UNIT 11 MICROWAVE COMMUNICATION

Structure

11.1 INTRODUCTION

You have already learnt in the first block of this course that the electromagnetic waves with wavelengths from about 1m to 1 cm (300 MHz – 300 GHz) are categorised as microwaves. In Unit 3 we have taken an overview of electromagnetic wave propagation through various media and you learnt that for microwave frequencies, hollow metallic conduits called *waveguides* are used. Microwave can penetrate through the upper atmospheric layers and hence can reach the geostationary satellites placed beyond the earth's atmosphere. This makes microwaves a good candidate for broad area communication.

The generation and detection of microwaves needs special devices different from those used in medium and short wave communication. Due to the electromagnetic field nature of microwave propagation, the signal diversion, splitting etc. also require some special devices, which are quite different from simple '*T*' joints of wires used at lower frequencies.

Use of microwaves for communication has been an inevitable part of long and short distance communication. These systems can be classified into:

(a) Analogue Microwave Communication system and
(b) Digital Microwave Communication system.

Presently major operational terrestrial and satellite communication systems use analog FM modulation and analog multiplexing techniques. However they are shifting to digital modulation and multiplexing techniques in a big way.

In this Unit we will discuss the different aspects related to microwave communication. In Sec. 11.2 we recapitulate the modes of signal propagation through earth's atmosphere, which you have already learnt in Unit 1 and 3 of this course. Some characteristics related to microwave antenna will be discussed in Sec. 11.3.
In Sec. 11.4 you will learn about different microwave components including sources, detectors and other components used in a typical microwave system. The historic

application of microwaves has been in Radar systems used in navigation and military applications. In Sec. 11.5 we take brief review of the radar system. Sec. 11.6 is devoted to the discussion about satellite communication.

Objectives

After studying this unit, you should be able to:

- define and calculate space attenuation;
- describe various geometries of antenna used in microwave;
- compare the performance of different antenna feeds;
- describe the working of reflex klystron microwave generator;
- explain the working of tunnel and Gunn diode;
- compare the performance of various microwave detectors;
- enumerate the functions of various microwave components;
- explain the working of radar;
- describe the signal handling in satellite systems; and
- explain the communication related systems on a satellite.

11.2 TERRESTRIAL MICROWAVE PROPAGATION

The telecommunication links can be broadly classified according to the information and the carrier frequencies.

- Telephony: long distance voice transmission using single sideband transmission in the 30 kHz to 100 MHz frequency range.
- Sound broadcasting system: Frequency ranges upto 30 MHz in the most familiar long wave, medium wave and short wave bands of radio receiver. The frequency extends to VHF region (30 – 300 MHz) if frequency modulation is used.
- TV broadcasting uses VHF and UHF bands of frequencies (30 MHz to 3 GHz).
- Satellite communication: Television Broadcasting is now done with the use of satellite i.e. the ground station beams the signal to the space satellite which transmits to a large part of the earth where the signals can be received by individual direct receiver sets. Because of the intervening absorbing atmospheric medium, the frequency range for satellite broadcasting is limited to 3GHz to 30 GHz.

The terrestrial communication system mainly deals with the radiation and propagation of electromagnetic waves from one point to another point on the surface of the earth through the atmospheric medium. You have learnt in Unit 1 that in view of the behaviour of the propagating medium at different frequencies, the radio communication can be classified into the following categories:

1. Ground wave communication;
2. Sky wave or Ionospheric communication;
3. Tropo-scatter communication; and
4. Line-of-sight (LOS) communication.

Out of these, the most common form of microwave terrestrial communication is the line-of-sight communication. This line-of-sight communication system has a range limited by the horizon due to the curvature of the earth. The transmission distance is determined by the height of the antenna above the earth. Normally microwave horizon is more than the optical horizon due to bending or refraction of microwaves beyond the optical horizon. The height of the microwave towers (transmitter and receiver or repeater) should be such that the microwave beam is not obstructed by high rise buildings, mountains or forest trees. These communication systems are very popularly used in multi-channel telephony and TV broadcasting. They operate in the 1 GHz to 10 GHz range of frequencies.

Consider a transmitter radiating a power of P_0 W. There are two receivers at a distance R_1 and R_2 in the same direction from the source. The power density at R_1 is $[P_0/ 4\pi R_1^2]$ and at the distance R_2 is $[P_0/ 4\pi R_2^2]$. The ratio $P_1/P_2 = [R_2/R_1]^2$ is the attenuation of the power density in the distance between R_1 and R_2, and is called the **space attenuation.**

If the medium is microwave absorbing, having power attenuation of α dB km^{-1}, then the net attenuation (A_{Total}) in the space between the two receivers is the addition of space attenuation and absorption in the medium.

i.e. $$A_{Total} = 10 \log [R_2/R_1]^2 + \alpha [R_2 - R_1] \quad \text{dB} \qquad (11.1)$$

$$= 20 \log [R_2/R_1] + \alpha [R_2 - R_1] \quad \text{dB} \qquad (11.2)$$

The earth's atmosphere consists of oxygen and water vapour, which are strong absorbers of microwaves. Oxygen is a magnetic dipole and water molecule is an electric dipole. They both absorb microwaves strongly. The narrow resonance lines of these molecules get broadened by collision broadening in the atmosphere hence they exhibit broad absorption lines as shown in Fig. 11.1. The absorption becomes prominent at frequencies above 10 GHz. The galactic and extragalactic sources of

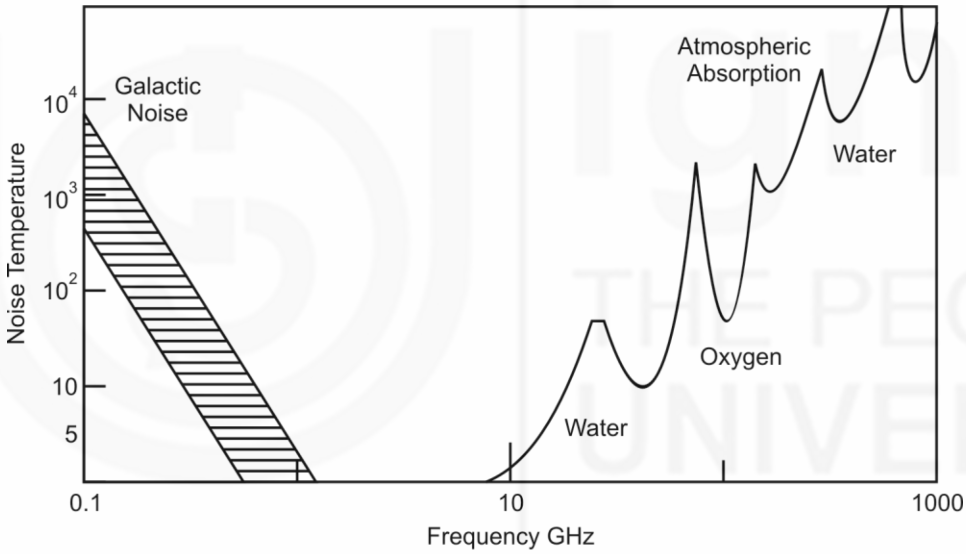

Fig. 11.1: Atmospheric absorption with frequency

radio frequency noise affecting the communication system are very effective up to 1 GHz. The sky is pretty cool at frequencies above 1 GHz. It is therefore apparent that the best region for space communication is from 1 to 10 GHz. At higher frequencies power requirement of the terrestrial transmitter may be increased due to atmospheric losses. In addition, losses due to the rain and snowfall have to be taken into account.

SAQ 1

Spend 2 Min.

Power of a microwave source is P_0=100 W. There are two receivers at distances R_1= 1 km and R_2= 10 km and the attenuation of the absorbing medium is 0.1 dB km^{-1}. Calculate the net loss.

In any microwave communication system the intelligent signals (voice, images or data) are first modulated, multiplexed and then transmitted via terrestrial or satellite communication. In order to increase the distance covered by the signal, repeater stations are installed, which have a receiving antenna, booster amplifiers and a transmitting antenna. At the receiver end the signal received by an antenna is frequency down converted, demultiplexed, decoded and converted into user tangible form. The repeaters can be installed on the earth or sometimes satellites can be used as repeaters, which receive and retransmit the signals.

In all these systems, highly directional antenna plays an important role. In the following section we discuss some antenna characteristics relevant to microwave signal communication.

11.3 MICROWAVE ANTENNA

11.3.1 Horns

Microwaves can be radiated directly from the end of a waveguide in the same way as from the end of an open transmission line as we discussed in Unit 4. The end of the waveguide represents an abrupt transition from the characteristic impedance of the waveguide into that of free space, and the resulting radiation is neither efficient nor very directive. This problem can be tackled by flaring out the end of the waveguide to form a horn-like structure. This provides a gradual transition as the wave passes from the mouth of the horn.

Narrow-mouthed horns with long flare sections produce sharper beams than shorter, wide-mouthed horns. Also, the wider-mouthed horns tend to produce a wavefront with a distinct curvature, which is undesirable. The ideal would be for the wave to leave the horn with a completely planar wavefront, and to accomplish this a focussing mechanism, such as a curved reflector or a lens, may be used with the horn.

Three types of horns are shown in Fig. 11.2. The first is the sectoral horn, which is flared in only one plane (Fig. 11.2a); the second is the pyramidal horn, which is flared in both the planes (Fig. 11.2b). Both of these are used with rectangular waveguides. The third type is conical (Fig. 11.2c) used with a circular waveguide to produce a circularly polarized beam. Horn-type antennas do not provide very high directivity but are of simple, rugged construction. This makes them ideal as primary feed antennas for parabolic reflectors and lenses.

The choice of horn dimensions is dependent on the desired beam angle and directive gain and involves specification of the ratio of flare length to wavelength L/λ and flare angle ϕ, as shown in Fig. 11.2d.

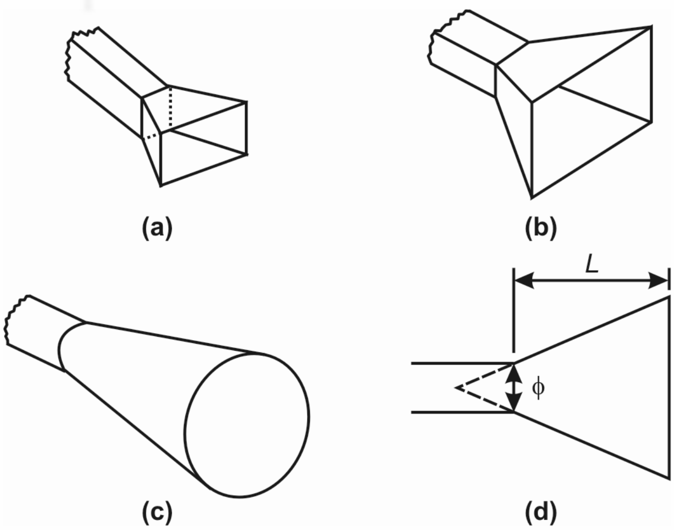

Fig. 11.2: Microwave horn antennas: a) sectoral horn; b) pyramidal horn; c) conical horn; and d) horn flare dimensions

11.3.2 Paraboloidal Reflector Antenna

The most widely used antenna for microwaves is the *paraboloidal reflector antenna*, which consists of a primary (feed) antenna such as a dipole or horn situated at the focal point of a paraboloidal reflector, as shown in Fig. 11.3a. The mouth, or physical aperture, of the reflector is circular. The cross section of the reflector containing the focal point *F* and the vertex *V* forms a parabola as shown in Fig. 11.3b. For this geometry, all the beams falling on the reflector travel parallel to each other after reflection.

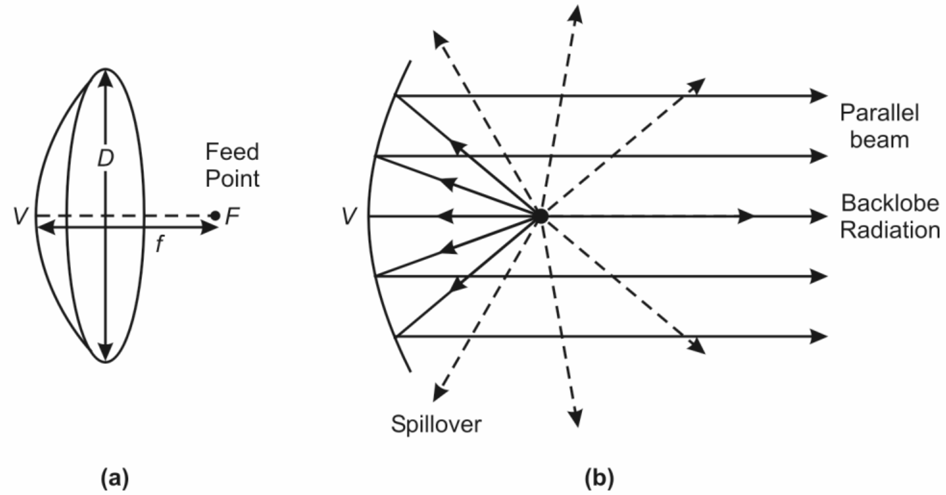

Fig. 11.3: a) Paraboloidal reflector; and b) radiation from the paraboloidal reflector fed by an isotropic primary radiator

An isotropic point source is assumed to be situated at the focal point. It radiates in all directions however only the beams captured by the reflector traverse the parallel path. The rays not captured by the reflector result into *spillover*. In the receive mode, spillover increases noise pickup. Also, some radiation from the primary radiator occurs in the forward direction in addition to the desired parallel beam. This is called a *backlobe* radiation since it is from the backlobe of the primary radiator. Backlobe radiation is undesirable because it can interfere destructively with the reflection beam, and practical radiators are designed to eliminate or minimise this. The isotropic radiator at the focal point radiates spherical waves, and the paraboloidal reflector converts these to plane waves. Thus, over the aperture of an ideal reflector, the wavefront is of constant amplitude and constant phase.

The directivity of the paraboloidal reflector is a function of the primary antenna directivity and the ratio of focal length to reflector diameter, f/D. This ratio, known as the **aperture number**, determines the **angular aperture** of the reflector, 2ψ (Fig. 11.4a), which in turn determines as to how much of the primary radiation is intercepted by the reflector. Assuming that radiation from the primary antenna is circularly symmetric about the reflector axis (*F*-*V*) and is confined to angles θ in the range $-\pi/2 < \theta < \pi/2$, it is found that the effective area is given by

$$A_{eff} = A\, I(\theta)\,, \tag{11.3}$$

where $A = \pi D^2/4$ is the physical area of the reflector aperture, and $I(\theta)$ is a function, termed the **aperture efficiency** (or *illumination efficiency*), which takes into account both the radiation pattern of the primary radiator and the effect of the angular aperture. With the focal point outside the reflector, as shown in Fig. 11.4a (which requires $f/D > 1/4$), the primary radiation at the perimeter of the reflector will not be much reduced from that at the centre, and the reflector illumination approaches a uniform value. This increases the aperture efficiency, but at the expense of spillover. Making f/D too large increases spillover and reduces aperture efficiency. Reducing f/D to less than 1/4 places the focal point inside the reflector, as shown in Fig. 11.4b. Here, no spillover occurs, but the illumination of the reflector tapers from a maximum at the centre to zero within the reflector region. This non-uniform illumination tends to reduce aperture efficiency. Also, placing the primary antenna too close to the reflector results in the reflector affecting the primary antenna impedance and radiation pattern that is difficult to take into account.

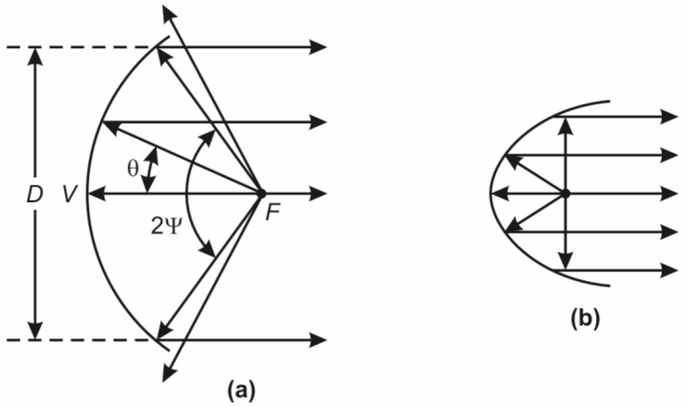

Fig. 11.4: a) Focal point outside the reflector; and b) focal point inside the reflector

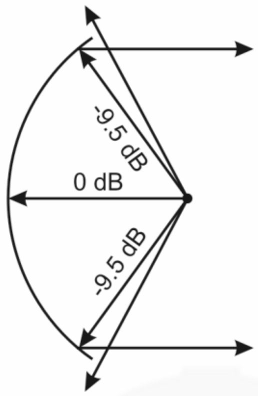

Fig. 11.5: Edge illumination from primary antenna is –9 to –10 dB below than at the centre

It can be shown that the aperture efficiency peaks at about 80%, with the angular aperture ranging from about 40° to 70° depending on the primary radiation pattern. The relationship between aperture number and angular aperture is

$$\frac{f}{D} = 0.25\cot\left(\frac{\Psi}{2}\right) \tag{11.4}$$

Typically, for an angular aperture of 60°, the aperture number is

$$\frac{f}{D} = 0.25 \times 1.73 = 0.43$$

This shows that the focal point should lie outside the mouth of the reflector, since f/D is then greater than 1/4. Satisfactory results are obtained in practice if the main lobe of the primary antenna intercepts the perimeter of the reflector at the –9 to –10 dB level as shown in Fig. 11.5.

The gain of an antenna is defined as

$$G = \frac{4\pi A_{eff}}{\lambda^2}. \tag{11.5}$$

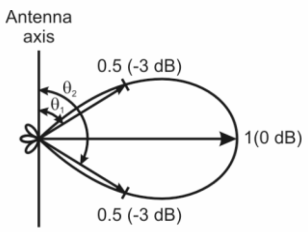

Fig. 11.6: –3 dB beamwidth of an antenna

On substituting $\pi D^2/4$ for A in Eq. (11.3) and using Eq. (11.5) for gain, we get

$$G = I(\theta)\left(\frac{\pi D}{\lambda}\right)^2. \tag{11.6}$$

In antennas the –3 dB beamwidth is defined as the angle subtended at the centre of the antenna radiation pattern by –3 dB gain lines as shown in Fig. 11.6. The beamwidth also depends on the primary radiator and its position. In practice, it is found that for most types of feed the –3 dB beamwidth is given approximately by

$$BW_{(-3db)} \cong \frac{70\lambda}{D}\,\text{degrees}. \tag{11.7}$$

The directivity of the beam is

$$D_o = \frac{4\pi}{\lambda^2} A_{eff} . \qquad (11.8)$$

This can be expressed in absolute value or in dB. You will observe that the expressions for gain (11.5) and directivity (11.8) are equivalent.

SAQ 2

Spend 5 *Min.*

Find the effective area, directivity and beamwidth for a paraboloidal reflector antenna for which the reflector diameter is 6 m and the illumination efficiency is 0.65. The frequency of operation is 10 GHz.

After getting conversant with the antenna geometry, let us now discuss the various types of primary feeds used in the paraboloidal reflector antenna.

11.3.3 Antenna Feeds

The paraboloidal reflector is most commonly used antenna for fixed point-to-point microwave communication systems. It is relatively simple in construction and quite inexpensive. Special feed systems are used so that the feed antenna is reduced in size or physically located out of the path of the incoming radiation. Two types of feed are shown in Fig. 11.7. The first of these (Fig. 11.7a) uses a dipole antenna, which normally radiates onto the parabola, and has a spherical reflector placed directly behind the dipole to prevent direct radiation. The backlobe radiation is reflected back at the parabola and is added to the main portion of the radiation.

The second method is known as the *Cassegrain feed system* (Fig. 11.7b). The horn feed antenna, the paraboloidal reflector, and the hyperboloid sub-reflector have a common axis of symmetry and the virtual focal point of the hyperboloid is

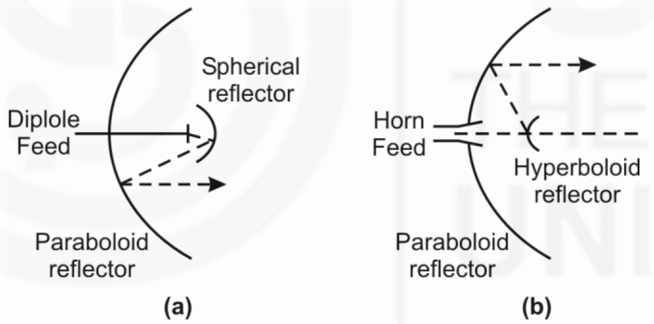

Fig. 11.7: Method of feeding microwave antenna. a) Parabolic reflector fed from a dipole and a small spherical reflector; and b) Cassegrain feed – paraboloid fed from a horn and a small hyperboloid reflector

coincident with the focal point of the paraboloid. Radiation reflected off the sub-reflector illuminates the main reflector uniformly. In this geometry, the spillover at the edges is low. This is of particular advantage in low-noise receiving systems, where large spillover results in high noise levels.

11.3.4 Dielectric Lens Antenna

Electromagnetic radiation is refracted when it passes through a surface separating a zone of lower dielectric constant from one of higher dielectric constant in exactly the same manner as light is refracted. The angles of incidence and refraction shown in Fig. 11.8a are related by the modified version of Snell's law, which states that

$$\frac{\sin \phi_r}{\sin \phi_i} = \sqrt{\frac{\varepsilon_{ri}}{\varepsilon_{rr}}} = \frac{1}{n}, \qquad (11.9)$$

where n is the refractive index of the material.

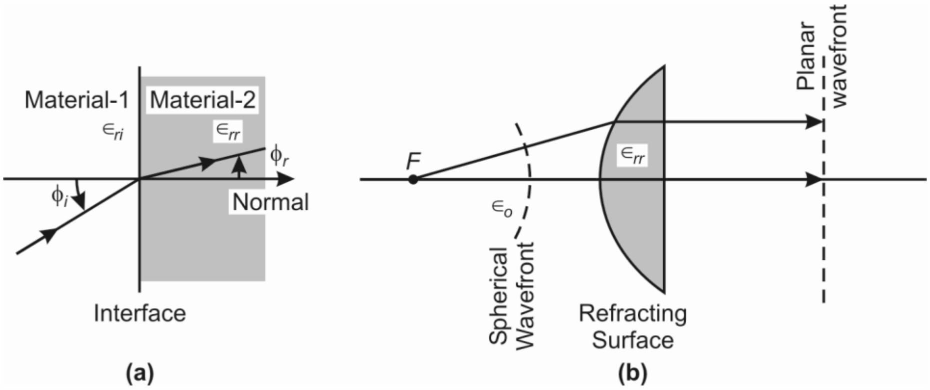

Fig. 11.8: Dielectric lens antenna: a) Snell's law of refraction; and b) principle of the collimating lens

The material used for the lens is usually one of the high-dielectric plastics, such as polystyrene or Teflon.

Fig. 11.8b illustrates the principle of the collimating (convex) lens, which is used to convert the diverging beam of radiation into a directional beam with a planar wavefront. Radiation along the axis of the lens passes through both surfaces at right angle, so no refraction takes place. Radiation at an inclined angle from the axis gets refracted toward the normal as it passes through the lens. The curvature of the lens is such that after refraction the rays emerging are all parallel to the axis.

Ideally, the lens should be fed with even illumination over its entire surface to achieve maximum efficiency and gain. The horn antenna is the most popular method of feed and most closely approximates the even-illumination requirement. For the higher microwave frequencies, the lens makes a very compact, highly directive narrow-beam antenna that is popular for applications such as portable communication links and mobile radar systems.

11.4 MICROWAVE COMPONENTS

Any typical microwave set up (may or may not be used for communication purposes) essentially contains a microwave source (or generator); a microwave detector and some intermediate microwave components. In this section we discuss these special microwave devices and components.

11.4.1 Microwave Sources

Microwave can be generated by using vacuum tube devices or solid-state devices. We will discuss two such sources here.

a. Reflex klystron

The microwave tube sources operate in vacuum and the microwave energy is coupled in/out through a waveguide window. A thermionic cathode generates electron beam which is formed into a cylindrical beam at constant velocity. This electron beam gets velocity modulated by interaction with dc and high frequency electric fields. The velocity modulation ends up in density modulation or bunching of electrons. These electron bunches pass through interaction space and give up energy.

Microwave tubes are classified into two types of tubes, the parallel or linear beam tubes (Klystrons and travelling wave tubes) and crossed field tubes (Magnetron). The linear beam tubes have the accelerating electric field parallel to the static focussing magnetic field. In this section we will consider the working of reflex klystron.

Principle of operation: A reflex klystron is a single cavity resonator klystron. It is a low power microwave source delivering about 10 to 500 mW output in the microwave frequency range of 1 to 30 GHz with an efficiency of 30% maximum. Reflex klystrons are used as local oscillators in commercial radars, military radars, Doppler radars and in most of microwave measurement systems.

A schematic diagram of the structure of a reflex klystron is shown in Fig.11.9. A reflex klystron consists of an electron gun for generating a cylindrical beam of electrons of constant speed. This beam is accelerated toward the cavity, which has a

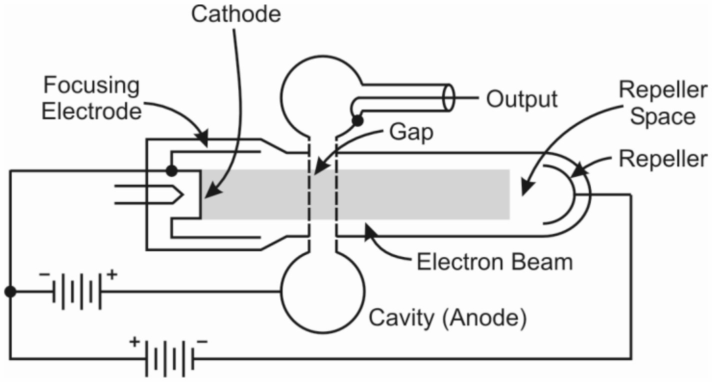

Fig. 11.9: Schematic sketch of reflect klystron

high positive voltage with respect to cathode and acts as an anode. The electrons overshoot the gap in this cavity and continue on to the next electrode, but never reach it because this *repeller* electrode has a fairly high negative voltage applied to it, and precautions are taken to ensure that it is not bombarded by the electrons. Accordingly, electrons in the beam reach some point in the repeller space and are then turned back, eventually to be dissipated in the anode cavity. If the cavity gap is excited with alternating voltage (at microwave frequency), the returning electrons give more energy to the gap than they took from it on the outward journey and sustained oscillations take place under specific conditions.

Working of reflex klystron: The operating mechanism of klystron is best understood by considering the behaviour of individual electrons. Let us take the electron that passes the cavity gap on its way to the repeller at the time when the gap voltage is zero and going negative as a reference electron. This electron, overshoots the cavity gap, traverses some distance in the repeller space and eventually returns back into the cavity gap. An electron passing the gap slightly earlier would have encountered a slightly positive voltage at the gap. The resulting acceleration would have propelled this electron slightly farther into the repeller space, and the electron would thus have taken a slightly longer time than the reference electron to return to the gap. Similarly, an electron passing the gap a little after the reference electron will encounter a slightly negative voltage. The resulting retardation will shorten its stay in the repeller space. It is seen that around the reference electron, earlier electron takes longer to return to the gap than later electrons, and so they reach the gap at the same time. These conditions are right for bunching to take place.

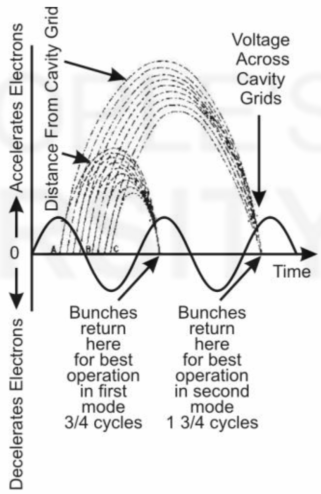

Fig. 11.10: Bunching of electrons in klystron

Hence, in the case of klystron, velocity modulation is converted to current modulation in the repeller space, and one bunch is formed per cycle of oscillations.

Transit time: For oscillations to be maintained, the transit time in the repeller space, or the time taken for the reference electron from the instant it leaves the gap to the instant of its return, must have the correct value. This is determined by investigating the best possible time for electrons to leave the gap and the best possible time for them to return.

The most suitable departure time is centred on the reference electron, at the 180° point of the sine-wave voltage across the resonator gap as shown in Fig. 11.10. It is also interesting to note that, ideally, no energy is spent in velocity-modulating the electron beam. It does take some energy to accelerate electrons, but just as much energy is gained from retarding electrons. Since as many electrons are retarded as accelerated by the gap voltage, the total energy spent is nil. This actually raises a most important point: *energy is spent in accelerating bodies* (*electron in this case*), *but energy is gained from retarding them*.

You will agree that the best possible time for electrons to return to the gap is when the voltage existing across the gap will apply maximum retardation to them. This is the time when the gap voltage is maximum positive (on the right side of the gap in Fig. 11.9). Electrons then fall through the maximum negative voltage between the gap grids, thus giving the maximum amount of energy to the gap. Therefore, the best time for electrons to return to the gap is at the 90° point of the sine-wave gap voltage. Returning after 1¾ cycles obviously satisfies these requirements; more generally it may be stated that

$$T = n + \frac{3}{4} \qquad (11.10)$$

where T is the transit time of electrons in repeller space and n is any integer.

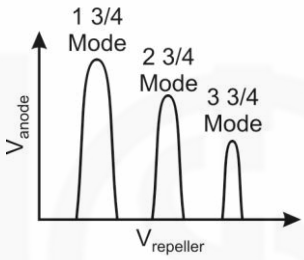

Fig. 11.11: Modes of klystron oscillations

Modes of oscillations: The transit time depends on the repeller and anode voltages, so that both must be carefully adjusted and regulated. Once the cavity is tuned to the correct frequency, both–the anode and the repeller–voltages are adjusted to give the correct value of T from data supplied by the manufacture. Each combination of acceptable anode-repeller voltages will provide conditions permitting oscillations for a particular value of n. In turn, each value of n is said to correspond to a different reflex klystron *mode* as shown in Fig. 11.11. In Fig. 11.10 you must have observed that the first mode occurs at T = 3/4 (n =0), however, the repeller voltage necessary to achieve this mode is very high. Hence, practical transit times correspond to the range from 1¾ to 6¾ cycles of gap voltage. Modes corresponding to n = 2 or n =3 are the ones used most often in practical klystron oscillators.

You will appreciate that a mode does not correspond to a fixed single frequency but is a narrow band of frequencies allowed by Q of the cavity. It means that the frequencies corresponding to a small range of repeller voltage are sustained by the cavity. Now, this allowed range of repeller voltage can be used beneficially in communication circuits. We can apply the modulating signal to the repeller overriding on the fixed negative bias. This modulates the oscillator frequency within a small frequency range determined by the Q of the cavity.

Though this is a straightforward method of modulating a microwave signal, usually it is not used in practice. The reasons for not modulating a source directly have been discussed already in Unit 9 while discussing the frequency modulation circuits.
In practice, the modulation of microwave signal is achieved by using a separate mixer, as will be discussed later in this section.

Another popular microwave tube source is **magnetron**. The first magnetron was developed in 1921and it became very important during the period of the Second World War as a high power microwave source. It is a crossed field tube with electric and magnetic fields perpendicular to each other. It is also called an M type tube. Magnetrons are capable of delivering megawatts of microwave peak power in the cm wavelength region and few kilowatts at 100 GHz with efficiencies greater than 30%. It has applications in military radars. A common commercial use of the magnetron is in the household microwave oven. We will not go into the details of this here. Let us now discuss the solid-state devices used as microwave sources.

b. Tunnel diode

A tunnel diode is nothing but a semiconductor *p-n* junction with heavily doped *p* and *n* sides. The doping density is of the order of 10^{20} atoms cm^{-3} and the junction depletion layer width is as narrow as 100 Å. This width narrows down as the doping density is increased. Because of heavy doping, the Fermi levels are located within the allowed (valence and conduction) bands.

The forward bias current-voltage characteristics of a typical tunnel diode are shown in Fig.11.12a. Let us understand them with the help of energy band diagram of the semiconductor *p-n* junction. In heavily doped junction, the Fermi energy level lies inside the conduction band on the *n*-side and inside the valence band on the *p*-side. At zero bias, the

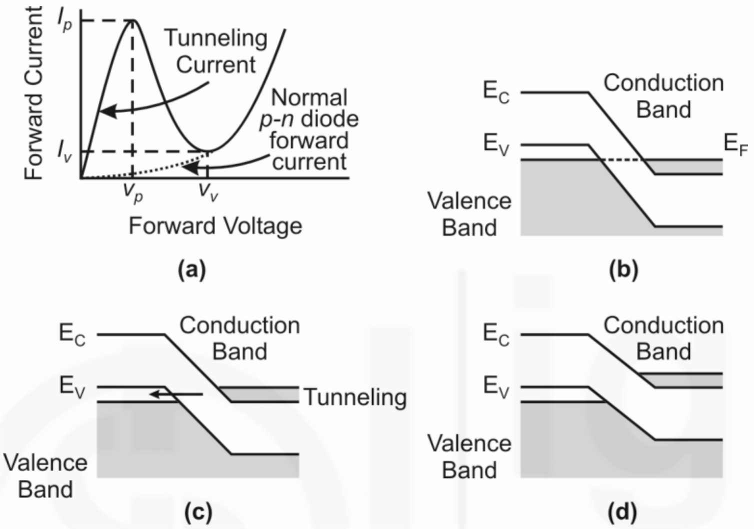

Fig.11.12: Quantum mechanical tunnelling in a heavily doped *p-n* junction: a) *I - V* characteristics; band diagrams at b) no bias; c) small forward bias; and d) large forward bias

two Fermi levels are at equal level, as shown in band diagram of Fig. 11.12b. Across the *p-n* junction there are no occupied electronic states above the Fermi level while all the states below the Fermi level are filled. In this condition the diode current is zero. When some forward bias is applied, the occupied states on the *n*-side face unoccupied states on the *p*-side with a narrow potential barrier in between (Fig. 11.12c). Proper conditions are now ready for quantum mechanical tunnelling of electrons from *n* to *p* side and the diode current increases. The tunnelling current increases to a peak and then decreases with increasing bias. The decreasing portion of the current is important in tunnel diode working. It is called **negative differential resistance (NDR)** region. In this region, the diode current decreases, while the forward bias voltage increases, i.e. $\left(\frac{dV}{dI}\right)$ is negative. Since the diode differential resistance is negative, the diode gives power instead of consuming power. The magnitude of the negative resistance increases with bias voltage.

Please remember that the absolute value of resistance is still positive even in the case of NDR region. This is evident from the fact that the *I-V* characteristics are in the first quadrant of the plot.

You will observe in Fig. 11.12a that in the forward direction the current first increases due to tunnelling to a peak value, I_p, at a voltage V_p and then decreases to a valley value I_v at a voltage V_v . At large forward bias, (Fig. 11.12d) the two bands go out of alignment and the tunnelling of electrons ceases and the diode current follows normal diode characteristics.

A tunnel diode shows a dynamic negative conductance, which enables it to function as a microwave oscillator and amplifier. An equivalent circuit of a tunnel diode is illustrated in Fig. 11.13 where $-R_p$ is the negative resistance; and C_p is the junction capacitance of the diode. R_s and L_s are the resistance and inductance of the packaging circuit.

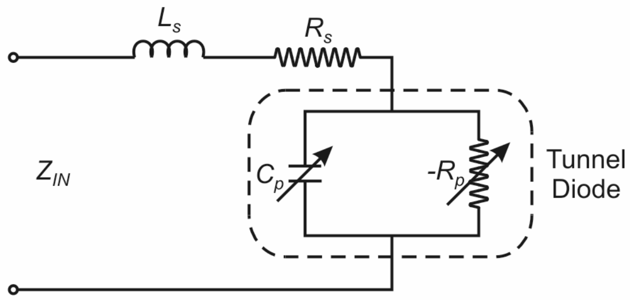

Fig. 11.13: Equivalent circuit diagram of a tunnel diode

The input impedance Z_{in} of the equivalent circuit is expressed as:

$$Z_{in} = R_s + j\omega L_s + \left(\frac{R_p \left(\frac{j}{\omega C_p} \right)}{-R_p - \frac{j}{\omega C_p}} \right) \tag{11.11}$$

The circuit involves two frequencies, the self resonance frequency when the imaginary part of the impedance is zero and the resistive cut-off frequency when the real part of the input impedance is zero. These frequencies are given by the following expressions:

Resistive cut-off, frequency, $f_c = \frac{1}{2\pi R_p C_p} \left[\frac{R_p}{R_s} - 1 \right]^{\frac{1}{2}}$ (11.12)

Self-resonance frequency, $f_r = \frac{1}{2\pi} \left[\frac{1}{L_s C_p} - \left(\frac{1}{R_p C_p} \right)^2 \right]^{\frac{1}{2}}$ (11.13)

When the diode is operated in the negative resistance region it is desirable to have the operating frequency much less than the maximum resistive cut off frequency; which in return should be less than the self-resonance frequency.

i.e. $f << f_c < f_r$

This shows that the series inductance must be lowered. A figure of merit of a tunnel diode is defined as the ratio of the peak current to the capacitance at the valley voltage, I_p/C_p. The negative resistance is inversely proportional to the peak current. The switching speed of the tunnel diode depends on the current and the RC product.

In tunnel diode, the values of $-R_p$ and C_p are governed by the forward bias voltage. A fixed value of the bias voltage fixes the operating point of this device. A modulation voltage superimposed on this bias changes values of C_p and $-R_p$ dynamically. This in turn varies the operating frequency of the device and it is possible to use this diode as a frequency modulation device in communication circuits at microwave frequencies. However, for the reasons already discussed, it is not preferable to modulate the frequency source directly. Usually tunnel diode is used as fixed frequency generator.

c. Gunn devices

Gunn device is a transferred electron device (TED) that can provide negative differential resistance by transfer of electrons and can be used as a microwave continuous wave (CW) and pulsed oscillator and amplifier. The device accomplishes a one step conversion from dc to microwave energy from a single low voltage supply. These are fabricated from direct band-gap compound semiconductors.

Gunn devices are some times called *diodes*; however, there is no junction present in these devices. These are bulk semiconductor materials. They achieve negative resistance via the transferred electron effect, a quantum property exhibited by direct band-gap semiconductors like GaAs. Negative resistance is obtained in the GaAs material itself under certain conditions and is a result of transfer of electrons in the conduction band of GaAs from high mobility low energy valley to low mobility higher energy satellite valleys, thus affecting the conduction mechanism of the semiconductor material.

The band structure of GaAs is shown in Fig.11.14. This is called *E*-*k* diagram in solid state physics. The *E*-*k* diagram of solids is plotted to depict energy bands in the reciprocal lattice space. In this diagram, we can see several minima in the conduction band. When the lowest minimum of the conduction band falls exactly above the valence band maximum (i.e. at same *k* value), then this type of semiconductors are of **direct band gap** type. In the case of silicon and germanium, the lowest minimum of the conduction band does not align with the maxima of valence band and such semiconductors are called **indirect band gap** materials.

The conventional energy levels drawn for the junctions have energy on the *y*-axis which represents the top of the valence band and the bottom of the conduction band. The *x*-axis on these diagrams is distance in real spaces with reference to some point (like junction).

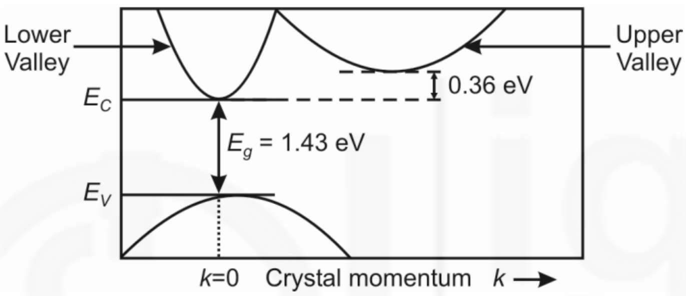

Fig. 11.14: *E*-*k* diagram of two conduction valleys of GaAs

The GaAs band diagram shows the presence of two valleys in the conduction band: (i) a lower sharp valley at $k = 0$ separated by the band-gap energy from the valence band, and (ii) a flat parabolic upper valley, separated by an added energy gap of 0.36 eV. It is situated at different value of momentum than the lower valley.

The effective mass and the mobility of electrons in the two valleys are different. The electrons in the upper valley are heavy (have greater effective mass) and hence their drift velocity is much slower. At low external voltage, the conduction involves the lower valley alone but at higher voltage (fields), the electrons make inter-valley transfer into the bands of the upper valley where they tend to remain there. Hence, as the applied voltage (field) exceeds a threshold value due to inter-valley transfer of electrons, the conduction current goes on decreasing and the material behaves like a negative resistance because of negative differential mobility. Gunn observed the generation of microwaves in an *n*-type GaAs specimen when the applied voltage exceeded a certain critical value. This critical voltage produced an electric field of 3.2 kV cm^{-1} across the active region of the diode.

We can sum up the conditions for the functioning of TEDs as:

- The two valleys must be separated by an energy which is at least several times the thermal energy, ($\Delta E >> kT$). The value of the thermal energy at room temperature is 0.026eV.
- The separation energy between the valleys must be smaller than the band-gap energy of the semiconductor. $\Delta E < E_g$. The band-gap is 1.43 eV for GaAs and 1.34 eV for InP.

- The electrons in the upper valley must have high effective mass, low mobility (μ_1) and a high density of state as compared to the lower valley. The electron drift velocities ($v_d = \mu E$) must be much smaller in the upper valleys than in the lower valley.

The common elemental semiconductors, silicon and germanium, are indirect band-gap semiconductors and they do not meet these criteria required for the operation of a transferred electron device, whereas compound semiconductors like InAs, InP, GaAs, CdTe qualify as TEDS.

Operations of Gunn devices: A dc bias is applied across the *n*-type GaAs piece. The positive supply can be called anode and the negative as cathode. As the electric field exceeds the threshold field, the electrons at the cathode end transfer to the upper valley; their mobility drops to the value corresponding to the upper valley, μ_2; and the electrons slow down. The electrons away from the cathode region drift at a faster velocity. This difference in the velocities groups the electrons into slowly moving bunches, called *domains*. As the domain grows in size, the domain field is increased and the field in the rest of the semiconductor is reduced. When the domain reaches the anode, the process is repeated and a new domain is formed at the cathode. This is known as the *transit time mode* of operation of Gunn diode and the frequency of operation is determined by the drift velocity and the length of the active region of the device.

The conditions for Gunn oscillations in GaAs material of length L and *n*-type doping density (n_o) sandwiched between two conductors are:

- frequency of oscillations: $f = v_d / L$; and
- product of carrier concentration and length $(= n_o L)$ should be greater than 10^{12}cm^{-2}. Here, v_d is the drift velocity of electrons in GaAs.

Gunn oscillators: The packaged Gunn diode is a rugged device. It is mounted inside a resonant cavity. The Fig. 11.15 shows the mounting in a cavity which is externally tuned to the oscillating frequency of the Gunn diode. In a half wavelength long cavity, a conductive *dowel* supports the Gunn device and connects it to the ends of the cavity. The output is taken through a short coupling loop parallel to dowel.

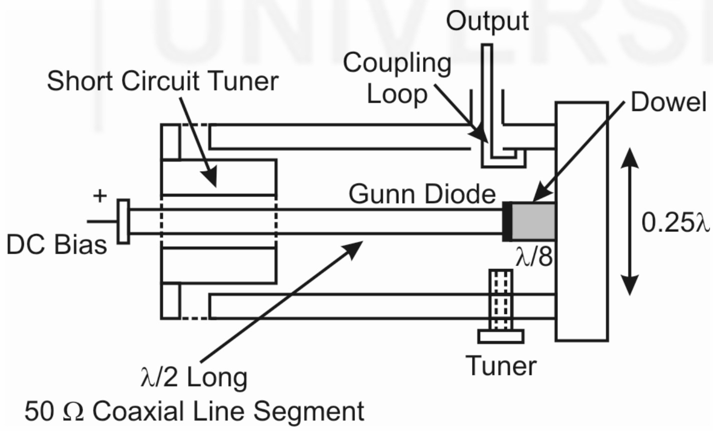

Fig. 11.15: Coaxial cavity Gunn oscillator

Low power GaAs Gunn diodes are useful as low noise sources in local oscillators, locking oscillators, low power radar applications and motion detection applications.

You may like to attempt an SAQ now.

SAQ 3

Spend 3 Min.

A typical *n*-GaAs Gunn device has length of 10 μm and its frequency of operation is 20 GHz. The threshold field is 2.8 kV cm^{-1}. If a field of 4 kV cm^{-1} is applied to this device, calculate the electron drift velocity and the applied voltage.

After discussing the microwave sources, let us discuss some microwave detectors.

11.4.2 Microwave Detectors

A semiconductor junction can function as a microwave signal detector in the forward biased non-linear resistor mode. When such a device is subjected to a microwave signal of applied voltage $V = V_1 \cos \omega t$ the device current can be expressed as:

Diode current $I = I_o (aV + a^2V^2/2! + ...)$ (11.14)

where I_o is the reverse saturation current and V is the applied voltage.

Substituting for V we get

$$I = I_o \,[a\,(V_1 \cos \omega t) + a^2 (V_1 \cos \omega t)^2/2! + ...]$$

$$= I_o \left[aV_1 \cos\omega t + \frac{a^2V_1^2}{4}\left(1 + \cos 2\,\omega t\right) + ... \right] \quad (11.15)$$

Hence, the dc component of the diode current is $= \left(\dfrac{a^2V_1^2 I_o}{4} \right)$

The diode current consists of dc component and components involving signal frequency and multiples of signal frequency. It is observed that the dc component of the forward current is proportional to the microwave power or the square of the voltage. The device acts as a square law detector. Detectors are nonlinear devices. The most common devices used in the microwave region for detection of signal are point contact diodes and Schottky barrier diodes.

Before proceeding further, you may like to attempt an SAQ.

SAQ 4

Spend 2 Min.

Calculate the dc component of current in a *p-n* diode detector having the reverse saturation current $=10^{-6}$ A, $a = 40\ V^{-1}$ and the amplitude of applied voltage is 1 V.

a. Schottky barrier diode

Schottky barrier diode is a metal-semiconductor junction diode. It uses metals like gold, silver, platinum, tungsten etc. on one side of the junction and an *n*-type semiconductor (Silicon or GaAs) on the other side as shown in Fig. 11.16a. A rectifying contact is formed by the potential barrier at the metal-semiconductor interface. The rectifying properties exist if the work function of the metal is greater than that of the semiconductor. When a junction is formed, electrons flow from the semiconductor side to the metal side to equalise the Fermi level, and the net positive space charge is produced on the semiconductor side as shown in Fig. 11.16b, which poses a barrier to the flow of electrons into the metal. When the diode is unbiased, electrons on the *n*-side have lower energy levels than electrons in the metal hence they cannot flow over the barrier.

In the forward biased condition, conduction electrons in the *n*-side have enough energy to cross the junction.

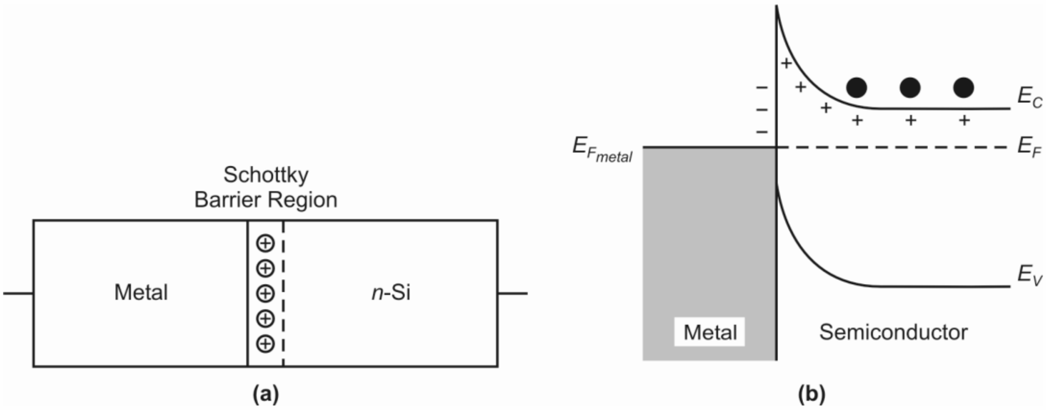

Fig. 11.16: a) Construction; and b) band diagram of Schottky diode

The unique features of Schottky barrier are:

- it has only electrons as majority carriers on both sides of the junction and hence is a unipolar device as against an ordinary *p-n* junction which has both electrons and holes as majority carriers;
- due to the unipolar nature of carriers, there is no depletion layer in the metal side. This results in faster switching characteristic as compared to a bipolar diode; and
- The Schottky diode has a very small junction capacitance hence it can be operated at very high frequencies (10 GHz).

Hence the Schottky diodes can be used as a detector of very high frequency signals; as a switching device in digital computers; as a clipper or clamper in electronic circuits; and as a mixer and detector in communication systems.

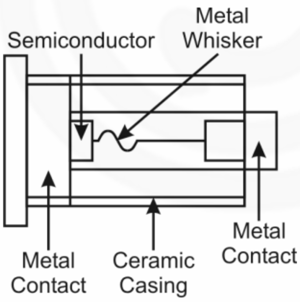

Fig. 11.17: Construction of a point contact diode

Historically a point contact diode has been used as a microwave detector, which uses a mechanical contact between a metal whisker and the semiconductor surface to make a rectifying contact. Selection of material for a whisker wire is critical. The materials used are tungsten, molybdenum or phosphor-bronze. The whisker point is etched before inserting it into the semiconductor wafer.

A schematic of typical point contact detector is shown in Fig. 11.17.

b. Varactor diode

A Varactor diode is a voltage dependent variable capacitor in the form of a semiconductor *p-n* junction. It is a short form of *variable reactance*. In a Varactor diode, the depletion layer created by the reverse bias (which has no free carriers) acts as a capacitor dielectric while *p* and *n* regions act as capacitor plates. The reversed biased *p-n* junction thus has an effective capacitance of:

$$C = \varepsilon A/W, \qquad (11.16)$$

where ε is the dielectric constant of the semiconductor material; *A* is the area of the junction and *W* is the width of the space charge region. The width *W* of the space charge region is approximately proportional to the square root of the reverse bias voltage while the area *A* and ε are constant. The value of the capacitances at different bias voltages, *V*, can be generally written as:

$$C = K/\sqrt{V}\,. \qquad (11.17)$$

When the reverse bias voltage increases, the depletion layer widens and the value of C decreases. This indicates that the variation of capacitance is maximum when bias is zero and it reduces in nonlinear manner as the bias is made more and more negative. Hence this capacitance can be varied in a controlled manner with a bias voltage, V, (either dc or low frequency RF).

The C-V relationship depends on the doping profile of the *p*-*n* device. The width of the depletion layer (which decides the value of the capacitance) varies inversely with the impurity concentration of *p* and *n* sides. Very high doping concentration results in very narrow depletion width and high capacitance.

As a non-linear device, a Varactor can be used as voltage detector; signal modulator and voltage control oscillator (VCO); parametric amplifier; and harmonic generator.

c. *p-i-n* Diode

A *p*-*i*-*n* diode is a three layer semiconductor structure with an intrinsic (undoped) layer (*i*) of pure silicon of width W sandwiched between heavily doped *p* and *n* type silicon layers as shown in Fig. 11.18. The intrinsic semiconductor is a very high resistive, undoped semiconductor. Since *i* region is a high resistance region or low doping region, it provides the following advantages over an ordinary *p*-*n* junction:

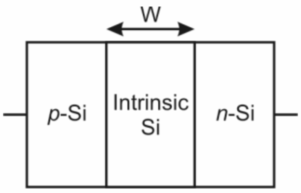

Fig. 11.18: Construction of a *p-i-n* diode

- The capacitance of the device, which is inversely proportional to the distance between *p* and *n* regions, decreases. This advantage allows *p*-*i*-*n* diode to have fast response time; hence the devices are useful at high frequencies.
- There is a greater electron-hole pair generation because of the increased electric field between the *p* and *n* regions. This advantage allows the *p*-*i*-*n* diode to process even weak signals.

The *p*-*i*-*n* diode acts as a variable resistance in the forward bias mode and in the reversed bias mode, it acts as *open* circuit with a constant capacitance value.

The *p*-*i*-*n* diode can be used as a dc controlled microwave switch; and in attenuator applications where the resistance can be controlled by the current.

d. Microwave mixers

To use the microwave source gainfully in a communication circuit, it is necessary to modulate its frequency, so that the signal can be sent on a microwave carrier. As we have already discussed in Unit 9, it is not, however, advisable to modulate the frequency source directly. For this purpose, usually the microwave sources are operated at a constant frequency and the modulating signal is superimposed on them by using a mixer. Fig. 11.19 shows a typical microwave mixer using a point contact diode. The source (local oscillator) frequency, f_{LO} is fed via a co-axial cable into the

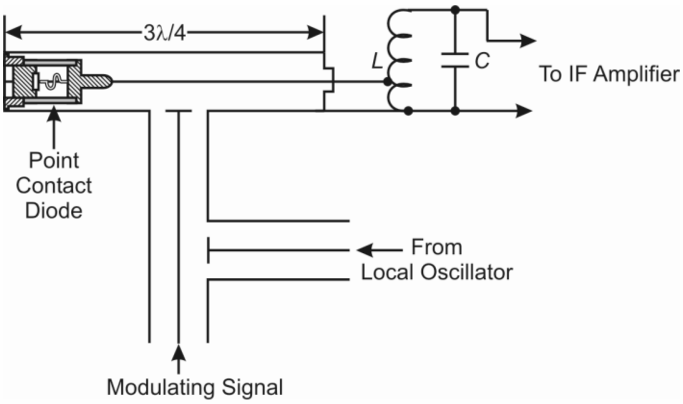

Fig. 11.19: Microwave mixer

cavity containing the point contact diode. The modulating signal (f_s) is also fed through a T-joint as shown in the figure. The output taken from the mixer diode via a co-axial contact contains the sum-difference frequency components like ($f_{LO} \pm f_s$). This mixer output is filtered using appropriate filters before transmitting.

11.4.3 Microwave Waveguide Components

You have already learnt about the basic conductor used for microwave signals in Unit 4 of this course. This is a hollow metallic conductor called *waveguide*. You have also learnt about the different modes of signal transmission in the waveguides. In this section we will discuss some microwave components (based mostly on waveguides) used for microwave signal processing and handling.

a. Waveguide attenuator

Waveguide attenuators are used to control the power level of a signal in a microwave circuit. They absorb microwave power in a waveguide component without reflection or very low reflection. Attenuators are available commercially as fixed value attenuators (attenuating pads) or calibrated variable attenuator units, where attenuation can be controlled by a calibrated knob.

An attenuator consists of a waveguide section with a tapered microwave attenuating strip placed inside the waveguide along the electric field direction as shown in Fig. 11.20a. The attenuating strip is a dielectric sheet coated with a microwave resistive film. In a TE_{10} mode, the strip placed in the middle of the waveguide along the E_{max} position gives maximum attenuation by causing power dissipation in the resistive film.

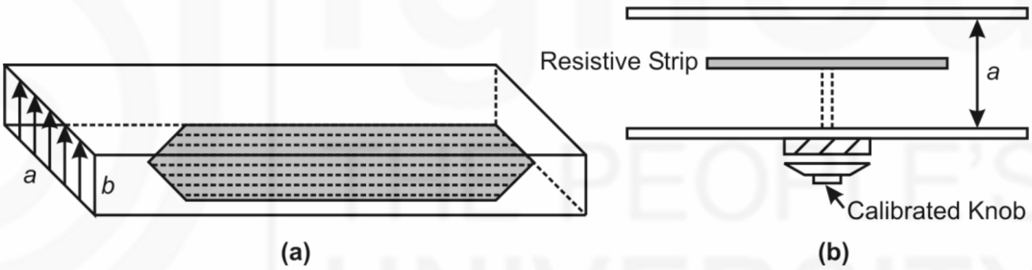

Fig. 11.20: Types of variable waveguide attenuators: a) fixed attenuator; and b) moving strip

The strip can also be moved sideways by thin dielectric rods and the movement of the strip can be controlled by a calibrated knob from outside as shown in a top cross-sectional view in Fig. 11.20b. This moving strip variable attenuator is calibrated against a known standard attenuator and the attenuation is marked on the knob in dB. An attenuator is a reciprocal device which gives the same amount of attenuation both ways.

b. Waveguide isolator

An isolator is a unidirectional or non-reciprocal microwave component which allows microwave power to flow freely in the forward direction while it is attenuated heavily in the reverse direction. This component is mostly used between a microwave source and the varying load so as to isolate the source from the reflected power. Non-reciprocal property is achieved by the use of a microwave ferrite material as shown in Fig. 11.21. If a rectangular waveguide is coupled to a circular waveguide and a ferrite rod is placed inside the circular waveguide in an external axial magnetic field then as the input wave passes through the ferrite rod in the direction of magnetic field, it suffers 45^o rotation of the plane of polarisation. The wave output is taken out by a 45^o twisted waveguide. There is no attenuation of the input power as it traverses through

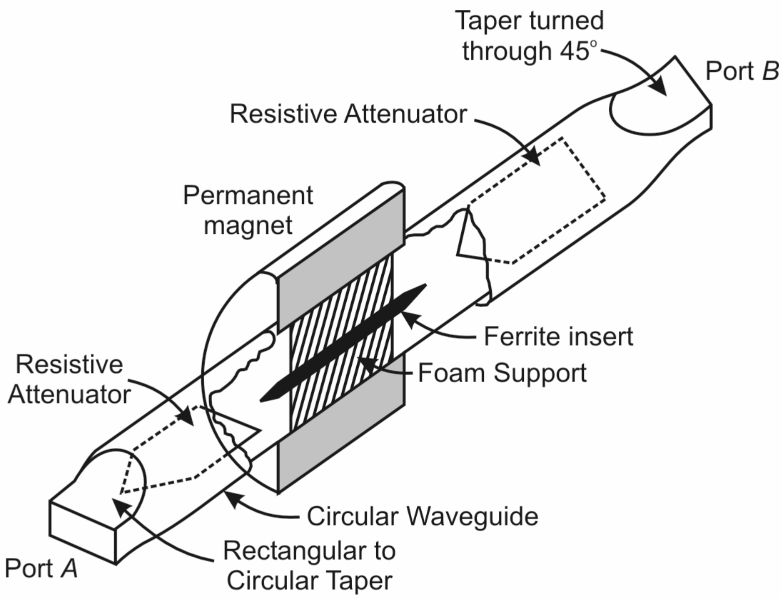

Fig. 11.21: Cutaway view of isolator

the isolator from Port *A* to Port *B*. Consider the power entering from the reverse direction (i.e. from Port *B*). The wave already has a 45^o twist provided by the waveguide and it suffers additional 45^o change of polarisation in the ferrite material. This means that the wave reaches Port *A* in orthogonal direction and the power is absorbed by the flat attenuating pad. In an ideal lossless and matched isolator, the input and output VSWR is 1.

c. Magic *T*

A microwave waveguide junction of three waveguide sections is called a *T* junction. Such a *T* junction has two forms.

(a) *E*-plane or series *T* and

(b) *H*-plane or shunt *T*.

E-plane *T* is formed by coupling a side arm on a waveguide whose axis is parallel to the *E* field of the main waveguide as shown in Fig.11.22. Port 1 is the *E*-plane arm. As the power enters the *E*-plane (Port 1), the power is equally divided into Port 2 and Port 3 of the main waveguide but in opposite phases (180^o out of phase). If the three ports are perfectly matched, the VSWR is unity.

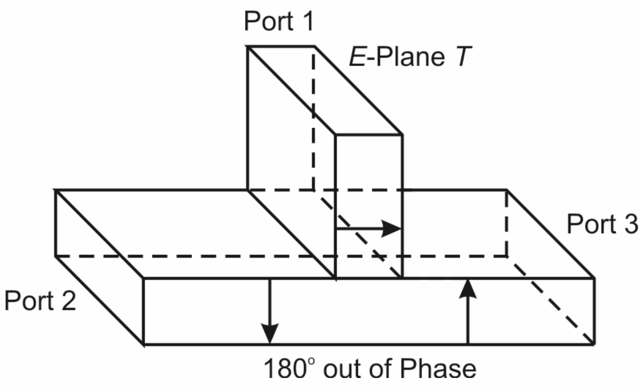

Fig.11.22: *E*-plane waveguide *T* structure

H-plane *T* or a shunt *T* is formed by coupling a side arm on the waveguide such that its axis is parallel to *H* field of the main waveguide. It is shown in Fig.11.23. The phase relationship of the divided signals of an *H*- plane *T* is different than that in the *E*- plane *T*. As the power is incident to the *H* arm (Port 1), the power is equally divided into Port 2 and Port 3 in phase.

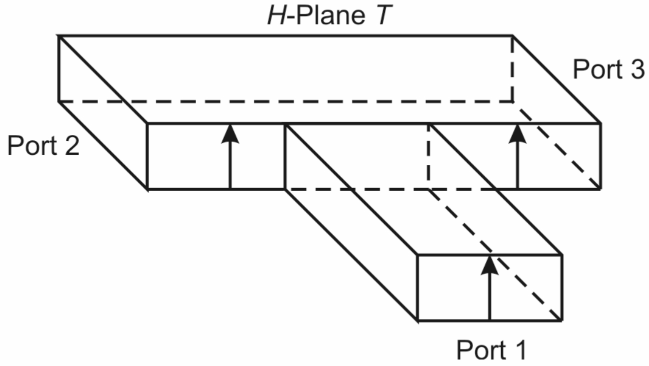

Fig. 11.23: *H*-plane waveguide *T* structure

You will observe that if two waves are fed in-phase to the ends of the main waveguide i.e. Ports 2 and 3, the output wave will be the sum of the two waves and in phase and will appear at the Port 1.

A combination of *E*- plane *T* and the *H*- plane *T* in one waveguide structure as shown in Fig.11.24 is a *Magic T* or a *Hybrid T*. A magic *T* is a well matched and tuned hybrid *T* which demonstrates similar electrical performance from all the four ports.

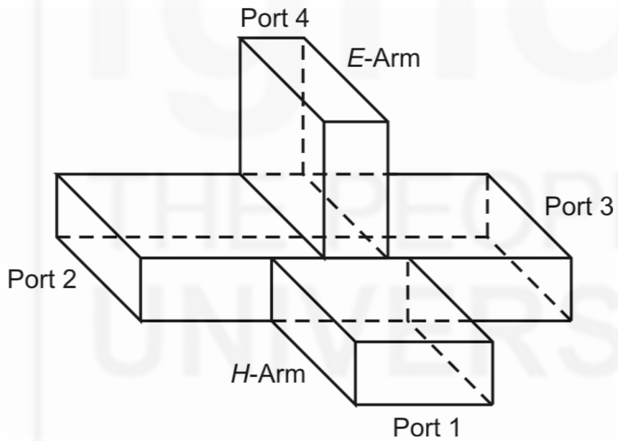

Fig. 11.24: Magic *T*

The electrical performance of a magic *T* with main waveguide as Port 2 and Port 3 , *H* arm as Port 1 and *E* arm as Port 4 are as follows:

- If a wave is incident to Port 1 (*H* arm), it is divided equally between Port 2 and Port 3 in phase and the output power at Port 4 is zero.
- If a wave is incident to Port 4 (*E* arm), it is divided equally between Port 2 and Port 3 in opposite phase and the output at Port 1 is zero.
- If a wave is incident to either Port 2 or Port 3, it will not appear at the other Port 3 or Port 2 respectively because of the junction of *E* and *H* arms.
- If two waves of equal magnitude and opposite phases are fed into Port 2 and Port 3, the output will be zero in Port 1 and additive at Port 4.
- If two waves of equal magnitude and same phase are fed into Port 2 and Port 3, the output will be zero in arm 4 and additive at Port 1.

Applications of magic T are multifarious. One such application is in impedance measurement. A magic T can be used as a bridge for measurement of unknown impedances. For this purpose the microwave source is connected to Port 1, a known variable impedance to Port 3, an unknown impedance to Port 2 and a meter at Port 4. The impedance value of the standard variable impedance is varied till the bridge is balanced and the meter reads null. Microwave power entering at Port 1 is equally divided to the unknown and the known impedance at Ports 2 and 3. If the reflected waves from Port 3 and Port 2 are equal (i.e. when both impedance values are perfectly matched), there will be no output at Port 4 and the null detector will show zero output. From the value of known impedance the unknown impedance can be accurately determined.

d. Circulator

A circulator is a three-port, four-port or a multi-port unidirectional (non-reciprocal) waveguide component in which the incident microwave power is freely transmitted to the adjacent port in one direction; but there is no transmission to the port in the other direction. In Fig. 11.25 a three-port circulator is shown. Here, there is perfect coupling of signal from Port 1 to Port 2, but perfect isolation from Port 2 to Port 1. The non-reciprocity is achieved by a magnetised ferrite material disc placed at the junction of the three waveguides in a three way circulator.

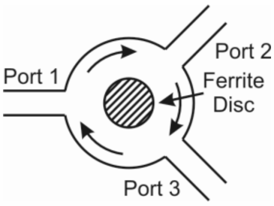

Fig. 11.25: Schematic diagram of a circulator

e. Directional coupler

A directional coupler is a four-port microwave component which couples a fraction of microwave power from the main waveguide to an auxiliary waveguide port. It consists of a main waveguide section with Port 1 and Port 2 and an auxiliary waveguide coupled to the main guide by a series of holes on the broad side and ends with Ports 3 and 4 as shown in Fig. 11.26. Power incident to Port 1 is freely transmitted to Port 2. A part of the power flowing in the main waveguide appears at Port 4 but there is no power output from Port 3. In a similar way, if the power is incident to Port 2, the power is partially coupled to Port 3 and no power flows out from Port 4.

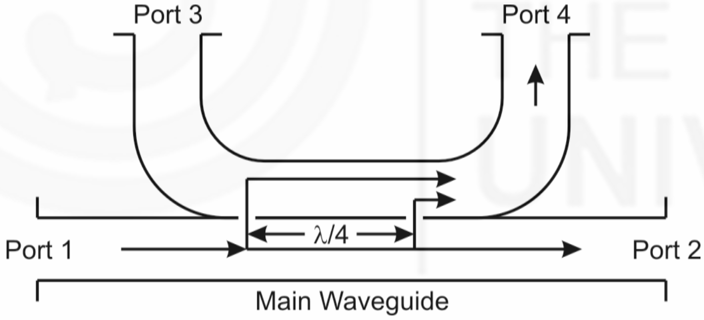

Fig.11.26: Directional coupler

The characteristic performance of a directional coupler is expressed in terms of i) Coupling Factor and ii) Directivity. Assuming that the wave is propagating from Port 1 to Port 2 in the main guide, these two factors are defined respectively by:

- Coupling factor is the ratio of the power input to Port 1 and the coupled power to output Port 4 and is expressed in dB.

$$C = 10 \log (P_1/P_4) \qquad (11.18)$$

- Directivity is the ratio of the power coupled in output Port 4 to the power output at Port 3. This ratio is very large for good directivity and is expressed in dB.

$$D = 10 \log (P_4/P_3) \qquad (11.19)$$

Spend 3 Min.

SAQ 5

Consider a symmetrical directional coupler with 10 dB coupling and infinite directivity. The ports are matched. If 1 W microwave signal is incident at the Port 1, what is the power distribution in the various ports of the coupler?

f. Microwave resonant cavities

Microwave waveguide cavity is a completely closed waveguide resonant structure storing microwave energy inside it. There are two main types of resonator cavity structures viz. a rectangular-cavity resonator and a circular waveguide resonator. A resonant cavity is equivalent to an LC parallel tuned circuit having a resonance frequency $f = 1/[2\pi\sqrt{LC}]$. A resonator normally can operate in a large number of resonant modes and each mode corresponds to a definite frequency. If the cavity forms a part of a transmission line and the frequency of the input microwave source is swept, the cavity will absorb power from the line at its resonant frequency. This results in a dip in the frequency response of the waveguide like a notch (band stop) filter.

g. MESFET

At microwave frequencies, there are some specially designed transistors with similar structure as conventional MOSFET. These are Metal Semiconductor Field Effect Transistors (MESFETs). You will observe that the structure of MESFET shown in Fig. 11.27 is similar to conventional MOSFET with the only difference that the oxide layer is absent. Because of this, the metal contact made for the gate terminal acts as metal – semi-conductor (Schottky) contact. This reduces the junction capacitances encountered in MOSFETS and increases the operating frequency to microwave range.

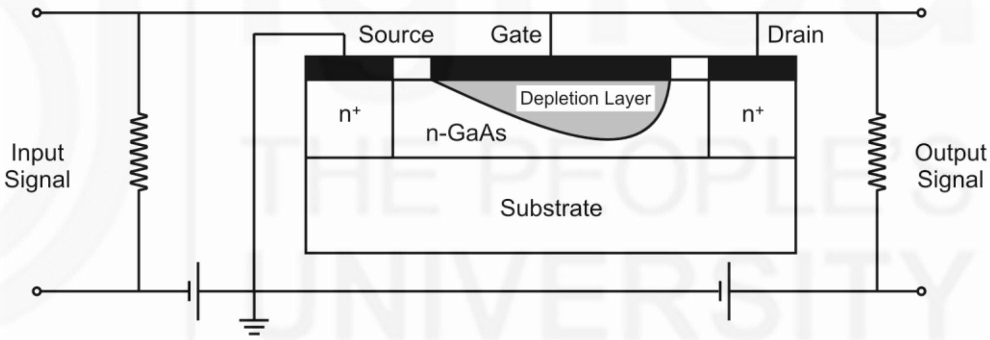

Fig. 11.27: Construction of MESFET with applied bias.

You will agree that this structure is just like any other semiconductor device used at low frequency and can be built on a semiconductor wafer in the form of an IC. The discrete components you have learnt (waveguide, directional couplers, isolators etc.), can also be built on the same base and a whole microwave circuit can be made on a single chip. These are called the *Monolithic Microwave Integrated Circuits* (MIMICs), Fig. 11.28 shows a typical MIMIC structure. MIMICs are the workhorses in the recent mobile cellular phones.

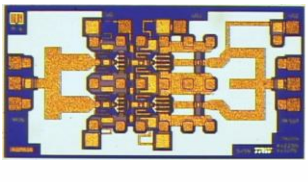

Fig. 11.28: A typical MIMIC structure

The microwave circuits like oscillators have a similar principle of operation as their low frequency counterparts, however here the discrete *L-C* components are replaced by distributed elements as discussed in Unit 4 of this course. These components can be formed on the IC chip by planar structures of microstriplines.

h. VSWR measurement set-up

VSWR measurement is one of the most important characterisation techniques in the microwave systems. It allows us to find out the matching of load and account for any losses in the system.

A typical setup of equipment used in VSWR measurements is shown in Fig. 11.29. The signal generator (Klystron or Gunn diode) in the set up is tuned at the frequency of the VSWR meter. The detector is a microwave field sensor (typically a piezoelectric crystal) which is inserted into the circuit via a slotted waveguide (slotted line). The probe is moved across the standing wave pattern over the length of the slotted waveguide and the positions and amplitudes of voltage maxima and minima are noted down. The distance between maxima and minima allows us to measure the signal wavelength, since they are separated by $\lambda/4$ distance. From the amplitude of maxima and minima we can calculate the value of VSWR.

A general rule in slotted line measurements is to use minimum penetration of the detector sampling probe. The power picked up by the sampling probe causes a distortion in the standing wave pattern. The penetration of the probe affects the standing wave pattern because the probe acts as an admittance shunting the line. The impedance in a standing wave pattern is greatest at a voltage maximum. A shunt admittance introduced by the probe lowers this impedance, causing an error in measured SWR.

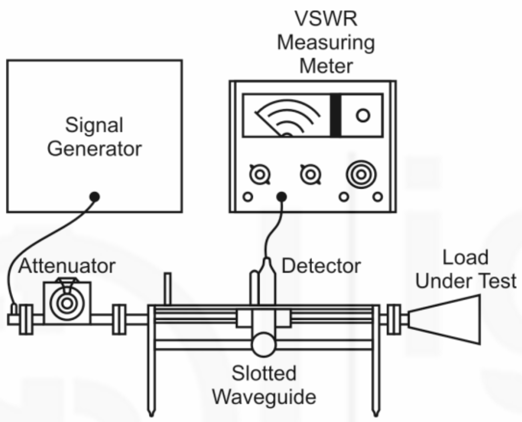

Fig. 11.29: VSWR measurement set-up

While measuring VSWR, it is more desirable to locate a voltage minimum than at the voltage maximum, since at minimum, the probe effects are lower. With a low SWR, the voltage minimum is fairly broad, and locating the exact point of the minimum is difficult with a single measurement. A more accurate procedure is to note the probe carriage position at two equal output readings on either side of the minimum, then average the two readings.

In this section have discussed various microwave components. A comprehensive waveguide microwave bench for experimental purposes consists of the following waveguide components arranged serially by coupling the flanges of the waveguides.

1. Klystron oscillator (with associated power supply) or Gunn diode oscillator;
2. Isolator;
3. Standard attenuator up to 20 or 30 dB maximum attenuation;
4. Variable screw tuner;
5. Directional coupler;
6. Tuner;
7. Standing wave probe with output to a VSWR meter;
8. Open termination to connect unknown impedance;
9. Wavemeter or a cylindrical waveguide cavity with calibrated head for measuring frequency connected to directional coupler;
10. Crystal detector with output to a meter.

The bench can be used to make measurement of frequency, VSWR, reflection coefficient, impedance and other properties of waveguide components and material.

You know that the microwaves are important carrier medium for carrying TV, telephony signals. Fig. 11.30 shows a block diagram of a typical microwave communication system, which uses multiple stages of IF.

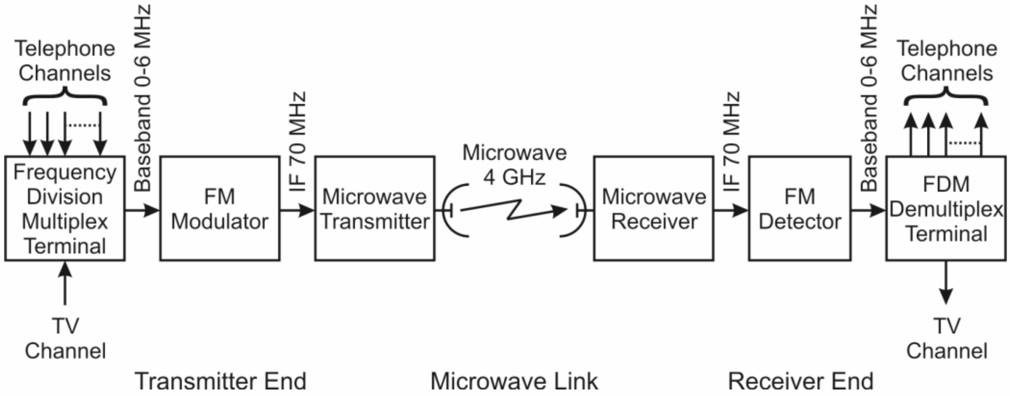

Fig. 11.30: A typical microwave communication system

After studying these essentials of microwaves, let us discuss some applications of microwaves. The first and historic application of microwave has been in Radars.

11.5 RADAR

Radar (Radio Detection and Ranging) is usually used all around us, although it is normally invisible. Air traffic control (ATC) uses radar to track planes both on the ground and in the air, and also to guide planes in for smooth landings. Police use radar to detect the speed of passing motorists. The military uses it to detect the enemy and to guide weapons. Meteorologists use radar to track storms, hurricanes and tornadoes. Radar can also be used to detect stationary objects buried underground. In some cases, radar can identify an object and its speed, if it is moving. Obviously, radar is an extremely useful technology.

The functioning of a Radar system is accomplished using two commonplace phenomena, viz. **echo** and **Doppler shift**. These two concepts are easy to understand in the realm of sound because your ears hear echo and Doppler shift every day. Radar makes use of the same techniques using radio waves.

You know that the echo of a sound can be used to determine how far an object is, and we can use the Doppler shift of the echo to determine how fast that object is moving. It is therefore possible to create “sound radar” or **sonar**. Submarines and boats use sonar all the time in water. You could use the same principles with sound in the air, but sound in the air has some of problems:

- Sound has a limited range of a kilometre or two at the most;
- Everyone can hear sounds, so a "sound radar" would disturb the other people (of course we can eliminate this problem by using ultrasound instead of audible sound); and
- Because the echo of the sound would be very faint, very good sensors would be required.

To overcome these problems, radar uses radio waves instead of sound. Radio waves travel far, do not disturb human beings around and are easy to detect even when they are faint.

Radar consists of a transmitter and a receiver, each connected to a directional antenna. The transmitter sends out UHF or microwave power through the antenna. The receiver collects the echoes reflected by the target in its direction and then after processing this information displays it in a suitable form. The receiving antenna very often also acts as the transmitting antenna. This is accomplished through a time-division multiplexing arrangement, since the radio energy is very often sent out in the form of pulses. Such radars are called the **Pulsed Radars**.

The block diagram of an elementary pulsed radar set is shown in Fig. 11.31. It consists of a transmitter and a receiver connected to the same antenna via a duplexer. You have learnt about the function of duplexer in the unit on telephony.

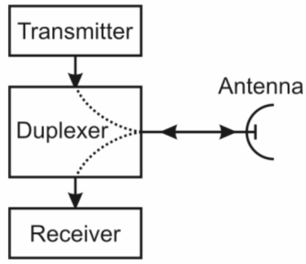

Fig. 11.31: Block diagram of an elementary pulsed radar

In response to an internally generated trigger signal, the transmitter generates a short, rectangular pulse. As soon as a small fraction of the pulse power is fed to the duplexer, this device disconnects the receiver from the antenna and connects the transmitter to it. In most radars, the antenna moves in a predetermined pattern, i.e. it *scans*. Since the antenna is highly directional, it sends out the generated pulse in the direction in which it is pointing at the time. The scanning speed may be mechanically high, but it is much slower compared with time taken by the pulses to return from a normal range of targets. Thus, when such echoes are received, the antenna still points in their direction to collect them.

As soon as the transmitted pulse terminates, the duplexer disconnects the transmitter from the antenna. The duplexer also reconnects the receiver to the antenna, allowing the returning echoes to be correctly processed. The received pulses are amplified and demodulated by the receiver. These pulses from the returning echoes are then processed and fed to the device on which they can be displayed. Once this cycle is completed, the set is once again ready for the transmission of the next pulse and the succeeding ones, while the antenna scans along its predetermined path.

The radar is able to show the position of the target, because information about the *azimuth* (horizontal direction) and the *elevation* (vertical direction) of the antenna is available. In addition, the distance to the target may be calculated from the total time taken by the pulse on its forward and return journeys. Since the velocity of electromagnetic radiations is approximately 300,000 km s^{-1}, and since the pulse must cover the distance to the target twice, it is easy to calculate that the return distance covered in 1μs is 300 m. Thus the target must be 150 m away if a pulse reflected by it is received 1 μs after sending. Nautical mile (nmi) (1852m) is the unit of distance often used in connection with radar distances, and it is seen that the time required for the return of a pulse is close to 12.4 μs nmi^{-1}. Finally, the velocity and direction of movement of the target relative to the radar set can be calculated from the positions of successive echoes.

In this discussion, you must have noted that the transmitted pulses have to be quite short compared to the intervals between successive pulses. A radar set must be capable of receiving echoes from targets located at the maximum distance for which it has been designed. Thus, if the range is to be 25 nmi, at least 310 μs must elapse from the time a pulse is sent out to the time at which the duplexer disconnects the receiver from the antenna for sending the next pulse. On the other hand, the pulse repetition frequency must be quite high compared to the scanning speed. Hence a limit exists on the minimum number of pulses that can be sent per second. Practically a typical system sends out several hundred pulses per second, having duration of about 1μs.

Pulse modulated magnetrons, klystrons, travelling wave tubes (another type of microwave source not discussed by us) are normally used as transmitter output tubes, and the first stage of the receiver is often a diode detector. The antenna generally uses a parabolic reflector, about which you have learnt in an earlier section.

Spend 2 Min.

SAQ 6

What is the maximum pulse width of a radar transmitter in order to capture a signal from an object 300 m away?

Let us now discuss another significant application of microwaves used extensively in modern days. This is in Satellite Communication.

11.6 SATELLITE COMMUNICATION

Developments in microwave technology made communication technology more and more complex, and thousands of independent signals now can be transmitted through a single channel. These systems, however, suffer from the handicap that their operation is restricted to line-of-sight (LOS) only. Two stations located beyond the horizon cannot communicate directly with each other.

In an LOS system if d is the distance between transmitting and receiving antennas (of heights h_1 and h_2 respectively) and the radius of the earth is 6370 km, then the LOS communication range is given by:

$$d = 3.57(\sqrt{h_1} + \sqrt{h_2}) \text{ Km.} \tag{11.20}$$

Repeater stations have to be placed at these distances for long-range, terrestrial communication. These difficulties are solved by employing a communication satellite orbiting around the earth to establish communication links between distant points.

The satellites can be used for various purposes like survey of earth surface using infrared camera to detect the natural resource; as weather satellites taking picture of cloud/storm structures; as surveillance satellites keeping track of enemy movements etc. However, the most important application of satellites is in the communication field. Depending on the assigned applications the satellites are placed in various orbits and they rotate around the earth at their assigned velocities.

The communication satellite is typically placed in a *geo-stationary* (or *geo-synchronous*) orbit so that it has a speed matching the speed of earth's rotation and hence such satellites keep pointing at the same point on the earth all the while. Such satellites are placed in the geo-synchronous orbits at 35,786 km above the earth's surface, so they take 24 hours to go around the earth.

The communication satellites have transponders which receive signals uplinked from the earth station and transmit them back to be picked up by the receivers on the earth.

11.6.1 Indian Satellites

India has been active in the field of satellite technology for more than three decades now. Under able leadership of Dr. Vikram Sarabhai, the Indian Satellite Programme kicked off. Government of India set up the Space Commission and Department of Space (DOS) in June 1972. Indian Space Research Organisation (ISRO) under DOS executes the Indian space programme through its establishments located in different places in India. The main objective of the space programme includes development of satellites, launch vehicles, Sounding Rockets and associated ground systems. The experiment phase of Indian space programme included Satellite Instructional Television Experiment (SITE), Satellite Telecommunication Experiment Programme (STEP), remote sensing application satellites like Aryabhata, Bhaskara, Rohini and some satellite launch vehicles like SLV-3 and ASLV.

Currently operational Indian space systems include Indian National Satellite (INSAT) for telecommunication, broadcasting, meteorology and disaster warning and Indian Remote Sensing Satellite (IRS) for natural resource management. ISRO has also developed the Polar Satellite Launch Vehicle (PSLV) used for launching IRS Satellites and Geosynchronous Satellite Vehicle (GSLV).

One of the greatest impacts of satellite technology in India has been in the field of education. You receive IGNOU's instructional programmes via Gyan Darshan Satellite Channels. Recently you must have also participated in the IGNOU programmes relayed from the EDUSAT, India's first satellite dedicated exclusively to educational purposes.

Let us now discuss the basic components and systems in a typical communication satellite.

11.6.2 Frequency Plans and Polarisations

There are well-defined frequency bands allocated for satellite use, depending on the type of service (for example, mobile communications and broadcast). The frequency bands also differ depending on the geographic region of the earth where the earth stations are located. Frequency allocations are made through the International Telecommunication Union (ITU). The most widely used bands at present are the *C* band and the *Ku* band.

The designation *Ku* arises from the fact that this frequency is under a microwave band known as the *K* band, and the *u* is sometimes shown as a subscript.

Uplink transmission in *C* band is nominally at 6 GHz (5.925 GHz - 6.425 GHz) and the downlink frequency is nominally at 4 GHz (3.7 GHz - 4.2 GHz). The band is sometimes referred to as the 6/4 GHz band. Uplink transmission in the *Ku* band takes place in the range of 14 GHz (14 GHz–14.5 GHz) and downlink in nominally 12 GHz (11.7GHz–12.2 GHz), this being referred to as the 14/12 GHz band. As you must have noted, the width of these bands is 500 MHz.

In all these cases, the higher frequency range is used for the uplink. The reason for using the higher frequency on the uplink is that losses tend to be greater at higher frequencies, and it is much easier to increase the power from an earth station rather than from a satellite to compensate for this.

To make the most of available bandwidth, *polarisation discrimination* scheme is used. A typical polarisation diagram is shown in Fig. 11.32.

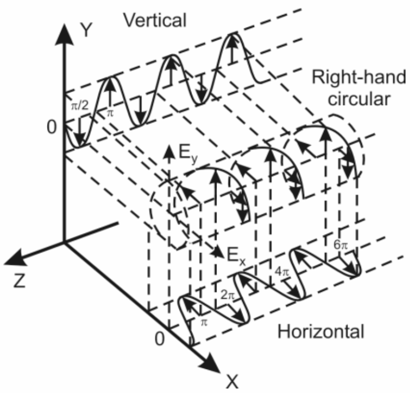

Fig. 11.32: Polarisation in electromagnetic waves

Adjacent transponder channels can be assigned alternate polarisations, for example horizontal and vertical. Fig.11.33 shows a typical frequency and polarisation plan for the *C* band in a satellite. The 24 transponder channels are first of all formed into two groups of 12, labelled *A* and *B* transponders. The downlink signals for group *A* are horizontally polarised and for group *B* vertically polarised. Thus, although there is some overlap in the transponder bandwidths, the different polarisations prevent interference from occurring. For example, transponder 2A has a centre frequency of 3760 MHz, and its bandwidth (including guard bands) extends from 3740 to 3780 MHz. It uses horizontal polarisation. Transponder 2B has a centre frequency of 3780 MHz, however the polarisation used here is vertical and hence the two signals do not interfere. The use of polarisation to increase the available frequency bandwidth is referred to as *frequency reuse*. You will also observe that the uplink signals in each group are polarised in the opposite sense to the downlink signals.

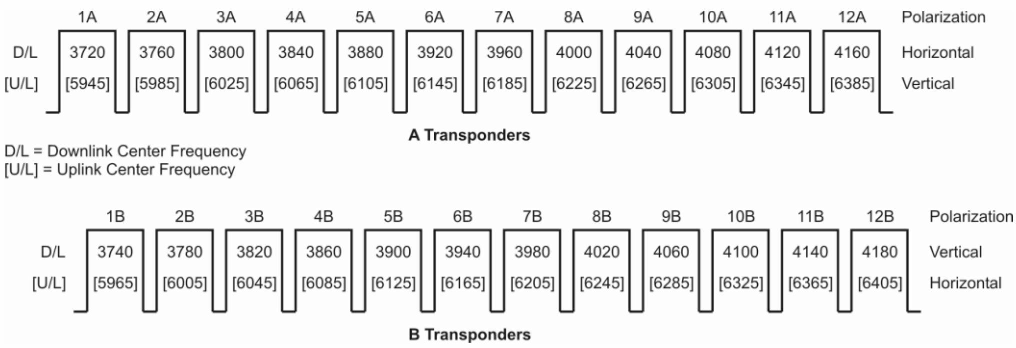

Fig. 11.33: A typical frequency and polarisation plan for the *C* band

11.6.3 Transponders

The word *transponder* is derived from *transmitter-responder*. This part of the satellite system connects the receiving antenna with the transmit antenna. The transponder itself is not a single unit of equipment, but consists of some units that are common to all transponder channels and others that can be identified with a particular channel. Fig. 11.34 shows a block schematic of typical transponder channels for *C* band satellite.

In majority of communication applications, a basic bandwidth of 500 MHz is available at the *C* band frequencies. The transponder caters to all channels encompassing the uplink frequency range (5.925 to 6.425 GHz). This input range of signals is passed through a wideband, band-pass filter (BPF) to limit any noise or interference; and then on to a wideband receiver, which provides a frequency down-conversion common to all channels. The output frequency range of this receiver is

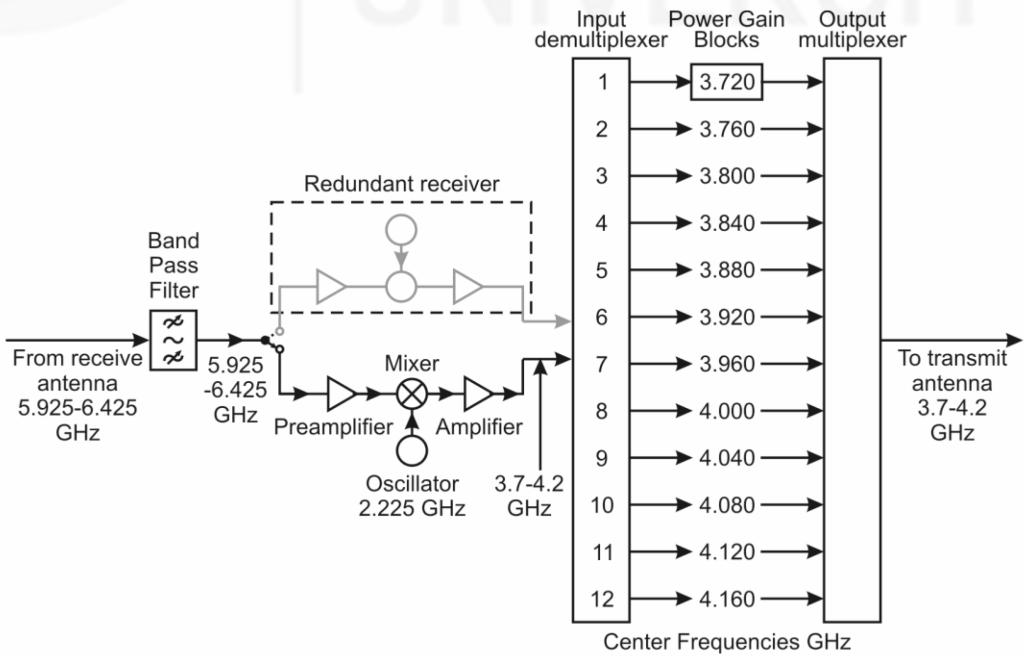

Fig. 11.34: *C* band satellite transponder channels

3.7 to 4.2 GHz, which is the downlink frequency band. The overall gain is provided in two sections, one at the input frequency range and the other at the output frequency range after the mixer. This makes for a more stable arrangement and prevents oscillation, which might arise if the total gain is provided at one frequency range. Solid-state amplifiers are used throughout.

Because the wideband receiver is critical to all transponders, a *redundant* receiver is provided. This is essentially a backup receiver that is switched on automatically if the other fails. An input *demultiplexer* following the wide-band receiver is a combination of microwave circulators and filters that separates the 500-MHz band into the separate transponder channel bandwidths. A typical transponder bandwidth is 36 MHz, or 40 MHz including guard bands. Following the demultiplexer, power amplifiers are provided for the individual transponder channels, which bring the power levels up to those required for retransmission on the downlink.

11.6.4 Power Systems

A satellite stays in orbit essentially as a result of natural forces and in the absence of external disturbances would orbit the earth indefinitely without having to carry fuel for propulsion. In practice disturbance forces exist, hence satellites must carry fuel on board so that corrective forces can be applied from time to time, usually through thruster jets. The finite capacity to carry fuel imposes one of the major limitations on the useful life of a satellite which may range from a few years to decades.

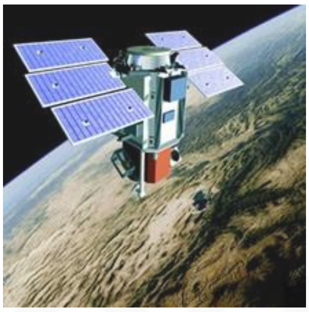

Fig. 11.35: Satellite with Solar Sails

In addition, the satellite must receive energy to power the electronic equipment on board. This is invariably supplied by solar cells. With cylindrically shaped satellites, these are arranged around the body of the satellite. The advantage of the cylindrical arrangement is that the satellite can be set spinning to maintain its position through the gyroscopic effect, but with this arrangement only about one-third of the satellite body is illuminated by the sun at any given time, and so the power available is limited.

An alternative arrangement is to employ solar sails, as shown in Fig. 11.35. With this type of construction, spin stabilisation cannot be used and other methods like thrusters jets have to be used. The orientation of the solar sails can be adjusted automatically for maximum solar illumination, so high power outputs can be obtained.
For a period of about 45 days around each equinox, the satellite is eclipsed by the earth, the eclipse lasting for a maximum period of around 70 min at its peak during each eclipse. Battery backup supplies must be provided during these periods, and long-life batteries have been especially developed for this purpose. You might have read that our national stock exchanges suspend their operations during the satellite eclipse period, since they use satellite communication to keep track of transactions of other international stock exchanges and this gets interrupted during the satellite eclipse period.

The advances in technology are bringing in better and better performance of satellite communication and you may read about them in advanced books later on.

Let us now summarise the points covered in this unit.

11.7 SUMMARY

- Microwave communication uses line-of-sight method for terrestrial transmission, while satellites are used for broad area transmission.
- Line-of-sight communication is restricted by space attenuation (atmospheric absorption) and curvature of earth's surface.

- Horns and paraboloidal reflectors are commonly used antenna at microwave frequencies.
- The reflector type antenna needs a primary feeder antenna which can be a dipole or a horn antenna.
- Geometry of reflector and placement of feeder decide the directivity of the antenna.
- Reflex klystron works on the principle of bunching of electrons caused by proper arrangement of potentials across the cavity and repeller plate.
- Tunnel diode generates microwave frequencies while operating in the negative differential resistance range.
- Gunn devices operate on the principle of inter-valley transfer of electrons within various unoccupied bands in direct band gap semiconductors. Typically GaAs is used in Gunn devices.
- Typical microwave detectors are Schottky barrier diodes, point contact diodes, varactor diodes, *p*-*i*-*n* diodes etc.
- Different microwave waveguide components include attenuator, isolator, magic *T*, circulator, directional coupler, resonant cavities etc.
- VSWR measurement is done by probing a standing wave pattern in a slotted waveguide.
- Pulsed radar uses same antenna for transmit and receive of pulses via a duplexer.
- Geostationary satellites are used for satellite communication. They are placed at 35,786 km orbit around the earth.
- Satellite uplink frequencies are higher than the downlink frequencies. *C* and *Ku* are commonly used frequency bands in satellites.
- A transponder in satellite down converts the frequency of received signal and transmits it back after proper amplification.
- Solar energy is the commonly used power source in satellites.

11.8 TERMINAL QUESTIONS *Spend* 25 *Minutes*

1. Calculate the angular aperture for a paraboloid reflector antenna for aperture number of 0.5. Given that the diameter of the reflector mouth is 20 m, calculate the position of focal point.

2. How is the velocity modulation achieved in a reflex klystron? How is it converted into current modulation?
3. Explain the process giving rise to negative differential resistance in a direct band gap semiconductor material.

4. Why is *n*-type GaAs used as a channel material in MESFETs?

5. Explain how a magic *T* can be used as a duplexer in a Radar.

11.9 SOLUTIONS AND ANSWERS

Self Assessment Questions

1. Total attenuation = 20 log [10/1] + 0.1[10 – 1] dB = 20.9 dB

2. $\lambda = \frac{c}{f} = \frac{300 \times 10^6}{10 \times 10^9} = 0.03\,\text{m} = 3\,\text{cm}$

   $A = \frac{\pi D^2}{4} = \frac{3.14 \times 6^2}{4} = 28.27\,\text{m}^2$

   $A_{eff} = 0.65A = 18.4\ \text{m}^2$

   $D_0 = \frac{4\pi}{\lambda^2} A_{eff} = 257{,}000$ i.e. 54.1 dB

   Beamwidth, $BW_{(-3\text{dB})} = \frac{70\lambda}{D} = \frac{70 \times 0.03}{6} = 0.35^\circ$

3. Electron drift velocity $v_d = f \times L = 20 \times 10^9 \times 10 \times 10^{-6} = 2 \times 10^5$ m s^{-1}.
   Applied voltage = $E \times L = (4 \times 10^3 \times 10^2)$ V m^{-1} $\times (10 \times 10^{-6})$ m = 4 V.

4. The dc component $= \frac{a^2 V_1^2 I_o}{4}$ = 0.25 (40)^2 (1)^2 10^{-6} = 0.4 mA

5. The input power P_1 = 1 W
   The coupled power at Port 4 = P_4 / P_1 = −10dB. ∴ P_4 = 0.1 W.
   The power output at Port 3 = 0, as the directivity is infinite, decoupled; and the power output at Port 2 = 0.9 W.

6. A pulse takes 2μs to return from an object 300 m away. Hence the radar pulse has to end before this period in order to be ready to receive the reflected signal.

Terminal Questions

1. Given that $\frac{f}{D} = 0.5$
   for D = 20 m, f = 10 m.

   From Eq. (11.4)

   $\frac{f}{D} = 0.25 \cot\left(\frac{\psi}{2}\right)$ and hence,

   $\psi = 2 \cot^{-1}\left(4\frac{f}{D}\right) = 2 \cot^{-1} 2$

2. The velocity modulation happens in klystron cavity when alternating (microwave frequency) voltage is applied across the cavity. At various amplitudes of this voltage, the electrons getting subjected to it pick up different amounts of energies, thereby varying their velocities. This is in effect the velocity modulation.

   When these electrons travel in repeller space, they reach different distances before getting repelled by the repeller. Hence they encounter bunching, which is effectively current modulation.

3. Refer to Sec 11.4.1c.

4. MESFET works on the principle of Schottky (metal-semiconductor) contact. The Schottky devices work efficiently as rectifiers when an *n*-type semiconductor and metal with proper work function are brought in contact, so that there is transfer of electrons from the semiconductor to metal with electrons as majority (and only majority) carriers.

5 A magic *T* can be used as a duplexer in the Radar circuit. The radar transmitter and the receiver are connected to the Ports 2 and 3. The antenna is connected to Port 1 and Port 4 is terminated in a matched load. During transmission, half of the power goes to the antenna and the other half is absorbed in the matched load at Port 4. No transmitter power reaches the receiver. During the reception time, the received power at Port 1 is divided between the receiver at Port 3 and the transmitter at Port 2. The transmitter can be protected from received signal by using as isolator.

Reference Material:

1. *Electronic Communications* by Roddy, Dennis and Coolen, John; (IV Edition) (Prentice-Hall of India)
2. *Electronic and Radio Engineering* by Terman, F.E.; (IV Edition) (McGraw-Hill International)
3. *Microwave Engineering* by Swarup, Prem; (I Edition) (Cyber Tech Publications)
4. www.isro.org