_2}{V_1}$$

Where V_1 and V_2 correspond to output voltage readings from the original fiber length and the cut-back fiber length respectively.

Fiber absorption loss measurement

It was indicated in the preceding section that there is a requirement for the optical fiber manufacturer to be able to separate the total fiber attenuation into the contributions from the major loss mechanisms. Material absorption loss measurements allow the level of impurity content within the fiber material to be checked in the manufacturing process. The measurements are based on calorimetric methods which determine the temperature rise in the fiber or bulk material resulting from the absorbed optical energy within the structure.

The apparatus shown in Figure is used to measure the absorption loss in optical fibers, was modified from an earlier version which measured the absorption losses in bulk glasses. This temperature measurement technique, illustrated diagrammatically in Figure, has been widely adopted for absorption loss measurements. The two fiber samples shown in Figure are mounted in capillary tubes surrounded by a low-refractive-index liquid (e.g. methanol) for good electrical contact, within the same enclosure of the apparatus shown in Figure. A thermocouple is wound around the fiber containing capillary tubes using one of them as a reference junction (dummy fiber).

Light is launched from a laser source (Nd: YAG or krypton ion depending on the wavelength of interest) through the main fiber (not the dummy), and the temperature rise due to absorption is measured by the thermocouple and indicated on a nanovoltmeter. Electrical calibration may be achieved by replacing the optical fibers with thin resistance wires and by passing known electrical power through one. Independent measurements can then be made using the calorimetric technique and with electrical measurement instruments.

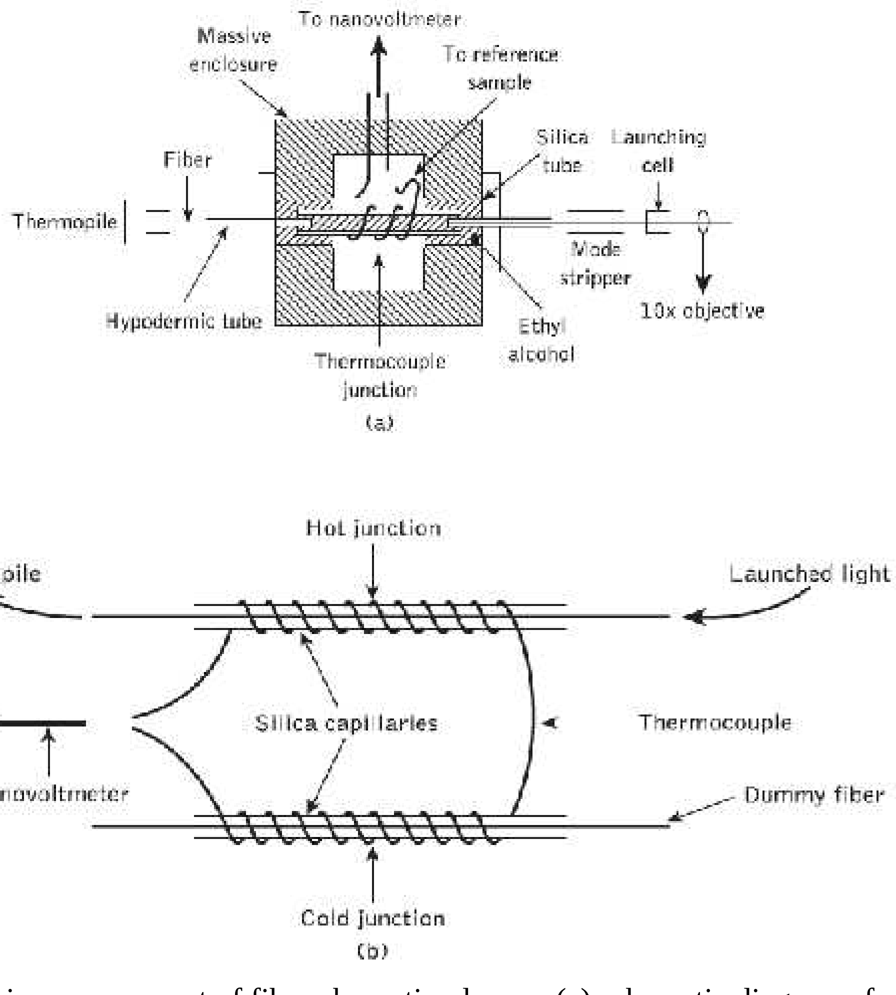


Figure Calorimetric measurement of fiber absorption losses: (a) schematic diagram of a version of the apparatus; (b) the temperature measurement technique using a thermocouple

The calorimetric measurements provide the heating and cooling curve for the fiber sample used. A typical example of this curve is illustrated in Figure. The attenuation of the fiber due to absorption α_{abs} may be determined from this heating and cooling characteristic. A time constant t_c can be obtained from a plot on a logarithmic scale against the time t, an example of which shown in Figure was obtained from the heating characteristic displayed in Figure. This corresponds to the maximum temperature rise of the fiber under test and T_t is the temperature rise at a time t. It may be observed from Figure that corresponds to a steady-state temperature for the fiber when the heat loss to the surroundings balances the heat generated in the fiber resulting from absorption at a particular optical power level.

The time constant t_c may be obtained from the slope of the straight line plotted in Figure

$$t_c = \frac{t_2 - t_1}{\ln(T_\infty - T_{t_1}) - \ln(T_\infty - T_{t_2})}$$

$$\alpha_{abs} = \frac{CT_\infty}{P_{opt} t_c} \text{ dB km}^{-1}$$

Where C is proportional to the thermal capacity per unit length of the silica capillary and the low- refractive-index liquid surrounding the fiber, and P_{opt} is the optical power propagating in the fiber under test. The thermal capacity per unit length may be calculated, or determined by the electrical calibration utilizing the thin resistance wire.

Fiber scattering loss measurement

The usual method of measuring the contribution of the losses due to scattering within the total fiber attenuation is to collect the light scattered from a short length of fiber and compare it with the total optical power propagating within the fiber. Light scattered from the fiber may be detected in a scattering cell as illustrated in the experimental arrangement shown in Figure. This may consist of a cube of six square solar cells or an integrating sphere and detector. The solar cell cube which contains index-matching fluid surrounding the fiber gives measurement of the scattered light, but careful balancing of the detectors is required in order to achieve a uniform response.

This problem is overcome in the integrating sphere which again usually contains index matching fluid but responds uniformly to different distributions of scattered light. However, the integrating sphere does exhibit high losses from internal reflections. Other variations of the scattering cell include the internally reflecting cell and the sandwiching of the fiber between two solar cells. A laser source (i.e. He–Ne, Nd: YAG, krypton ion) is utilized to provide sufficient optical power at a single wavelength together with a suitable instrument to measure the response from the detector. In order to avoid inaccuracies in the measurement resulting from scattered light which may be trapped in the fiber, cladding mode strippers are placed before and after the scattering cell.

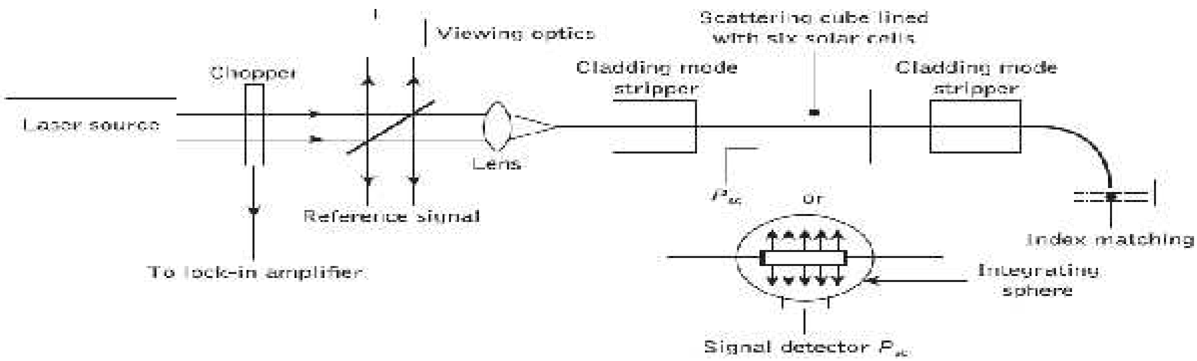


Figure An experimental setup for measurement of fiber scattering loss illustrating both the solar cell cube and integrating sphere scattering cells

These devices remove the light propagating in the cladding so that the measurements are taken only using the light guided by the fiber core. Also, to avoid reflections contributing to the optical signal within the cell, the output fiber end is index matched using either a fluid or suitable surface.

The loss due to scattering α_{sc} is given by:

$$\alpha_{sc} = \frac{10}{l(\text{km})} \log_{10}\left(\frac{P_{opt}}{P_{opt} - P_{sc}}\right) \text{dB km}^{-1}$$

where l(km) is the length of the fiber contained within the scattering cell, P_{opt} is the optical power propagating within the fiber at the cell and P_{sc} is the optical power scattered from the short length of fiber l within the cell. As $P_{opt} >> P_{sc}$, then the logarithm in equation may be expanded to give:

Since the measurements of length are generally in centimeters and the optical power is normally registered in volts, Equation can be written as:

$$\alpha_{sc} = \frac{4.343 \times 10^5}{l(\text{cm})} \left(\frac{V_{sc}}{V_{opt}}\right) \text{dB km}^{-1}$$

Where V_{sc} and V_{opt} are the voltage readings corresponding to the scattered optical power and the total optical power within the fiber at the cell. The relative experimental accuracy (i.e. repeatability) for scatter loss measurements is in the range ±0.2 dB using the solar cell cube and around 5% with the integrating sphere. However, it must be noted that the absolute accuracy of the measurements is somewhat poorer, being dependent on the calibration of the scattering cell and the mode distribution within a multimode fiber.

Fiber dispersion measurements

Dispersion measurements give an indication of the distortion to optical signals as they propagate down optical fibers. The delay distortion which, for example, leads to the broadening of transmitted light pulses limits the information-carrying capacity of the fiber. The measurement of dispersion allows the bandwidth of the fiber to be determined. Therefore, besides attenuation, dispersion is the most important transmission characteristic of an optical fiber. As discussed, there are three major mechanisms which produce dispersion in optical fibers (material dispersion, waveguide dispersion and intermodal dispersion). The importance of these different mechanisms to the total fiber dispersion is dictated by the fiber type. For instance, in multimode fibers (especially step index), intermodal dispersion tends to be the dominant mechanism, whereas in single-mode fibers intermodal dispersion is nonexistent as only a single mode is allowed to propagate. In the single-mode case the dominant dispersion mechanism is chromatic. The dominance of intermodal dispersion in multimode fibers makes it essential that dispersion measurements on these fibers are performed only when the equilibrium mode distribution has been established within the fiber, otherwise inconsistent results will be obtained. Therefore devices such as mode scramblers or filters must be utilized in order to simulate the steady state mode distribution.

Dispersion effects may be characterized by taking measurements of the impulse response of the fiber in the time domain, or by measuring the baseband frequency response in the frequency domain. If it is assumed that the fiber response is linear with regard to power, a mathematical description in the time domain for the optical output power $P_o(t)$ from the fiber may be obtained by convoluting the power impulse response h(t) with the optical input power Pi(t) as:

$$P_o(t) = h(t) * P_i(t)$$

where the asterisk * denotes convolution. The convolution of h(t) with Pi(t) shown in Equation may be evaluated using the convolution integral where:

$$P_o(t) = \int_{-\infty}^{\infty} P_i(t-x)\,h(x)\,\mathrm{d}x$$

In the frequency domain the power transfer function *H*(w) is the Fourier transform of *h*(*t*) and therefore by taking the Fourier transforms of all the functions in Equation

$$\mathcal{P}_o(\omega) = H(\omega)\mathcal{P}_i(\omega)$$

Time domain measurement

The most common method for time domain measurement of pulse dispersion in multimode optical fibers is illustrated in Figure. Short optical pulses (100 to 400 ps) are launched into the fiber from a suitable source (e.g. A1GaAs injection laser) using fast driving electronics. The pulses travel down the length of fiber under test (around 1 km) and are broadened due to the various dispersion mechanisms. However, it is possible to take measurements of an isolated dispersion mechanism by, for example, using a laser with a narrow spectral width when testing a multimode fiber. In this case the chromatic dispersion is negligible and the measurement thus reflects only intermodal dispersion.

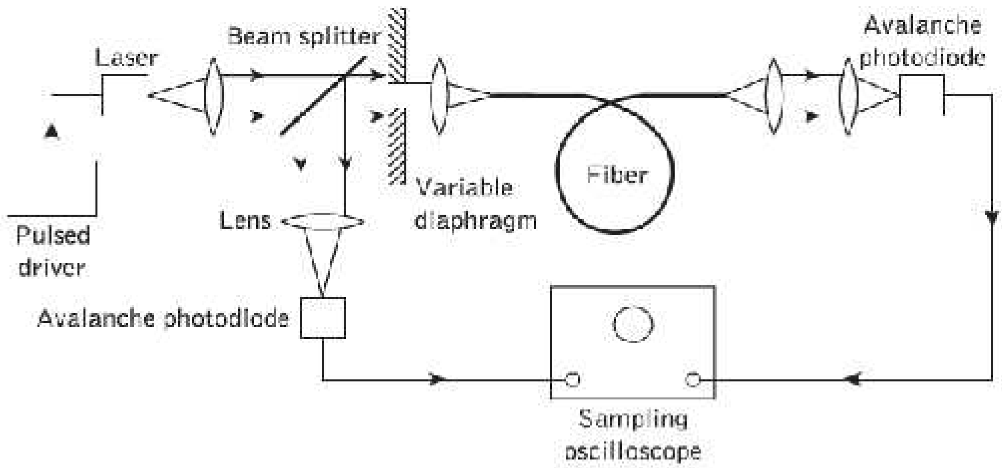


Figure Experimental arrangement for making multimode fiber dispersion measurements in the time domain.

The pulses are received by a high-speed photodetector (i.e. avalanche photodiode) and are displayed on a fast sampling oscilloscope. A beam splitter is utilized for triggering the oscilloscope and for input pulse measurement. After the initial measurement of output pulse width, the long fiber length may be cut back to a short length and the measurement repeated in order to obtain the effective input pulse width. The fiber is generally cut back to the lesser of 10 m or 1% of its original length. As an alternative to this cut-back technique, the insertion or substitution method similar to that used in fiber loss measurement can be employed. This method has the benefit of being nondestructive and only slightly less accurate than the cut-back technique.

The fiber dispersion is obtained from the two pulse width measurements which are taken at any convenient fraction of their amplitude. If $P_i(t)$ and $P_o(t)$ of Equation are assumed to have a Gaussian shape then Equation may be written in the form:

$$\tau_o^2(3\text{ dB}) = \tau^2(3\text{ dB}) + \tau_i^2(3\text{ dB})$$

where τ(3 dB),τ_i (3dB) and τ_o (3 dB) are the 3 dB pulse widths at the fiber input and output, respectively, and τ (3 dB) is the width of the fiber impulse response again measured at half the maximum amplitude. Hence the pulse dispersion in the fiber (commonly referred to as the pulse broadening when considering the 3 dB pulse width) in ns km−1 is given by:

$$\tau(3\text{ dB}) = \frac{(\tau_o^2(3\text{ dB}) - \tau_i^2(3\text{ dB}))^{\frac{1}{2}}}{L}\text{ ns km}^{-1}$$

Where τ(3 dB),τ_i (3dB) and τ_o (3 dB) are measured in ns and L is the fiber length in Km. It must be noted that if a long length of fiber is cut back to a short length in order to take the input pulse width measurement, then *L* corresponds to the difference between the two fiber lengths in km.

Frequency domain measurement

Frequency domain measurement is the preferred method for acquiring the bandwidth of multimode optical fibers. This is because the baseband frequency response of the fiber may be obtained directly from these measurements using Equation without the need for any assumptions of Gaussian shape, or alternatively, the mathematically complex deconvolution of Equation which is necessary with measurements in the time domain. Thus the optical bandwidth of a multimode fiber is best obtained from frequency domain measurements.

One of two frequency domain measurement techniques is generally used. The first utilizes a similar pulsed source to that employed for the time domain measurements shown in Figure. However, the sampling oscilloscope is replaced by a spectrum analyzer which takes the Fourier transform of the pulse in the time domain and hence displays its constituent frequency components.

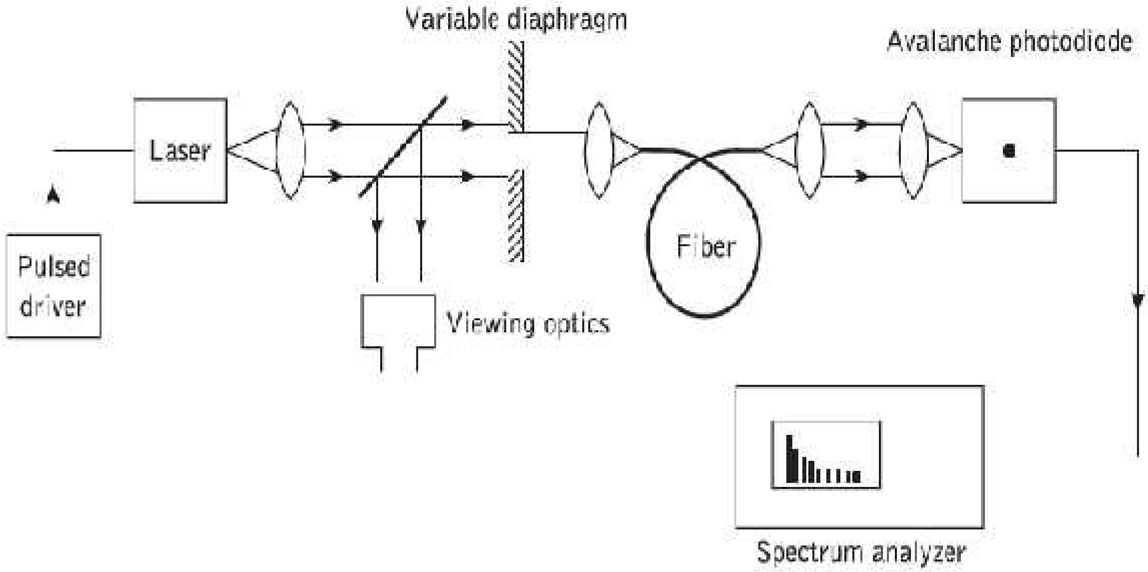


Figure Experimental setup for making fiber dispersion measurements in the frequency domain using a pulsed laser source

The experimental arrangement is illustrated in Figure. Comparison of the spectrum at the fiber output $P_o(w)$ with the spectrum at the fiber input $P_i(w)$ provides the baseband frequency response for the fiber under test where:

$$H(\omega) = \frac{\mathcal{P}_o(\omega)}{\mathcal{P}_i(\omega)}$$

The second technique involves launching a sinusoidally modulated optical signal at different selected frequencies using a sweep oscillator. Therefore the signal energy is concentrated in a very narrow frequency band in the baseband region, unlike the pulse measurement method where the signal energy is spread over the entire baseband region.

Fiber refractive index profile measurements

The refractive index profile of the fiber core plays an important role in characterizing the properties of optical fibers. It allows determination of the fiber's numerical aperture and the number of modes propagating within the fiber core, while largely defining any intermodal and/or profile dispersion caused by the fiber. Hence a detailed knowledge of the refractive index profile enables the impulse response of the fiber to be predicted.

Also, as the impulse response and consequently the information-carrying capacity of the fiber is strongly dependent on the refractive index profile, it is essential that the fiber manufacturer is able to produce particular profiles with great accuracy, especially in the case of graded index fibers (i.e. optimum profile). There is therefore a requirement for accurate measurement of the refractive index profile. These measurements may be performed using a number of different techniques each of which exhibit certain advantages and drawbacks.

Interferometric methods

Interference microscopes (e.g. Mach–Zehnder, Michelson) have been widely used to determine the refractive index profiles of optical fibers. The technique usually involves the preparation of a thin slice of fiber (slab method) which has both ends accurately polished to obtain square (to the fiber axes) and optically flat surfaces. The slab is often immersed in an index-matching fluid, and the assembly is examined with an interference microscope. Two major methods are then employed, using either a transmitted light interferometer or a reflected light interferometer.

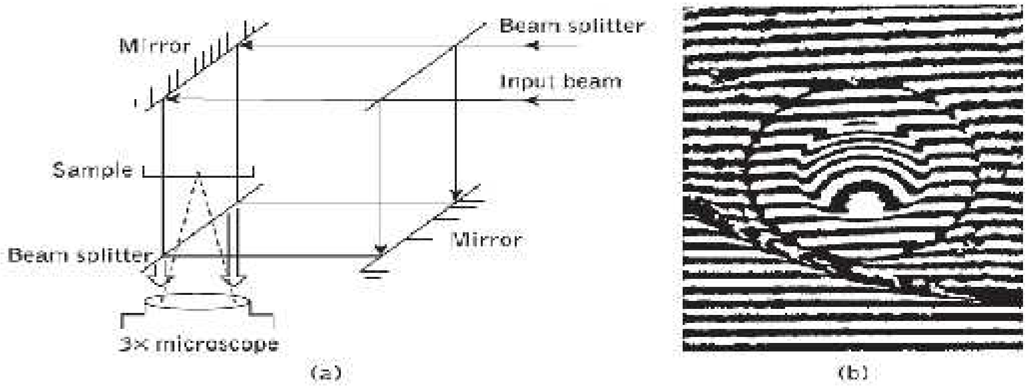


Figure (a) The principle of the Mach–Zehnder interferometer. (b) The interference fringe pattern obtained with an interference microscope from a graded index fiber

In both cases light from the microscope travels normal to the prepared fiber slice faces (parallel to the fiber axis), and differences in refractive index result in different optical path lengths. This situation is illustrated in the case of the Mach–Zehnder interferometer in Figure. When the phase of the incident light is compared with the phase of the emerging light, a field of parallel interference fringes is observed. The fringe displacements for the points within the fiber core are then measured using as reference the parallel fringes outside the fiber core (in the fiber cladding). The refractive index difference between a point in the fiber core (e.g. the core axis) and the cladding can be obtained from the fringe shift q, which corresponds to a number of fringe displacements.

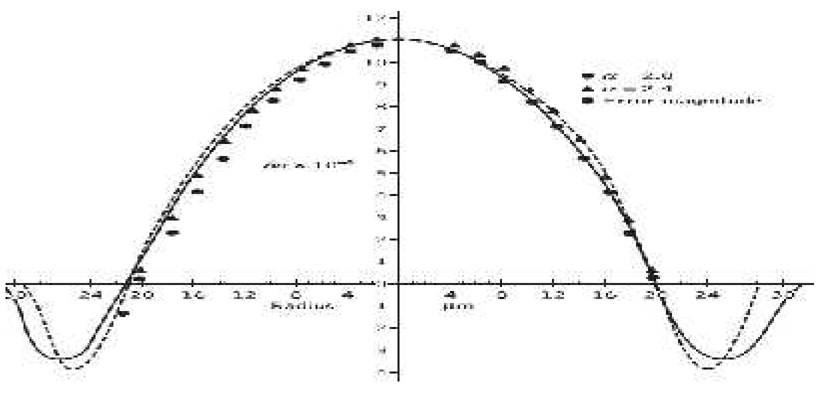


Figure The fiber refractive index profile computed from the interference pattern

This difference in refractive index is given by:

$$\delta n = \frac{q\lambda}{x}$$

Where *x* is the thickness of the fiber slab and λ is the incident optical wavelength. The slab method gives an accurate measurement of the refractive index profile, although computation of the individual points is somewhat tedious unless an automated technique is used. Figure shows the refractive index profile obtained from the fringe pattern indicated in Figure

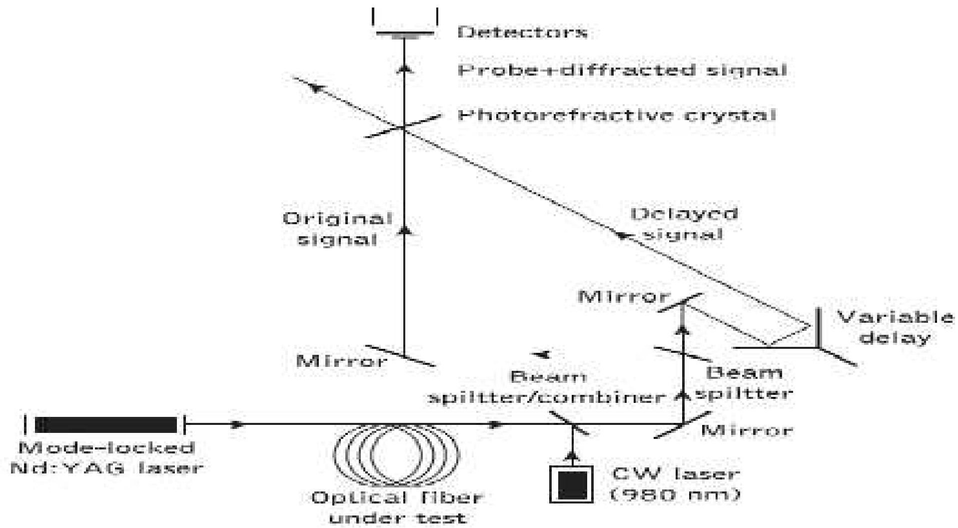


Figure Experimental setup for the measurement of the refractive index of silica fiber using the induced-grating autocorrelation function technique

Figure shows the experimental setup used to observe an IGA response using a nonlinear optical loop mirror interferometer. It consists of a laser source and a combination of optical lenses and mirrors where a beam splitter separates the signal creating the delayed path. The two optical signals (i.e. original and delayed signals) combine at a point where a photorefractive crystal is placed which is the mixing element employed in this method. Several crystalline material systems, known as photorefractive crystals, can be used to produce a diffraction grating in order to implement IGA. Photorefraction is however, an electro-optic phenomenon in which the local index of refraction is modified by spatial variations of the light intensity.

Near-field scanning method

The near-field scanning or transmitted near-field method utilizes the close resemblance that exists between the near-field intensity distribution and the refractive index profile, for a fiber with all the guided modes equally illuminated. It provides a reasonably straightforward and rapid method for acquiring the refractive index profile.

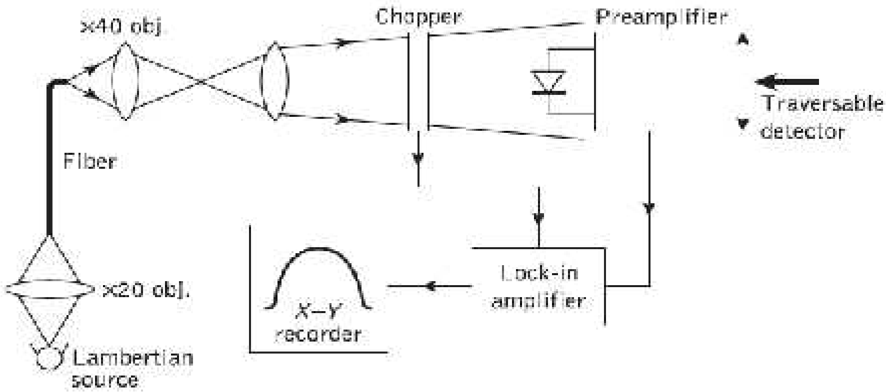


Figure Experimental setup for the near-field scanning measurement of the refractive index profile

When a diffuse Lambertian source (e.g. tungsten filament lamp or LED) is used to excite all the guided modes then the near-field optical power density at a radius *r* from the core axis $P_D(r)$ may be expressed as a fraction of the core axis near-field optical power density $P_D(0)$ following:

$$\frac{P_D(r)}{P_D(0)} = C(r, z)\left[\frac{n_1^2(r) - n_2^2}{n_1^2(0) - n_2^2}\right]$$

where $n_1(0)$ and $n_1(r)$ are the refractive indices at the core axis and at a distance *r* from the core axis respectively, *n*2 is the cladding refractive index and $C(r, z)$ is a correction factor. The correction factor which is incorporated to compensate for any leaky modes present in the short test fiber may be determined analytically.

The transmitted near-field approach is, however, not similarly recommended for single-mode fiber. The output from a Lambertian source is focused onto the end of the fiber using a microscope objective lens. A magnified image of the fiber output end is displayed in the plane of a small active area photodetector (e.g. silicon *p*–*i*–*n* photodiode). The photodetector which scans the field transversely receives amplification from the phase-sensitive combination of the optical chopper and lock-in amplifier. Hence the profile may be plotted directly on an *X*–*Y* recorder.

The test fiber is generally 2m in length to eliminate any differential mode attenuation and mode coupling. A typical refractive index profile for a practical step index fiber measured by the near-field scanning method is shown in Figure. It may be observed that the profile dips in the center at the fiber core axis. This dip was originally thought to result from the collapse of the fiber preform before the fiber is drawn in the manufacturing process but has been shown to be due to the layer structure inherent at the deposition stage.

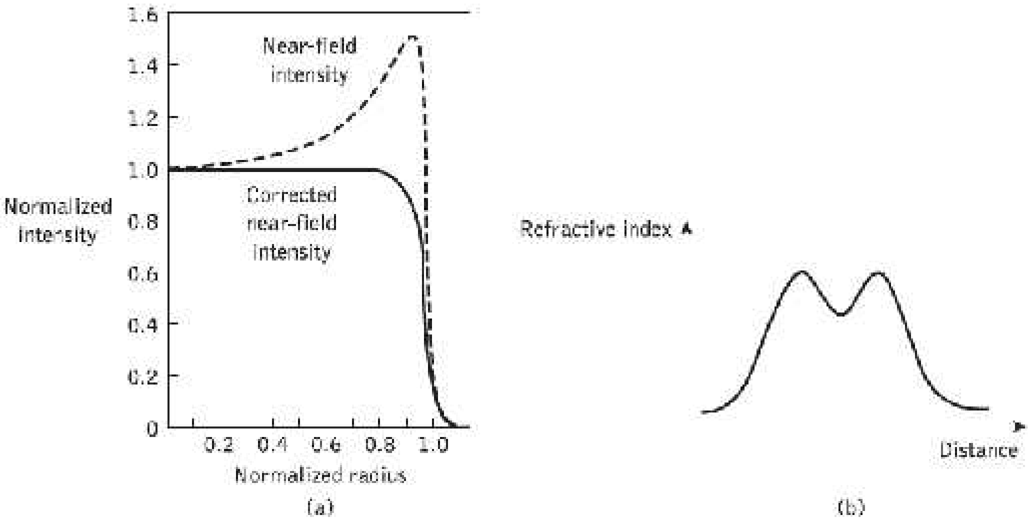


Figure (a) The refractive index profile of a step index fiber measured using the near-field scanning method, showing the near-field intensity and the corrected near field intensity. (b) The refractive index profile of a practical step index fiber measured by the near-field scanning method

Fiber cutoff wavelength measurements

A multimode fiber has many cutoff wavelengths because the number of bound propagating modes is usually large. For example, considering a parabolic refractive index graded fiber, the number of guided modes M_g is:

$$M_g = \left(\frac{\pi a}{\lambda}\right)^2 (n_1^2 - n_2^2)$$

Where a is the core radius and $n1$ and $n2$ are the core peak and cladding indices respectively. It may be observed from Equation that operation at longer wavelengths yields fewer guided modes. Therefore it is clear that as the wavelength is increased, a growing number of modes are cutoff where the cutoff wavelength of a LPlm mode is the maximum wavelength for which the mode is guided by the fiber.

Usually the cutoff wavelength refers to the operation of single-mode fiber in that it is the cutoff wavelength of the LP11 mode (which has the longest cutoff wavelength) which makes the fiber single moded when the fiber diameter is reduced to 8 or 9 μm. Hence the cutoff wavelength of the LP11 is the shortest wavelength above which the fiber exhibits single-mode operation and it is therefore an important parameter to measure. The theoretical value of the cutoff wavelength can be determined from the fiber refractive index profile because of the large attenuation of the LP11 mode near cutoff. However, the parameter which is experimentally determined is called the effective cutoff wavelength, which is always smaller than the theoretical cutoff wavelength by as much as 100 to 200 nm. It is this effective cutoff wavelength which limits the wavelength region for which the fiber is 'effectively' single-mode.

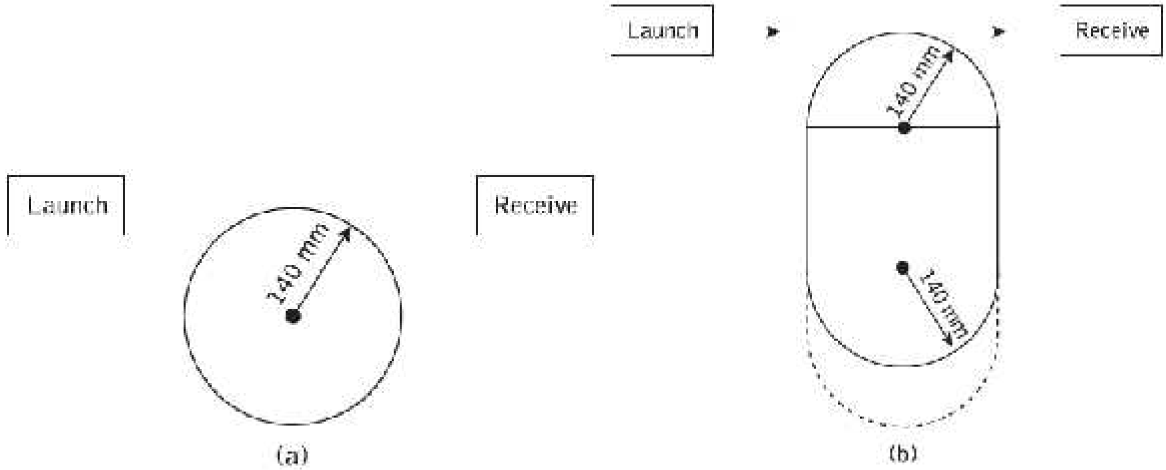


Figure Configurations for the measurement of uncabled fiber cutoff wavelength: (a) single turn; (b) split mandrell

In the bending-reference technique the power $P_s(\lambda)$ transmitted through the fiber sample in the configurations shown in Figure is measured as a function of wavelength. Thus the quantity $P_s(\lambda)$ corresponds to the total power, including launched higher order modes, of the ITU-T definition for cutoff wavelength. Then keeping the launch conditions fixed, at least one additional loop of sufficiently small radius (60 mm or less) is introduced into the test sample to act as a mode filter to suppress the secondary LP11 mode without attenuating the fundamental mode at the effective cutoff wavelength. In this case the smaller transmitted spectral power $P_b(\lambda)$ is measured which corresponds to the fundamental mode power referred to in the definition. The bend attenuation a_b(λ comprising the level difference between the total power and the fundamental power is calculated as:

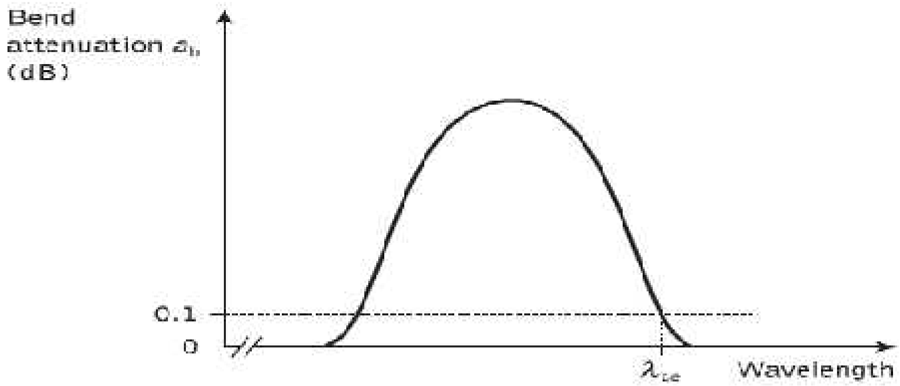


Figure Bend attenuation against wavelength in the bending method

The end attenuation characteristic exhibits a peak in the wavelength region where the radiation losses resulting from the small loop are much higher for the LP11 mode than for the LP01 fundamental mode, as illustrated in figure. It should be noted that the shorter wavelength side of the attenuation maximum corresponds to the LP11 mode, being well confined in the fiber core, and hence negligible loss is induced by the 60 mm diameter loop, whereas on the longer wavelength side the LP11 mode is not guided in the fiber and therefore, assuming that the loop diameter is large enough to avoid any curvature loss to the fundamental mode, there is also no increase in loss.

The relative attenuation $a_m(\lambda)$ or level difference between the powers launched into the multimode and single-mode fibers may be computed as:

$$a_m(\lambda) = 10 \log_{10} \frac{P_s(\lambda)}{P_m(\lambda)}$$

Fiber numerical aperture measurements

The numerical aperture is an important optical fiber parameter as it affects characteristics such as the light-gathering efficiency and the normalized frequency of the fiber (*V*). This in turn dictates the number of modes propagating within the fiber (also defining the single mode region) which has consequent effects on both the fiber dispersion (Intermodal) and, possibly, the fiber attenuation (Differential attenuation of modes). The numerical aperture (*NA*) is defined for a step index fiber as:

$$NA = \sin \theta_a = (n_1^2 - n_2^2)^{\frac{1}{2}}$$

Where Θ_a is the maximum acceptance angle, n_1 is the core refractive index and n_2 is the cladding refractive index. Although equation may be employed with graded index fibers, the numerical aperture thus defined represents only the local *NA* of the fiber on its core axis (the numerical aperture for light incident at the fiber core axis). The graded profile creates a multitude of local *NA*s as the refractive index changes radially from the core axis.

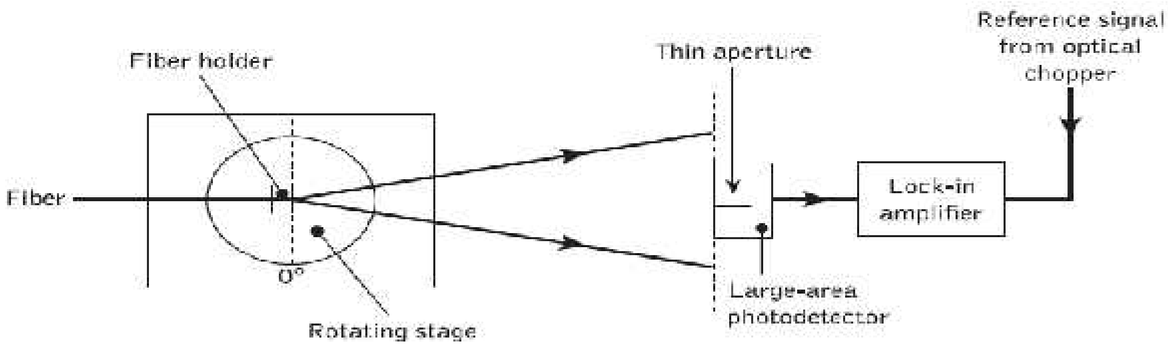


Figure Fiber numerical aperture measurement using a scanning photodetector and a rotating stage

It is assumed that the light is incident on the fiber end face from air with a refractive index (n_0) of unity. For the general case of a graded index fiber these local numerical apertures $NA(r)$ at different radial distances r from the core axis may be defined by:

$$NA(r) = \sin \theta_a(r) = (n_1^2(r) - n_2^2)^{\frac{1}{2}}$$

Therefore, calculations of numerical aperture from refractive index data are likely to be less accurate for graded index fibers than for step index fibers unless the complete refractive index profile is considered. The numerical aperture may be determined by calculation.

An example of an experimental arrangement with a rotating stage is shown in Figure. A 2m length of the graded index fiber has its faces prepared in order to ensure square smooth terminations.

The fiber output end is then positioned on the rotating stage with its end face parallel to the plane of the photodetector input, and so that its output is perpendicular to the axis of rotation. Light at a wavelength of 0.85 μm is launched into the fiber at all possible angles (overfilling the fiber) using an optical system similar to that used in the spot attenuation measurements.

The photodetector, which may be either a small-area device or an apertured large-area device, is placed 10 to 20 cm from the fiber and positioned in order to obtain a maximum signal with no rotation (0°). Hence when the rotating stage is turned the limits of the far-field pattern may be recorded. The output power is monitored and plotted as a function of angle, the maximum acceptance angle being obtained when the power drops to 5% of the maximum intensity. Thus the numerical aperture of the fiber can be obtained. A less precise measurement of the numerical aperture can be obtained from the far-field pattern by trigonometric means. The experimental apparatus is shown in Figure.

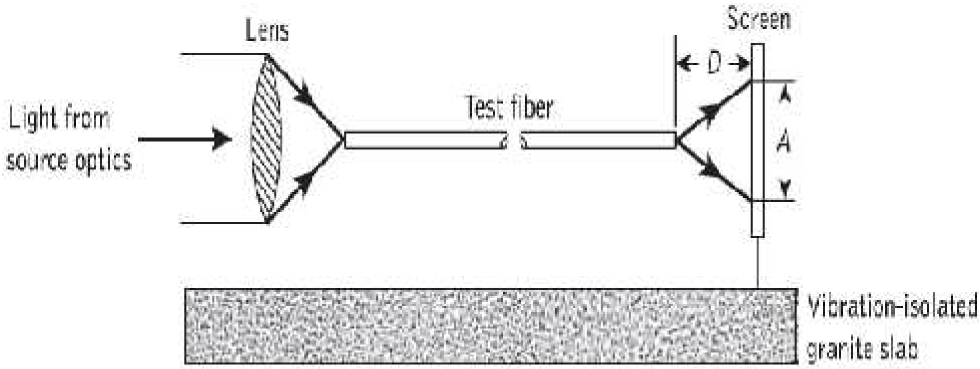


Figure Apparatus for trigonometric fiber numerical aperture measurement

Where the end prepared fiber is located on an optical base plate or slab. Again light is launched into the fiber under test over the full range of its numerical aperture, and the far field pattern from the fiber is displayed on a screen which is positioned a known distance *D* from the fiber output end face. The test fiber is then aligned so that the optical intensity on the screen is maximized. Finally, the pattern size on the screen *A* is measured using a calibrated vernier caliper. The numerical aperture can be obtained from simple trigonometrical relationships where

$$NA = \sin\theta_a = \frac{A/2}{[(A/2)^2 + D^2]^{\frac{1}{2}}} = \frac{A}{(A^2 + 4D^2)^{\frac{1}{2}}}$$

It must be noted that the accuracy of this measurement technique is dependent upon the visual assessment of the far-field pattern from the fiber. The above measurement techniques are generally employed with multimode fibers only, as the far-field patterns from single-mode fibers are affected by diffraction phenomena.

Fiber diameter measurements

Outer diameter

It is essential during the fiber manufacturing process (at the fiber drawing stage) that the fiber outer diameter (cladding diameter) is maintained constant to within 1%. Any diameter variations may cause excessive radiation losses and make accurate fiber–fiber connection difficult. Hence on-line diameter measurement systems are required which provide accuracy better than 0.3% at a measurement rate greater than 100 Hz (i.e. a typical fiber drawing velocity is 1 m/s). Use is therefore made of non contacting optical methods such as fiber image projection and scattering pattern analysis.

The most common on-line measurement technique uses fiber image projection (shadow method) and is illustrated in Figure. In this method a laser beam is swept at a constant velocity transversely across the fiber and a measurement is made of the time interval during which the fiber intercepts the beam and casts a shadow on a photodetector.

In the apparatus shown in Figure the beam from a laser operating at a wavelength of 0.6328 μm is collimated using two lenses (*G*1 and *G*2). It is then reflected off two mirrors (*M*1 and *M*2), the second of which (*M*2) is driven by a galvanometer which makes it rotate through a small angle at a constant angular velocity before returning to its original starting position. Therefore, the laser beam which is focused in the plane of the fiber by a lens (*G*3) is swept across the fiber by the oscillating mirror and is incident on the photodetector unless it is blocked by the fiber. The velocity d*s*/d*t* of the fiber shadow thus created at the photodetector is directly proportional to the mirror velocity following:

$$\frac{\mathrm{d}s}{\mathrm{d}t} = l\frac{\mathrm{d}\phi}{\mathrm{d}t}$$

Where *l* is the distance between the mirror and the photodetector. Furthermore, the shadow is registered by the photodetector as an electrical pulse of width W_e which is related to the fiber outer diameter d_o as:

$$d_o = W_e\frac{\mathrm{d}s}{\mathrm{d}t}$$

Thus the fiber outer diameter may be quickly determined and recorded on the printer. The measurement speed is largely dictated by the inertia of the mirror rotation and its accuracy by the rise time of the shadow pulse. Other on-line measurement methods, enabling faster diameter measurements, involve the analysis of forward or backward far-field patterns which are produced when a plane wave is incident transversely on the fiber. These techniques generally require measurement of the maxima in the center portion of the scattered pattern from which the diameter can be calculated after detailed mathematical analysis. They tend to give good accuracy (e.g. ±0.25 μm) even though the theory assumes a perfectly circular fiber cross-section. Also, for step index fibers the analysis allows determination of the core diameter, and core and cladding refractive indices.

Measurements of the fiber outer diameter after manufacture (off-line) may be performed using a micrometer or dial gage. These devices can give accuracies of the order of ±0.5 μm. Alternatively, off-line diameter measurements can be made with a microscope incorporating a suitable calibrated micrometer eyepiece

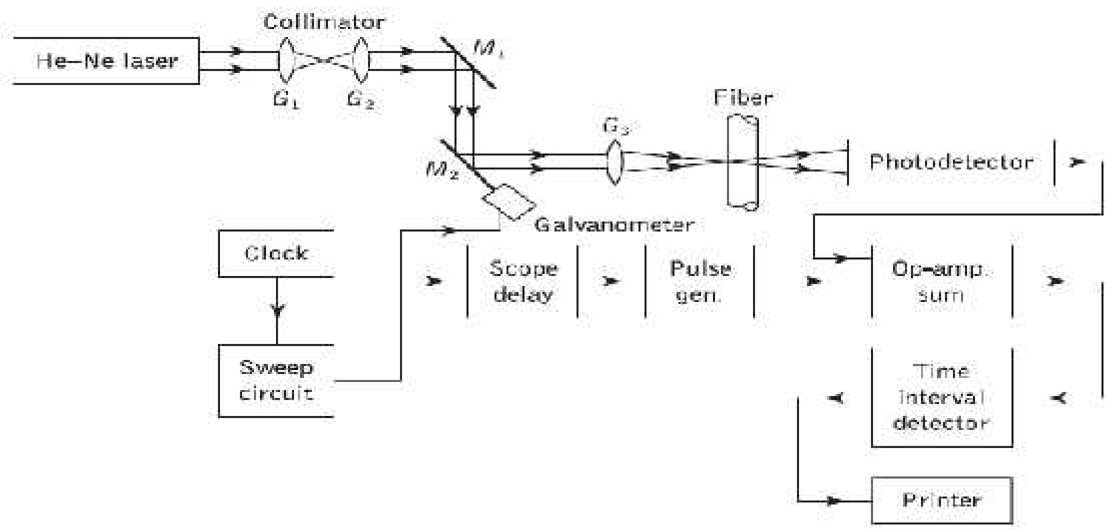


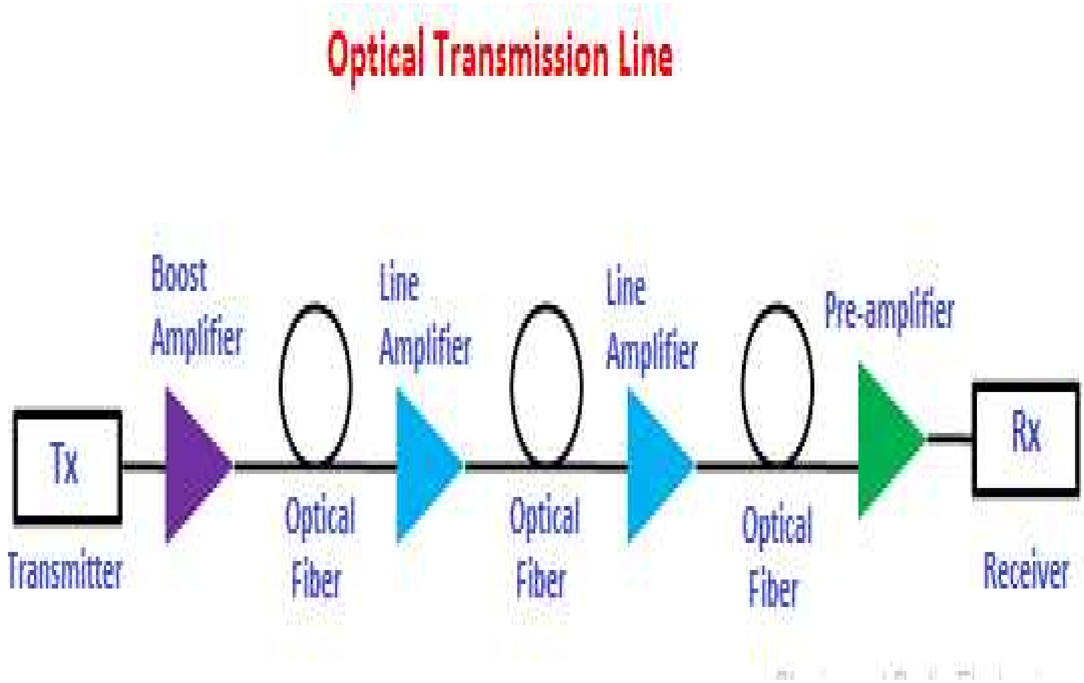
Figure the shadow method for the on-line measurement of the fiber outer diameter.

Core diameter

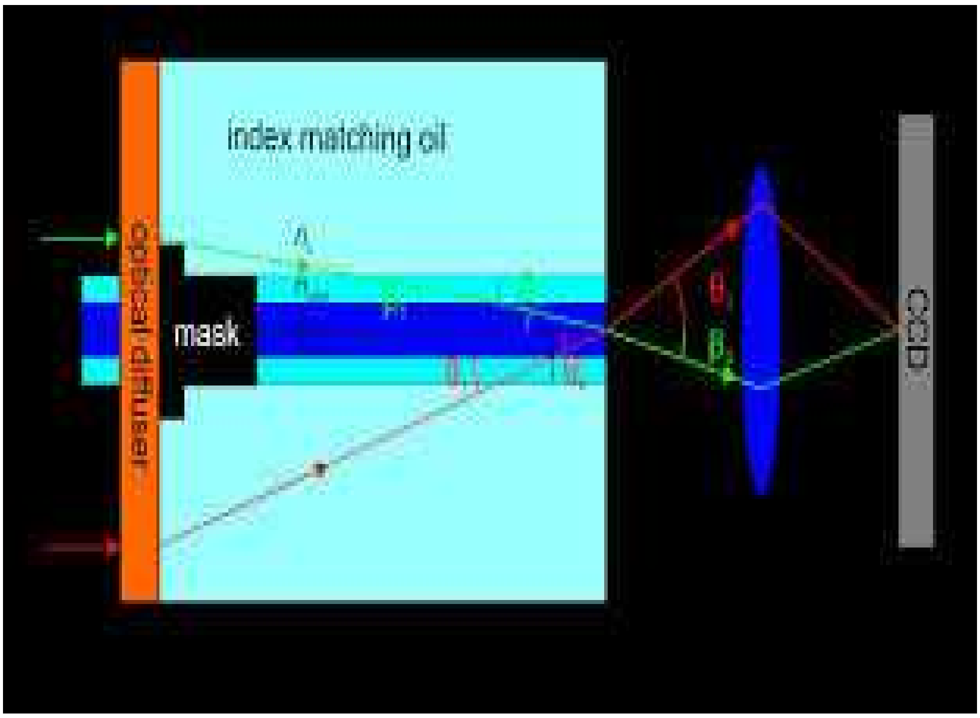
The core diameter for step index fibers is defined by the step change in the refractive index profile at the core–cladding interface. Therefore the techniques employed for determining the refractive index profile (interferometric, near-field scanning, refracted ray, etc.) may be utilized to measure the core diameter. Graded index fibers present a more difficult problem as, in general, there is a continuous transition between the core and the cladding.

In this case it is necessary to define the core as an area with a refractive index above a certain predetermined value if refractive index profile measurements are used to obtain the core diameter. Core diameter measurement is also possible from the near-field pattern of a suitably illuminated (all guided modes excited) fiber. The measurements may be taken using a microscope equipped with a micrometer eyepiece similar to that employed for off-line outer diameter measurements.

APPLICATIONS



Example of pre-amplifier in EDFA



Examples of 2Dimensional refractive index profiling of optical fibers

POST TEST-MCQ TYPE

1. Which refers to any spurious or undesired disturbances that mask the received signal in a communication system?
a) Attenuation
b) Noise
c) Dispersion
d) Bandwidth

2. How many types of noise are observed because of the spontaneous fluctuations in optical fiber communication systems?
a) One
b) Four
c) Two
d) Three

3.Which is caused due to thermal interaction between the free electrons and the vibrating ions in the conduction medium.
a) Thermal noise
b) Dark noise
c) Quantum noise
d) Gaussian noise

4. The minimum pulse energy needed to maintain a given bit-error-rate (BER) which any practical receiver must satisfy is known as
a) Minimal energy
b) Quantum limit
c) Point of reversed
d) Binary signaling

5. Which is used in the specification of optical detectors.
a) Noise equivalent power
b) Polarization
c) Sensitivity
d) Electron movement

6. Which of the following APDs are recognized for their high gain-bandwidth products?
a) GaAs
b) Alloy-made
c) Germanium
d) Silicon

7. How many circuits are present in an equivalent circuit for the digital optical fiber receiver?
a) Four
b) One
c) Three
d) Two

8. Which compensates for distortion of the signal due to the combined transmitter, medium and receiver characteristics?
a) Amplification
b) Distortion
c) Equalization
d) Dispersion

9. The phase frequency response of the system should be _____________ in order to minimize inter-symbol interference.
a) Non-Linear
b) Linear
c) More
d) Less

10. How many amplifier configurations are frequently used in optional fiber communication receivers?
a) One
b) Two
c) Three
d) Four

12. The major advantage of the trans-impedance configuration over the high-impedance front end is
a) Greater bandwidth
b) Less bandwidth
c) Greater dynamic range
d) Less dynamic range

13. The trans-impedance front end configuration operates as a ___________ with negative feedback.
a) Current mode amplifier
b) Voltage amplifier
c) Attenuator
d) Resonator

14. Which is the lowest noise amplifier device?
a) Silicon FET
b) Amplifier-A
c) Attenuator
d) Resonator-B

15. FET device has extremely high input impedance greater than
a) 10^7 Ohms and less than 10^8
b) 10^6 Ohms and less than 10^7
c) 10^{14} Ohms
d) 10^{23} Ohms

16. High-performance microwave FETs are fabricated from
a) Silicon
b) Germanium
c) Gallium arsenide
d) Zinc

17. Which receiver can be fabricated using PIN-FET hybrid approach?
a) Trans-impedance front end receiver
b) Gallium arsenide receiver
c) High-impedance front-end
d) Low-impedance front-end

18. What is usually required by FETs to optimize the figure of merit?
a) Attenuation of barrier
b) Matching with the depletion region
c) Dispersion of the gate region
d) Matching with the detector

19. A technique used for determining the total fiber attenuation per unit length is ________ method.
a) Frank
b) Cut-off
c) cut-back
d) Erlangen

20. The system designer finds greatest interest in the
a) Overall fiber attenuation
b) Fiber dispersion
c) Latitude of the fiber
d) Durability

21. How many parameters are usually worked upon by the measurement techniques in attenuation?
a) Three
b) Two
c) One
d) Five

22. What type of light source is usually present in the cut-back method?
a) Tungsten or xenon
b) LED
c) Laser
d) Photo-sensor

23. The device used to remove any scattered optical power from the core is
a) Mode setup terminator
b) Nodal spectrum
c) Mode stripper
d) Attenuator

24. What is the unit of measurement of the optical attenuation per unit length?
a) dB-km
b) dB/km
c) km/dB
d) V

25. What are used to allow measurements at a selection of different wavelengths?
a) Diaphragms
b) Spot attenuators
c) Belts
d) Interference filters

26. Which technology is used by the backscatter measurement method?
a) Refraction
b) Francis flat recovery
c) Optical time domain reflectometry
d) Optical frequency

27. Which measurements checks the impurity level in the manufacturing process?
a) Material reflectometry
b) Material absorption loss
c) Material attenuation loss
d) Calorimetric loss

28. Which removes the light propagating in the cladding?
a) Cladding mode strippers
b) Core strippers
c) Mode enhancers
d) Attenuators

29. Which measurements give an indication of the distortion to the optical signals as they propagate down optical fibers?
a) Attenuation
b) Dispersion
c) Encapsulation
d) Frequency

30. The measurement of dispersion allows the __________ of the fiber to be determined.
a) Capacity
b) Frequency
c) Bandwidth
d) Power

31. How many types of mechanisms are present which produce dispersion in optical fibers?
a) Three
b) Two
c) One
d) Four

32. Intermodal dispersion is nonexistent in ________ fibers.
a) Multimode
b) Single mode
c) Step index- multimode
d) Al-GU

33. In the single mode fibers, the dominant dispersion mechanism is
a) Intermodal dispersion
b) Frequency distribution
c) Material dispersion
d) Intra-modal dispersion

34. How many domains support the measurements of fiber dispersion?
a) One
b) Three
c) Four
d) Two

35. The time domain dispersion measurement setup involves ______________ as the photo detector.
a) Avalanche photodiode
b) Oscilloscope
c) Circulator
d) Gyrator

36. The detailed knowledge of the refractive index profile predicts the ___________ of the fiber.
a) Nodal response
b) Variation in frequency
c) Impulse response
d) Amplitude

37. Which of the fiber is strongly dependent on the refractive index profile?
a) Amplitude
b) Tuning frequency
c) Diameter
d) Information carrying capacity

38. What is required in case of graded index fibers?
a) High amplitude
b) High frequency
c) High impulse response
d) Optimum profile

39. Which of the following have been widely used to determine the refractive index profiles of optical fibers?
a) Interference microscopes
b) Gyro meters
c) Mode-diameter device
d) Tunable microscopes

40. Which of the following is the main drawback of the slab technique?
a) Efficiency
b) Amplitude
c) Time required
d) Accuracy

41. What does ‘a’ stands for in the given equation?
$M_g=(\pi a/\lambda)^2(n_1^2 - n_2^2)$
a) Radius of the core
b) Constant
c) Coefficient of refraction
d) Density

42. What is the name of the test used to determine the efficient values of the effective cutoff wavelength?
a) Round robin test
b) Mandarin test
c) Hough Werner test
d) Fulton test
43. How many bend effects are produced in the fiber?
a) One
b) Three
c) Two
d) Four

44. Which method is the most commonly used method for the determination of the fiber refractive index profile?
a) Refracted near-field method
b) Bending-reference
c) Power step method
d) Alternative test method
45. The numerical aperture for a step index fiber is sine angle of the
a) Efficient angle
b) Aperture
c) Acceptance angle
d) Attenuation

46. Far field pattern measurements with regard to multimode fibers are dependent on the ______________ of the fiber.
a) Amplitude
b) Frequency
c) Diameter
d) Length

47. Which of the following is a non-contacting optical method of on-line diameter measurement?
a) Brussels’s method
b) Velocity differentiator method
c) Photo detector method
d) Image projection method

48. Which affects both the fiber attenuation and dispersion?
a) Refractive index
b) Micro-bending
c) Connectors
d) Splices

49. Which of the following is not included in the optical fiber link measurement test?
a) Attenuation measurement
b) Dispersion measurement
c) Splice loss measurement
d) Receiver sensitivity

50. The handheld optical power meter has a measurement accuracy of ________
a) 0.01 dB
b) 0.25 dB
c) 0.8 dB
d) 1 dB

51. Which may be used for measurement of the absolute optical attenuation on a fiber link?
a) Silicon photodiodes
b) InGaAsP photodiodes
c) Optical power meters
d) Gyrators

52. During the fiber drawing process, the fiber outer diameter is maintained constant to within
a) 2%
b) 1%
c) 5%
d) 10%

53. What is the minimum value of accuracy in diameter is needed to avoid radiation losses in the fiber?
a) 0.1%
b) 0.2%
c) 0.3%
d) 0.03%

54. Which of the following is a non-contacting optical method of on-line diameter measurement?
a) Brussels's method
b) Velocity differentiator method
c) Photo detector method
d) Image projection method

CONCLUSION

In this unit, the characteristics of fiber optic receivers were discussed. The Fiber Attenuation measurements, Dispersion measurements, Fiber Refractive index profile measurements, Fiber cut- off Wave length Measurements, Fiber Numerical Aperture Measurements, Fiber diameter measurements and their techniques were learnt.

REFERENCES

1. Gerd Keiser, “Optical Fiber Communication” McGraw-Hill International,4th Edition. 2010.
2. John M. Senior, “Optical Fiber Communication”, Second Edition, Pearson Education, 2007.
3. Ramaswami, Sivarajan and Sasaki “Optical Networks”, Morgan Kaufmann, 2009.
4. J.Senior, Optical Communication, Principles and Practice,Prentice Hall of India, 3rd Edition, 2008.
5. J.Gower, “Optical Communication System”, Prentice Hall of India, 2001.

ASSIGNMENT

1. With diagram explain the measurement of Cutoff wavelength of optical fiber?
2. With diagram explain the measurement of numerical aperture of a fiber and measurement of refractive index fiber?
3. Explain detail in fiber attenuation measurement and Dispersion measurements.
4. Explain the error sources and device the Probability of error for fiber optic system?
5. Explain the fiber optic receiver operation?
6. Write short notes on quantum limit?

AIM & OBJECTIVES

- ❖ To learn the basic elements of optical fiber transmission link, fiber modes configurations and structures.
- ❖ To understand the different kind of losses, signal distortion, SM fibers.
- ❖ To learn the various optical sources, materials and fiber splicing.
- ❖ To learn the fiber optical receivers and noise performance in photo detector.
- ❖ To explore link budget, WDM, solitons and SONET/SDH network.

PRE TEST-MCQ TYPE

1. The network structure formed due to the interconnectivity patterns is known as a
a) Network
b) Struck
c) Topology
d) D-pattern

2. In the ___________ topology, the data generally circulates bi-directionally.
a) Mesh
b) Bus
c) Star
d) Ring

3. The ring and star topologies are combined in a _________ configuration.
a) Mesh
b) Fringe
c) Data
d) Singular

4. Packet switching is also called as
a) Frame switching
b) Cell switching
c) Trans-switching
d) Buffer switching

5. A ________________ is a series of logical connections between the source and destination nodes.
a) Cell circuit
b) Attenuation circuit
c) Virtual circuit
d) Switched network

UNIT V OPTICAL NETWORKS AND SYSTEM TRANSMISSION

Basic Networks – SONET / SDH – Broadcast – and –select WDM Networks –Wavelength Routed Networks – Non linear effects on Network performance –Link Power budget -Rise time budget- Noise Effects on System Performance-Operational Principles of WDM Performance of WDM + EDFA system – Solutions – Optical CDMA – Solitons in Optical Fiber -Ultra High Capacity Networks.

THEORY

Introduction

SONET/SDH

- SONET is the TDM optical network standard for North America
- SONET is called Synchronous Digital Hierarchy (SDH) in the rest of the world
- SONET is the basic physical layer standard
- Other data types such as ATM and IP can be transmitted over SONET
- OC-1 consists of 810 bytes over 125 us; OC-n consists of 810n bytes over 125 us
- Linear multiplexing and de-multiplexing is possible with Add-Drop-Multiplexers
- The SONET/SDH standards enable the interconnection of fiber optic transmission equipment from various vendors through multiple-owner trunk networks.
- The basic transmission bit rate of the basic SONET signal is
- In SDH the basic rate is 155.52 Mb/s.

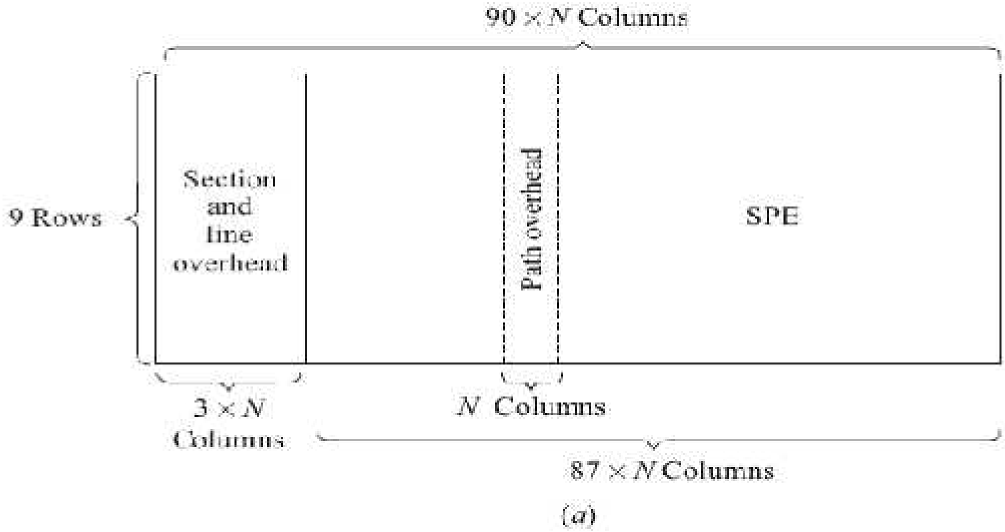


Figure Basic formats of an STS-N SONET frame

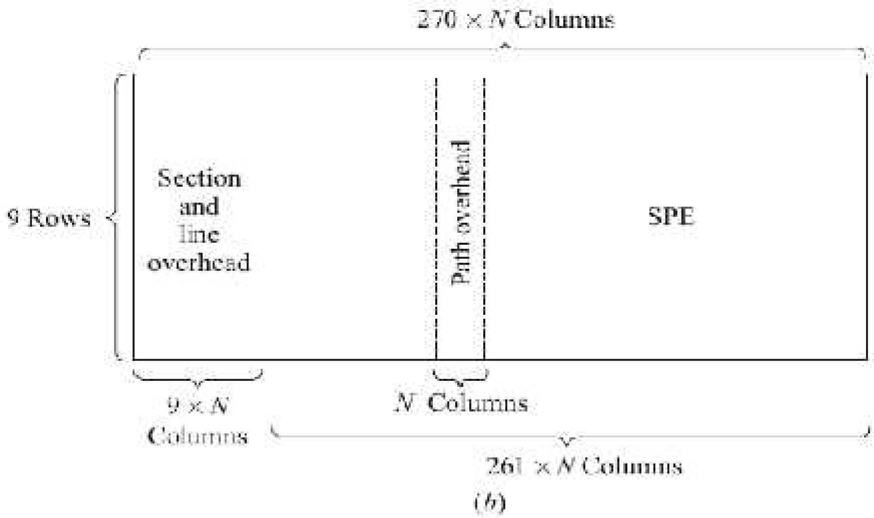


Figure Basic formats of an STM-N SDH frame

Common values of OC-N and STM-N:

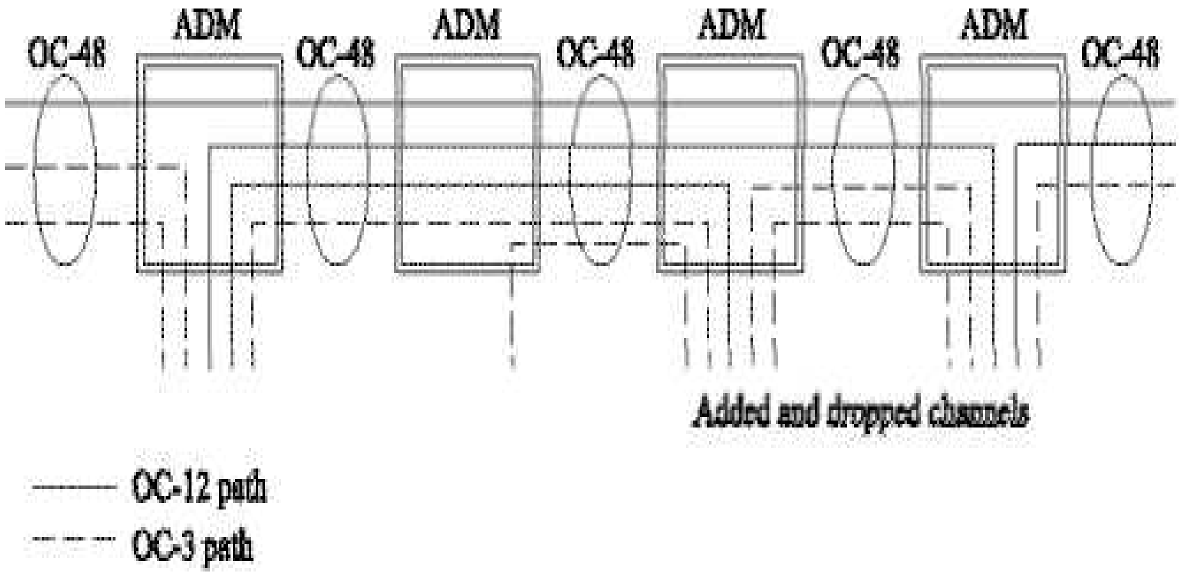
- OC stands for optical carrier. It has become common to refer to SONET links as OC- N links.
- The basic SDH rate is 155.52 Mb/s and is called the synchronous transport module-level 1 (STM 1).

SONET Add Drop Multiplexers:

SONET ADM is a fully synchronous, byte oriented device, that can be used add/drop OC sub- channels within an OC-*N* signal

Ex: OC-3 and OC-12 signals can be individually added/ dropped from an OC-48 carrier

SONET level	Electrical level	SDH level	Line rate (Mb/s)	Common rate name
OC-N	STS-N	—	N × 51.84	—
OC-1	STS-1	—	51.84	—
OC-3	STS-3	STM-1	155.52	155 Mb/s
OC-12	STS-12	STM-4	622.08	622 Mb/s
OC-48	STS-48	STM-16	2488.32	2.5 Gb/s
OC-192	STS-192	STM-64	9953.28	10 Gb/s
OC-768	STS-768	STM-256	39813.12	40 Gb/s



SONET/SDH Rings:

- SONET and SDH can be configured as either a ring or mesh architecture
- SONET/SDH rings are self-healing rings because the traffic flowing along a certain path can be switched automatically to an alternate or standby path following failure or degradation of the link segment
- Two popular SONET and SDH networks:
  - 2-fiber, unidirectional, path-switched ring (2-fiber UPSR)
  - 2-fiber or 4-fiber, bidirectional, line-switched ring (2-fiber or 4-fiber BLSR)

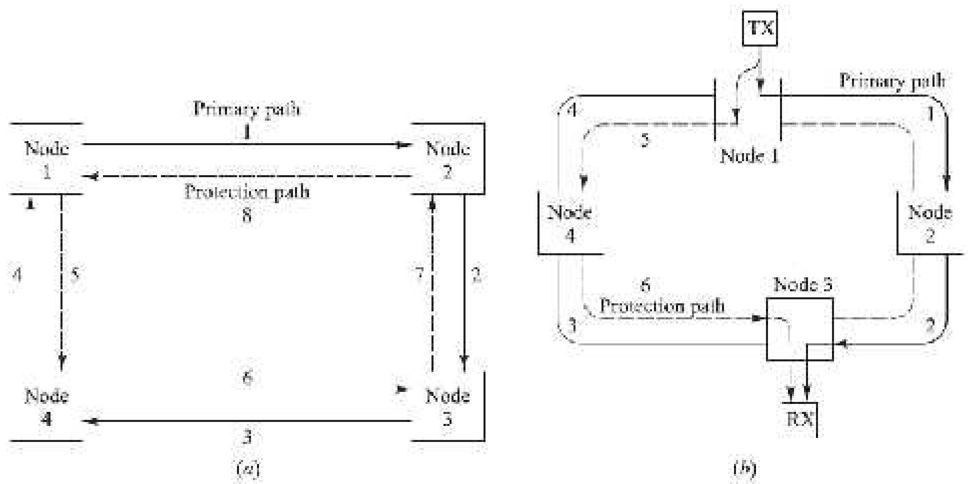
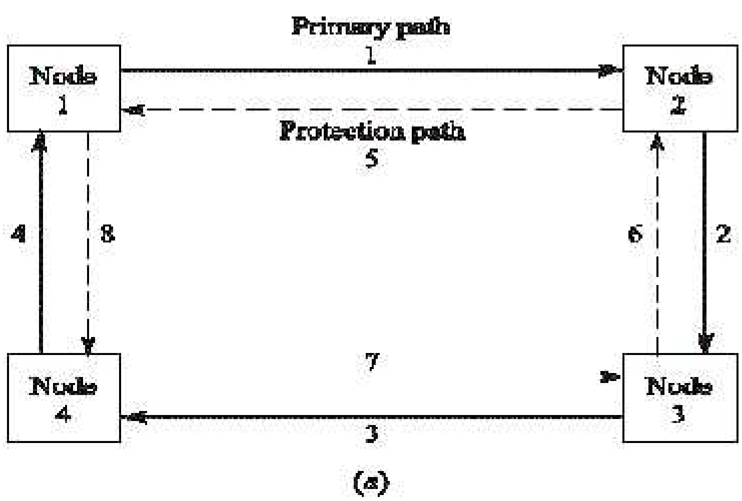


Figure Generic 2-fiber UPSR with a counter-rotating protection path

2-Fiber UPSR Basics:



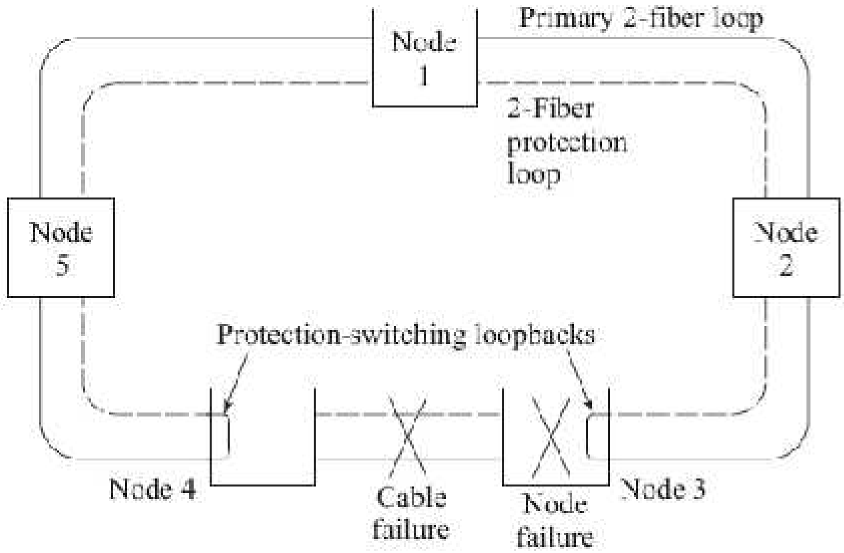
Ex: Total capacity OC-12 may be divided to four OC-3 streams, the OC-3 is called a path here

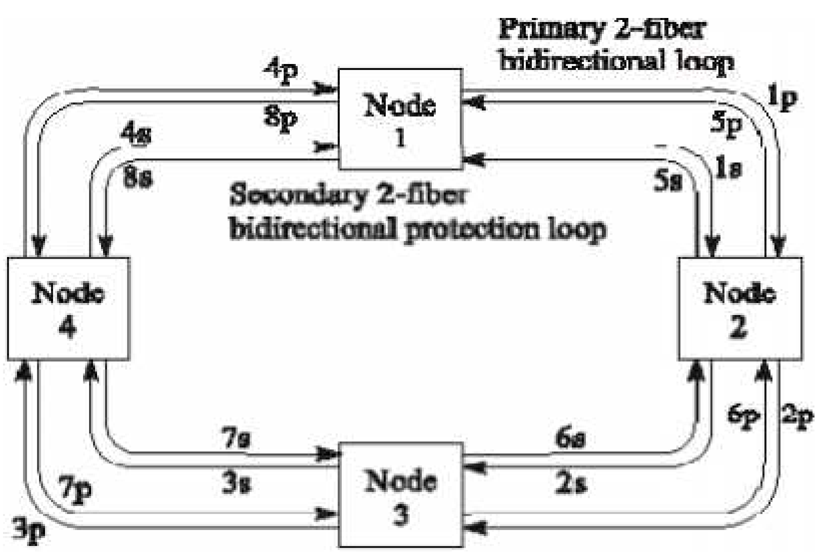
2-Fiber UPSR Protection:

- ❖ Rx compares the signals received via the primary and protection paths and picks the best one
- ❖ Constant protection and automatic switching

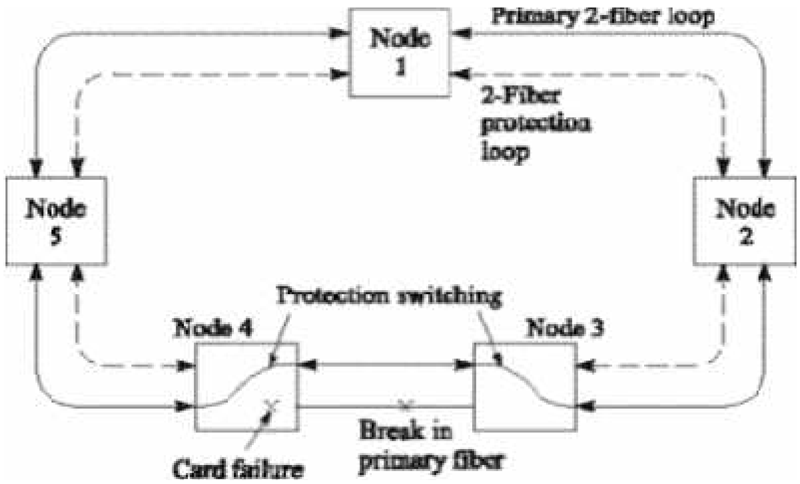
4-Fiber BLSR Basics:

Node 1 3; 1p, 2p Node3 1;3p, 4p



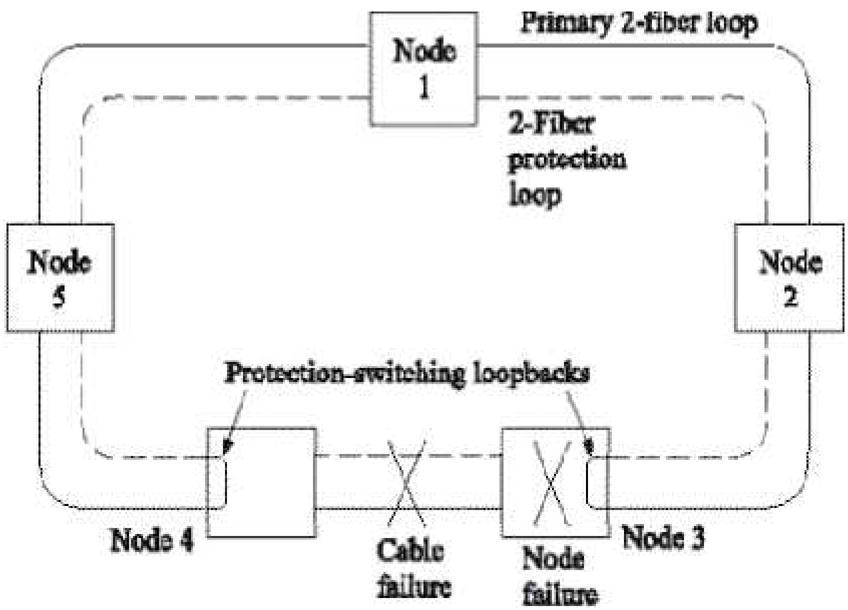


BLSR Fiber-Fault Reconfiguration:



In case of failure, the secondary fibers between only the affected nodes (3 & 4) are used, the other links remain unaffected

BLSR Node-Fault Reconfiguration



If both primary and secondary are cut, still the connection is not lost, but both the primary and secondary fibers of the entire ring is occupied

BLSR Recovery from Failure Modes:

- If a primary-ring device fails in either node 3 or 4, the affected nodes detect a loss-of-signal condition and switch both primary fibers connecting these nodes to the secondary protection pair

- If an entire node fails or both the primary and protection fibers in a given span are severed, the adjacent nodes switch the primary-path connections to the protection fibers, in order to loop traffic back to the previous node.

Broadcast and Select WDM Networks

Optical signals of different wavelength can propagate without interfering with each other. The scheme combining a number of wavelengths over a single fiber is wavelength division multiplexing.

Two categories of broadcast and select WDM networks

1. Single hop networks

2. Multihop networks

Broadcast and select single hop networks

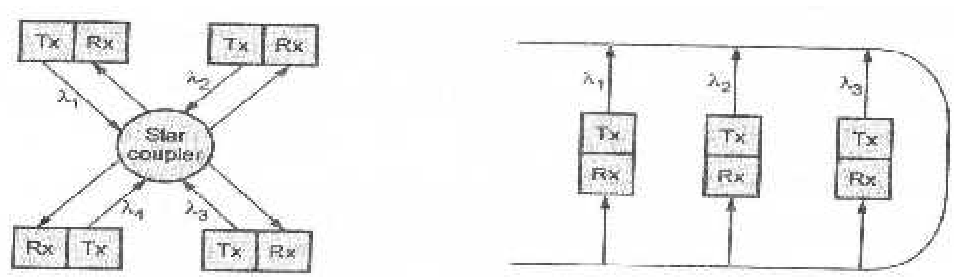


Fig: star configuration and bus configuration

In single hop network, data transmitted reaches its destination without being converted to electrical energy at any intermediate point

Two physical configurations: star and bus

Each transmitter sends its information at different wavelengths. All transmissions from various nodes are combined in a passive star coupler or coupled onto a bus. The result is sent to all receivers. A coupler is a device which is used to combine and split signals in an optical network. Each receiver sees all wavelengths and uses a tunable filter to select the particular wavelength.

Passive star topology is attractive

- No tapping or insertion loss

- Logarithmic splitting loss in the coupler Advantages of single hop networks

- Simple network architecture

- Protocol transparent

Disadvantages

- It needs rapidly tunable lasers

Broadcast and select multihop networks

Intermediate electro optical conversion may take place. Each node has fixed tuned optical transmitters and receivers. Each node transmits signals on its wavelengths and presented to WDM mux.

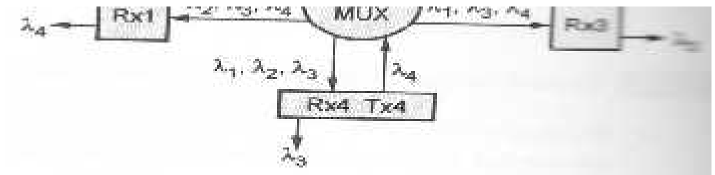


Figure multihop broad cast and select networks

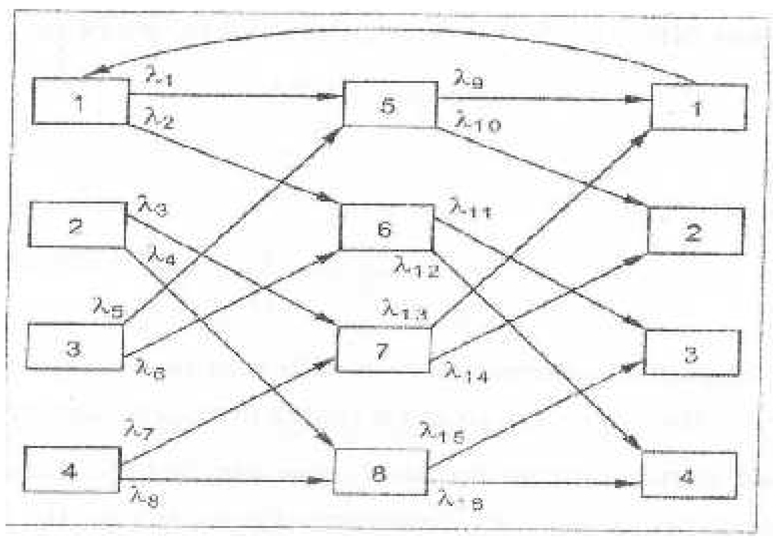


Figure logical interconnection pattern and wavelength assignment of a(p,) = (2,2) shuffle net

A WDM mux is a passive device that does wavelength division multiplexing and transmits a multiplexed signal further along the fiber

Shuffle net Multi hop network

One of the topologies for multihop networks is shuffle net. A cylindrical arrangement of ‘k’ columns, each having ‘P_k’ nodes where P is the number of fixed transceiver per node.

Total number of nodes, $N = k P_k$ with k = 1,2,3… and P =1,2,3..

Each node requires P wavelengths to transmit information, the total number of wavelengths $N_\lambda = PN = k P_k+1$

Maximum no. of hops = Hmax = 2^{k-1}. Consider the connections between node 1 and 5 and between nodes 1 and 7. First case hop number is one. Second case, three hops are needed. Per user throughput 'S' = C/N, where C= total network capacity

Advantages of multihop networks

1. No packet collision within the network

2. Rapidly tunable lasers are not required

Disadvantages

There is a throughput per delay of 1/H for H hops between nodes.

Wavelength Routed Networks

Three network nodes are interconnected using two wavelength channels where the solid line connecting the nodes represents the available wavelength channel and the dashed line identifies that the wavelength channel is in use.

If the network node 1 is required to connect with node 3 then as indicated in figure. There is no single wavelength channel available to establish a light path between them. When a light path cannot be established on a link using a single wavelength channel it is referred to as a wavelength continuity constraint.

To reduce this wavelength continuity constraint is to switch the wavelength channel at node 2 by converting the incoming wavelength $\lambda 2$ to $\lambda 1$ (which is available between nodes 2 and 3) to enable a link between node 2 and 3 to be established. The newly set up path uses two wavelength stages (i.e. two hops) to interconnect nodes 1 and 3. Such networks which employ wavelength conversion devices (or switches) are known as wavelength convertible networks. Three different WDM network architectures employing the wavelength conversion function are Full wavelength conversion, where each network link utilizes a dedicated wavelength converter, is depicted in Figure. All the wavelength channels at the output port of the optical switch will be converted into their compliant wavelength channel by the appropriate wavelength converter (WC).

It is more cost effective to implement networks with fewer and hence shared wavelength converters. The arrangement of wavelength converters organized in a WCB is illustrated in the inset to Figure. This figure depicts a WCB servicing the optical fiber links where only the required wavelength channels are switched through the WCB. By contrast two optical switches are required to construct the shared per node wavelength convertible network architecture indicated in Figure. Optical switch 2 switches the converted wavelength channels to their designated nodes. In dense WDM networks a light path is established by reserving a particular wavelength on the physical links between the source and destination edge nodes.

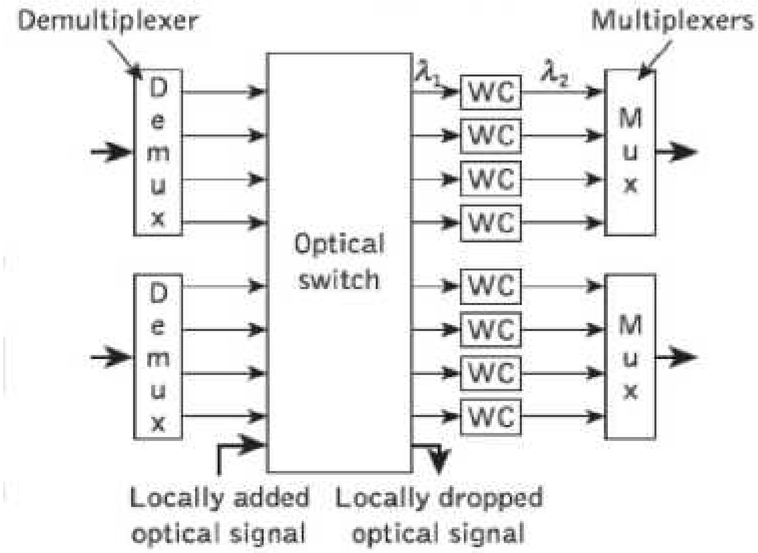


Figure Wavelength convertible routing network architectures: full or rededicated wavelength converters;

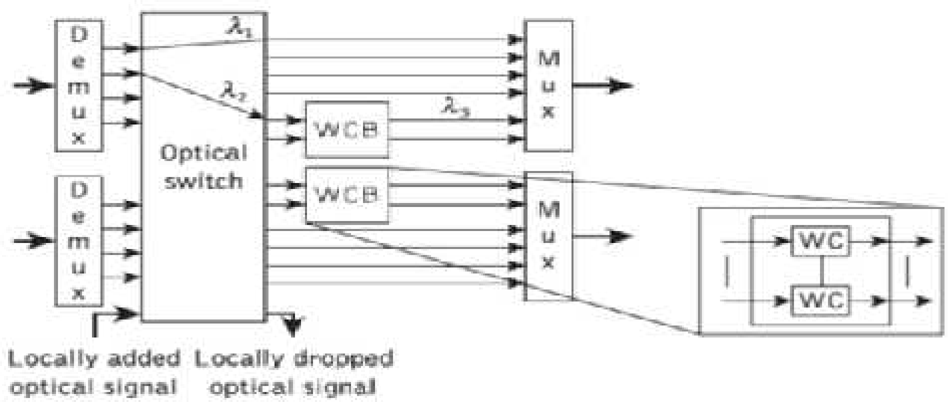


Figure Wavelength convertible routing network architectures: shared per link

It is a two-stage search and-select process related to both routing (i.e. searching/selecting a suitable path) and wavelength assignment (i.e. searching/selecting or allocating an available wavelength for the connection). The overall process is often referred to as the routing and wavelength assignment (RWA) problem. The implementation of RWA can be static or dynamic depending upon the traffic patterns in the network. Static RWA techniques are employed to provide a set of semi permanent connections, which remain active for a relatively longer time.

Dynamic RWA deals with establishing the light path in frequently varying traffic patterns. The traffic patterns are not known and therefore the connection requests are initiated in a random fashion, depending on the network state at the time of a request each time a request is made, an algorithm must be executed in real time to determine whether it is feasible to accommodate the request and, if so, to perform RWA.

A five-node network with fixed connections where node 1 requested to establish a link with node 5 is illustrated in Figure. Although there is no direct physical connection or path available, there are four possibilities to establish the link between nodes 1 and 5, depending on the available or assigned wavelengths between each of the network nodes. These are: via node 2 using a single hop; nodes 4 and 2 comprising two hops; nodes 2 and 3 with two hops; and the longest possible route stretching over three hops via nodes 4, 2 and 3. Considering these four routes, the single hop remains the shortest path between nodes 1 and 5.

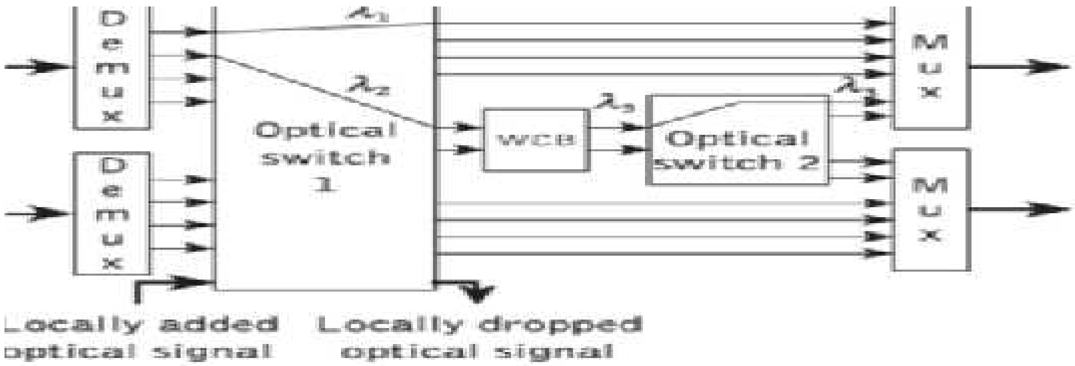


Figure wavelength routing and selection of a path

Non linear effects on Network performance

There are two categories of nonlinear effects.

The first arises due to the interaction of light waves with phonon (molecular vibrations) in the silica medium Rayleigh scattering. The two main effects in this category are stimulated Brillouin scattering (SBS) and stimulated Raman scattering (SRS). The second set of nonlinear effects arises due to the dependence of the refractive index on the intensity of the applied electric field, which in turn is proportional to the square of the field amplitude. The most important nonlinear effects in this category are self-phase modulation (SPM) and four-wave mixing (FWM). Modeling the nonlinear processes can be quite complicated, since they depend on the transmission length, the cross-sectional area of the fiber, and the optical power level in the fiber.

Stimulated Raman scattering

Stimulated Raman scattering is an interaction between light waves and the vibrational modes of silica molecules. If a photon with energy hv1 is incident on a molecule having a vibrational frequency vm, the molecule can absorb some energy from the photon.

In this interaction the photon is scattered, thereby attaining a lower frequency v2 and a corresponding lower energy hv2. The modified photon is called a Stokes photon. Because the optical signal wave that is injected into a fiber is the source of the interacting photons, it is often called the pump wave , since it supplies power for the newly generated wave.

This process generates scattered light at a wavelength longer than that of the incident light. If another signal is present at this longer wavelength, the SRS light will amplify it and the pump wavelength signal will decrease in power

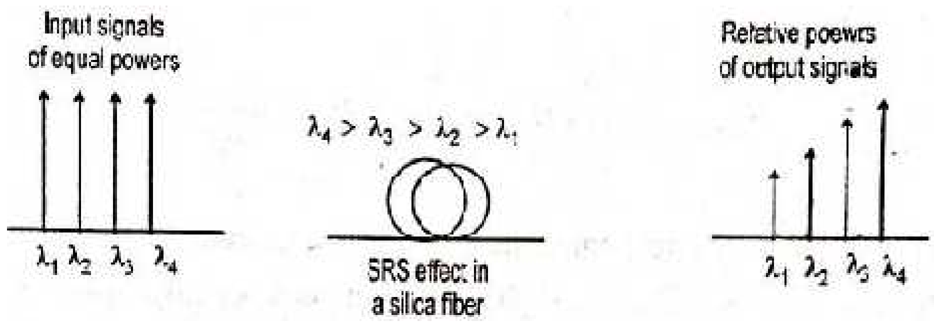


Figure SRS generates scattered light at a longer wavelength, thereby decreasing the power in the pump wavelength signal.

Stimulated Brillouin scattering

Stimulated Brillouin scattering arises when light waves scatter from acoustic waves. The resultant scattered wave propagates principally in the backward direction in single-mode fibers. This backscattered light experiences gain from the forward-propagating signals, which leads to depletion of the signal power. The frequency of the scattered light experiences a Doppler shift

given by $V_B = 2nV_s / \lambda$

where n is the index of refraction and Vs is the velocity of sound in the material.

The effects of SBS accumulate individually for each channel, and consequently they occur at the same power level in each channel as occurs in a single-channel system.

Self-Phase Modulation (SPM)

SPM arises because the refractive index of the fiber has an intensity- dependent component. This nonlinear refractive index causes an induced phase shift that is proportional to the intensity of the pulse. Thus different parts of the pulse undergo a different phase shift which gives rise to chirping of the pulses. Pulse chirping in turn enhances the pulse-broadening effects of chromatic dispersion. This chirping effect is proportional to the transmitted signal power so that SPM effects are more pronounced in systems using high transmitted powers.

In WDM systems, the refractive index nonlinearity gives rise to cross-phase modulation (XPM), which converts power fluctuations in a particular wavelength channel to phase fluctuations in other co propagating channels. This can be mitigated greatly in WDM systems operating over standard non- dispersion shifted single-mode fiber, but can be a significant problem in WDM links operating at 10 Gbps and higher over dispersion-shifted fiber.

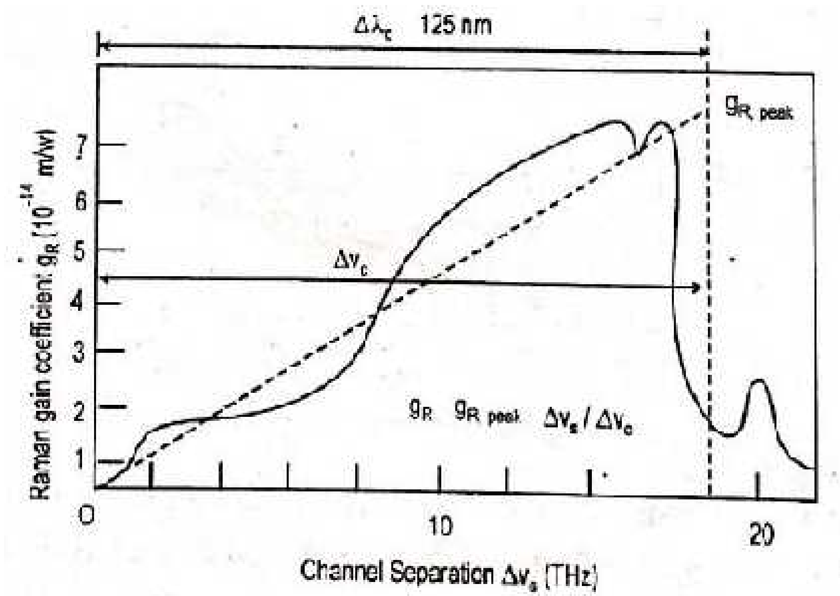


Figure Wavelength Channel separation

Four-wave mixing

Four-wave mixing is a third-order nonlinearity in silica fibers that is analogous to inter modulation distortion in electrical systems. When wavelength channels are located near the zero-dispersion point, three optical frequencies (ν_i, ν_j, ν_k) will mix to produce a fourth inter modulation product νijk given by

$V_{ij}k = V_i + V_j - V_K$ with i,j # k

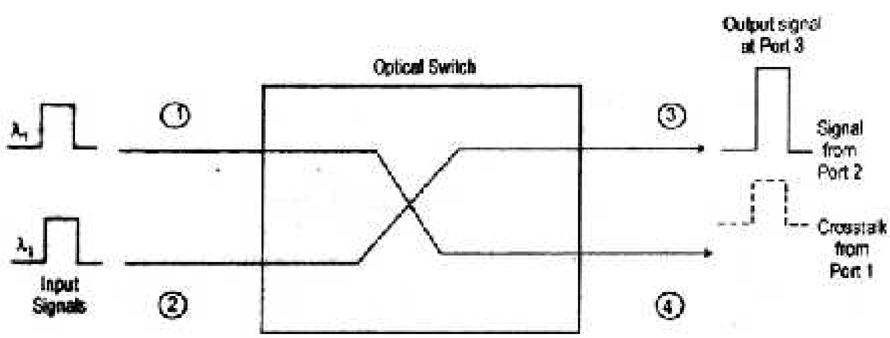


Figure Origin of interchannel crosstalk

Figure shows a simple example for two waves at frequencies ν_1 and ν_2. As these waves co propagate along a fiber, they mix and generate sidebands at

$2V_1 – V_2$ and $2V_2 – V_1$

When this new frequency falls in the transmission window of the original frequencies, it can cause severe crosstalk

LINK POWER BUDGET:

For optimizing link power budget an optical power loss model is to be studied as shown in Figure. Let

- l_c denotes the losses occur at connector.
- L_{sp} denotes the losses occur at splices.
- α_f denotes the losses occur in fiber.

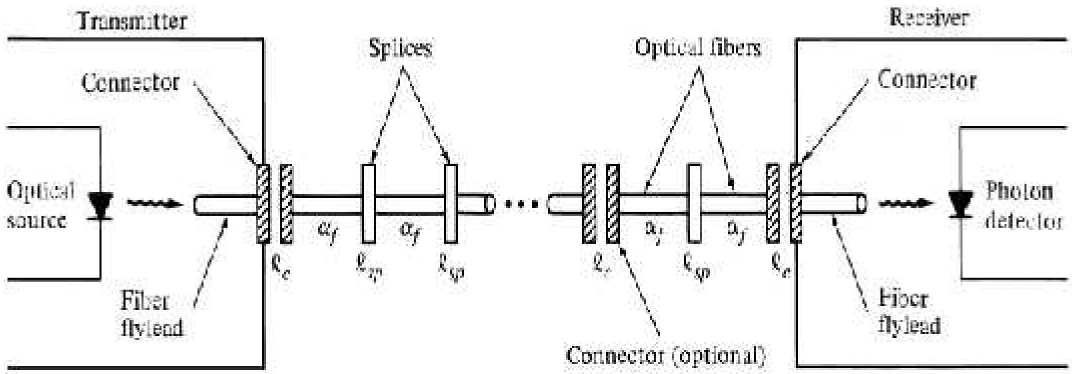


Figure Link power budget

All the losses from source to detector comprises the total loss (P_T) in the system. Link power margin considers the losses due to component aging and temperature fluctuations.

Usually a link margin of 6-8 dB is considered while estimating link power budget. Total optical loss = Connector loss + (Splicing loss + Fiber attenuation) + System margin (P_m)

$$P_T = 2lc + \alpha_{fL} + \text{System margin } (P_m)$$

Where, L is transmission distance.

Rise Time Budget

Rise time gives important information for initial system design. Rise-time budget analysis determines the dispersion limitation of an optical fiber link. Total rise time of a fiber link is the root-sum-square of rise time of each contributor to the pulse rise time degradation.

$$t_{sys} = \sqrt{t_{r1}^2 + t_{r2}^2 + t_{r3}^2 + \cdots}$$

$$t_{sys} = \left(\sum_{i=1}^{N} t_{ri}^2 \right)^{1/2}$$

The link components must be switched fast enough and the fiber dispersion must be low enough to meet the bandwidth requirements of the application adequate bandwidth for a system can be assured by developing a rise time budget. As the light sources and detectors has a finite response time to inputs. The device does not turn-on or turn-off instantaneously. Rise time and fall time determines the overall response time and hence the resulting bandwidth.

Connectors, couplers and splices do not affect system speed, they need not be accounted in rise time budget but they appear in the link power budget. Four basic elements that contributes to the rise-time are, Transmitter rise-time (ttx)

Group Velocity Dispersion (GVD) rise time (t_{GVD}) Modal dispersion rise time of fiber (t_{mod})

Receiver rise time (trx) Where,

Rise time due to modal dispersion is given as

$$t_{mod} = \frac{440}{B_M} = \frac{440\,Lq}{B_0}$$

Where,

B_M is bandwidth (MHz) L is length of fiber (km)

q Is a parameter ranging between 0.5 and 1.

$$t_{sys} = \left[t_{tx}^2 + t_{mod}^2 + t_{GVD}^2 + t_{rx}^2\right]^{1/2}$$

B_0 is bandwidth of 1 km length fiber

Rise time due to group velocity dispersion is

$$t_{GVD} = D^2 \sigma_\lambda^2 L^2$$

Where, D is dispersion [ns/(nm.km)] Σλ is half-power spectral width of source L is length of fiber

Receiver front end rise-time in nanoseconds is

$$t_{rx} = \frac{350}{B_{rx}}$$

B_{rx} is 3 dB – bW of receiver (MHz).

Equation can be written as

$$t_{sys} = \left[t_{tx}^2 + t_{mod}^2 + t_{GVD}^2 + t_{rx}^2\right]^{1/2}$$

$$t_{sys} = \left| t_{tx}^2 + \left(\frac{440\,Lq}{\text{[illegible]}}\right)^2 + D^2 \sigma_\lambda^2 L^2 + \left(\frac{350}{\text{[illegible]}}\right) \right|^{1/2}$$

WDM

WDM (Wavelength-division Multiplexing) is the technology of combing a number of wavelengths onto the same fiber simultaneously. A powerful aspect of WDM is that each optical channel can carry any transmission format. WDW increases the capacity of a fiber network dramatically. Thus it is recognized as the Layer 1 transport technology in all tiers of the network. The purpose of this article is to give a brief overview of WDM technology and its applications.

NEED OF WDM

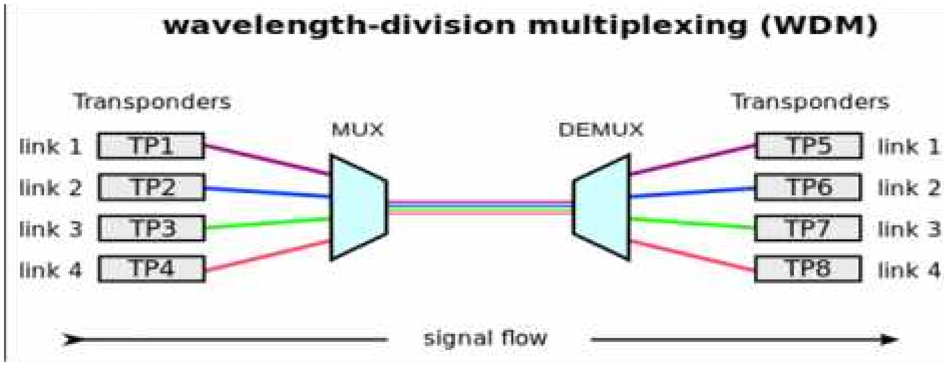
Due to the rapid growth in telecommunication links, high capacity and faster data transmission rates over farther distances are required. To meet these demands, network managers are relying more and more on fiber optics. Typically, there are three methods for expanding capacity: installing more cables, increasing system bit rate to multiplex more signals and wavelength division multiplexing. The first method, installing more cables, will be preferred in many cases, especially in metropolitan areas, since fiber has become incredibly inexpensive and installation methods more efficient. But when conduit space is not available or major construction is necessary, this may not be the most cost-effective.

Another way for capacity expansion is to increase system bit rate to multiplex more signals. But increasing system bit rate may not prove cost effective either. Since many systems are already running at SONET OC-48 rates (2.5 GB/s) and upgrading to OC-192 (10 GB/s) is expensive, requires changing out all the electronics in a network, and adds 4 times the capacity, may not be necessary. Thirdly, the WDM has been proved to be the more cost-effective technology. It does not only support current electronics and fibers but also can share fibers by transmitting channels at different wavelengths (colors) of light. Besides, systems are already using fiber optic amplifiers as repeaters also do not require upgrading for most WDM.

From the above comparison of three methods for expanding capacity, it can easily draw a conclusion that WDM is the best solution to meet the demand for more capacity and faster data transmission rates. Actually, it is not difficult to understand the operating principle of WDM. Consider the fact that you can see many different colors of light: red, green, yellow, blue, etc. The colors are transmitted through the air together and may mix, but they can be easily separated by using a simple device like a prism. It’s like separating the “white” light from the sun into a spectrum of colors with the prism. WDM is equivalent to the prism in the operating principle. A WDM system uses a multiplexer at the transmitter to joint the several signals together. At the same time, it uses a demultiplexer at the receiver to split them apart, as shown in the following diagram. With the right type of fiber, it is possible to function as an optical add-drop multiplexer.

This technique was originally demonstrated with optical fiber in the early 80s. The first WDM systems combined only two signals. Modern systems can handle up to 160 signals and can thus expand a basic 10 Gbit/s system over a single fiber pair to over 1.6 Tbit/s.

Because WDM systems can expand the capacity of the network and accommodate several generations of technology development in optical infrastructure without having to overhaul the backbone network, they are popular with telecommunications companies.



CWDM VS DWDM

WDM systems are divided into different wavelength patterns: CWDM (Coarse Wavelength Division Multiplexing) and DWDM (Dense Wavelength Division Multiplexing). There are many differences between CWDM and DWDM: spacings, DFB lasers, and transmission distances. The channel spacings between individual wavelengths transmitted through the same fiber serve as the basis for defining CWDM and DWDM. Typically, the spacing in CWDM systems is 20 nm, while most DWDM systems today offer 0.8 nm (100 GHz) wavelength separation according to the ITU standard.

Due to wider CWDM channel spacing, the number of channels (lambdas) available on the same link is significantly reduced, but the optical interface components do not have to be as precise as DWDM components. CWDM equipment is thus significantly cheaper than DWDM equipment. Both CWDM and DWDM architectures utilize the DFB (Distributed Feedback Lasers). However, CWDM systems use DFB lasers that are not cooled. These systems typically operate from 0 to 70℃ with the laser wavelength drifting about 6 nm over this range. Coupled with the laser wavelength of up to ±3 nm, the wavelength drift yields a total wavelength variation of about ±12 nm.

DWDM systems, on the other hand, require the larger cooled DFB lasers, because a semiconductor laser wavelength drifts about 0.08 nm/℃ with temperature. DFB lasers are cooled to stabilize the wavelength from outside the passband of the multiplexer and demultiplexer filters as the temperature fluctuates in DWDM systems. Due to the unique attributes of CWDM and DWDM, they are deployed for different transmission distances. Typically, CWDM can travel anywhere up to about 160 km. If this needs to transmit the data over a long range, the DWDM system is the best choice. DWDM supports 1550 nm wavelength size, which can be amplified to extend transmission distance to hundreds of kilometers.

OPERATIONAL PRINCIPLES OF WDM

Since the spectral width of a high-quality source occupies only a narrow slice of optical bandwidth, there are many independent operating regions across the spectrum, ranging from the a-band through the L-band,that can be used simultaneously. The original use of WDM was to upgrade the capacity of installed point-to-point transmission links. This was achieved with wavelengths that were separated from several tens up to 200 nm in order not to impose strict wavelength-tolerance requirements on the different laser sources and the receiving wavelength splitters.

Subsequently, the development of lasers that have extremely narrow spectral emission widths allowed wavelengths to be spaced less than a nanometer apart. This is the basis of wavelength-division multiplexing, which simultaneously uses a number of light sources, each emitting at a slightly different peak wavelength.

Each wavelength carries an independent signal, so that the link capacity is increased greatly. The main trick is to ensure that the peak wavelength of a source is spaced sufficiently far from its neighbor so as not to create interference between their spectral extents. Equally important is the requirement that during the operation of a system these peak wavelengths do not drift into the spectral territory occupied by adjacent channels. In addition to maintaining strict control of the wavelength, system designers include an empty guard band between the channels as an operations safety factor. Thereby the fidelities of the independent messages from each source are maintained for subsequent conversion to electrical signals at the receiving end.

WDM Operating Regions

The possibility of having an extremely high-capacity link by means of WDM can be seen by examining the characteristics of a high-quality optical source. As an example, a distributed- feedback (DFB) laser has a frequency spectrum on the order of I MHz, which is equivalent to a spectral line width of 10-5 nm. With such spectral widths, simplex systems make use of only a tiny portion of the transmission bandwidth capability of a fiber. This can be seen from Figure which depicts the attenuation of light in a silica fiber as a function of wavelength. The curve shows that the two low-loss regions of a standard G.652 single-mode fiber extend over the O- band wavelengths ranging from about 1270 to 1350 nm (originally called the second window) and from 1480 to 1600nm (originally called the third window). This can view these regions either in terms of spectral width (the wavelength band occupied by the light signal) or by means of optical bandwidth (the frequency band occupied by the light signal).

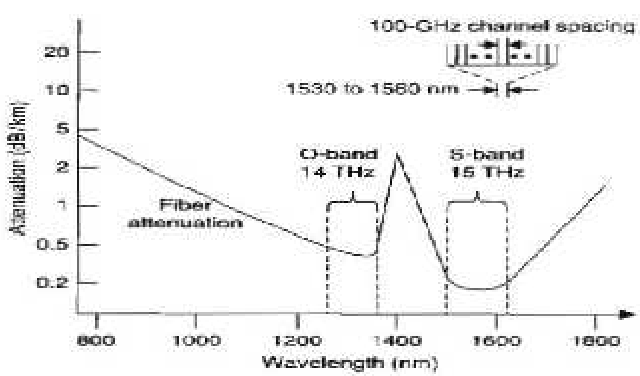


Figure Generic representation of attenuation of light in a silica fiber as a function of wavelength

To find the optical bandwidth corresponding to a particular spectral width in these regions, This uses the fundamental relationship c=Lamda*v, which relates the wavelength Laamda to the carrier frequency v, where c is the speed of light. Differentiating this,

$$\Delta\nu = \frac{c}{\lambda^2}\Delta\lambda$$

Where the frequency deviation corresponds to wavelength deviation around the wavelength

If fiber has the attenuation characteristic shown in Figure. The optical bandwidth is .Delta v= 14THz for a usable spectral band. Delta Lamda= 80 nm in the center of the O-band. Similarly, .Delta v= 15 THz for a usable spectral band Delta Lamda= 120 nm in the low-loss region running from near the beginning of the S-band to almost the end of the L-band. This yields a total available fiber bandwidth of about 30THz in the two low-loss windows.

Prior to about 2000, the peak wavelengths of adjacent light sources typically were restricted to be separated by 0.8 to 1.6 nm (100 to 200 GHz) in a WDM system. This was done to take into account possible drifts of the peak wavelength due to aging or temperature effects, and to give both the manufacturer and the user some leeway in specifying and choosing the precise peak emission wavelength. The next generation of WDM systems specified both narrower and much wider channel spacings depending on the application and on the wavelength region being used. The much narrower spacings thus require strict wavelength control of the optical source. On the other hand, the wider wavelength separations offer inexpensive WDM implementations since wavelength control requirements are relaxed significantly.

Generic WDM Link

The implementation of WDM networks requires a variety of passive and/or active devices to combine, distribute, isolate, add, drop, attenuate, and amplify optical power at different wavelengths. Passive devices require no external electric power or control for their operation, so they have a fixed application in WDM networks. These passive components are used to separate and combine wavelength channels, to divide optical power onto a number of fiber lines, or to tap off part of an optical signal for monitoring purposes. The performance of active devices can be controlled electronically, thereby providing a large degree of network flexibility, Active WDM components include tunable optical filters, tunable light sources, configurable add/drop multiplexers, dynamic gain equalizers, and optical amplifiers.

The transmitting side has a series of independently modulated fixed-wavelength light sources, each of which emits signals at a unique wavelength. Here a multiplexer (popularly called a mux) is needed to combine these optical outputs into a continuous spectrum of signals and couple them onto a single fiber. Within a standard telecommunication link there may be various types of optical amplifiers, a variety of specialized active components (not shown), and passive optical power splitters. The operations and maintenance benefits of PONs are that no active devices are used between the transmitting and receiving endpoints.

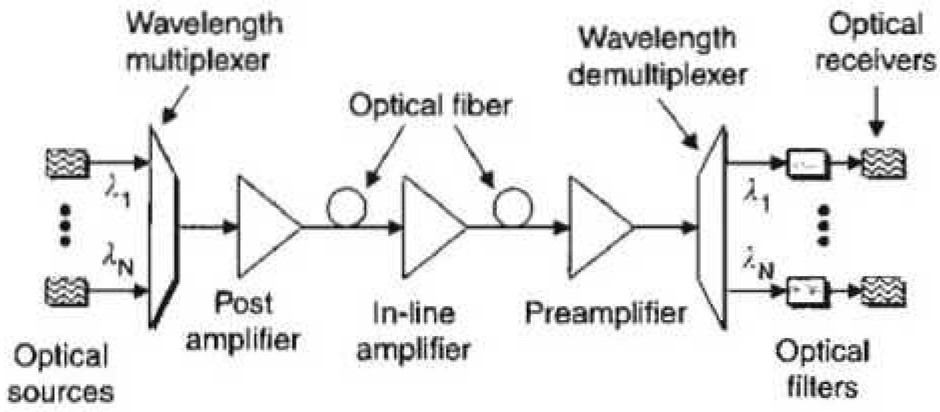


Figure Implementation of a simple WDM link

At the receiving end a demultiplexer is required to separate the individual wavelengths of the independent optical signals into appropriate detection channels for signal processing. At the transmitter the basic design challenge is to have the multiplexer provide a low-loss path from each optical source to the multiplexer output. A different requirement exists for the demultiplexer, since photodetectors usually are sensitive over a broad range of wavelengths, which could include all the WDM channels. To prevent spurious signals from entering a receiving channel, that is, to give good channel isolation of the different wavelengths being used, the demultiplexer must exhibit narrow spectral operation or very stable optical filters with sharp wavelength cutoffs must be used.

The tolerable crosstalk levels between channels can vary widely depending on the application. In general, a -10 dB level is not sufficient, whereas a level of - 30 dB is acceptable. In principle, any optical demultiplexer can also be used as a multiplexer. For simplicity, the word multiplexer is used as a general term to refer to both combining and separating functions, except when it is necessary to distinguish the two devices or functions.

Wavelength Division Multiplexing (WDM)

Optical signals of different wavelength (1300-1600 nm) can propagate without interfering with each other. The scheme of combining a number of wavelengths over a single fiber is called wavelength division multiplexing (WDM). Each input is generated by a separate optical source with a unique wavelength. Optical multiplexer couples light from individual sources to the transmitting fiber. At the receiving station, an optical demultiplexer is required to separate the different carriers before photodetection of individual signals. To prevent spurious signals to enter into receiving channel, the demultiplexer must have narrow spectral operation with sharp wavelength cut-offs. The acceptable limit of crosstalk is – 30 dB.

Features of WDM

- ❖ Capacity upgrade: Since each wavelength supports independent data rate in Gbps.
- ❖ Transparency: WDM can carry fast asynchronous, slow synchronous, synchronous analog and digital data.

- ❖ Wavelength routing: Link capacity and flexibility can be increased by using multiple wavelength.
- ❖ Wavelength switching: WDM can add or drop multiplexers, cross connects and wavelength converters.

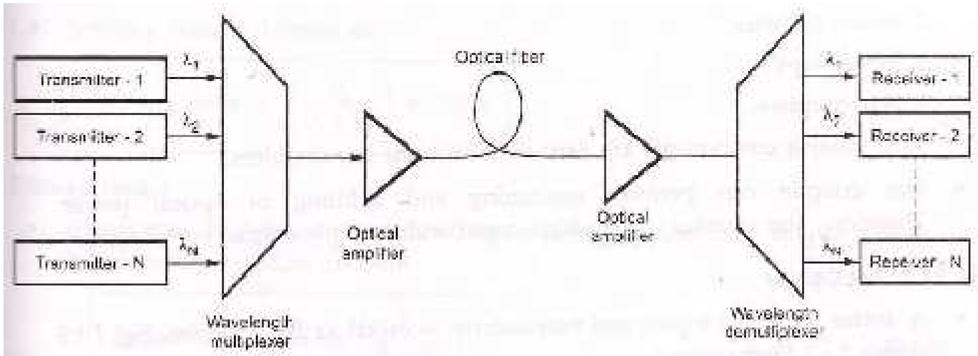


Figure WDM scheme

Passive Components

For implementing WDM various passive and active components are required to combine, distribute, isolate and to amplify optical power at different wavelength. Passive components are mainly used to split or combine optical signals. These components operate in optical domains. Passive components don't need external control for their operation. Passive components are fabricated by using optical fibers by planar optical waveguides. Commonly required passive components are

- ❖ N x N couplers
- ❖ Power splitters
- ❖ Power taps
- ❖ Star couplers.

Most passive components are derived from basic stat couplers. Star coupler can person combining and splitting of optical power. Therefore, star coupler is a multiple input and multiple output port device.

Dense Wavelength Division Multiplexing (DWDM)

- ❖ DWDM (Dense wavelength – division multiplexing) is a data transmission technology having very large capacity and efficiency.
- ❖ Multiple data channels of optical signals are assigned different wavelengths, and are multiplexed onto one fiber.
- ❖ DWDM system consists of transmitters, multiplexers, optical amplifier and demultiplexer.

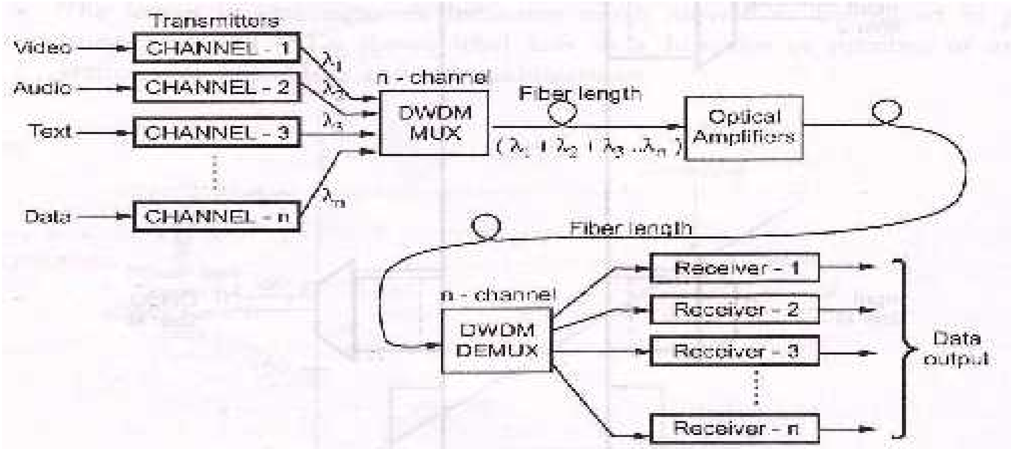


Figure DWDM System

- ❖ DWDM used single mode fiber to carry multiple light waves of different frequencies.
- ❖ DWDM system uses Erbium – Doped Fiber Amplifiers (EDFA) for its long haul applications, and to overcome the effects of dispersion and attenuation channel spacing of 100 GHz is used.

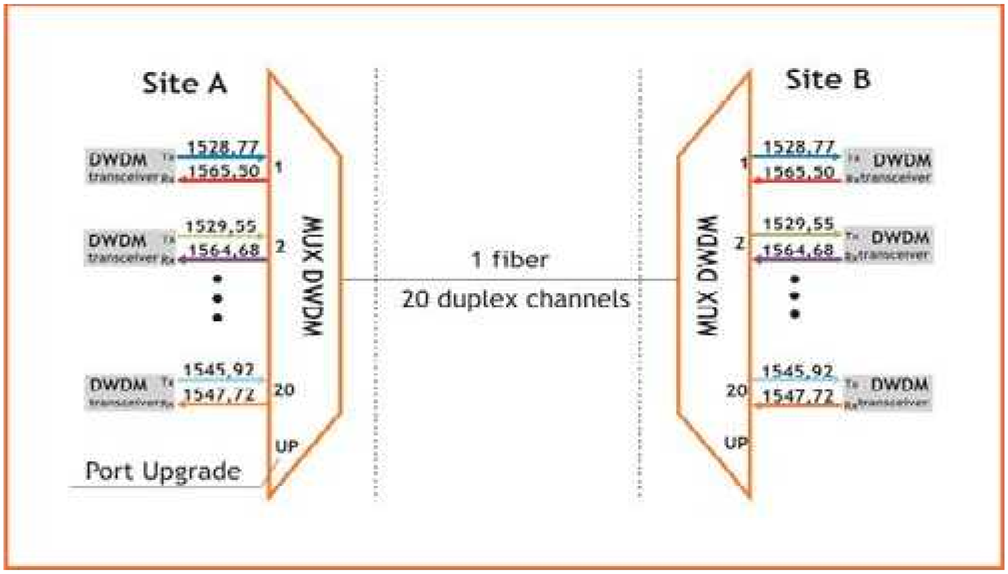
DWDW is short for dense wavelength division multiplexing. It is an optical multiplexing technology used to increase bandwidth over existing fiber networks. DWDM works by combining and transmitting multiple signals simultaneously at different wavelengths on the same fiber. It has revolutionized the transmission of information over long distances. DWDM can be divided into passive DWDM and active DWDM. This article will detail these two DWDM systems.

Passive DWDM

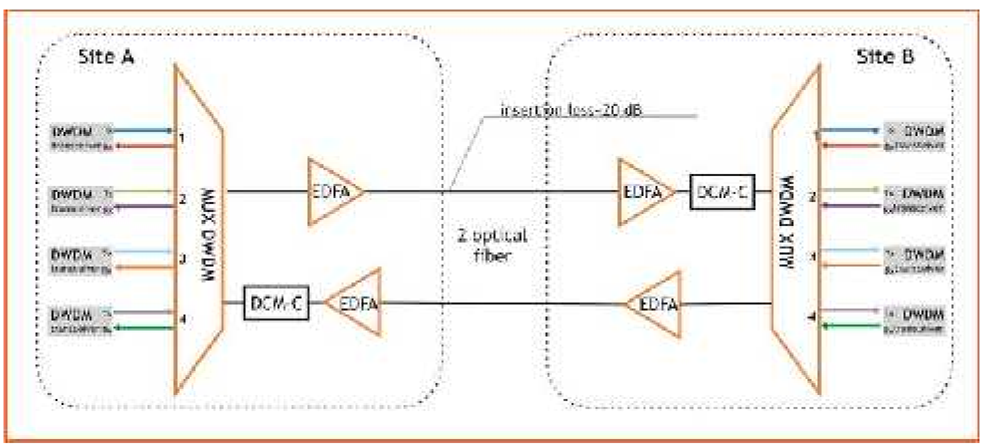
Passive DWDM systems have no active components. The line functions only due to the optical budget of transceivers used. No optical signal amplifiers and dispersion compensators are used. Passive DWDM systems have a high channel capacity and potential for expansion, but the transmission distance is limited to the optical budget of transceivers used. The main application of passive DWDM system is metro networks and high speed communication lines with a high channel capacity.

Active DWDM

Active DWDM systems commonly refer to as a transponder-based system. They offer a way to transport large amounts of data between sites in a data center interconnect setting. The transponder takes the outputs of the SAN or IP switch format, usually in a short wave 850nm or long wave 1310nm format, and converts them through an optical-electrical-optical (OEO) DWDM conversion. When creating long-haul DWDM networks, several EDFA amplifiers are installed sequentially in the line. The number of amplifiers in one section is limited and depends on the optical cable type, channel count, data transmission rate of each channel, and permissible OSNR value.



The possible length of lines when using active DWDM system is determined not only with installed optical amplifiers and the OSNR value, but also with the influence of chromatic dispersion—the distortion of transmitted signal impulses, on transmitted signals. At the design stage of the DWDM network project, permissible values of chromatic dispersion for the transceivers are taken into account, and, if necessary, chromatic dispersion compensation modules (DCM) are included in the line. DCM introduces additional attenuation into the line, which leads to a reduction of the amplified section length. At this stage, a basic DWDM system contains several main components:



WDM multiplexer for DWDM communications

DWDM terminal multiplexer- The terminal multiplexer contains a wavelength- converting transponder for each data signal, an optical multiplexer and where necessary an optical amplifier (EDFA). Each wavelength-converting transponder receives an optical data signal from the client-layer, such as Synchronous optical networking [SONET /SDH] or another type of data signal, converts this signal into the electrical domain and re- transmits the signal at a specific wavelength using a 1,550 nm band laser. These data signals are then combined together into a multi-wavelength optical signal using an optical multiplexer, for transmission over a single fiber (e.g., SMF-28 fiber).

The terminal multiplexer may or may not also include a local transmit EDFA for power amplification of the multi-wavelength optical signal. In the mid-1990s DWDM systems contained 4 or 8 wavelength-converting transponders; by 2000 or so, commercial systems capable of carrying 128 signals were available.

An intermediate line repeater is placed approximately every 80–100 km to compensate for the loss of optical power as the signal travels along the fiber. The 'multi-wavelength optical signal' is amplified by an EDFA, which usually consists of several amplifier stages. An intermediate optical terminal, or optical add-drop multiplexer. This is a remote amplification site that amplifies the multi-wavelength signal that may have traversed up to 140 km or more before reaching the remote site. Optical diagnostics and telemetry are often extracted or inserted at such a site, to allow for localization of any fiber breaks or signal impairments.

In more sophisticated systems (which are no longer point-to-point), several signals out of the multi-wavelength optical signal may be removed and dropped locally. A DWDM terminal demultiplexer at the remote site, the terminal de-multiplexer consisting of an optical de-multiplexer and one or more wavelength-converting transponders separates the multi-wavelength optical signal back into individual data signals and outputs them on separate fibers for client-layer systems. Originally, this de-multiplexing was performed entirely passively, except for some telemetry, as most SONET systems can receive 1,550 nm signals.

However, in order to allow for transmission to remote client-layer systems (and to allow for digital domain signal integrity determination) such de-multiplexed signals are usually sent to O/E/O output transponders prior to being relayed to their client-layer systems. Often, the functionality of output transponder has been integrated into that of input transponder, so that most commercial systems have transponders that support bi-directional interfaces on both their 1,550 nm (i.e., internal) side, and external (i.e., client-facing) side. Transponders in some systems supporting 40 GHz nominal operation may also perform forward error correction (FEC) via digital wrapper technology, as described in the ITU-T G.709 standard.

Optical Supervisory Channel (OSC). This is data channel which uses an additional wavelength usually outside the EDFA amplification band (at 1,510 nm, 1,620 nm, 1,310 nm or another proprietary wavelength). The OSC carries information about the multi-wavelength optical signal as well as remote conditions at the optical terminal or EDFA site. It is also normally used for remote software upgrades and user (i.e., network operator) Network Management information. It is the multi-wavelength analogue to SONET's DCC (or supervisory channel). ITU standards suggest that the OSC should utilize an OC-3 signal structure, though some vendors have opted to use 100 megabit Ethernet or another signal format. Unlike the 1550 nm multi-wavelength signal containing client data, the OSC is always terminated at intermediate amplifier sites, where it receives local information before re-transmission.

Erbium-Doped Fiber Amplifiers

An important class of fiber amplifiers makes use of rare-earth elements as a gain medium by doping the fiber core during the manufacturing process. Although doped-fiber amplifiers were studied as early as 1964, their use became practical only 25 years later, after the fabrication and characterization techniques were perfected. Amplifier properties such as the operating wavelength and the gain bandwidth are determined by the dopants rather than by the silica fiber, which plays the role of a host medium. Many different rare-earth elements, such as erbium, holmium, neodymium, samarium, thulium, and ytterbium, can be used to realize fiber amplifiers operating at different wavelengths in the range 0.5–3.5 μm.

Erbium-doped fiber amplifiers (EDFAs) have attracted the most attention because they operate in the wavelength region near 1.55 μm. Their deployment in WDM systems after 1995 revolutionized the field of fiber-optic communications and led to light wave systems with capacities exceeding 1 Tb/s. This section focuses on the main characteristics of EDFAs.

Pumping Requirements

The design of an EDFA looks similar to that in Figure with the main difference that the fiber core contains erbium ions (Er^{3+}). Pumping at a suitable wavelength provides gain through population inversion. The gain spectrum depends on the pumping scheme as well as on the presence of other dopants, such as germania and alumina, within the fiber core.

The amorphous nature of silica broadens the energy levels of Er^{3+} into bands. Figure (a) shows a few energy levels of Er^{3+} in silica glasses. Many transitions can be used to pump an EDFA. Early experiments used the visible radiation emitted from argon-ion, Nd:YAG, or dye lasers even though such pumping schemes are relatively inefficient. From a practical standpoint the use of semiconductor lasers is preferred.

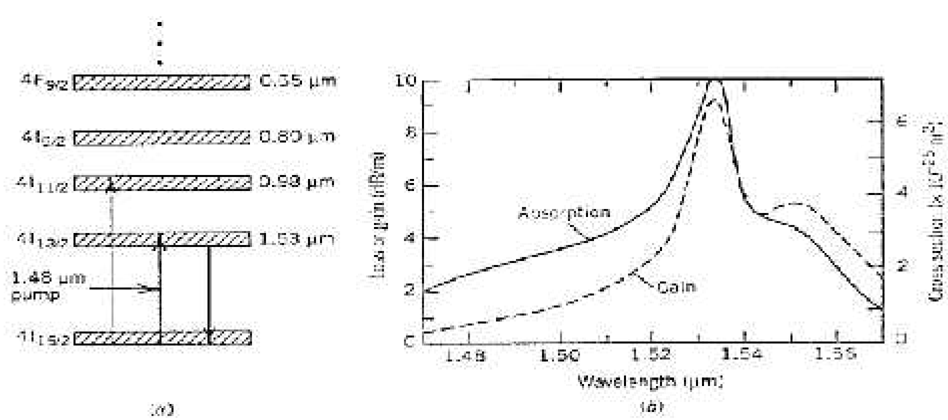


Figure (a) Energy-level diagram of erbium ions in silica fibers; (b) absorption and gain spectra of an EDFA whose core was co doped with germania

Efficient EDFA pumping is possible using semiconductor lasers operating near 0.98 and 1.48μm wavelengths. Indeed, the development of such pump lasers was fueled with the advent of EDFAs. It is possible to realize 30-dB gain with only 10–15 mW of absorbed pump power. Efficiencies as high as 11 dB/mW were achieved by 1990 with 0.98-μmpumping. The pumping transition 4I15/2 → 4I9/2 can use high power GaAs lasers, and the pumping efficiency of about 1 dB/mW has been obtained at 820 nm. The required pump power can be reduced by using silica fibers doped with aluminum and phosphorus or by using fluorophosphate fibers. With the availability of visible semiconductor lasers, EDFAs can also be pumped in the wavelength range 0.6–0.7 μm. In one experiment, 33-dB gain was realized at 27mW of pump power obtained from an AlGaInP laser operating at 670 nm. The pumping efficiency was as high as 3 dB/mW at low pump powers. Most EDFAs use 980-nm pump lasers as such lasers are commercially available and can provide more than 100 mW of pump power. Pumping at 1480 nm requires longer fibers and higher powers because it uses the tail of the absorption band shown in Figure.

EDFAs can be designed to operate in such a way that the pump and signal beams propagate in opposite directions, a configuration referred to as backward pumping to distinguish it from the forward-pumping configuration shown in Figure. The performance is nearly the same in the two pumping configurations when the signal power is small enough for the amplifier to remain unsaturated.

In the saturation regime, the power-conversion efficiency is generally better in the backward-pumping configuration, mainly because of the important role played by the amplified spontaneous emission (ASE). In the bidirectional pumping configuration, the amplifier is pumped in both directions simultaneously by using two semiconductor lasers located at the two fiber ends. This configuration requires two pump lasers but has the advantage that the population inversion, and hence the small-signal gain, is relatively uniform along the entire amplifier length.

The foregoing analysis assumes that both pump and signal waves are in the form of CW beams. In practice, EDFAs are pumped by using CW semiconductor lasers, but the signal is in the form of a pulse train (containing a random sequence of 1 and 0 bits), and the duration of individual pulses is inversely related to the bit rate.

The question is whether all pulses experience the same gain or not. As discussed the gain of each pulse depends on the preceding bit pattern for SOAs because an SOA can respond on time scales of 100 ps or so. Fortunately, the gain remains constant with time in an EDFA for even microsecond-long pulses. The reason is related to a relatively large value of the fluorescence time associated with the excited erbium ions ($T_1 \sim 10$ ms). When the time scale of signal-power variations is much shorter than T_1, erbium ions are unable to follow such fast variations. As single-pulse energies are typically much below the saturation energy (~10 μJ), EDFAs respond to the average power. As a result, gain saturation is governed by the average signal power, and amplifier gain does not vary from pulse to pulse even for a WDM signal.

In some applications such as packet-switched networks, signal power may vary on a time scale comparable to T_1. Amplifier gain in that case is likely to become time dependent, an undesirable feature from the standpoint of system performance. A gain control mechanism that keeps the amplifier gain pinned at a constant value consists of making the EDFA oscillate at a controlled wavelength outside the range of interest (typically below 1.5 μm). Since the gain remains clamped at the threshold value for a laser, the signal is amplified by the same factor despite variations in the signal power. In one implementation of this scheme, an EDFA was forced to oscillate at 1.48 μm by fabricating two fiber Bragg gratings acting as high-reflectivity mirrors at the two ends of the amplifier.

Multichannel Amplification

The bandwidth of EDFAs is large enough that they have proven to be the optical amplifier of choice for WDM applications. The gain provided by them is nearly polarization insensitive. Moreover, the interchannel crosstalk that cripples SOAs because of the carrier-density modulation occurring at the channel spacing does not occur in EDFAs. The reason is related to the relatively large value of T_1 (about 10 ms) compared with the carrier lifetime in SOAs (<1 ns). The sluggish response of EDFAs ensures that the gain cannot be modulated at frequencies much larger than 10 kHz.

A second source of interchannel crosstalk is cross-gain saturation occurring because the gain of a specific channel is saturated not only by its own power (self saturation) but also by the power of neighboring channels. This mechanism of crosstalk is common to all optical amplifiers including EDFAs. It can be avoided by operating the amplifier in the unsaturated regime. Experimental results support this conclusion. In a 1989 experiment negligible power penalty was observed when an EDFA was used to amplify two channels operating at 2 Gb/s and separated by 2 nm as long as the channel powers were low enough to avoid the gain saturation. The main practical limitation of an EDFA stems from the spectral non uniformity of the amplifier gain. Even though the gain spectrum of an EDFA is relatively broad, as seen in Figure, the gain is far from uniform (or flat) over a wide wavelength range. As a result, different channels of a WDM signal are amplified by different amounts.

This problem becomes quite severe in long-haul systems employing a cascaded chain of EDFAs. The reason is that small variations in the amplifier gain for individual channels grow exponentially over a chain of in-line amplifiers if the gain spectrum is the same for all amplifiers. Even a 0.2-dB gain difference grows to 20 dB over a chain of 100 in-line amplifiers, making channel powers vary by a factor of 100, an unacceptable variation range in practice. To amplify all channels by nearly the same amount, the double-peak nature of the EDFA gain spectrum forces one to pack all channels near one of the gain peaks. In a simple approach, input powers of different channels were adjusted to reduce power variations at the receiver to an acceptable level.

This technique may work for a small number of channels but becomes unsuitable for dense WDM systems.

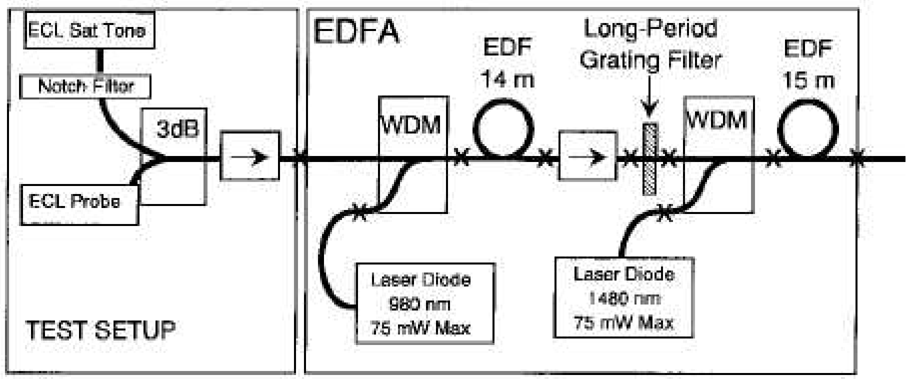


Figure Schematic of an EDFA designed to provide uniform gain over the 1530–1570- nm bandwidth using an optical filter containing several long-period fiber gratings. The two stage design helps to reduce the noise level.

The entire bandwidth of 35–40 nm can be used if the gain spectrum is flattened by introducing wavelength-selective losses through an optical filter. The basic idea behind gain flattening is quite simple. If an optical filter whose transmission losses mimic the gain profile (high in the high-gain region and low in the low-gain region) is inserted after the doped fiber, the output power will become constant for all channels. Although fabrication of such a filter is not simple, several gain-flattening techniques have been developed. For example, thin-film interference filters, Mach–Zehnder filters, acousto-optic filters, and long-period fiber gratings have been used for flattening the gain profile and equalizing channel gains .

The gain-flattening techniques can be divided into active and passive categories. Most filter-based methods are passive in the sense that channel gains cannot be adjusted in a dynamic fashion. The location of the optical filter itself requires some thought because of high losses associated with it. Placing it before the amplifier increases the noise while placing it after the amplifier reduces the output power. Often a two-stage configuration shown in Figure is used. The second stage acts as a power amplifier while the noise figure is mostly determined by the first stage whose noise is relatively low because of its low gain. A combination of several long-period fiber gratings acting as the optical filter in the middle of two stages resulted by 1977 in an EDFA whose gain was flat to within 1 dB over the 40-nm bandwidth in the wavelength range of 1530–1570 nm.

Ideally, an optical amplifier should provide the same gain for all channels under all possible operating conditions. This is not the case in general. For instance, if the number of channels being transmitted changes, the gain of each channel will change since it depends on the total signal power because of gain saturation.

The active control of channel gains is thus desirable for WDM applications. Many techniques have been developed for this purpose. The most commonly used technique stabilizes the gain dynamically by incorporating within the amplifier a laser that operates outside the used bandwidth. Such devices are called gain-clamped EDFAs (as their gain is clamped by a built-in laser) and have been studied extensively.

WDM light wave systems capable of transmitting more than 80 channels appeared by 1998. Such systems use the C and L bands simultaneously and need uniform amplifier gain over a bandwidth exceeding 60 nm. Moreover, the use of the L band requires optical amplifiers capable of providing gain in the wavelength range 1570–1610 nm. It turns out that EDFAs can provide gain over this wavelength range, with a suitable design. An L-band EDFA requires long fiber lengths (>100 m) to keep the inversion level relatively low. Figure shows an L-band amplifier with a two-stage design.

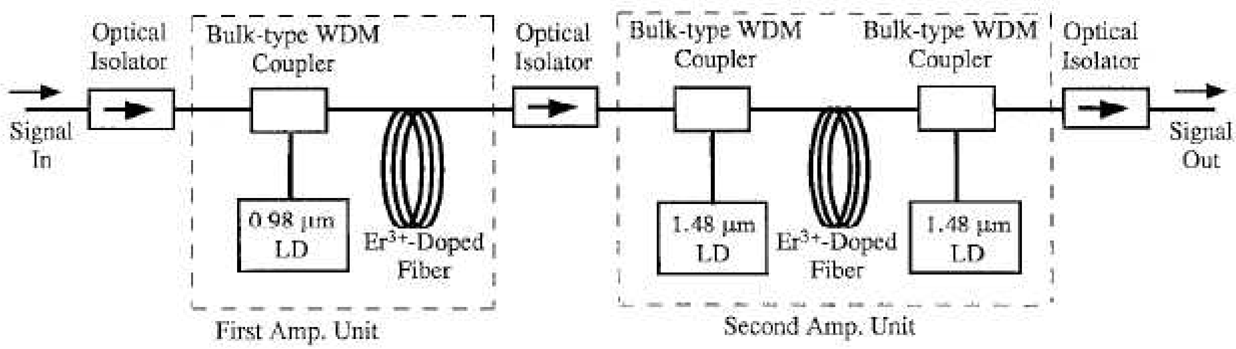


Figure Schematic of an L-band EDFA providing uniform gain over the 1570–1610-nm bandwidth with a two-stage design

The first stage is pumped at 980 nm and acts as a traditional EDFA (fiber length 20–30 m) capable of providing gain in the range 1530–1570 nm. In contrast, the second stage has 200-m-long doped fiber and is pumped bidirectionally using 1480-nm lasers. An optical isolator between the two stages passes the ASE from the first stage to the second stage (necessary for pumping the second stage) but blocks the backward propagating ASE from entering the first stage. Such cascaded, two-stage amplifiers can provide flat gain over a wide bandwidth while maintaining a relatively low noise level. As early as 1996, flat gain to within 0.5 dB was realized over the wavelength range of 1544–1561 nm.

The second EDFA was co doped with ytterbium and phosphorus and was optimized such that it acted as a power amplifier. Since then, EDFAs providing flat gain over the entire C and L bands have been made. Raman amplification can also be used for the L band. Combining Raman amplification with one or two EDFAs, uniform gain can be realized over a 75nm bandwidth covering the C and L bands. A parallel configuration has also been developed for EDFAs capable of amplifying over the C and L bands simultaneously.

In this approach, the incoming WDM signal is split into two branches, which amplify the C-band and L-band signals separately using an optimized EDFA in each branch. The two-arm design has produced a relatively uniform gain of 24 dB over a bandwidth as large as 80 nm when pumped with 980-nm semiconductor lasers while maintaining a noise figure of about 6 dB.

The two-arm or two-stage amplifiers are complex devices and contain multiple components, such as optical filters and isolators, within them for optimizing the amplifier performance. An alternative approach to broadband EDFAs uses a fluoride fiber in place of silica fibers as the host medium in which erbium ions are doped. Gain flatness over a 76-nm bandwidth has been realized by doping a tellurite fiber with erbium ions. Although such EDFAs are simpler in design compared with multistage amplifiers, they suffer from the splicing difficulties because of the use of non silica glasses. Starting in 2001, high-capacity light wave systems began to use the short-wavelength region the so-called S band extending from 1470 to 1520 nm.

Optical CDMA

In OCDMA, each user has a unique code as an assignment address that spreads over a relatively wide bandwidth. This specific code is modulated and then a message signal is transmitted at an arbitrary time to an intended receiver, which can match the correct code to recover the encoded information. The principle of OCDMA multiplexing leads to support of a larger channel count than other techniques, allows asynchronous transmission with efficient access and enhances information security potentially in the network.

Furthermore, it has employment of simplified network control and management, multi-class traffic with different formats and bit rates and can be easily upgraded in terms of its architecture. Each user has been assigned to some chips of the code sequences to share the same transmission line using power splitters or combiners. This operation can be performed in the optical-domain and/or in the space-domain as well.

Decoders at the receiver recognize a target code by employing match filtering. Six types of coding

Direct-sequence or temporal coding optical CDMA systems

• Spectral Amplitude Coding (SAC) Optical CDMA systems

• Spectral Phase Coding (SPC) optical CDMA systems

• Temporal phase coding optical CDMA systems;

• Two-Dimensional (2-D) spatial or spread space coding optical CDMA systems

• Hybrid coding optical CDMA systems

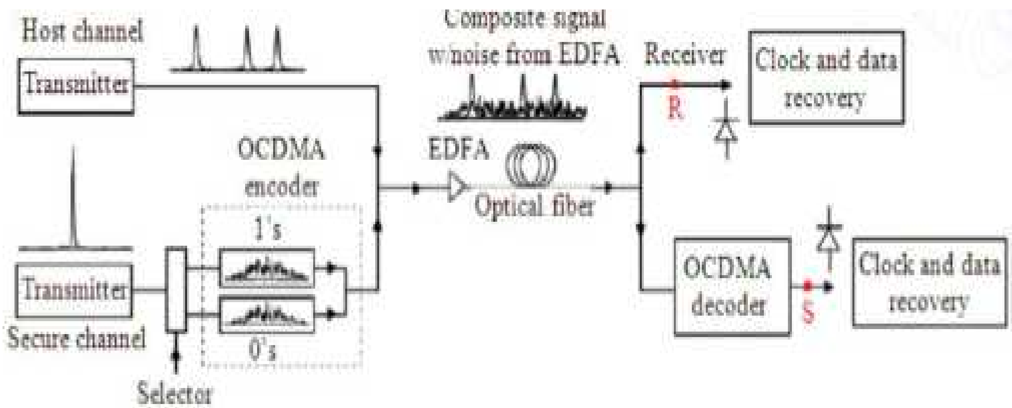


Figure Hybrid System

Two signals are used as shown in figure, a secure signal is encoded and temporally spread to be hidden under a host channel. The purpose of the host channel in this scheme is to provide an ad hoc security enhancement for an encoded signal. The OCDMA en/decoder consists of a coherent spectral phase with direct detection.

Solitons:

Solitons are narrow pulses with high peak powers and special shapes. The most commonly used soliton pulses are called fundamental solitons. The shape of these pulses is shown in Figure. The soliton pulses take advantage of nonlinear effects in silica, specifically self-phase modulation, to overcome the pulse-broadening effects of group velocity dispersion. These pulses can propagate for long distances with no change in shape.

The pulse shapes for which this balance between pulse compression and broadening occurs so that the pulse either undergoes no change in shape or undergoes periodic changes in shape only are called solitons. The family of pulses that undergo no change in shape are called fundamental solitons, and those that undergo periodic changes in shape are called higher-order solitons.

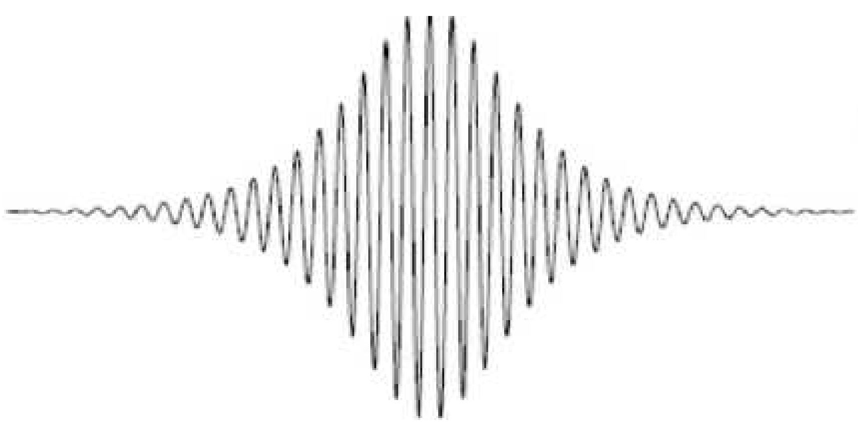


Figure soliton pulse -envelope

The significance of solitons for optical communication is that they overcome the detrimental effects of chromatic dispersion completely Solitons and optical amplifiers, when used together, offer the promise of very high-bit- rate, repeaterless data transmission over very large distances. By the combined use of solitons and erbium-doped fiber amplifiers repeaterless data transmission at a bit rate of 80 Gb/s over a distance of 10,000 km.

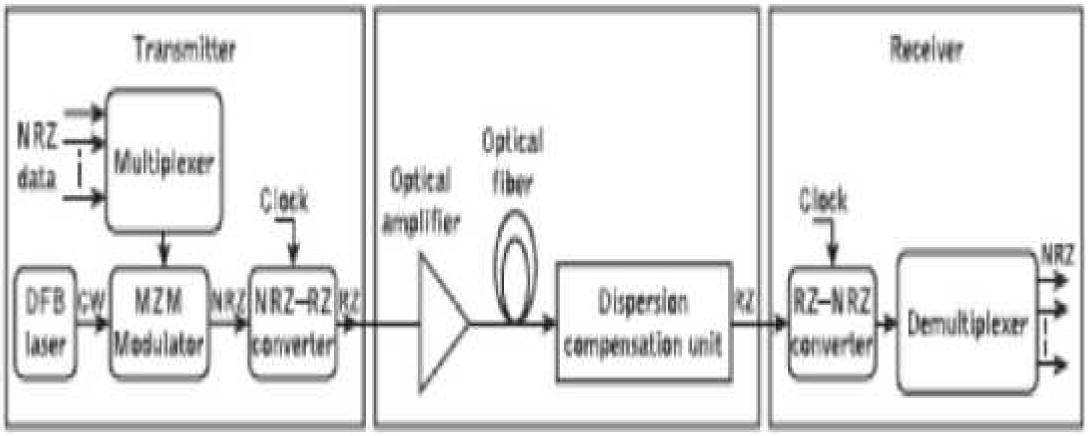


Figure Block schematic of optical fiber soliton transmission system

The use of soliton pulses is key to realizing the very high bit rates required in OTDM systems. The main advantage of soliton systems is their relative immunity to fiber dispersion, which in turn allows transmission at high speeds of a few tens of gigabits per second. The major element in the transmitter section is a return-to-zero pulse generator. A simple approach to generate RZ pulses is to employ an optical modulator and an NRZ-to-RZ converter which is driven by a DFB laser source.

Instead of using a single NRZ data stream, however, it is useful to modulate an optical NRZ signal incorporating several multiplexed NRZ data streams before the conversion into RZ pulses takes place. At the receiving end the incoming signal requires conversion back from RZ to NRZ and then finally a demultiplexer separates the specific NRZ data for each channel. The transmission bit rate of a soliton communication system is dependent on mainly two factors: namely, the soliton pulse width τ and the duration of the bit period to

$$B_T = 1/T_o = 1/2q_o\tau$$

$$q_o = T_o / 2\tau$$

The ratio of T_o / τ determines the nature of the nonlinear propagation for soliton pulses.

Ultra High Capacity Networks

In long haul transmission links the capacity can be improved by ultrafast optical TDM scheme. Two forms of optical TDM schemes are used. Bit interleaved optical TDM, packet interleaved TDM. Optical signals representing data streams from multiple sources are interleaved in time to produce a single data stream. The interleaving can be done on a bit-by-bit basis as shown in Figure

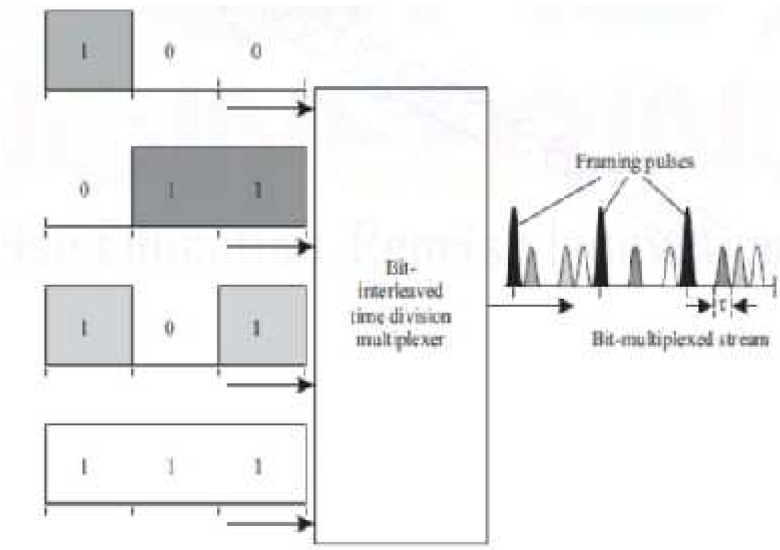


Figure Bit interleaved optical TDM

In the bit-interleaved case, if n input data streams are to be multiplexed, a framing pulse is used every n bits. The periodic pulse train generated by a mode-locked laser is split, and one copy is created for each data stream to be multiplexed. The pulse train for the ith data stream, i = 1, 2, . . .,n, is delayed by τ . This delay can be achieved by passing the pulse train through the appropriate length of optical fiber.

Thus the delayed pulse streams are non overlapping in time. The undelayed pulse stream is used for the framing pulses. Each data stream is used to externally modulate the appropriately delayed periodic pulse stream. The outputs of the external modulator and the framing pulse stream are combined to obtain the bit-interleaved optical TDM stream.

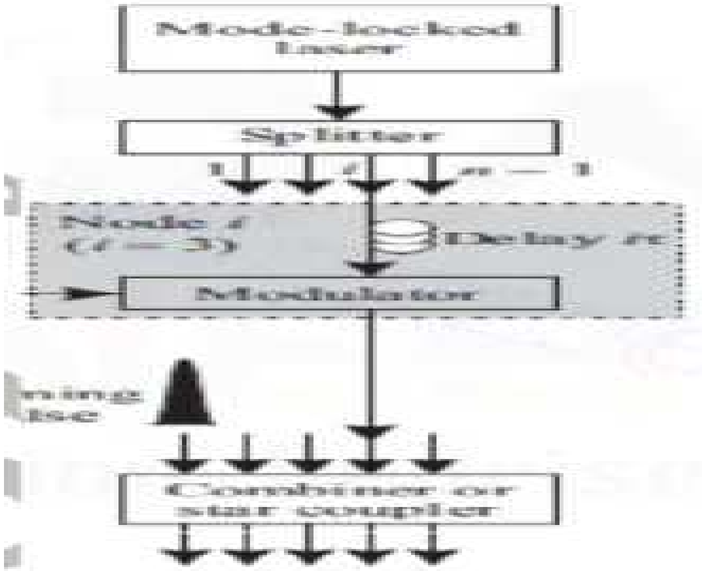


Figure optical multiplexer to create nit interleaved TDM Stream Since the velocity of light in silica fiber is about 2 × 108 m/s, 1 meter of fiber provides a delay of about 5 ns.

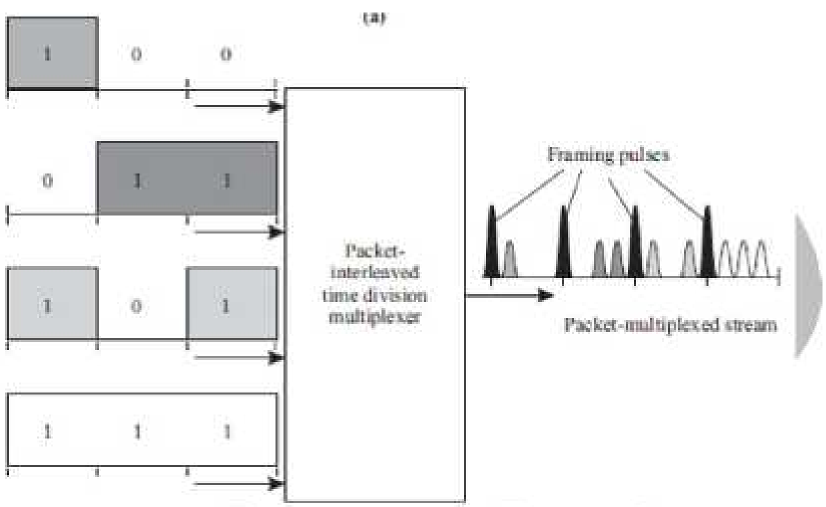
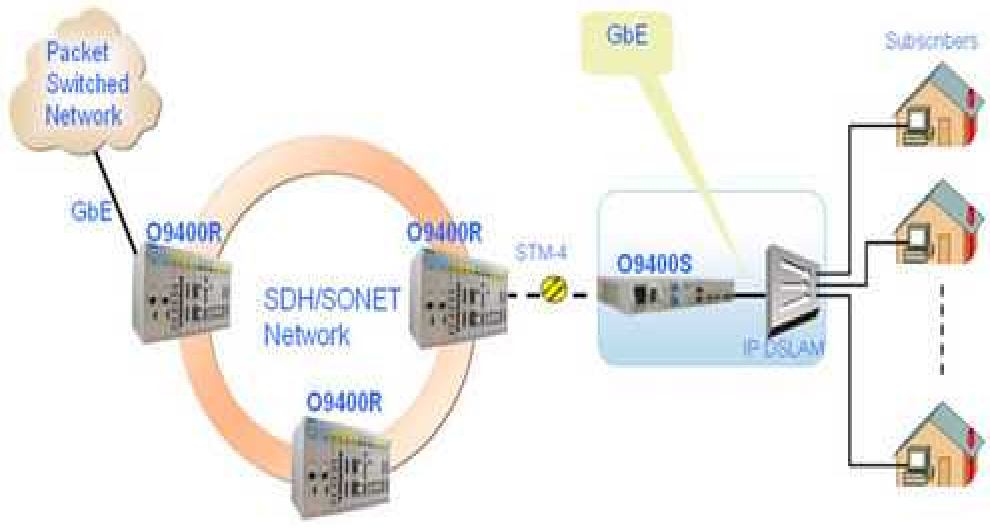


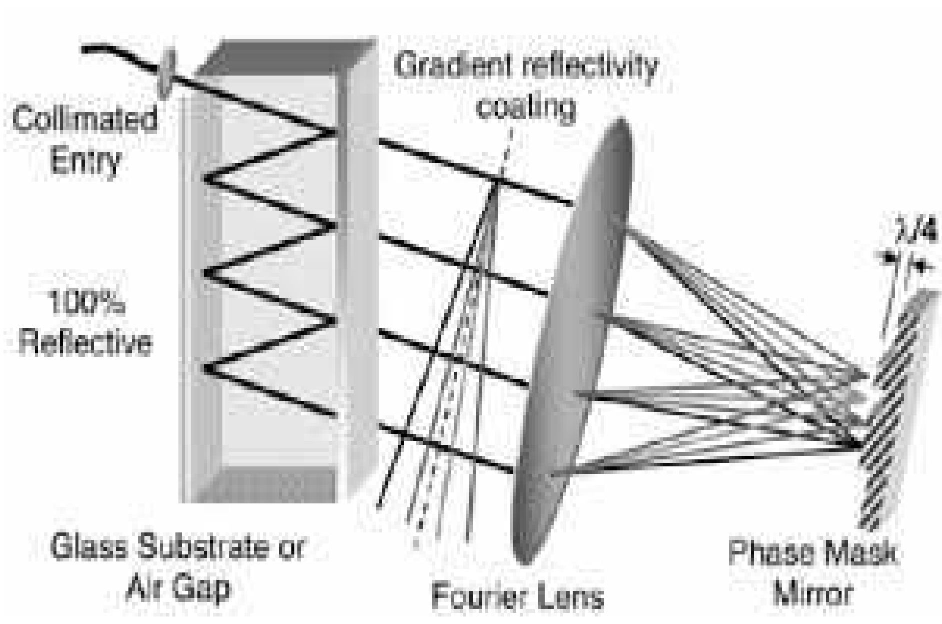
Figure packet interleaved optical TDM

In both the bit-interleaved and the packet-interleaved case, framing pulses can be used In the packet-interleaved case, framing pulses mark the boundary between packets. The j th compression stage is shown in Figure Each compression stage consists of a pair of 3 dB couplers, two semiconductor optical amplifiers (SOAs) used as on-off switches, and a delay line. the output pulses are separated by a time interval of τ .

APPLICATIONS



Examples of Remote Device Management & Optical SDH Application Distributor



Examples of Emerging Optical CDMA application

POST TEST-MCQ TYPE

1. Each stage of information transfer is required to follow the fundamentals of
a) Optical interconnection
b) Optical hibernation
c) Optical networking
d) Optical regeneration

2. What is a multi-functional element of optical network?
a) Hop
b) Optical node
c) Wavelength
d) Optical attenuation

3. A signal carried on a dedicated wavelength from source to destination node is known as a
a) Light path
b) Light wave
c) Light node
d) Light source

4. The fundamentals of optical networking are divided into _______ areas.
a) Two
b) One
c) Four
d) Three

5. The optical networking fundamentals are _____________ of the transmission techniques.
a) Dependent
b) Independent
c) Similar
d) Dissimilar

6. What are the array of switches which forms circuit switching fabrics?
a) Packet arrays
b) Optical cross connects
c) Circuit arrays
d) Optical networks

7. Which of the following is an example of a static circuit-switched network?
a) OXC
b) Circuit regenerator
c) Packet resolver
d) SDH/SONET

8. What is the main disadvantage of OCS?
a) Regenerating mechanism
b) Optical session
c) Time permit
d) Disability to handle burst traffic

9. Optical electro-conversions takes place in __________________ networks.
a) Sessional
b) Optical packet-switched
c) Optical circuit-switched
d) Circular

10. How many functions are performed by an optical packet switch?
a) Four
b) Three
c) Two
d) One

11. Which provides data storage for packets to resolve contention problems?
a) Switching
b) Routing
c) Buffering
d) Reversing

12. What is usually required by a packet to ensure that the data is not overwritten?
a) Header
b) Footer
c) Guard band
d) Payload

13. Which of the following provides efficient designation, routing, forwarding, switching of traffic through an optical packet-switched network?
a) Label correlation
b) Multiprotocol label switching
c) Optical correlation
d) Routing

14. Electrical devices in optical network are basically used for ______________
a) Signal degradation
b) Node transfer
c) Signal control
d) Amplification

15. A ____________ digital hierarchy was required to enable the international communications network to evolve in the optical fiber era.
a) Asynchronous
b) Dedicated
c) Seismic
d) Synchronous

16. Which is a packetized multiplexing and switching technique which combines the benefits of circuit and packet switching?
a) Synchronous mode
b) Asynchronous transfer mode
c) Circuit packet
d) Homogeneous mode

17. An advanced type of reconfigurable OTN is referred to as an _______________
a) Automatic OTN
b) Auto-generated photon
c) Automatically switched optical network
d) Optical reimbursement

18. The mapping of IP frames in SDH/SONET is accomplished in ___________ stages.
a) Four
b) Two
c) Three
d) One

19. Which supports a great number of wavelength channels and reduces the number of switches within the optical network?
a) Waveband switching
b) Optical remuneration
c) Optical genesis
d) Wavelength multiplexing

20. The routing and wavelength assignment problem addresses the core issue of
a) Traffic patterns in a network
b) Wavelength adjustment
c) Wavelength continuity constraint
d) Design problem

21. How many techniques of implementation are there for routing wavelength assignment (RWA)?
a) Two
b) Six
c) Three
d) Four

22. The ____________ provides information about the physical path and wavelength assignment for all active light paths.
a) Network state
b) RWA
c) LAN topology
d) Secluded communication protocol

23. Which is a network that connects several regional or national networks together?
a) Long-haul network
b) Domain network
c) Short-haul network
d) Erbium network

24. What is the range of transmission of extended long haul network?
a) 200-400 km
b) 600-1000 km
c) 1000-2000 km
d) 2000-4000 km

25. What is the range of transmission of ultra-long haul network?
a) 200-400 km
b) 600-1000 km
c) 1000-2000 km
d) 2000-4000 km

26. Which feature plays an important role in making the longer haul networks feasible?
a) Channeling
b) Forward error control
c) Backward error control
d) Interconnection

27. Which of the following is not an element of a submerged cable system?
a) Repeater
b) Branching unit
c) Gain equalizer
d) Attenuator

28. Which provides interconnection between the United States and European countries?
a) TAT
b) WTE
c) PFE
d) POP

29. A single fiber in TAT-14 can carry __________ wavelength channels.
a) One
b) Twelve
c) Sixteen
d) Ten

30. Optical MAN'S are usually structured in ________ topologies.
a) Ring
b) Bus
c) Mesh
d) Star

31. What is the exception in the similarities between the optical Ethernet and the Ethernet LAN?
a) Physical layer
b) Data-link layer
c) Refractive index
d) Attenuation mechanism

32. Which technology is used by optical Ethernet?
a) GP-technology
b) HJ-technology
c) IP-technology
d) GB-technology

33. Optical Ethernet can operate at the transmission rates as low as
a) 10 M bits per second
b) 40 M bits per second
c) 100 M bits per second
d) 1000 M bits per second

34. How many types of optical Ethernet connections are developed?
a) Two
b) One
c) Four
d) Three

35. Which type of connection can be used as an Ethernet switch?
a) Point-to-point
b) Multipoint-to-multipoint
c) Multipoint-to-point
d) Point-to-multipoint

36. The ________________ provides point-to-point access to a bidirectional single-mode optical fiber.
a) Optical regenerator
b) Optical session
c) Optical distribution node
d) Optical buffer

37. The ____________ protocol is not used when the Ethernet connections are configured for a full duplex operation.
a) TCP/IP
b) MAC
c) CSMA/CD
d) DTH

38. Optical Ethernet provides switching capabilities in layers
a) 1 and 2
b) 2 and 3
c) 3 and 4
d) 1 and 4

39. The more advantages optical amplifier is
a) Fiber amplifier
b) Semiconductor amplifier
c) Repeaters
d) Mode hooping amplifier

40. Which of the following cannot be used for wideband amplification?
a) Semiconductor optical amplifier
b) Erbium-doped fiber amplifier
c) Raman fiber amplifier
d) Brillouin fiber amplifier

41. Which of the following is used preferably for channel selection in a WDM system?
a) Semiconductor optical amplifier
b) Erbium-doped fiber amplifier
c) Raman fiber amplifier
d) Brillouin fiber amplifier

42. For used in single-mode fiber ___________ are used preferably.
a) Semiconductor optical amplifier
b) Erbium-doped fiber amplifier
c) Raman fiber amplifier
d) Brillouin fiber amplifier

43. _______________ is superior as compared to _________________

a) TWA, FPA

b) FPA, TWA

c) EDFA, FPA

d) FPA, EDFA

44. Signal amplification is obtained in

a) Erbium-doped fluoro-zir-carbonate fiber multimode

b) Rare-earth-doped fiber amplifiers

c) Raman fiber systems

d) Brillouin fiber amplifier

45. Which is constructed using erbium-doped glass?

a) Erbium-based micro fiber amplifier

b) Rare-earth-doped fiber amplifiers

c) Raman fiber systems

d) Brillouin fiber amplifier

46. In _____________ Rayleigh scattering can be reduced.

a) An erbium-based micro fiber amplifier

b) Rare-earth-doped fiber amplifiers

c) Raman fiber systems

d) Distributed Raman amplification

47. Which is defined as a process by which the wavelength of the transmitted signal is changed without altering the data carried by the signal?

a) Wavelength conversion

b) Attenuation

c) Sigma management

d) Wavelength dispersion

48. The device which is used to perform wavelength conversion is called as

a) Attenuator

b) Wavelength Gyrator

c) Wavelength Circulator

d) Wavelength translator

49. The ___________ converters cannot process different modulation formats.

a) Shifting

b) Optoelectronic wavelength

c) Opt-circular

d) Magnetic simulating

50. Which of the following is NOT an application of optical amplifier?
a) Power amplifier
b) In-line repeater amplifier
c) Demodulator
d) Preamplifier

51. What is the typical range of the noise figure?
a) 1 – 2 dB
b) 3 – 5 dB
c) 7 – 11 dB
d) 12 – 14 dB

52. A major attribute of coherent optical transmission was its ability to provide ________________ for future multicarrier systems and networks.
a) Attenuation
b) Dispersion
c) Frequency selectivity
d) Noisy carriers

53. A multicarrier modulation format in which there has been growing interest to compensate for impairments in optical fiber transmission systems is
a) OFDM
b) EDM
c) WDM
d) ADM

54. WDM stands for?
a) Wave division multiplexing
b) Wavelength division multiplexing
c) Wavelength dependent multiplexing
d) Wave dependent multiplexing

55. A technique that can be a solution to the problem of synchronizing data sources.
a) framing
b) data link control
c) full link control
d) pulse stuffing

56. Wavelength Division Multiplexing (WDM) is an analog multiplexing technique to combine
a) magnetic signals
b) electromagnetic signals
c) digital signals
d) optical signals

57. EDFAs generally operate in the wavelength region near
a) 1550 nm and can offer capacities exceeding 1000 Gbps.
b) 850 nm and can offer capacities around 100 Gbps.
c) 1150 nm and can offer capacities around 1000 Gbps.
d) 1550 nm and can offer capacities around 100 Gbps.

58. Which wavelength is the most appropriate one for pumping an EDFA?
a) 850 nm
b) 980 nm
c) 1300 nm
d) 1550 nm

59. The main difference between an SOA and an EDFA is that
a) An SOA operates in the electrical domain, whereas the EDFA operates in the optical domain.
b) An SOA is pumped electrically, whereas the EDFA is pumped optically.
c) An SOA is pumped optically, whereas the EDFA is pumped electrically.
d) An SOA amplifies 1300 nm wavelength, whereas the EDFA amplifies 1550 nm.

60. Which band specifies the operation of EDFA?
a) O band
b) X band
c) C band
d) S band

61. Basically, Solitons are pulses which propagate through the fiber without showing any variation in
a) Amplitude
b) Shape
c) Velocity
d) All of the above

62. DWDM stands for
a) Digital Wavelength-Division Modulation
b) Dense Wavelength-Division Modulation
c) Double Wavelength-Division Modulation
d) Dense Wavelength-Division Multiplexing

63. In SONET, OC-1 stands for
a) Optical Carrier level one
b) Optical Coupler unidirectional
c) Optical Channel one
d) Optical Cable type 1

CONCLUSION

In this unit, Basic Networks of SONET / SDH, Broadcast and select WDM Networks, Wavelength Routed Networks, Power budget, Noise Effects on System Performance, EDFA system, Solitons, Optical CDMA, Ultra High Capacity Networks were discussed in detailed.

REFERENCES

1. Gerd Keiser, "Optical Fiber Communication" McGraw-Hill International,4th Edition. 2010.
2. John M. Senior, "Optical Fiber Communication", Second Edition, Pearson Education, 2007.
3. Ramaswami, Sivarajan and Sasaki "Optical Networks", Morgan Kaufmann, 2009.
4. J.Senior, Optical Communication, Principles and Practice,Prentice Hall of India, 3rd Edition, 2008.
5. J.Gower, "Optical Communication System", Prentice Hall of India, 2001.

ASSIGNMENT

1. Discuss the concepts of Media Access Control protocols in Broadcast and select networks.
2. Explain the basics of optical CDMA systems.
3. Explain the features of Ultra High Capacity networks
4. Explain the principle of WDM networks.
5. Discuss the nonlinear effects on optical network performance.
6. With suitable example explain the conditions and constraints in the formulation and solution of routing and wavelength assignment in an optimal way
7. Write a note on Solitons
8. Explain in detail different types of broadcast and select WDM networks
9. Explain SONET layers and frame structure with diagram. What is a four fiber BLSR ring in a SONET?

UNIT-1

BASICS OF RADAR

- **Introduction**
- **Maximum Unambiguous Range**
- **Simple form of Radar Equation**
- **Radar Block Diagram and Operation**
- **Radar Frequencies and Applications**
- **Prediction of Range Performance**
- **Minimum Detectable Signal**
- **Receiver Noise**
- **Modified Radar Range Equation**

RADAR EQUATION

- **SNR**
- **Envelop Detector**
- **False Alarm time and Probability**
- **Integration of Radar Pulses**
- **Radar Cross Section of Targets (simple targets: sphere and cone sphere)**
- **Transmitter Power**
- **PRF and Range Ambiguities**
- **System Losses (qualitative treatment)**
    - **Important Formulae**
    - **Illustrative Problems**
    - **Questions from Previous Year Examinations**

BASICS OF RADAR

Introduction:

Basic principles and features:

- Radar is a contraction of the words **Ra**dio **D**etection **A**nd **R**anging. Radar is an electromagnetic system for the detection and location of objects. It operates by transmitting a particular type of waveform, a pulse-modulated sine wave for example, and detects the nature of the echo signal.
- Radar can see through conditions such as darkness, haze, fog, rain, and snow which is not possible for human vision. In addition, radar has the advantage that it can measure the distance or range to the object.
- An elementary form of radar consists of a transmitting antenna emitting electromagnetic Radiation generated by an oscillator of some sort, a receiving antenna, and a signal receiver. A portion of the transmitted signal is intercepted by a reflecting object (target) and is reradiated in all directions. The receiving antenna collects the returned signal and delivers it to a receiver, where it is processed to detect the presence of the target and to extract its location and relative velocity. The distance to the target is determined by measuring the time taken for the Radar signal to travel to the target and back. The direction, or angular position, of the target is determined from the direction of arrival of the reflected wave front. The usual method of measuring the direction of arrival is with narrow antenna beams.
- If relative motion exists between target and radar, the shift in the carrier frequency of the reflected wave (Doppler Effect) is a measure of the target's relative (radial) velocity and may be used to distinguish moving targets from stationary objects. In radars which continuously track the movement of a target, a continuous indication of the rate of change of target position is also available.
- It was first developed as a detection device to warn the approach of hostile aircraft and for directing antiaircraft weapons. A well designed modern radar can extract more information from the target signal than merely range.

Measurement of Range:

- The most common radar waveform is a train of narrow, rectangular-shape pulses modulating a sine wave carrier.
- The distance, or range, to the target is determined by measuring the time T_R taken by the pulse to travel to the target and return.
- Since electromagnetic energy propagates at the speed of light **c (3 x 10^8 m/s)** the range **R** is given by : $R = cT_R/2$
- The factor 2 appears in the denominator because of the two-way propagation of radar. With the range **R** in kilometers or nautical miles, and T_R in microseconds, the above relation becomes: $R(km) = 0.15 \times T_R\ (\mu S)$

(1 mile = 0.8689 nautical mile or 1.6 km
1 nautical mile = 1.15078 miles or 1.8412 km)

Maximum unambiguous range:

Once the transmitter pulse is emitted by the radar, sufficient time must elapse to allow any echo signals to return and be detected before the next pulse is transmitted. Therefore, the rate at which the pulses may be transmitted is determined by the longest range at which targets are expected. If the pulse repetition frequency is too high, echo signals from some targets might arrive after the transmission of the next pulse, and ambiguities in measuring range might result. Echoes that arrive after the transmission of the next pulse are called second-time-around (or multiple-time-around) echoes. Such an echo would appear to be at a much shorter range than the actual. The range beyond which targets appear as second-time-around echoes (or the farthest target range that can be detected by a Radar without ambiguity) is called the ***maximum unambiguous range*** and is given by: $R_{unambig.} = C/2f_p$
Where f_p = pulse repetition frequency, in Hz. **(PRF)**

This can also be explained with the following simple relations.

- T_R is the time elapsed between transmission pulse and Echo pulse.
- $T_R = 2R/C$ where **R**= Range of target
- T_R increases with Range **R** and in extreme case Echo pulse merges with next Transmitted Pulse. Then T_R becomes equal to T_P Where T_P= Pulse repetition period
- $T_{R\,max} = T_P = 2R_{max}/C$ **and so** $R_{max} = CT_P/2 = C/2f_P = R_{unambig}$
- Therefore $R_{unambig}$ is directly proportional to the Pulse period T_P (or Inversely proportional to the PRF f_p)

Simple form of Radar Equation:

The radar equation

- Relates the range of a Radar to the characteristics of the transmitter, receiver, antenna, target, and environment.
- Useful as a means for determining the maximum measurable distance from the radar to the target
- It serves both as a tool for understanding radar operation and as a basis for radar design.

Derivation of the simple form of radar equation:

- If the power of the radar transmitter is denoted by P_t and if an isotropic antenna is used (one which radiates uniformly in all directions) the ***power*** density (watts per unit area) at a distance ***R*** from the radar is equal to the transmitter power divided by the surface area $4\pi R^2$ of an imaginary sphere of radius **R** with radar at its centre, or

Power density from anisotropic antenna = $P_t/4\pi R^2$

- Radars employ directive antennas to direct the radiated power P_t into some particular direction. The **gain G** of an antenna is a measure of the increased power radiated in the direction of the target as compared with the power that would have been radiated from an isotropic antenna. *It may be defined as the ratio of the maximum radiation intensity from the given antenna to the radiation intensity from a lossless, isotropic antenna with the same power input.* (The radiation intensity is the power radiated per unit solid angle in a given direction.)Then the power density at the target from an antenna with a transmitting gain **G** is given by

$$\text{Power density from directive antenna} = P_t . G/4\pi R^2$$

- The target intercepts a portion of the incident power and reradiates it in various directions. The measure of the amount of incident power intercepted by the target and reradiated back in the direction of the radar is denoted as the radar cross section **σ**, and is defined by the relation

$$\text{Power density of echo signal at radar} = (P_t . G/4\pi R^2)(\sigma)/4\pi R^2)$$

- The radar cross section **σ** has units of area. It is a characteristic of the particular target and is a measure of its size as seen by the radar. The radar antenna captures a portion of the echo power. If the effective area of the receiving antenna is denoted as A_e, then the power P_r.received by the radar is given by

$$P_r = (P_t . G/4\pi R^2).(\sigma/4\pi R^2).A_e$$

$$= (P_t . G. A_e. \sigma)/(4\pi)^2.R^4$$

- The maximum radar range R_{max} is the distance beyond which the target cannot be detected. It occurs when the received echo signal power P_r just equals the minimum detectable signal S_{min}.

Therefore

$$R_{max} = [(P_t . G. A_e. \sigma)/(4\pi)^2.S_{min}]^{1/4} \quad .....(1)$$

This is the fundamental form of the radar equation. Note that the important antenna parameters are the ***transmitting gain*** and the ***receiving effective area***.
Antenna theory gives the relationship between the transmitting gain and the receiving effective area of an antenna as:

$$G = 4\pi A_e/\lambda^2$$

Since radars generally use the same antenna for both transmission and reception, the above relation between gain **G** and affective aperture area $\mathbf{A_e}$ can be substituted into the above equation, first for A_e and then for G, to give two other forms of the radar equation.

$$R_{max} = [(P_t . G^2 . \lambda^2 . \sigma)/ (4\pi)^3 . S_{min}]^{1/4} \quad ......(2)$$

$$R_{max} = [(P_t . A_e^2 . \sigma)/ 4\pi . \lambda^2 . S_{min}]^{1/4} ....(3)$$

These three forms (eqs.1, 2, and 3) illustrate the need to be careful in the interpretation of the radar equation. For example, from Eq. (2) it might be thought that the range of radar varies as $\lambda^{1/2}$, but Eq. (3) indicates a $\lambda^{-1/2}$relationship, and Eq. (1) shows the range to be independent of λ. The correct relationship depends on whether it is assumed the gain is constant or the effective area is constant with wavelength.

Limitations of the simple form of Radar equation:

- Does not adequately describe the performance of practical radar.
- Many important factors that affect range are not explicitly included.
- In practice, the observed maximum radar ranges are usually much smaller than what would be predicted by the above equations, sometimes by as much as a factor oftwo.

There are many reasons for the failure of the simple radar equation to correlate with actual performance and these will be explained subsequently in the ***modified Radar range equation***.

Radar block diagram and operation:

The operation of a typical pulse radar is described with the help of a simple block diagram shown in the figure below. The transmitter is an oscillator, such as a magnetron, that is "pulsed" (turned on and off) by the modulator to generate a repetitive train of pulses. The magnetron has been the most widely used of the various microwave generators for radar. A typical radar for the detection of aircraft at ranges of 100 or 200 nmi employs a peak power of the order of one megawatt, an average power of several kilowatts, a pulse width of several microseconds, and a pulse repetition frequency of several hundred pulses per second. The waveform generated by the transmitter travels via a transmission line to the antenna, where it is radiated into space.

A single antenna is generally used for both transmitting and receiving. The receiver must be protected from damage caused by the high power of the transmitter. This is the function of the duplexer. The duplexer also serves to channel the returned echo signals to the receiver and not to the transmitter. The duplexer consists of two gas-discharge devices, one known as a TR (transmit-receive) and the other as ATR (anti-transmit-receive). The TR protects the receiver during transmission and the ATR directs the echo signal to the receiver during reception. Solid-state ferrite circulators and receiver protectors with gas-plasma TR devices and/or diode limiters are also employed as duplexers. The receiver is usually of the super heterodyne type. The first stage normally is a low-noise RF amplifier, such as a parametric amplifier or a low-noise transistor. The mixer and local oscillator (LO) convert the RF signal to an intermediate frequency IF. Typical IF amplifier center frequencies are 30 or 60 MHz and will have a bandwidth of the order of one megahertz.

The IF amplifier should be designed as a ***matched filter*** i.e.***, its frequency-response function H(f) should maximize the peak-signal-to-mean-noise-power ratio at the output. This occurs when the magnitude of the frequency-response function |H(f|is equal to the magnitude of the echo signal spectrum |S(f)|, and the phase spectrum of the matched filter is the negative of the phase spectrum of the echo signal.*** In a radar whose signal waveform approximates a rectangular pulse, the conventional IF filter band pass characteristic approximates a matched filter when the product of the IF bandwidth **B** and the pulse width **τ** is of the order of unity, that is, ***Bτ = 1***.

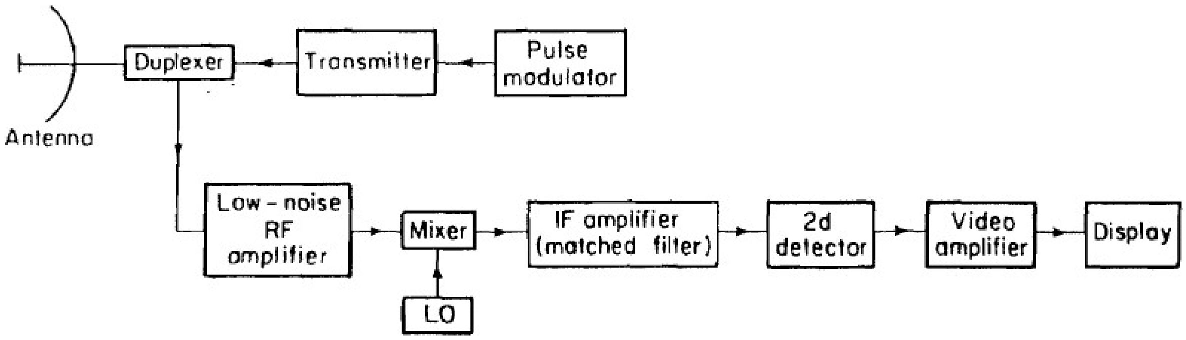


Fig 1.2: Block diagram of a pulse radar.

After maximizing the signal-to-noise ratio in the IF amplifier, the pulse modulation is extracted by the second detector and amplified by the video amplifier to a level where it can be properly displayed, usually on a cathode-ray tube (CRT). Timing signals are also supplied to the indicator to provide the range zero. Angle information is obtained from the pointing direction of the antenna.

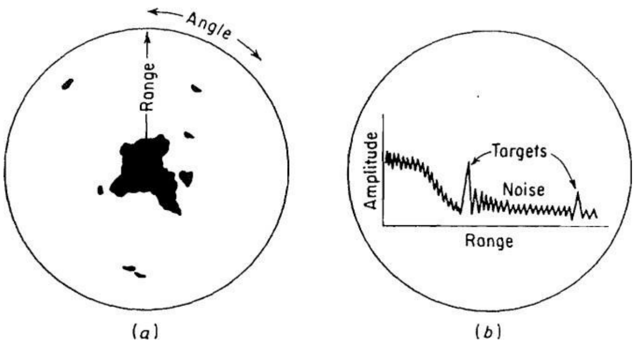


Fig 1.3(a) PPI presentation displaying Range vs. Angle (intensity modulation)
(b) A-scope presentation displaying Amplitude vs. Range (deflection modulation)

The most common form of cathode-ray tube display is the **P**lan **P**osition **I**ndicator, or **PPI (Fig. a)** which maps in polar coordinates the location of the target in azimuth and range. This is an intensity-modulated display in which the amplitude of the receiver output modulates the electron-beam intensity (z axis) as the electron beam is made to sweep outward from the center of the tube. The beam rotates in angle in response to the antenna position. Another form of display is the ***A-scope,*** shown in Fig. **b,** which plots target amplitude **(y axis)** vs. range **(x axis)**, for some fixed direction. This is a deflection-modulated display. It is more suited for tracking-radar application than for surveillance radar.

A common form of radar antenna is a reflector with a parabolic shape, fed (illuminated) from a point source at its focus. The parabolic reflector focuses the energy into a narrow beam, just as a searchlight or an automobile headlamp. The beam is scanned in space by mechanical pointing of the antenna.

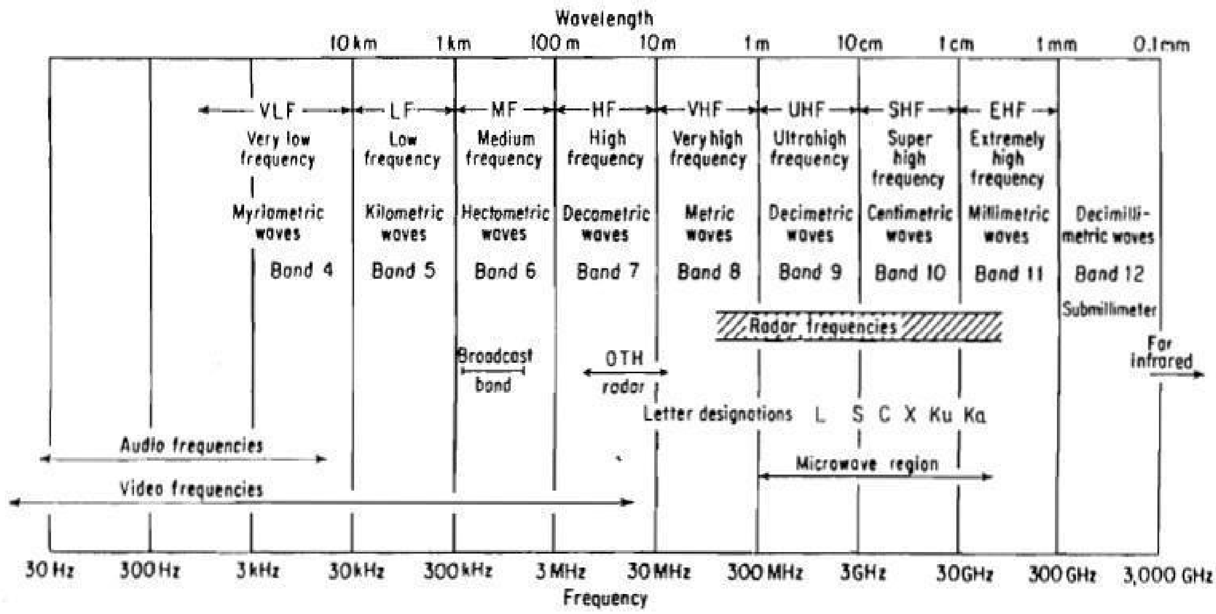
Radar frequencies and applications:

Radar frequencies:

Conventional radars are operated at frequencies extending from about 220 MHz to 35 GHz**,** a spread of more than seven octaves. These are not necessarily the limits, since radars can be, and have been, operated at frequencies outside either end of this range.

The place of radar frequencies in the electromagnetic spectrum is shown in the figure below. Some of the nomenclature employed to designate the various frequency regions is also shown in this figure.

ELECROMAGNETIC SPECTRUM



Letter code designation of Radar frequencies:

Early in the development of radar, a letter code such as S, X, L, etc., was employed to designate Radar frequency bands. Although it's original purpose was to guard military secrecy, the designations were maintained, probably out of habit as well as the need for some convenient short nomenclature. This usage has continued and is now an accepted practice of radar engineers. The table below lists the radar-frequency letter-band nomenclature adopted by the **IEEE.** These are related to the specific bands assigned by the **I**nternational **T**elecommunications Union for radar. For example, although the nominal frequency range for L band is 1000 to 2000 MHz, a L-band radar is thought of as being confined within the region from 1215 to1400MHz since that is the extent of the assigned band.

Table 1.1: Standard radar-frequency letter-band nomenclature

Band designation	Nominal frequency range	Specific radiolocation (radar) bands based on ITU assignments for region 2
HF	3–30 MHz	
VHF	30–300 MHz	138–144 MHz
		216–225
UHF	300–1000 MHz	420–450 MHz
		890–942
L	1000–2000 MHz	1215–1400 MHz
S	2000–4000 MHz	2300–2500 MHz
		2700–3700
C	4000–8000 MHz	5250–5925 MHz
X	8000–12,000 MHz	8500–10,680 MHz
K_u	12.0–18 GHz	13.4–14.0 GHz
		15.7–17.7
K	18–27 GHz	24.05–24.25 GHz
K_a	27–40 GHz	33.4–36.0 GHz
mm	40–300 GHz	

Applications of Radar:

1. Military Use: Initial and important user of Radar

(i)Early warning of intruding enemy aircraft & missiles

(ii) Tracking hostile targets and providing location information to Air Defense systems consisting of Tracking Radars controlling guns and missiles.

(iii) Battle field surveillance

(iv)Information Friend or Foe IFF

(v)Navigation of ships, aircraft, helicopter etc.

2. Civilian Use:

(i) Air Traffic Control (ATC)

All airports are equipped with ATC Radars, for safe landing and take-off and guiding of aircraft in bad weather and poor visibility conditions.

(ii) Aircraft Navigation

(a) All aircrafts fitted with weather avoidance radars. These Radars give warning information to pilot about storms, snow precipitation etc. lying ahead of aircraft's path.

(b) Radar is used as an altimeter to indicate the height of the aircraft or helicopter.

3. Maritime ship's safety and Navigation:

(i)Radar used to avoid collision of ships during poor visibility conditions (storms, cyclones etc.)

(ii)Guide ships into seaports safely.

4. Meteorological Radar:

Used for weather warnings and forecasting. Provides sufficient advance information to civilian administration for evacuation of population in times cyclones, storms etc.

Prediction of Range Performance:

The simple form of Radar equation derived earlier expresses the maximum radar range R_{max} in terms of radar and target parameters:

$$R_{max} = \left[(P_t . G. A_e. \sigma) / (4\pi)^2 . S_{min} \right]^{1/4}$$

Where P_t = transmitted power, watts

G = antenna gain

A_e = antenna effective aperture, m^2

σ = radar cross section, m^2

S_{min} = minimum detectable signal, watts

All the parameters are to some extent under the control of the radar designer, except for the target cross section **σ.** The radar equation states that if long ranges are desired,

- The transmitted power must be large,
- The radiated energy must be concentrated into a narrow beam (high transmitting antenna gain),
- The received echo energy must be collected with a large antenna aperture (also synonymous with high gain) and
- The receiver must be sensitive to weak signals.

In practice, however, the simple radar equation does not predict the range performance of actual radar equipment to a satisfactory degree of accuracy. The predicted values of radar range are usually optimistic. In some cases, the actual range might be only half of that is predicted.

Part of this discrepancy is due to

- The failure of the above equation to explicitly include the various losses that can occur throughout the system or
- The loss in performance usually experienced when electronic equipment is operated in the field rather than under laboratory-type conditions and

- Another important factor i.e the statistical or unpredictable nature of several of the parameters in the radar equation.

The minimum detectable signal S_{min} and the target cross section **σ** are both statistical in nature and must be expressed in statistical terms.

- Other statistical factors which do not appear explicitly in the simple radar equation but which have an effect on the radar performance are the meteorological conditions along the propagation path and the performance of the radar operator, if one is employed.

The statistical nature of these several parameters does not allow the maximum radar range to be described by a single number. Its specification must include a statement of the probability that the radar will detect a certain type of target at a particular range.

- *Hence in order to cover these aspects, the simple radar equation will be **modified** to include most of the important factors that influence radar range performance*.

Minimum detectable signal:

- The ability of a radar receiver to detect a weak echo signal is limited by the noise energy that occupies the same portion of the frequency spectrum as does the signal energy and accompanies the signal.
- The weakest signal the receiver can detect is called the ***minimum detectable signal.*** It is difficult to define ***minimum detectable signal*** (MDS) because of its statistical nature and because the criterion for deciding whether a target is present or not is not too well defined.
- Detection is normally based on establishing a threshold level at the output of the receiver (as shown by the dotted line in the figure below.)Whenever Rx output signal which is a mixture of echo and noise crosses this threshold, then it is detected as a target. This is called ***threshold detection.***
- Consider the output of a typical radar receiver as a function of time as shown in the figure below which typically represents one sweep of the video output displayed on an A-scope.

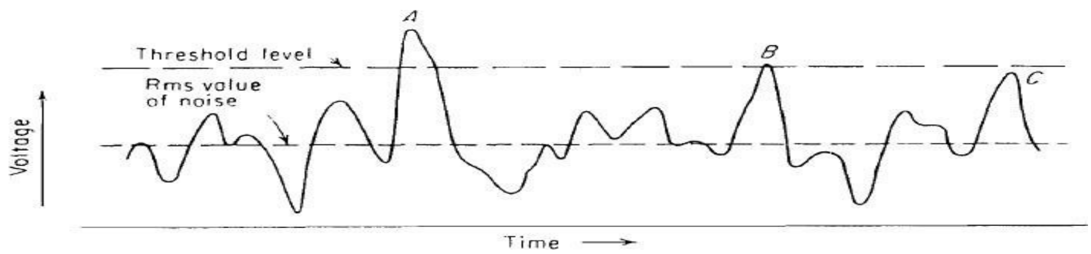


Fig 1.4: Typical envelope of the radar receiver output as a function of time. *A,* B, and *C*are three targets representing signal plus noise. A and B are valid detections, but C is a missed detection.

- The envelope has a fluctuating appearance due to the random nature of noise and consists of three targets A, B and C of different signal amplitudes.
- The signal at **A** is large which has much larger amplitude than the noise. Hence target detection is possible without any difficulty and ambiguity.
- Next consider the two signals at **B** and **C**, representing target echoes of equal amplitude. The noise voltage accompanying the signal at **B** is large enough so that the combination of signal plus noise exceeds the threshold and target detection is still possible.
- But for the target **C,** the noise is not as large and the resultant signal plus noise does not cross the threshold and hence target is not detected.
- ***Threshold Level setting:*** Weak signals such as **C** would not be lost if the threshold level were lower. But too low a threshold increases the likelihood that noise alone will rise above the threshold and is taken as target. Such an occurrence is called a ***false alarm***. Therefore, if the threshold is set too low, false target indications are obtained, but if it is set too high, targets might be missed. The selection of the proper threshold level is a compromise that depends upon how important it is if a mistake is made either by
  1. Failing to recognize a signal that is present (***probability of a miss***) or by
  2. Falsely indicating the presence of a signal when it does not exist (***probability of a false alarm***)
- The signal-to noise ratio necessary to provide adequate detection is one of the important parameters that must be determined in order to compute ***the minimum detectable signal***.
- Although the detection decision is usually based on measurements at the video output, it is easier to consider maximizing the signal-to-noise ratio at the output of the IF amplifier rather than in the video. The receiver may be considered linear up to the output of the IF. It is shown that maximizing the signal-to-noise ratio at the output of the IF is equivalent to maximizing the video output. The advantage of considering the signal-to-noise ratio at the IF is that the assumption of linearity may be made. It is also assumed that the IF filter characteristic approximates the matched filter, so that the output signal-to-noise ratio is maximized.

Receiver noise:

- Noise is unwanted electromagnetic energy which interferes with the ability of the receiver to detect the wanted signal thus limiting the receiver sensitivity.

It may originate within the receiver itself, or it may enter via the receiving antenna along with the desired signal. If the radar were to operate in a perfectly noise-free environment so that no external sources of noise accompanied the desired signal, and if the receiver itself were so perfect that it did not generate any excess noise, ***there would still exist an unavoidable component of noise generated by the thermal motion of the conduction electrons in the ohmic portions of the receiver input stages. This is called thermal noise, or Johnson's noise, and is directly proportional to the temperature of the ohmic portions of the circuit and the receiver band width***. The available noise power generated by a receiver of bandwidth B_n (in hertz) at a temperature **T** (degrees Kelvin) is given by :

Available thermal-noise power = **kTB_n**

where **k** = Boltzmann's constant =**1.38×10^{-23}J/deg**. If the temperature **T** is taken to be290 K, which corresponds approximately to room temperature (62^0F), the factor **kT** is **4×10^{-21}W/Hz** of bandwidth. If the receiver circuitry were at some other temperature, the thermal-noise power would be correspondingly different.

- A receiver with a reactance input such as a parametric amplifier need not have any significant ohmic loss. The limitation in this case is the thermal noise seen by the antenna and the ohmic losses in the transmission line.
- For radar receivers of the super heterodyne type (the type of receiver used for most radar applications), the receiver bandwidth is approximately that of the **intermediate-frequency** stages. It should be cautioned that the bandwidth **B_n** mentioned above is not the 3-dB, or half-power, bandwidth commonly employed by electronic engineers. It is an integrated bandwidth and is given by:

$$B_n = \frac{\int_{-\infty}^{\infty} |H(f)|^2 \, df}{|H(f_0)|^2}$$

where **H (f)** = frequency-response characteristic of IF amplifier (filter) and **f_o** = frequency of maximum response (usually occurs at mid band).

- ***The bandwidth B_n is called the noise bandwidth and is the bandwidth of an equivalent rectangular filter whose noise-power output is same as the filter with characteristic H(f).***It is not theoretically same as the 3-dB bandwidth. The 3-dB bandwidth is widely used since it is easy to measure. The measurement of noise bandwidth however involves a complete knowledge of the response characteristic **H(f).** The frequency-response characteristics of many practical radar receivers are such that the 3 dB and the noise bandwidths do not differ appreciably. Therefore the 3-dB bandwidth may be used in many cases as an approximation to the noise bandwidth.
- The noise power in practical receivers is often greater than can be accounted for by thermal noise alone and is due to mechanisms other than the thermal agitation of the conduction electrons. The exact origin of the extra noise components is not important except to know that it exists. Whether the noise is generated by a thermal mechanism or by some other mechanism the total noise at the output of the receiver may be considered to be equal to the thermal-noise power obtained from an “ideal“ receiver multiplied by a factor called the ***noise figure.***
- The noise figure **F_n** of a receiver is defined by the equation:

$$F_n = N_o / kT_o B_n G_a$$

= (Noise output of practical receiver) / (Noise output of ideal receiver at std. temp **T_o**)
Where **N_o** = noise output from receiver, and **G_a** = available gain. The standard temperature **T_o** is taken to be 290 **K,** according to the Institute of Electrical and Electronics Engineers definition. The noise **N_o** is measured over the linear portion of the receiver input-output characteristic, usually at the output of the IF amplifier before the nonlinear second detector. The receiver bandwidth ***B_n*** is that of the IF amplifier in most

receivers. The available gain G_a is the ratio of the signal out S_o to the signal in S_i and $kT_o B_n$ is the input noise N_i in an ideal receiver. The above equation may be rewritten as:

$$F_n = \frac{S_i/N_i}{S_o/N_o}$$

Therefore, the ***noise figure*** may be interpreted, as a measure of the degradation of signal-to noise-ratio as the signal passes through the receiver.

Modified radar equation:

Rearranging the above two equations for F_n, the input signal may be expressed as

$$S_i = \frac{kT_0 B_n F_n S_o}{N_o}$$

If the minimum detectable signal S_{min} is that value of S_i corresponding to the minimum ratio of output (IF) signal-to-noise ratio $(S_o/N_o)_{min}$ necessary for detection, then

$$S_{min} = kT_0 B_n F_n \left(\frac{S_o}{N_o}\right)_{min}$$

Substituting this expression for S_{min} into the earlier basic Radar equation results in the following form of the **modified radar equation**:

$$R_{max}^4 = \frac{P_t G A_e \sigma}{(4\pi)^2 k T_0 B_n F_n (S_o/N_o)_{min}} \quad \text{.........(4)}$$

RADAR EQUATION

Signal to Noise Ratio (SNR):

The results of statistical noise theory will be applied to obtain:

- ***The signal-to-noise ratio at the output of the IF amplifier necessary to achieve a specified probability of detection without exceeding a specified probability of false alarm***.

The output signal-to-noise ratio thus obtained is substituted into the final modified radar equation, we have obtained earlier.

The details of system that is considered:

- **IF** amplifier with bandwidth B_{IF} followed by a second detector and a video amplifier with bandwidth B_V as shown in the figure below.
- The second detector and video amplifier are assumed to form an envelope detector, that is, one which rejects the carrier frequency but passes the modulation envelope.
- To extract the modulation envelope, the video bandwidth must be wide enough to pass the low-frequency components generated by the second detector, but not so wide as to pass the high-frequency components at or near the intermediate frequency.

- The video bandwidth **B_V** must be greater than **$B_{IF}/2$** in order to pass all the videomodulation.

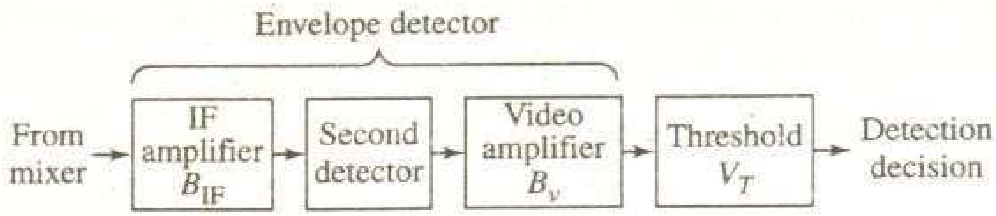


Figure 1.6: Envelope detector.

Step 1: To determine the Probability of false alarm when noise alone is assumed to be present as input to the receiver:

The noise entering the IF filter (the terms filter and amplifier are used interchangeably) is assumed to be Gaussian, with probability-density function given by

$$p(v) = \frac{1}{\sqrt{2\pi\psi_0}} \exp \frac{-v^2}{2\psi_0}$$

Where:

- **p(v) dv** is the probability of finding the noise voltage ***v*** between the values of **v**and**v + dv**
- ***ψ_0***is the variance, or mean-square value of the noise voltage, and the mean value of **v** is taken to be zero.

(Compare this with the Standard Probability density function of Gaussian noise

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{-(x-x_0)^2}{2\sigma^2}$$

*With σ^2 replaced by $\psi_0$and **$(x-x_0)$** replaced by **v** with mean value of zero)*

If Gaussian noise were passed through a narrowband IF filter whose Bandwidth is small compared with its mid band frequency-the probability density of the envelope of the noise voltage output is shown by **Rice** to be of the form of **Rayleigh** probability-density function

$$p(R) = \frac{R}{\psi_0} \exp\left(-\frac{R^2}{2\psi_0}\right) \quad \text{..........(6)}$$

where ***R is the amplitude of the envelope of the filter output***. The probability that the envelope of the noise voltage will lie between the values of $V_1$and$V_2$is

$$\text{Probability } (V_1 < R < V_2) = \int_{V_1}^{V_2} \frac{R}{\psi_0} \exp\left(-\frac{R^2}{2\psi_0}\right) dR$$

The probability that the noise voltage envelope will exceed the voltage threshold **V_T** is

$$\text{Probability } (V_T < R < \infty) = \int_{V_T}^{\infty} \frac{R}{\psi_0} \exp\left(-\frac{R^2}{2\psi_0}\right) dR$$

$$= \exp\left(-\frac{V_T^2}{2\psi_0}\right) = P_{fa} \qquad \textbf{…………(7)}$$

Whenever the voltage envelope exceeds the threshold **V_T**, a target is considered to have been detected. Since the probability of a false alarm is the probability that noise will cross the threshold, the above equation gives the probability of a false alarm, denoted by ***P_{fa}***.

The probability of false alarm as given above by itself does not indicate that Radar is troubled by the false indications of Target. The time between the false alarms **T_{FA}** is a better measure of the effect of Noise on the Radar performance. (Explained with reference to the figure below)
The average time interval between crossings of the threshold by noise alone is defined as the ***false-alarm time T_{FA}***

$$T_{fa} = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} T_k$$

Where **T_K** is the time between crossings of the threshold ***V_T*** by the noise envelope, when the slope of the crossing is positive.
The false-alarm probability may also be defined as the ratio of the duration of time the envelope is actually above the threshold to the total time it ***could have been*** above the threshold, i.e.

$$P_{fa} = \frac{\sum_{k=1}^{N} t_k}{\sum_{k=1}^{N} T_k} = \frac{\langle t_k \rangle_{av}}{\langle T_k \rangle_{av}} = \frac{1}{T_{fa} B} \qquad \textbf{…………. (8)}$$

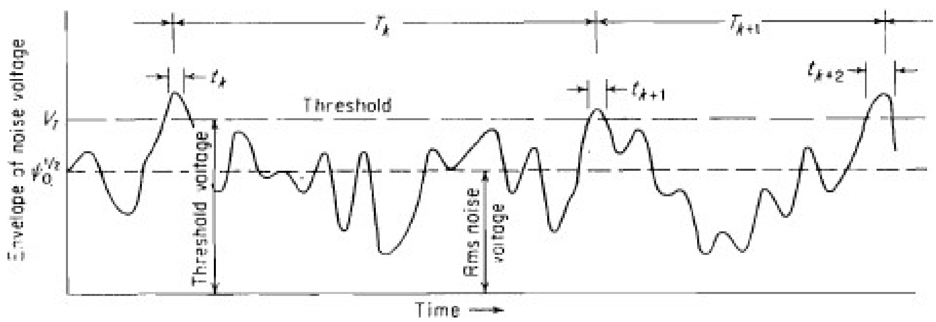


Fig 1.7: Envelope of receiver output illustrating false alarms due to noise.

Where ***t_K*** and **T_K** are shown in the Figure above. ***The average duration of a noise pulse is approximately the reciprocal of the bandwidth B,*** which in the case of the envelope detector is **B_{IF}**. Equating eqs. **7** and **8** we get

$$T_{fa} = \frac{1}{B_{IF}} \exp \frac{V_T^2}{2\psi_0} \qquad \text{……….(9)}$$

A plot of the above equation is shown in the figure below with **(V_T^2/2 ψ_0)** as the abscissa. As can be seen, average time between false alarms T_{fa} is directly proportional to the Threshold to noise ratio and inversely proportional to the Bandwidth.

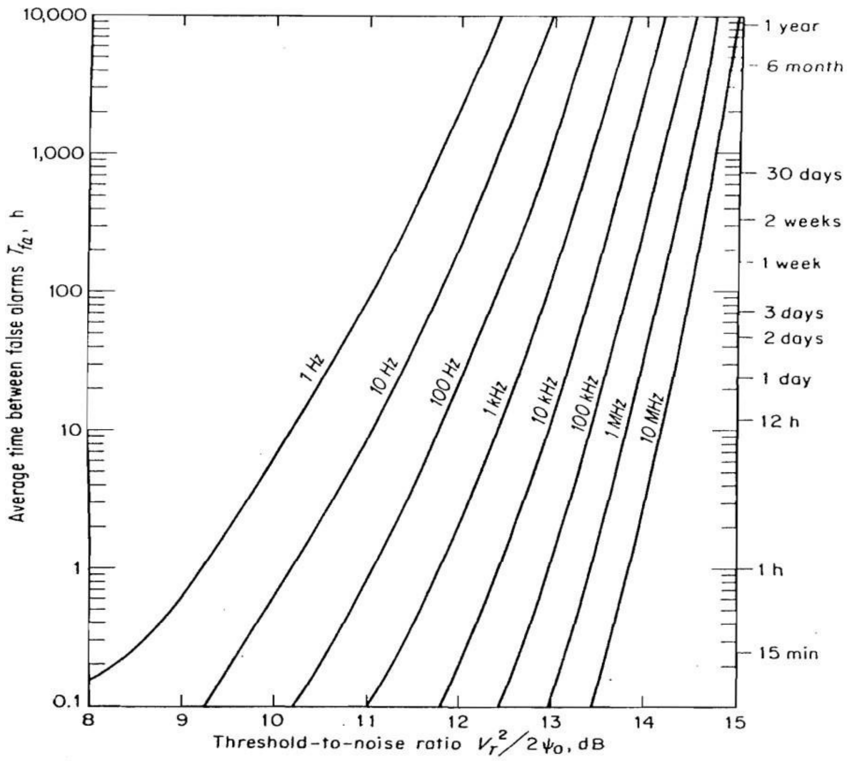


Fig. 1.8: Average time between false alarms as a function of the threshold level V_T and the receiver Bandwidth B. ψ_0 is the mean square noise voltage

Step 2 :

To determine Probability of detection when a sine wave signal is present along with noise:

Thus far, a receiver with only a noise input was discussed. Next, consider a sine-wave signal of amplitude **A** to be present along with noise at the input to the IF filters. The frequency of the signal is the same as the IF mid band frequency f_{IF}. The output of the envelope detector has a probability-density function given by

$$p_s(R) = \frac{R}{\psi_0} \exp\left(-\frac{R^2 + A^2}{2\psi_0}\right) I_0\left(\frac{RA}{\psi_0}\right) \qquad \text{………….(10)}$$

where $I_0(Z)$ is the modified Bessel function of zero order and argument **Z.**

When the signal is absent, A = 0 and the above equation for **PDF** for signal plus noise reduces to the probability-density function for noise alone. This Equation is sometimes called the **Rice** probability-density function.

The probability that the signal will be detected (which is the ***probability of detection)*** is the same as the probability that the envelope **R** will exceed the predetermined threshold V_T. The probability of detection P_d is therefore:

$$P_d = \int_{V_T}^{\infty} p_s(R)\, dR = \int_{V_T}^{\infty} \frac{R}{\psi_0} \exp\left(-\frac{R^2 + A^2}{2\psi_0}\right) I_0\left(\frac{RA}{\psi_0}\right) dR \quad \text{……….(11)}$$

(After the expression of **PDF** for **$P_s(R)$[Eq. 10]** is substituted into the first part of the above equation we get the probability of detection as in **[eqn.11]).** But this equation cannot be evaluated by simple means, and numerical & empirical techniques or a series approximation must be used.

The expression for P_d given by equation **(11)** after series expansion is a function of the signal amplitude A, threshold voltage V_T ,and mean noise power Ψ_0 . In Radar systems analysis, it is more convenient to use Signal to Noise power ratio (S/N) rather than signal to noise voltage ratio $A/\Psi_0^{1/2}$.These are related by:

$$\frac{A}{\psi_0^{1/2}} = \frac{\text{signal amplitude}}{\text{rms noise voltage}} = \frac{\sqrt{2}(\text{rms signal voltage})}{\text{rms noise voltage}} = \left(2\,\frac{\text{signal power}}{\text{noise power}}\right)^{1/2} = \left(\frac{2S}{N}\right)^{1/2}$$

The probability of detection P_d can then be expressed in terms of S/N , and Threshold- Noise ratio $V_T^2/2\Psi_0$The probability of false alarm is also a function of $V_T^2/2\Psi_0$ as given by : **P_{FA} = Exp($-V_T^2/2\psi_0$).**

The two expressions for P_d and P_{FA} can now be combined by eliminating the Threshold- Noise ratio **$V_T^2/2\Psi_0$**that is common in both expressions so as to get a single expression relating the probability of detection **P_d** ,Probability of false alarm **P_{FA}** and signal to Noise ratio **S/N**.The result is plotted in the figure below.

A much easier empirical formula developed by ***Albersheim*** for the relationship between **S/N,P_{FA}** and **P_d** is also given below :

S/N = A + 0.12AB + 1.7 B

Where A = ln [0.62/ P_{FA}] and B = ln [P_d /(1- P_d)]

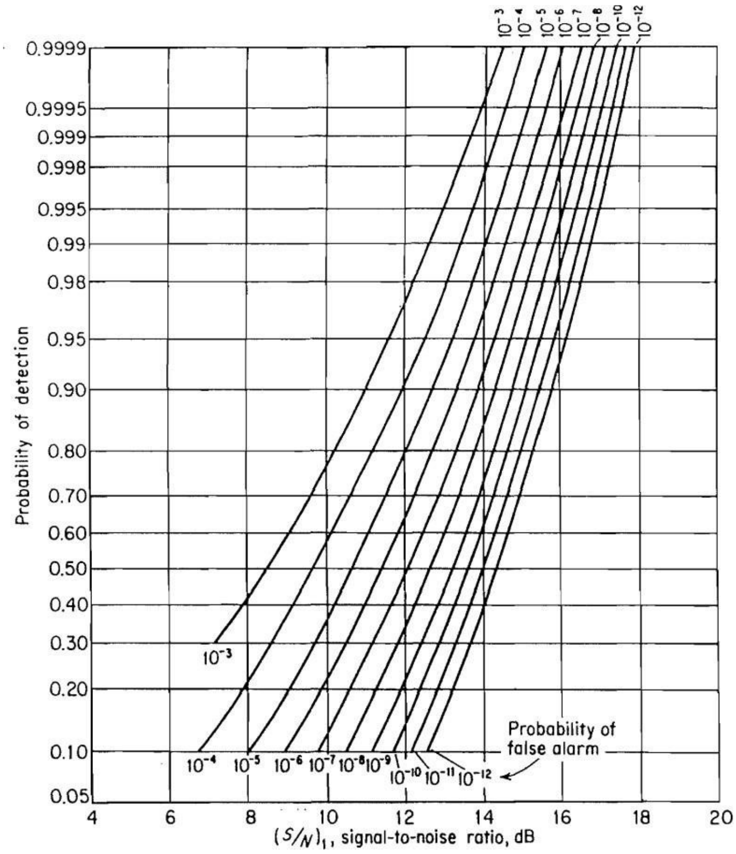


Fig. 1.9: Probability of detection for a sine wave in noise as a function of the signal-to-noise (power) ratio and the probability of false alarm

System design sequence:

- Both the false-alarm time$\mathbf{T_{FA}}$ and the detection probability$\mathbf{P_d}$are specified by the system requirements.
- The radar designer computes the probability of the false alarm using the above $\mathbf{T_{fa}}$& the relation

$$\mathbf{P_{fa} = 1/ T_{fa} . B}$$

- Then from the figure above or using the ***Albersheim's*** empirical equation given above the required signal-to-noise ratio to achieve the above $\mathbf{P_{fa}}$ & $\mathbf{P_d}$ is determined.

For example, suppose that the desired false-alarm time was **15 min** and the IF bandwidth was **1** MHz. This gives a false-alarm probability of $\mathbf{1.11 \times 10^{-9}}$**.** Figure above indicates that a signal-to-noise ratio of **13.1** dB is required to yield a **0.50** probability of detection, **14.7 dB** for **0.90**, and **16.5 dB** for **0.999**.

This is the signal-to-noise ratio that is to be used in the final modified Radar Equation we have obtained earlier.

$$R_{max}^4 = \frac{P_t G A_e \sigma}{(4\pi)^2 k T_0 B_n F_n (S_o/N_o)_{min}}$$

Integration of Radar Pulses:

The relation between the signal to noise ratio, the probability of detection and the probability of false alarm as shown in the figure or as obtained using the ***Albersheim's*** empirical equation applies for a single pulse only. However, many pulses are usually returned from any target on each radar scan and can be used to improve detection. The number of pulses n_B returned from a point target as the radar antenna scans through its beam width is

$$n_B = \theta_B \cdot f_P / \theta'_S = \theta_B \cdot f_P / 6\ \omega_m$$

where θ_B= antenna beam width, deg
f_P= pulse repetition frequency, **Hz**
θ'_S = antenna scanning rate, deg/s
ω_m= antenna scan rate, rpm

The process of summing all the radar echo pulses for the purpose of improving detection is called integration.
Integration may be accomplished in the radar receiver either before the second detector (in the IF) or after the second detector (in the video).

- Integration before the detector is called pre detection or coherent integration. In this the phase of the echo signal is to be preserved if full benefit is to be obtained from the summing process
- Integration after the detector is called post detection or non coherent integration. In this phase information is destroyed by the second detector. Hence post detection integration is not concerned with preserving RF phase. Due to this simplicity it is easier to implement in most applications, but is not as efficient as pre detection integration.

If ***n*** pulses, all of the same signal-to-noise ratio, were integrated by an ideal pre detection integrator, the resultant or integrated signal-to-noise (power) ratio would be exactly **n** times that of a single pulse. If the same ***n*** pulses were integrated by an ideal post detection device, the resultant signal-to-noise ratio would be less than **n** times that of a single pulse. ***This loss in integration efficiency is caused by the nonlinear action of the second detector, which converts some of the signal energy to noise energy in the rectification process.***

Due to its simplicity, Post detection integration is preferred many a times even though the integrated signal-to-noise ratio may not be as high as that of Pre-detection. An alert, trained operator viewing a properly designed cathode-ray tube display is a close approximation to the theoretical post detection integrator.
The efficiency of post detection integration relative to ideal pre-detection integration has been computed by ***Marcum*** when all pulses are of equal amplitude. The integration efficiency may be defined as follows:

$$E_i(n) = \frac{(S/N)_1}{n(S/N)_n}$$

Where ***n*** = number of pulses integrated

(S/N)₁= value of signal-to-noise ratio of a single pulse required to produce a given probability of detection(for **n = 1**)

(S/N)ₙ = value of signal-to-noise ratio per pulse required to produce the same probability of detection when **n** pulses (of equal amplitude) are integrated

The improvement in the signal-to-noise ratio when ***n*** pulses are integrated post detection is $n.E_i(n)$ and is the integration-improvement factor. ***It may also be thought of as the effective number of pulses integrated by the post detection integrator***. The improvement with ideal pre detection integration would be equal to **n**. Integration loss in decibels is defined as $L_i(n) = 10 \log [1/E_i(n)]$.
The integration-improvement factor (or the integration loss) is not a sensitive function of either the probability of detection or the probability of false alarm.

The radar equation with ***n*** pulses integrated can be written

$$R_{max}^4 = \frac{P_t G A_e \sigma}{(4\pi)^2 k T_0 B_n F_n (S/N)_n}$$

where the parameters are the same as in the earlier Radar equation except that $(S/N)_n$, is the signal-to-noise ratio of one of the **n** equal pulses that are integrated to produce the required probability of detection for a specified probability of false alarm. Substituting the equation for integration efficiency

$$E_i(n) = \frac{(S/N)_1}{n(S/N)_n}$$

into the above Radar equation gives the final modified Radar equation including integration efficiency.

$$R_{max}^4 = \frac{P_t G A_e \sigma n E_i(n)}{(4\pi)^2 k T_0 B_n F_n (S/N)_1}$$

Radar Cross Section of Targets:

The radar cross section of a target is the (fictional) area intercepting that amount of power which when scattered equally in all directions, produces an echo at the radar equal to that from the target. Or in other terms

$$\sigma = \frac{\text{power reflected toward source/unit solid angle}}{\text{incident power density}/4\pi}$$

$$= \lim_{R \to \infty} 4\pi R^2 \left| \frac{E_r}{E_i} \right|^2$$

Where **R**= distance between radar and target

E_r= strength of reflected field at radar

E_i= strength of incident field at target

For most common types of radar targets such as aircraft, ships, and terrain, the radar cross section does not necessarily bear a simple relationship to the physical area, except that the larger the target size, the larger will be the cross section.

Scattering and ***diffraction:*** are variations of the same physical process. When an object scatters an electromagnetic wave, the scattered field is defined as the difference between the total field in the presence of the object and the field that would exist if the object were absent (but with the sources unchanged). On the other hand, the diffracted field is the total field in the presence of the object. With radar backscatter, the two fields are the same, and one may talk about scattering and diffraction interchangeably.

Radar cross section of a simple sphere: is shown in the figure below as a function of its circumference measured in wavelengths.(**$2\pi a/\lambda$** where ***a*** is the radius of the sphere and **λ** is the wavelength). The plot consists of three regions.

1. Rayleigh Region:

- The region where the size of the sphere is small compared with the wavelength (**$2\pi a/\lambda < 1$**) is called the **Rayleigh** region.
- The Rayleigh scattering region is of interest to the radar engineer because the cross sections of raindrops and other meteorological particles fall within this region at the usual radar frequencies. **2. Optical region**:
- It is at the other extreme from the **Rayleigh** region where the dimensions of the sphere are large compared with the wavelength **($2\pi a/\lambda > 1$)**. For large **$2\pi a/\lambda$**, the radar cross section approaches the optical cross section **πa^2**.

3. Mie or Resonance region:

- Between the optical and the Rayleigh region is the ***Mie,*** or resonance, region. The cross section is oscillatory with frequency within this region. The maximum value is 5.6 dB greater than the optical value, while the value of the first null is 5.5 dB below the optical value. (The theoretical values of the maxima and minima may vary according to the method of calculation employed.

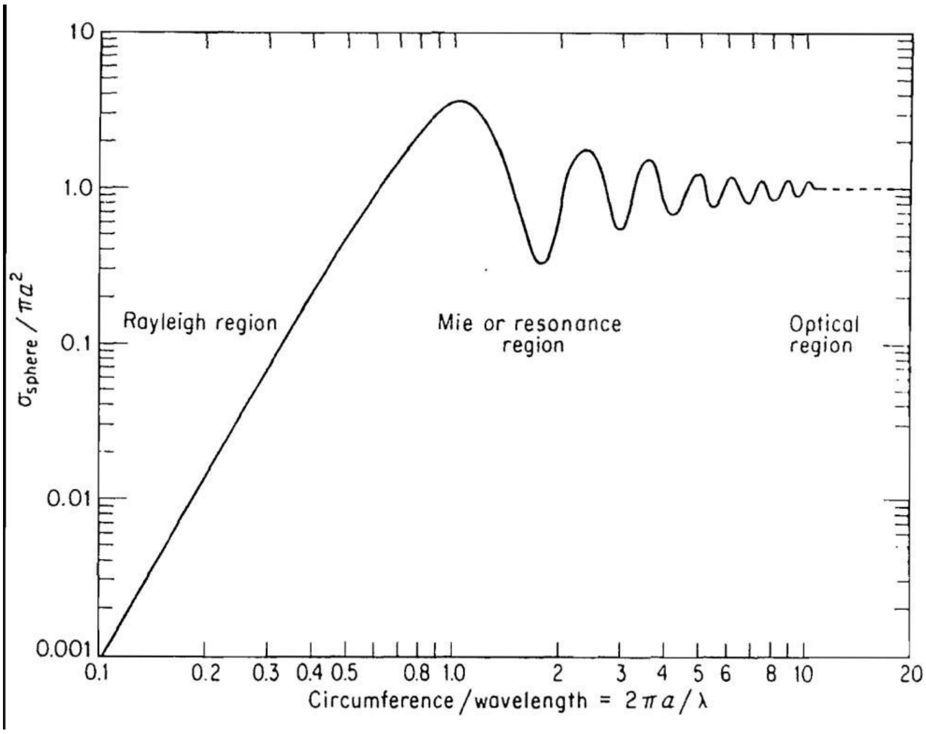


Figure 1.10: Radar cross section of the sphere. a = radius; λ= wavelength.

Since the sphere is a sphere no matter from what aspect it is viewed, its cross section will not be aspect-sensitive. The cross section of other objects, however, will depend upon the direction as viewed by the radar. (Aspect angle)

Radar cross section of a cone-sphere:

- An interesting radar scattering object is the cone-sphere, a cone whose base is capped with a sphere such that the first derivatives of the contours of the cone and sphere are equal at the joint. Figure below is a plot of the nose-on radar cross section. The cross section of the cone-sphere from the vicinity of the nose-on direction is quite low.
- Scattering from any object occurs from discontinuities. The discontinuities, and hence the backscattering, of the cone-sphere are from the tip and from the join between the cone and the sphere.
- The nose-on radar cross section is small and decreases as the square of the wavelength. The cross section is small over a relatively large angular region. **A** large **specular***(having qualities of a mirror)*return is obtained when the cone-sphere is viewed at near perpendicular incidence to the cone surface, i.e., when **θ**= 90 - **α**, where **α**= cone half angle. From the rear half of the cone-sphere, the radar cross section is approximately that of the sphere.

- The nose-on cross section of the cone-sphere varies, but its maximum value is approximately $\mathbf{0.4\lambda^2}$ and its minimum is $\mathbf{0.01\lambda^2}$ for a wide range of half-angles for frequencies above the Rayleigh region. The null spacing is also relatively insensitive to the cone half-angle.

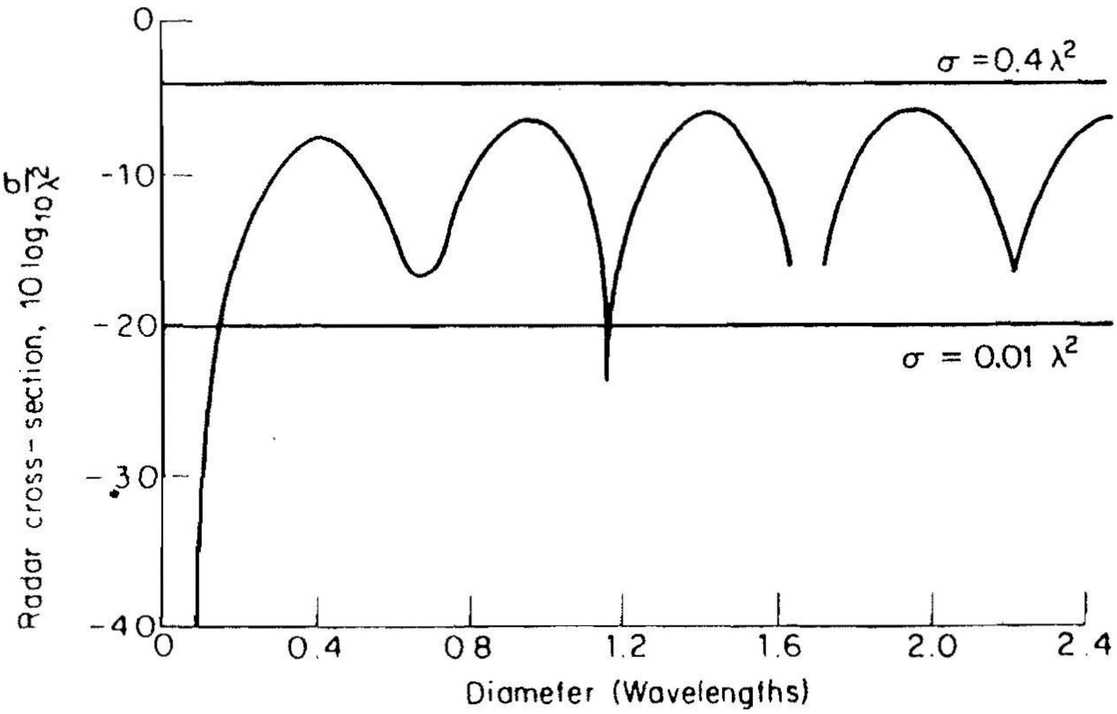


Figure 1.11: Radar cross section of a cone sphere with 15^0 half angle as a function of the diameter in Wave lengths.

- In order to realize in practice the very low theoretical values of the radar cross section for a cone sphere, the tip of the cone must be sharp and not rounded, the surface must be smooth (roughness small compared to a wavelength), the join between the cone and the sphere must have a continuous first derivative, and there must be no holes, windows, or protuberances on the surface.

- Shaping of the target, as with the cone-sphere, is a good method for reducing the radar cross section. Materials such as carbon-fiber composites, which are sometimes used in aerospace applications, can further reduce the radar cross section of targets as compared with that produced by highly reflecting metallic materials.

Transmitter Power:

The peak power: The power P_t in the radar equation is called the ***peak*** power. This is not the instantaneous peak power of a sine wave. It is the power averaged over that carrier-frequency cycle which occurs at the maximum power of the pulse.

The average radar power P_{av}: It is defined as the average transmitter power over the pulse-repetition period. If the transmitted waveform is a train of rectangular pulses of width **τ** and pulse-repetition period $T_p = 1/f_p$, then the average power is related to the peak power by

$$P_{av} = \frac{P_t \tau}{T_p} = P_t \tau f_p$$

Duty cycle: The ratio P_{av}/P_t, τ/T_P, or $\tau . f_P$ is called the duty cycle of the radar. A pulse radar for detection of aircraft might have typically a duty cycle of 0.001, while a CW radar which transmits continuously has a duty cycle of unity.

Writing the radar equation in terms of the average power rather than the peak power, we get

$$R_{max}^4 = \frac{P_{av} G A_e \sigma n E_i(n)}{(4\pi)^2 k T_0 F_n (B_n \tau)(S/N)_1 f_p}$$

The bandwidth and the pulse width are grouped together since the product of the two is usually of the order of unity in most pulse-radar applications.

Pulse Repetition Frequencies and Range Ambiguities:

- The pulse repetition frequency (**prf**) is determined primarily by the maximum range at which targets are expected. If the **prf** is made too high, the likelihood of obtaining target echoes from the wrong pulse transmission is increased. Echo signals received after an interval exceeding the pulse-repetition period are called ***multiple time around echoes.***
- Consider the three targets labeled ***A, B,*** and **C** in the figure(a) below. Target **A** is located within the maximum unambiguous range R_{unamb} **[=** $C.T_P/2$**]** of the radar, target **B** is at a distance greater than R_{unamb} but less than $2R_{unamb}$ and the target **C** is greater than $2R_{unamb}$ but less than $3R_{unamb}$ The appearance of the three targets on an A-scope is shown in the figure (b)below. The multiple-time-around echoes on the A-scope cannot be distinguished from proper target echoes actually within the maximum unambiguous range. Only the range measured for target ***A*** is correct; those for **B** and **C** are not.
- One method of distinguishing multiple-time-around echoes from unambiguous echoes is to operate with a varying pulse repetition frequency. The echo signal from an unambiguous range target will appear at the same place on the A-scope on each sweep no matter whether the **prf** is modulated or not. However, echoes from multiple-time-around targets will be spread over a finite range as shown in the figure (c) below. The number of separate pulse repetition frequencies will depend upon the degree of the multiple time around targets. Second-time targets need only two separate repetition frequencies in order to be resolved.

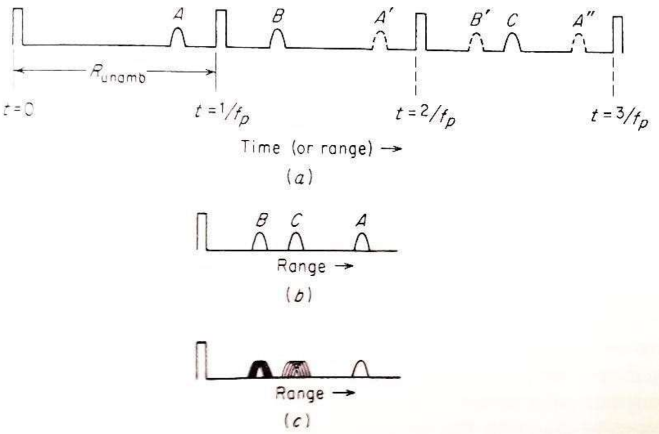


Fig. 1.12: Multiple-time-around echoes that give rise to ambiguities in range. (a) Three targets A, B and C, where A is within R_{unamb}, and B and Care multiple-time-around targets(b)the appearance ofthe three targets on the A-scope (c) appearance of the three targets on the A-scope with a changingprf.

System Losses:

- The losses in a radar system reduce the signal-to-noise ratio at the receiver output. They are two kinds, predictable with certain precision beforehand and unpredictable. The antenna beam-shape loss, collapsing loss, and losses in the microwave plumbing are examples of losses which are predictable if the system configuration is known. These losses are real and cannot be ignored.
- Losses not readily subject to calculation and which are less predictable include those due to field degradation and to operator fatigue or lack of operator motivation. They are subject to considerable variation and uncertainty.

Plumbing loss: This is loss in the transmission lines which connects the transmitter output to the antenna. (Cables and waveguides). At the lower radar frequencies the transmission line introduces little loss, unless its length is exceptionally long. At higher radar frequencies, loss/attenuation will not be small and has to be taken into account.

Connector losses: In addition to the losses in the transmission line itself, additional losses occurs at each connection or bend in the line and at the antenna **rotary joint** if used. Connector losses are usually small, but if the connection is poorly made, it can contribute significant attenuation. If the same transmission line is used for both receiving and transmission, the loss to be inserted in the radar equation is ***twice*** the one-way loss.

Duplexer loss: The signal suffers attenuation as it passes through the duplexer. Generally, the greater the isolation required from the duplexer on transmission, the larger will be the insertion loss. Insertion loss means the loss introduced when the component is inserted into the transmission line. For a typical duplexer it might be of the order of 1 dB.
In S-band (3000 MHz) radar, for example, the typical plumbing losses will be as follows:

100 ft of RG-113/U A1 waveguide transmission line (two-way):	*1.0 dB*
Loss due to poor connections (estimate):	*0.5 dB*
Rotary-joint loss:	*0.4 dB*
Duplexer loss:	*1.5 dB*
Total plumbing loss:	*3.4 dB*

Beam-shape loss: The antenna gain that appears in the radar equation was assumed to be a constant equal to the maximum value. But in reality the train of pulses returned from a target with scanning radar is modulated in amplitude by the shape of the antenna beam. To properly take into account the pulse-train modulation caused by the beam shape, the computations of the probability of detection (as explained earlier) would have to be performed assuming a modulated train of pulses rather than constant-amplitude pulses. But since this computation is difficult, a beam-shape loss is added to the radar equation and a maximum gain is employed in the radar equation rather than a gain that changes pulse to pulse.

Scanning loss: When the antenna scans rapidly enough , the gain on transmit is not the same as the gain on receive. An additional loss has to be computed, called the scanning loss. The technique for computing scanning loss is similar in principle to that for computing beam-shape loss. Scanning loss is important for rapid-scan antennas or for very long range radars such as those designed to view extraterrestrial objects.

Collapsing loss: If the radar were to integrate additional noise samples along with the wanted Signal-to-noise pulses, the added noise results in degradation called the ***collapsing loss.***

Non ideal equipment: The ***transmitter power*** in the radar equation was assumed to be the specified output power (either peak or average). However, all transmitting tubes are not uniform in quality, and even any individual tube performance will not be same throughout its useful life. Also, the power is not uniform over the operating band of frequencies. Thus, for one reason or another, the transmitted power may be other than the design value. To allow for this variation, a loss factor of about 2 dB is introduced.

Receiver noise figure also varies over the operating frequency band. Thus, if the best noise figure over the band is used in the radar equation, a loss factor has to be introduced to account for its poorer value elsewhere in the frequency band.
If the receiver is not the exact matched filter for the transmitted waveform, a loss in Signal-to-noise ratio will occur.
A typical value of loss for a non-matched receiver might be about **1 db.** Because of the exponential relation between the false-alarm time and the threshold level a slight change in the threshold can cause a significant change in the false alarm time. In practice, therefore, it may be necessary to set the threshold level slightly higher than calculated so as to insure a tolerable false alarm time in the event of circuit instabilities. This increase in the threshold is equivalent to a loss.

Operator loss: An alert, motivated, and well-trained operator performs as described by theory. However, when distracted, tired, overloaded, or not properly trained, operator performance will decrease. The resulting loss in system performance is called operator loss.

Field degradation: When a radar system is operated under laboratory conditions by engineering personnel and experienced technicians, the above mentioned losses give a realistic description of the performance of the radar. However, when a radar is operated under field conditions the performance usually deteriorates even more than that can be accounted for by the above losses. ***To minimize field degradation*** Radars should be designed with built-in automatic performance-monitoring equipment. Careful observation of performance-monitoring instruments and timely preventative maintenance will ***minimize field degradation.***

There are many causes of loss and inefficiency in a Radar. Although each of them may be small, the sum total can result in a significant reduction in radar performance. It is important to understand the origins of these losses, not only for better predictions of radar range, but also for the purpose of keeping them to a minimum by careful radar design.

Important formulae:

- **Range of a Radar $R = cT_R / 2$**
- **Maximum unambiguous Range $R_{unambig.} = C / 2f_p$**
- **Basic Radar equation : $R_{max} = [(Pt .G. A_e. \sigma)/ (4\pi)^2. S_{min}]^{1/4}$**
- **Gain of an antenna $G = 4\pi A_e / \lambda^2$**
- **Noise figure of a receiver $F_n = N_o / kT_o B_n G_a$**
  Also given by

$$F_n = \frac{S_i/N_i}{S_o/N_o}$$

- **If the minimum detectable signal S_{min} is that value of S_i corresponding to the minimum ratio ofoutput (IF) signal-to-noise ratio $(S_o/N_o)_{min}$ necessary for detection, then**

$$S_{min} = kT_0 B_n F_n \left(\frac{S_o}{N_o}\right)_{min}$$

- **And modified Maximum possible range in terms of the IF amplifier output signalto noise ratio and noise figure**

$$R_{max}^4 = \frac{P_t G A_e \sigma}{(4\pi)^2 k T_0 B_n F_n (S_o/N_o)_{min}}$$

- **Relationships between Probability of false alarm,Probability of detection and S/N :**
  - **$P_{fa} = 1/ T_{fa} . B$**
  - **$P_{FA} = Exp(-V_T^2/2\psi_0)$.**
  - $$T_{fa} = \frac{1}{B_{IF}} \exp \frac{V_T^2}{2\psi_0}$$
  - ***Albersheim*** empirical relationship between **S/N, P_{FA}** and **P_d**:

S/N = A + 0.12AB + 1.7 B

Where $A = \ln [0.62/ P_{FA}]$ and $B = \ln [P_d /(1- P_d)]$ •

The radar equation with *n* pulses integrated :

$$R_{max}^4 = \frac{P_t G A_e \sigma}{(4\pi)^2 k T_0 B_n F_n (S/N)_n}$$

- **Integration efficiency :**

$$E_i(n) = \frac{(S/N)_1}{n(S/N)_n}$$

- **Final Radar equation including Integration efficiencywith n pulses integrated:**

$$R_{max}^4 = \frac{P_t G A_e \sigma n E_i(n)}{(4\pi)^2 k T_0 B_n F_n (S/N)_1}$$

Illustrative Problems:

Example1: A certain Radar has PRF of 1250 pulses per second. What is the maximum unambiguous range? Max. Unambiguous Range is given by

$$R_{unambig.} = C/2f_p$$

$$R_{unambig.} = 3x10^8/2x1250 \text{ mtrs} = 120X10^3 \text{mts} = 120\text{Kms}$$

Example 2: A ship board radar has 0.9 micro sec transmitted pulse width. Two small boats in the same direction are separated in range by 150 mts. Will the radar detect the two boats as two targets?

***Radar Range Resolution: The range resolution of a Radar is its ability to distinguish two closely spaced targets along the same line of sight (LOS). The Range resolution is a function of the pulse length, where the pulse length L_P= c x τ/2*(Two way range corresponding to the pulse width)**

Radar Range resolution = $3x10^8 x0.9x10^{-6}/2$ = 135 mtrs.

Since the boats are at 150 Mts. apart, which is greater than the range Resolution of 135mtrs., the radar can detect the 2 boats as 2 separate targets.

Example 3: A Pulse Radar transmits a peak power of 1 Mega Watt. It has a PRT equal to 1000 micro sec and the transmitted pulse width is 1 micro sec. Calculate (i)Maximum unambiguous range (ii) Average Power (iii)Duty Cycle (iv) Energy transmitted & (v)Bandwidth

(i) Maximum unambiguous range = $c.T_P/2 = 3x10^8 x1000x10^{-6}/2 = 150x10^3$ mtrs = 150 Kms

(ii) Average Power = P_P x τ / T_P= $1x10^6$ $1x10^{-6}$ /$1000x10^{-6}$= 1000 watts = 1kw

(iii)Duty Cycle = τ / T_P= $1x10^{-6}$ /$1000x10^{-6}$= 0.001

(iv) Energy transmitted = P_P x τ (Peak power xTime) = $1x10^6$ $x1x10^{-6}$= 1

Joule (v)Bandwidth = 1/τ = $1/10^{-6}$= 1Mhz

Example4: The Bandwidth of I.F. Amplifier in a Radar Receiver is 1 MHZ. If the Threshold to noise ratio is 12.8 dB determine the False Alarm Time.

T_{fa} = False Alarm Time T_{fa}= [$1/B_{IF}$] Exp $\mathbf{V_T^2/2\psi_0}$ where B_{IF} = 1X 10^6 HZ

Threshold to Noise Ratio = 12.8 dB

i.e. 10 $Log_{10}[V_T^2/2\psi_0]$ **= 12.8db**

$\therefore V_T^2/2\psi_0$= $Antilog_{10}$ [12.8/10] = 19.05

$\therefore T_{fa}$= 1/(1X 10^6) Exp 19.05 = 187633284/10^6= 187.6 Seconds

Example5: The probability density of the envelope of the noise voltage output is given by the Rayleigh probability-density function

$$p(R) = \frac{R}{\psi_0} \exp\left(-\frac{R^2}{2\psi_0}\right)$$

where R is the amplitude of the envelope of the filter output for $R \geq 0$. If P_{fa} needed is $\leq 10^{-5}$. Determine the Threshold Level.

The probability of false alarm P_{FA} in terms of the threshold voltage level is given by :

$$P_{FA}= Exp(-V_T^2/2\psi_0) = 10^{-5}$$

Taking logarithms on both the sides we get

$$- 5 \, Log_e 10 = (-V_T^2/2\psi_0)$$

$$5 \times 2.3026 = (V_T^2/2\psi_0)$$

$$V_T^2 = 11.5 \times 2 \, \psi_0$$

$$V_T = \sqrt{23} \times \sqrt{\psi_0} = 4.8 \times \sqrt{\psi_0}$$

Example 6: The bandwidth of an IF amplifier is 1 MHz and the average false-alarm time that could be tolerated is 15 min. Find the probability of a false alarm.
The relationship between average false-alarm time T_{FA}, probability of a false alarm P_{FA} and the IF bandwidth B is given by :

$$P_{fa} = 1/ T_{fa} . B$$

Substituting **B** = 1 MHz ie 10^6 and $\mathbf{T_{fa}}$=15 mnts. i.e. 900 secs. we get PFA=1.11 X 10^{-9}

Example 7: What is the ratio of threshold voltage to the rms value of the noise voltage necessary to achieve this false-alarm time?
This is found out using the relationship $\mathbf{P_{FA} = Exp (-V_T^2/2\psi_0)}$
from which the ratio of Threshold voltage to rms value of the noise voltage is given by

$$V_T/\sqrt{\psi_0} = \sqrt{2 \ln (1/P_{fa})} = \sqrt{2 \ln 9X10^8} = 6.45 = 16.2dB$$

Example 8: Typical parameters for a ground-based search radar are : 1. Pulse repetition frequency :300 Hz, 2. Beam width : 1.5^0, and 3. Antenna scan rate: 5 rpm (30^0/s). Find out the number of pulses returned from a point target on each scan.
Solution : The number of pulses retuned from a point target on each scan n_B is given by:

$$n_B = \theta_B . f_P / \theta'_S = \theta_B . f_P / 6\ \omega_m$$

Substituting the above values we get : $n_B = 1.5 \times 300 / 30 = 15$

Questions from Previous Year Examinations:

1.(a) Derive Radar range equation in terms of MDS (minimum detectable signal) (b) What is maximum unambiguous range? How is it related with PRF?

2.(a) Explain the various system losses in a Radar (b) The bandwidth of The IF amplifier in a Radar is 1 Mhz and the threshold noise ratio is 13 db. Determine the false alarm time.

3.(a) Explain the basic principles of Radar and discuss about various parameters which improve the performance of the Radar (b) Discuss about Radar frequencies and list out the Applications of Radars.

4.(a) In a Radar receiver the mean noise voltage is 80 mv and the IF BW is 1 Mhz. If thetolerable false alarm time is 25 mnts., calculate the threshold voltage level and the probability of false alarm. (b) Bring out the advantages of Integration of Radar pulses.

5 (a) Discuss about the factors that influence the prediction of Radar range. (b) Define noise bandwidth of a radar receiver. How does it differ from 3-dB band width? Obtain the expression for minimum detectable signal in terms of noise bandwidth, noise figure and other relevant parameters. [8+8]

6. (a) Write the simplified version of radar range equation and explain how this equation does not adequately describe the performance of practical radar? (b) What are the specific bands assigned by the ITU for the Radar? What the corresponding frequencies? [8+8]

7. (a) Explain how the Radar is used to measure the range of a target? (b) Draw the block diagram of the pulse radar and explain the function of each block. [8+8]

8. (a) A low power, short range radar is solid-state throughout, including a low-noise RF amplifier which gives it an overall noise figure of 4.77dB. If the antenna diameter is 1m, the IF bandwidth is 500kHz, the operating frequency is 8 GHz and the radar set is supposed to be capable of detecting targets of $5m^2$ cross sectional area at a maximum distance of 12 km, what must be the peak transmitted pulse power? (b) The average false alarm time is a more significant parameter than the false alarm probability. Give the reasons. (c) Why post detection integration is not as efficient as pre-detection integration of radar pulses? [8+4+4]

9. (a) Explain the functioning and characteristics of PPI displayand A-Scope. [8]

10. (a) Explain how the Radar is used to measure the direction and position of target? (b) What are the peak power and duty cycle of a radar whose average transmitter power is 200W, pulse width of 1μs and a pulse repetition frequency of 1000Hz? [8+8]

11. (a) Explain how a threshold level is selected in threshold detection? (b) How to find the number of pulses that returned from a point target as the radar antenna scans through its beam width? (c) Why most of the radar receivers are considered as envelop detectors while calculating the SNR? [6+4+6]

12. (a) Obtain the SNR at the output of IF amplifier of Radar Receiver for a specified probability of detection without exceeding a specified probability of false alarm. (b) Explain how system losses will affect on the Radar Range? [8+8]

13. (a) What are the different range of frequencies that a radar can operate and give their applications? (b) What are the basic functions of radar? In indicating the position of a target, what is the difference between azimuth and elevation? [8+8]

14. (a) Describe how pulse repetition frequency of a Radar system controls the range of it's detection? (b) Explain how the Transmitted power affects the range. *8+8+

15. (a) Draw the block diagram of a pulsed radar and explain it's operation. (b) What are the desirable pulse characteristics and the factors that govern them in a Radar system? [10+6]

16. (a) Explain the radar cross section of the sphere. (b) Discuss in brief about pulse repetition frequency and range ambiguities.

17. (a) Define Range resolution and explain the parameters which affect the range resolution. (b)Distinguish between Monostatic and Bistatic Radars (c) Explain RCS of target. [6+5+4]

UNIT-2

CW AND FREQUENCY MODULATED RADAR

- **Doppler effect**
- **CW radar block diagram**
- **Isolation between Transmitter and receiver**
- **Nonzero IF receiver**
- **Receiver Bandwidth requirements**
- **Applications of CW Radar**
- **Illustrative problems**

FM-CW RADAR

- **Introduction**
- **Range and Doppler Measurement**
- **Block Diagram and characteristics (Approaching and Receding targets)**
- **FM-CW Altimeter**
- **Multiple frequency CW Radar**

    - **Important Formulae**
    - **Illustrative Problems**
    - **Questions from Previous Year Examinations**

CW AND FREQUENCY MODULATED RADAR

Doppler Effect:

A technique for separating the received signal from the transmitted signal when there is relative motion between radar and target is based on recognizing the change in the echo-signal frequency caused by the Doppler effect.

It is well known in the fields of optics and acoustics that if either the source of oscillation or the observer of the oscillation is in motion, an apparent shift in frequency will result. This is the ***Doppler effect*** and is the basis of CW radar. If **R** is the distance from the radar to target, the total number of wavelengths **λ** contained in the two-way path between the radar and the target are ***2R/λ.*** The distance **R** and the wavelength ***λ*** are assumed to be measured in the same units.

Since one wavelength corresponds to an phase angle excursion of **2π**radians, the total phase angle excursion **Ø** made by the electromagnetic wave during its transit to and from the target is **4πR/λ** radians. If the target is in motion, **R** and the phase **Ø** are continually changing. A change in **Ø** with respect to time is equal to frequency. This is the Doppler angular frequency ω_d and is given by:

$$\omega_d = 2\pi f_d = d\O/dt = d(4\pi R/\lambda)/dt = (4\pi/\lambda) \cdot dR/dt = (4\pi/\lambda) . V_r = 4\pi . V_r / \lambda$$

where f_d is the Doppler frequency shift in Hz, and V_r = relative velocity of the target with respect to the Radar. The Doppler frequency shift f_d is given by

$$f_d = 2V_r / \lambda = 2V_r f_0 / c$$

where f_0 is the transmitted frequency and **c** is the velocity of propagation of the electromagnetic waves (same as that of light) = 3×10^8 m/s. If f_d is **in** hertz. V_r in knots, and **λ** in meters then the Doppler frequency f_d is given by

$$f_d = 1.03\, V_r / \lambda$$

A plot of this equation is shown in the figure below

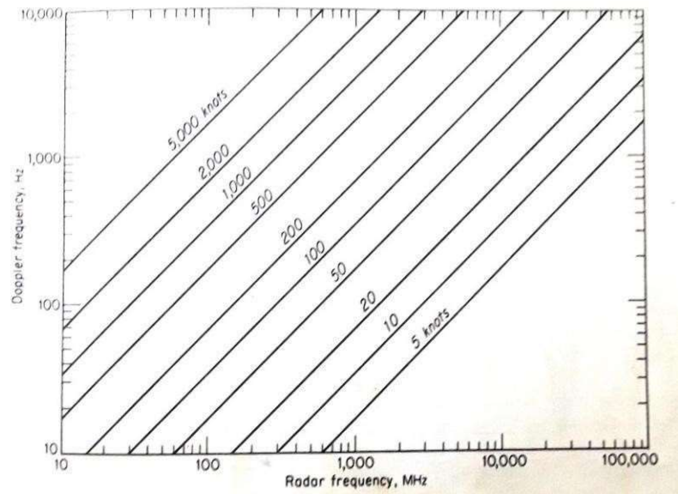


Figure: Doppler frequency f_d as a function of radar frequency and target relative velocity.

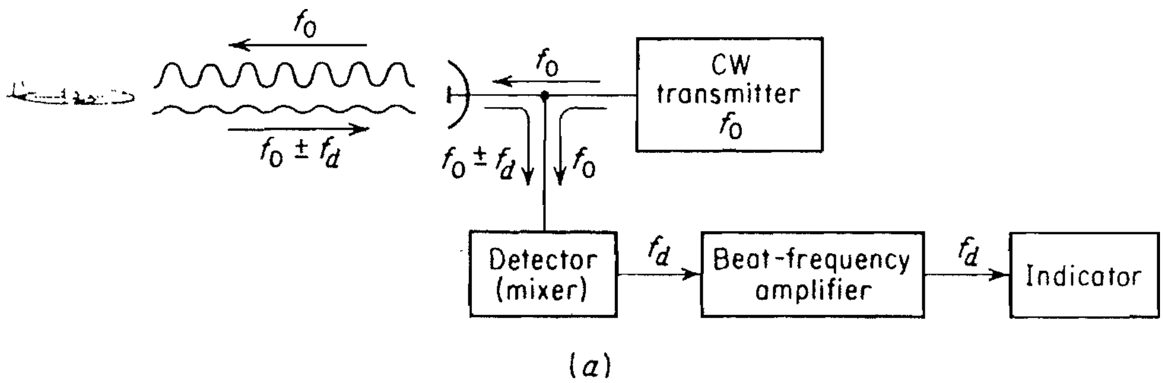
The relative velocity may be written as **V_r= V .cos θ** where **V** is the target speed and **θ** is angle made by the target trajectory and the line joining radar and target. When **θ=0** the Doppler frequency is maximum. The Doppler is zero when the trajectory is perpendicular to the radar line of sight **(θ= 90^0).**

The CW radar is of interest not only because of its many applications, but its study also serves as a means for better understanding the nature and use of the Doppler information contained in the echo signal, whether in a CW or a pulse radar (MTl) application. In addition to allowing the received signal to be separated from the transmitted signal, the CW radar provides a measurement of relative velocity which may be used to distinguish moving targets from stationary objects or clutter.

CW radar:

Consider the simple CW radar as illustrated by the block diagram of Figure below. **T**he transmitter generates a continuous (unmodulated) oscillation of frequency **fo**, which is radiated by the antenna. A portion of the radiated energy is intercepted by the target and is scattered, some of it in the direction of the radar, where it is collected by the receiving antenna. If the target is in motion with a velocity **V_r** relative to the radar, the received signal will be shifted in frequency from the transmitted frequency **f_o** by an amount **+/-f_d** as given by the equation : **$f_d = 2V_r / \lambda = 2 V_r f_0 / c$.** The plus sign associated with the Doppler frequency applies if the distance between target and radar is decreasing (approaching target) that is, when the received signal frequency is greater than the transmitted signal frequency. The minus sign applies if the distance is increasing (receding target). The received echo signal at a frequency **f_o +/- f_d** enters the radar via the antenna and is heterodyned in the detector (mixer) with a portion of the transmitter signal **f_o** to produce a Doppler beat note of frequency **f_d.** The sign of **f_d** is lost in this process.

The purpose of the Doppler amplifier is to eliminate echoes from stationary targets and to amplify the Doppler echo signal to a level where it can operate an indicating device. It's frequency response characteristic is shown in the figure (b) below. The low-frequency cutoff must be high enough to reject the d-c component caused by stationary targets, but yet it must be low enough to pass the smallest Doppler frequency expected. Sometimes both conditions cannot be met simultaneously and a compromise is necessary. The upper cutoff frequency is selected to pass the highest Doppler frequency expected.



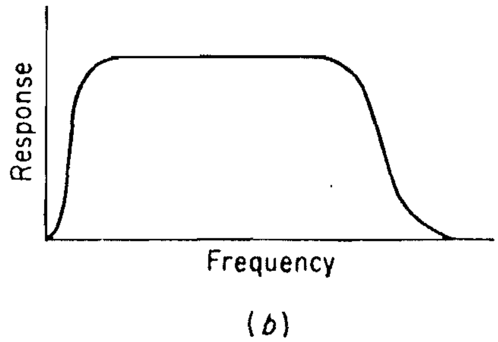


Figure : Simple CW radar block diagram (b) response characteristic of beat-frequency amplifier

Isolation between transmitter and receiver:

Isolation between transmitter and receiver is an important aspect to be studied and addressed in simple CW radars where a single antenna serves the purpose of both transmission and reception as described above. The related important aspects are explained below.

- In principle, a single antenna may be employed since the necessary isolation between the transmitted and the received signals is achieved via separation in frequency as a result of the Doppler Effect. In practice, it is not possible to eliminate completely the transmitter leakage. However, transmitter leakage is neither always undesirable. A moderate amount of leakage entering the receiver along with the echo signal supplies the reference necessary for the detection of the Doppler frequency shift. If a leakage signal of sufficient magnitude were not present, a sample of the transmitted signal has to be deliberately introduced into the receiver to provide the necessary reference frequency.
- There are two practical effects which limit the amount of transmitter leakage power which can be tolerated at the receiver. These are:
  (1) The maximum amount of power the receiver input circuitry can withstand before it is physically damaged or its sensitivity reduced (burnout) and

(2) The amount of transmitter noise due to hum, microphonics, stray pick-up &instability which enters the receiver from the transmitter and affects the receiver sensitivity.

Hence additional isolation is usually required between the transmitter and the receiver if the sensitivity is not to be degraded either by burnout or by excessive noise. The amount of isolation required depends on the transmitter power and the accompanying transmitter noise as well as the ruggedness and the sensitivity of the receiver. For example, If the safe value of power which might be applied to a receiver is 10 mW and if the transmitter power is 1 kW, the isolation between transmitter and receiver must be at least 50 dB.

- The amount of isolation needed in a long-range CW radar is more often determined by the noise that accompanies the transmitter leakage signal rather than by any damage caused by high power. For example, suppose the isolation between the transmitter and receiver is such that **10** mW of leakage signal appeared at the receiver. If the minimum detectable signal is 10^{-l3}watt (100 dB below 1mW), the transmitter noise must be at least **110** dB (preferably 120 or 130 dB) below the transmitted carrier.
- The transmitter noise of concern in Doppler radar includes those noise components that lie within the same frequency range as the Doppler frequencies. If complete elimination of the direct leakage signal at the receiver could be achieved, it might not entirely solve the isolation problem since echoes from nearby fixed targets (clutter) can also contain the noise components of the transmitted signal.
- The receiver of a pulsed radar is isolated and protected from the damaging effects of the transmitted pulse by the duplexer, which short-circuits the receiver input during the transmission period. Turning off the receiver during transmission with a duplexer is not possible in a CW radar since the transmitter is operated continuously.
- In CW Radars Isolation between transmitter and receiver might be obtained with a single antenna by using a hybrid junction, circulator, turnstile junction, or with separate polarizations. Separate antennas for transmitting and receiving might also be used.
  - The amount of isolation which can be readily achieved between the arms of practical hybrid junctions such as the magic-T, rat race, or short-slot coupler is of the order of 20 to 30 dB. In some instances, when extreme precision is exercised, an isolation of perhaps 60 dB or more might be achieved. But one limitation of the hybrid junction is the 6-dB loss in overall performance which results from the inherent waste of half the transmitted power and half the received signal power. Both the loss in performance and the difficulty in obtaining large isolations have limited the application of the hybrid junction to short-range radars.
  - Ferrite isolation devices such as the circulator do not suffer the 6-dB loss inherent in the hybrid junction. Practical devices have isolation of the order of 20 to 50 dB. Turnstile junctions achieve isolations as high as 40 to 60 dB.
  - The use of orthogonal polarizations for transmitting and receiving is limited to short range radars because of the relatively small amount of isolation that can be obtained.
- An important factor which limits the use of isolation devices with a common antenna is the reflections produced in the transmission line by the antenna. The reflection

coefficient from a mismatched antenna with a voltage standing-wave ratio **ς** is |**ρ**|=**(ς-l) / (ς+1).** Therefore, if an isolation of **20 dB** is to be obtained, the VSWR must be less than **1.22.** If **40 dB** of isolation is required, the VSWR must be less than **1.02.**

- The largest isolations are obtained with two antennas: one for transmission, the other for reception-physically separated from one another. Isolations of the order of 80 dB or more are possible with high-gain antennas. The more directive the antenna beam and the greater the spacing between antennas, the greater will be the isolation. **A** common radome enclosing the two antennas should be avoided since it limits the amount of isolation that can be achieved.
- Additional isolation can be obtained by properly introducing a controlled sample of the transmitted signal directly into the receiver. The phase and amplitude of this "buck-off" signal are adjusted to cancel the portion of the transmitter signal that leaks into the receiver. An additional 10 dB of isolation might be obtained.
- The transmitter signal is never a pure CW waveform. Minute variations in amplitude (AM) and phase (FM) can result in sideband components that fall within the Doppler frequency band. These can generate false targets or mask the desired signals. Therefore, both AM and FM modulations can result in undesired sidebands. AM sidebands are typically 120 dB below the carrier, as measured in a 1 kHz band, and are relatively constant across the usual Doppler spectrum of interest. The normal antenna isolation plus "feed through nulling" usually reduces the AM components below receiver noise in moderate power radars. FM sidebands are usually significantly greater than AM, but decrease with increasing offset from the carrier. These can be avoided by stabilizing the output frequency of the CW transmitter and by feeding back the extracted FM noise components so as to reduce the original frequency deviation.

Intermediate-frequency receiver:

Limitation of Zero IF receiver:

The receiver in the simple CW radar *shown earlier* is in some respects analogous to a super heterodyne receiver. Receivers of this type are called homodyne receivers, or super heterodyne receivers with zero IF. The function of the local oscillator is replaced by the leakage signal from the transmitter. Such a receiver is simpler than the one with a more conventional intermediate frequency since no IF amplifier or local oscillator is required. However, this simpler receiver is not very sensitive because of increased noise at the lower intermediate frequencies caused by flicker effect. Flicker-effect noise occurs in semiconductor devices such as diode detectors and cathodes of vacuum tubes. The noise power produced by the flicker effect varies as $1/f^{\alpha}$ where α is approximately unity. This is in contrast to shot noise or thermal noise, which is independent of frequency. Thus, at the lower range of frequencies (audio or video region), where the Doppler frequencies usually are found, the detector of the CW receiver can introduce a considerable amount of flicker noise, resulting in reduced receiver sensitivity. For short-range, low-power, applications this decrease in sensitivity might be tolerated since it can be compensated by a modest increase in antenna aperture and/or additional transmitter power.

But for maximum efficiency with CW radar, the reduction in sensitivity caused by the simple Doppler receiver with zero IF cannot be tolerated.

Non zero IF Receiver:

The effects of flicker noise are overcome in the normal super heterodyne receiver by using an intermediate frequency high enough to make the flicker noise small compared with the normal receiver noise. This results from the inverse frequency dependence of flicker noise. Figure below shows the block diagram of a CW radar whose receiver operates with a nonzero IF. Separate antennas are shown for transmission and reception. Instead of the usual local oscillator found in the conventional super heterodyne receiver, the local oscillator (or reference signal) is derived in the receiver from a portion of the transmitted signal mixed with a locally generated signal of frequency equal to that of the receiver IF. Since the output of the mixer consists of two sidebands on either side of the carrier plus higher harmonics, a narrowband filter selects one of the sidebands as the reference signal. The improvement in receiver sensitivity with an intermediate-frequency super heterodyne might be as much as 30 dB over the simple zero IF receiver discussed earlier.

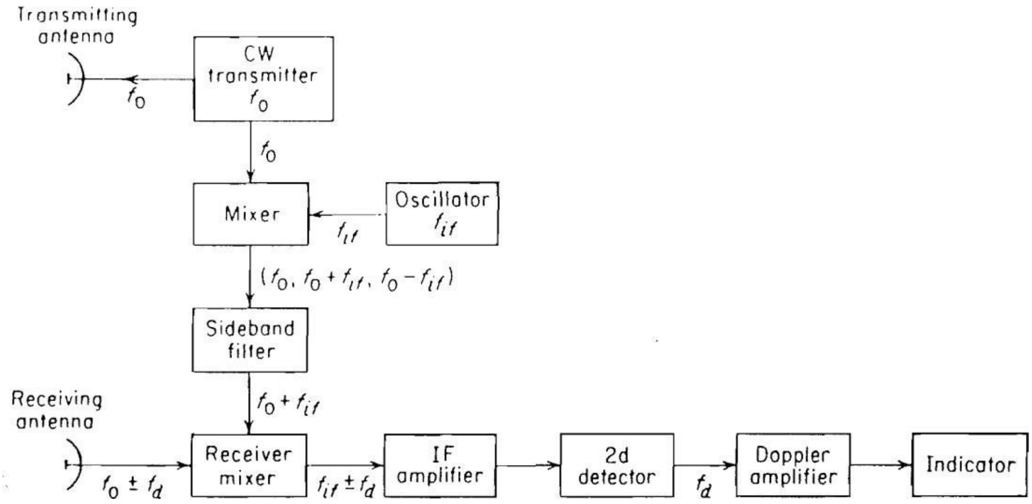


Figure: Block diagram of a CW Doppler radar with nonzero IF receiver, also called sideband super heterodyne Receiver.

Receiver bandwidth requirements:

One of the requirements of the Doppler-frequency amplifier in the simple CW radar (Zero IF) or the IF amplifier of the sideband super heterodyne (Non Zero IF) is that it has to be wide enough to pass the expected range of Doppler frequencies. In most cases of practical interest the expected range of Doppler frequencies will be much wider than the frequency spectrum occupied by the signal energy. Consequently, the use of a wideband amplifier covering the expected Doppler range will result in an increase in noise and a lowering of the receiver sensitivity. If the frequency of the Doppler-shifted echo signal were known beforehand, narrowband filter-that is just wide enough to reduce the excess noise without eliminating a significant amount of signal energy might be used. If the waveforms of the echo signal are known, as well as its carrier frequency, a matched filter could also be considered.

Several factors tend to spread the CW signal energy over a finite frequency band. These must be known if an approximation to the bandwidth required for the narrowband Doppler filter is to be obtained.

If the received waveform were a sine wave of infinite duration, its frequency spectrum would be a delta function as shown in the figure (a) below and the receiver bandwidth would be infinitesimal. But a sine wave of infinite duration and an infinitesimal bandwidth cannot occur in nature. The more normal situation is an echo signal which is a sine wave of finite duration. The frequency spectrum of a finite-duration sine wave has a shape of the form **$[\sin\pi(f-f_0)\delta]/\pi(f-f_0)]$** where **$f_0$** and **$\delta$** are the frequency and duration of the sine wave, respectively, and **f** is the frequency variable over which the spectrum is plotted (Fig b).

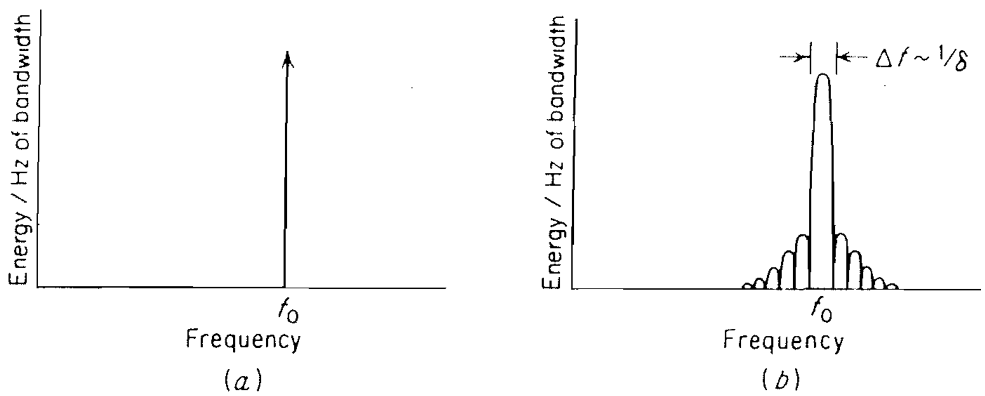


Figure: Frequency spectrum of CW oscillation of (a) infinite duration and (b) finite duration

Note that this is the same as the spectrum of a pulse of sine wave, the only difference being the relative value of the duration **δ.** In many instances, the echo is not a pure sine wave of finite duration but is perturbed by fluctuations in cross section, target accelerations, scanning fluctuations, etc., which tend to broaden the bandwidth still further. Some of these spectrum broadening-effects are considered below.

Causes for Spectrum broadening:

- **Spread due to finite time on target:** Assume a CW radar with an antenna beam width of **θ_B** deg. scanning at the rate of **θ'_S** deg/s. The time on target (duration of the received signal) is **$\delta = \theta_B/\theta_S$** sec. Thus, the signal is of finite duration and the bandwidth of the receiver must be of the order of the reciprocal of the time on target **(θ'_S/θ_B).** Although this is not an exact relation, it is a good enough approximation for purposes of the present discussion. If the antenna beam width is 2^0 and the scanning rate is 36^0/s (6 rpm), the spread in the spectrum of the received signal due to the finite time on target would be equal to 18 Hz, independent of the transmitted frequency.

- In addition to the spread of the received signal spectrum caused by the finite time on target, the spectrum gets widened due to target cross section fluctuations. The fluctuations widen the spectrum by modulating the echo signal. The echo signal from a propeller-driven aircraft can also contain modulation components at a frequency proportional to the propeller rotation. The frequency range of propeller modulation depends upon the shaft-rotation speed and the number of propeller blades. It is usually in the vicinity of 50 to 60 Hz for World War 2 aircraft engines. This could be a potential source of difficulty in a **CW** radar since it might mask the target's Doppler signal or it might cause an erroneous measurement of Doppler frequency. In some instances, propeller modulation can be of advantage. It might permit the detection of propeller-driven aircraft passing on a tangential trajectory, even though the Doppler frequency shift is zero.
- The rotating blades of a helicopter and the compressor stages of a jet engine can also result in a modulation of the echo and a widening of the spectrum that can degrade the performance of a **CW** Doppler radar.

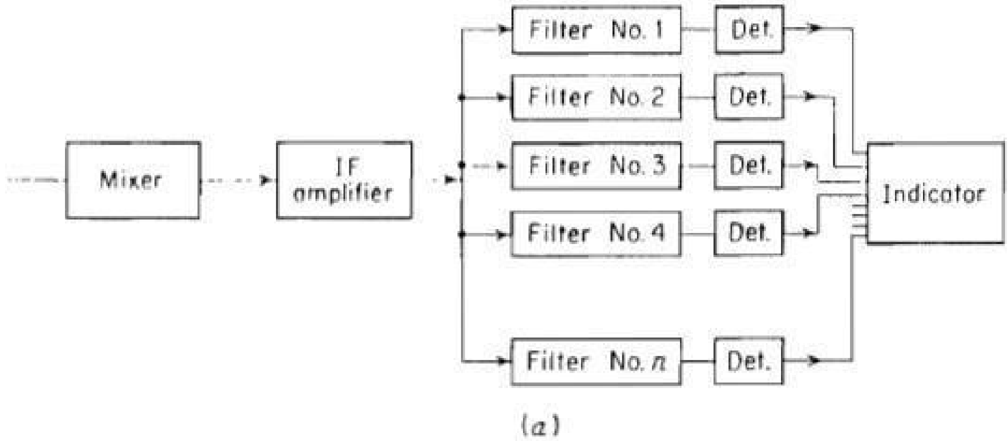
- **If** the target's relative velocity is not constant, a further widening of the received signal spectrum occurs. If $\mathbf{a_r}$ is the acceleration of the target with respect to the radar, the signal will occupy a bandwidth

$$\Delta f_d = \left(\frac{2a_r}{\lambda}\right)^{1/2}$$

If, for example, $\mathbf{a_r}$ is twice the acceleration due to gravity, the receiver bandwidth is approximately 20 Hz when the Radar wavelength is 10 cm.

When the Doppler-shifted echo signal is known to lie somewhere within a relatively wideband of frequencies, a bank of narrowband filters as shown below spaced throughout the frequency range permits a measurement of frequency and improves the signal-to-noise ratio.

- The bandwidth of each individual filter should be wide enough to accept the signal energy, but not so wide as to introduce more noise. The center frequencies of the filters are staggered to cover the entire range of Doppler frequencies.



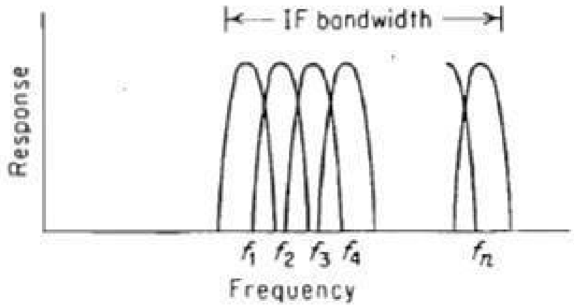


Figure: (a) Block diagram of IF Doppler filter bank (b) frequency-response characteristic of Doppler filter bank.

- A bank of narrowband filters may be used after the detector in the video of the simple CW radar instead of in the IF. The improvement in signal-to-noise ratio with a video filter bank is not as good as can be obtained with an IF filter bank, but the ability to measure the magnitude of Doppler frequency is still preserved. Because of fold over, a frequency which lies to one side of the IF carrier appears, after detection, at the same video frequency as one which lies an equal amount on the other side of the IF. Therefore the sign of the Doppler shift is lost with a video filter bank, and it cannot be directly determined whether the Doppler frequency corresponds to an approaching or to a receding target. (The sign of the Doppler maybe determined in the video by other means.) One advantage of the fold over in the video is that only half the number of filters are required than in the IF filter bank.
- A bank of overlapping Doppler filters, whether in the IF or video, increases the complexity of the receiver. When the system requirements permit a time sharing of the Doppler frequency range, the bank of Doppler filters may be replaced by a single narrowband tunable filter which searches in frequency over the band of expected Doppler frequencies until a signal is found.

Applications of CW radar:

- Measurement of the relative velocity of a moving target, as in the police speed monitor or in the rate-of-climb meter for vertical-take-off aircraft.
- Control of traffic lights, regulation of tollbooths, vehicle counting.
- As a sensor in antilock braking systems, and for collision avoidance.
- In railways, as a speedometer to replace the conventional axle-driven tachometer. In such an application it would be unaffected by errors caused by wheel slip on accelerating or wheel slide when braking.
- Monitoring the docking speed of large ships.
- Measurement of the velocity of missiles, ammunition, and baseballs.
-

Advantages and disadvantages of CW Radars:

- The principal advantage of CW Doppler radar over the other (non radar) methods of measuring speed is that there need not be any physical contact with the object whose speed is being measured. In industry this is used to measure turbine-blade vibration, the peripheral speed of grinding wheels, and the monitoring of vibrations in the cables of suspension bridges.
- Most of the above applications can be satisfied with a simple, solid-state CW source with powers in tens of milli watts
- High-power CW radars for the detection of aircraft and other targets have been developed and have been used in such systems as the Hawk missile systems. (Shown below)
- The difficulty of eliminating the leakage of the transmitter signals into the receiver has limited the utility of unmodulated CW radar for many long-range applications.
- The CW radar, when used for short or moderate ranges, is characterized by simpler equipment than a pulse radar. The amount of power that .can be used with a CW radar is dependent on the isolation that can be achieved between the transmitter and receiver since the transmitter noise that finds its way into the receiver limits the receiver sensitivity. (The pulse radar has no similar limitation to its maximum range because the transmitter is not operative when the receiver is turned on.)
- Major disadvantage of the simple CW radar is its inability to obtain a measurement of range. This limitation can be overcome by modulating the CW carrier, as in the frequency-modulated radar.
- Some anti-air-warfare guided missile systems employ semi active homing guidance in which a receiver in the missile receives energy from the target, the energy having been transmitted from an “illuminator" external to the missile. The illuminator will be at the launch platform. CW illumination has been used in many successful systems. An example is the Hawk tracking illuminator shown in the figure below. It is tracking radar as well as an illuminator since it must be able to follow the target as it travels through space.

CW radar allows operation in the presence of clutter and has been well suited for low altitude missile defense systems. A block diagram of a CW tracking illuminator is shown in the figure above. Note that following the wide-band Doppler amplifier is a speed ***gate,*** which is a narrow-band tracking filter that acquires the targets Doppler and tracks its changing Doppler frequency shift.

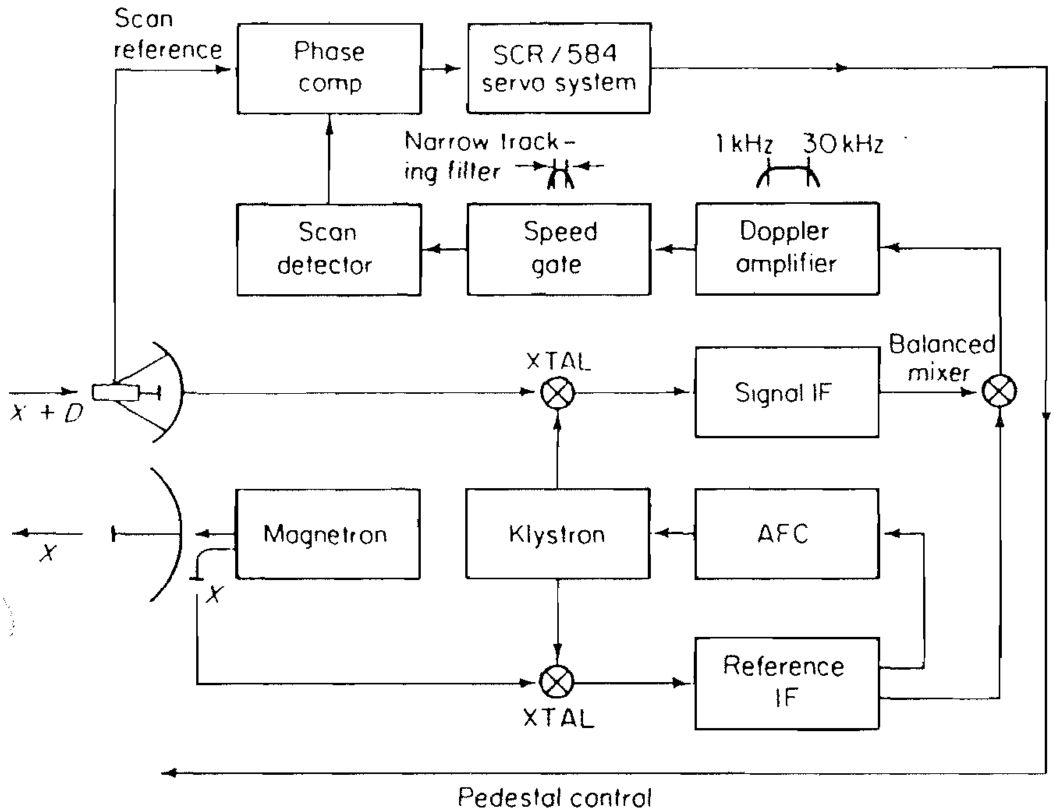


Figure: Block diagram of a CW tracking-illuminator

FM-CW RADAR

Introduction:

The inability of the simple CW radar to measure range is mainly due to the lack of a Timing mark. The timing mark permits the time of transmission and the time of return to be recognized but it increases the spectrum of the transmitted waveform. The sharper or more distinct the mark, the more accurate the measurement of the transit time. But the more distinct the timing mark, the broader will be the transmitted spectrum. This follows from the properties of the Fourier transform. Therefore a finite spectrum of necessity must be transmitted if transit time or range is to be measured. The spectrum of a CW transmission can be broadened by the application of a modulation - amplitude, frequency, or phase. An example of an amplitude modulation is the pulse radar. A widely used technique to insert a timing mark is to frequency-modulate the carrier. The timing mark is the changing frequency. The transit time is proportional to the difference in frequency between the transmitter signal and the echo signal. The greater the transmitter frequency deviation in a given time interval, the more accurate is the measurement of the transit time but the transmitted spectrum also becomes larger.

Range and Doppler measurement:

In the frequency-modulated CW radar (abbreviated FM-CW), the transmitter frequency is changed as a function of time in a known manner. Assume that the transmitter frequency increases linearly with time, as shown by the solid line in the figure below.

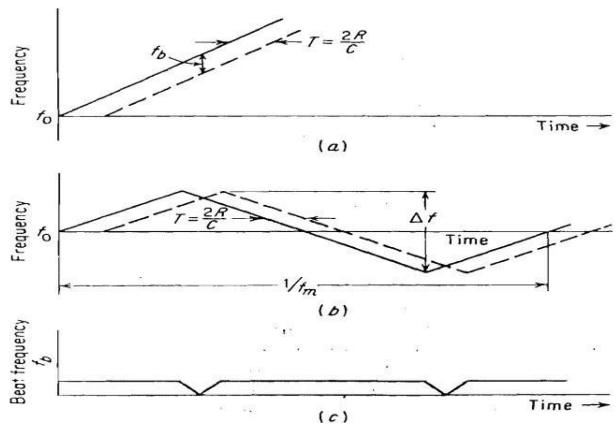


Figure: Frequency-time relation-ships in FM-CW radar. Solid curve represents transmitted signal; dashed curve represents echo. (a) Linear frequency modulation (b) triangular frequency modulation(c) beat note of (b).

If there is a reflecting object at a distance **R**, the echo signal will return after a time **T = 2R/c**. The dashed line in the figure represents the echo signal. When the echo signal is heterodyned with a portion of the transmitter signal in a nonlinear element such as a diode, a beat note f_b will be produced. If there is no Doppler frequency shift, the beat note (difference frequency) is a measure of the target's range and $f_b = f_r$ where f_r, is the beat frequency only due to the target's range. If the rate of change of the carrier frequency is **f_0(dot)** then the beat frequency is given by:

$$f_r = \dot{f}_0 T = \frac{2R}{c} \dot{f}_0$$

In any practical CW radar, the frequency cannot be continually changed in one direction only. Periodicity in the modulation is necessary, as in the triangular frequency-modulation waveform