**SRI VENKATESHWARAA COLLEGE OF ENGINEERING & TECHNOLOGY**
(Approved by AICTE, New Delhi & Affiliated to Pondicherry University, Puducherry.)
13-A, Villupuram – Pondy Main road, Ariyur, Puducherry – 605 102.
Phone: 0413-2644426, Fax: 2644424 / Website: www.svcetpondy.com

ELECTRONICS AND COMMUNICATION ENGINEERING

# ECT64-Antennas and Wave Propagation

YEAR-III SEMESTAR-VI

## UNIT-I

## ANTENNA FUNDAMENTALS

### INTRODUCTION

Antennas are device designed to radiate electromagnetic energy efficiently in a prescribed manner. It is the current distributions on the antennas that produce the radiation. Usually these current distributions are excited by transmission lines or waveguides. In two-way communication, the same antenna can be used for transmission and reception

An antenna is a circuit element that provides a transition form a guided wave on a transmission line to a free space wave and it provides for the collection of electromagnetic energy. The antenna is the transition between a guiding device (transmission line, waveguide) and free space (or another usually unbounded medium).Its main purpose is to convert the energy of a guided wave into the energy of a free-space wave (or vice versa) as efficiently as possible, while in the same time the radiated power has a certain desired pattern of distribution in space.

In transmit systems the RF signal is generated, amplified, modulated and applied to the antenna. In receive systems the antenna collects electromagnetic waves that are “cutting” through the antenna and induce alternating currents that are used by the receiver

a) transmission-line Thevenin equivalent circuit of a radiating
(transmitting) system

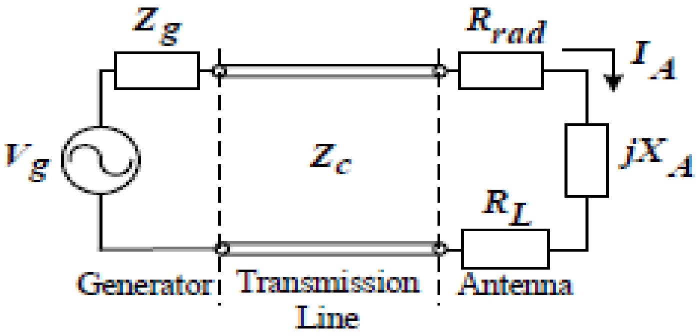


**Fig 1.1: Thevenin equivalent circuit of a transmitting antenna**

*Vg* - voltage-source generator (transmitter);
*Zg* - impedance of the generator (transmitter);

*Rrad* - radiation resistance (related to the radiated power

$$Prad = I A^{2} *Rrad$$

*RL* - loss resistance (related to conduction and dielectric losses);

*jX A* - antenna reactance.

Antenna impedance: ( ) $ZA = Rrad \; /RL + jX A$

b) transmission-line Thevenin equivalent circuit of a receiving antenna system

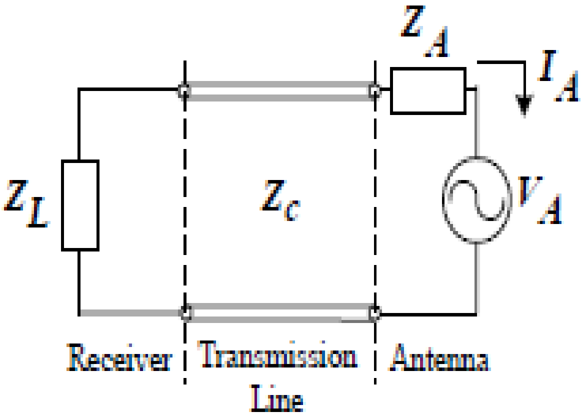


**Fig 1.2: Thevenin equivalent circuit of a receiving antenna**

The antenna is a critical component in a wireless communication system. A good design of the antenna can relax system requirements and improve its overall performance.

## RADIATION PATTERN

A radiation pattern (or field pattern) is a graph that describes the relative far field value, E or H, with direction at a fixed distance from the antenna. A field pattern includes an magnitude pattern |E| or |H| and a phase pattern ∠E or ∠H.Radiation pattern is an indication of radiated field strength around the antenna. Power radiated from a $\lambda/2$ dipole occurs at right angles to the antenna with no power emitting from the ends of the antenna. Optimum signal strength occurs at right angles or 180° from opposite the antenna. The radiation pattern or antenna pattern describes the relative strength of the radiated field in various directions from the antenna, at a constant distance.

The radiation pattern is a reception pattern as well, since it also describes the receiving properties of the antenna. The radiation pattern is three dimensional, but usually the measured radiation patterns are a two dimensional slice of the three-dimensional pattern, in the horizontal or vertical planes. These pattern measurements are presented in either a rectangular or a polar format. If the radiation from the antenna is expressed in terms of field strength then the radiation pattern is called field pattern. If the radiation from the antenna is expressed in terms of power then the radiation pattern is called power pattern.

**MAJOR LOBE-** It is also called as main beam and is defined as the radiation lobe containing the direction of maximum radiation. In some antennas, there may exist more than one major lobe.

**MINOR LOBE-** A minor lobe is any lobe except a major lobe (ie).,all the lobes except the major lobe are called minor lobe.

**SIDE LOBE-** It is adjacent to the main lobe and occupies the hemisphere in the direction of the main lobe.

**BACK LOBE-** It occupies the hemisphere in a direction opposite to that of the major lobe. Its axis makes an angle of approximately $180^0$ with respect to the beam of an antenna.

Minor lobes usually represent radiation in undesired directions and they should be minimized . Side lobes are normally the largest of minor lobes. The level of minor lobes is usually expressed as the ratio of power density in the lobe in question to that of the major lobe. This ratio is termed as side lobe ratio or side lobe level. By the reciprocity theorem, the radiation patterns of an antenna in the transmitting mode is same as the those for the antenna in the receiving mode.

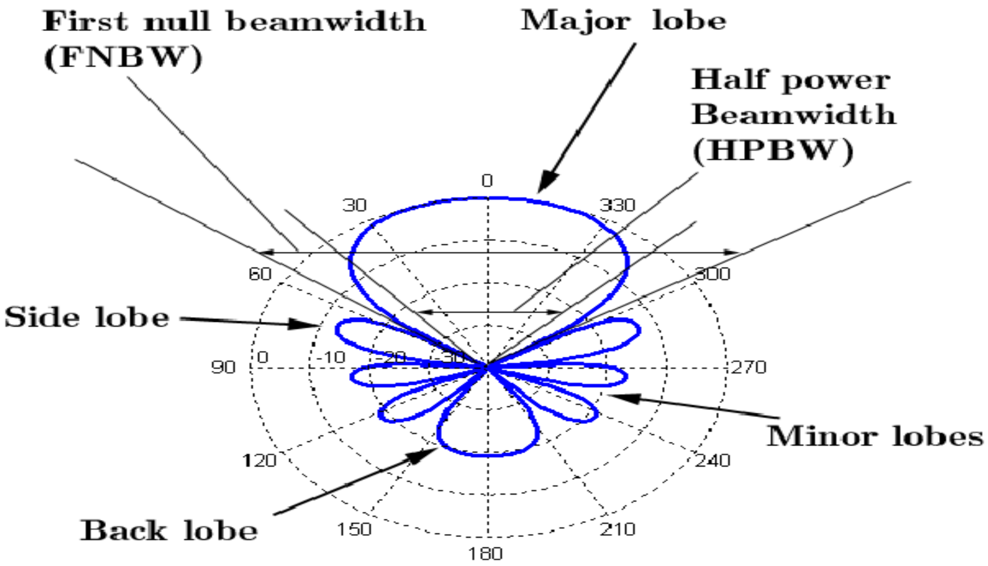


**Fig 1.3: Radiation pattern**

## RADIATION INTENSITY

The power radiated from an antenna per unit solid angle is called the radiation intensity U. The unit of radiation intensity in watts/ steradian or watts/ radian$^2$. The radiation intensity can also be obtained by simply multiplying the radiation density by the square of the distance.

In mathematical form, it can be expressed as,

$$U=W_{rad}\, r^2$$

Where U= radiation intensity (W/Sr)

$W_{rad}$ =radiation density (W/m$^2$)

The total power is obtained by integrating radiation intensity over the entire solid angle.

$P_{rad}=\iint U d\Omega = \iint U \sin\theta \, d\theta d\varphi$

Where $d\Omega$ =solid angle= $\sin\theta \, d\theta \, d\varphi$

The measure of plane angle is a radian. The measure of solid angle is sterdian. One steradian is defined as the solid angle with its vertex at the centre of sphere of radius r that is subtended by a spherical surface area equal to that of a square with each side of length r.

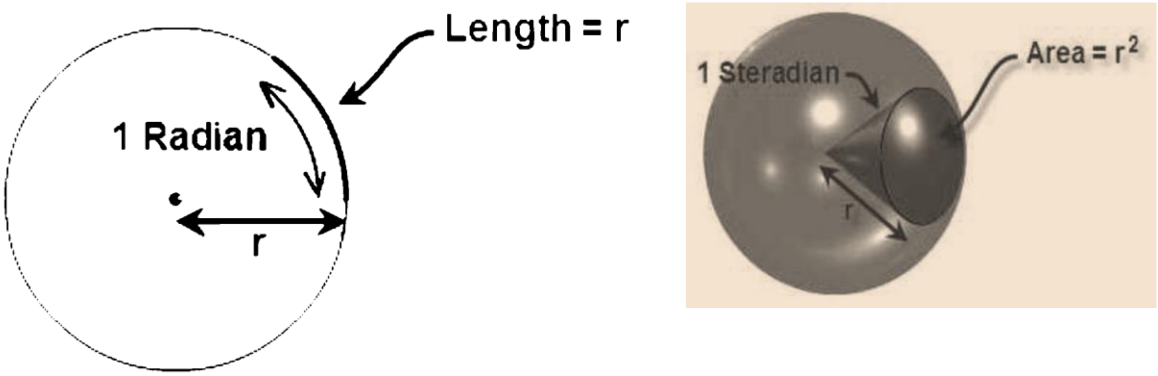


Fig 1.4: Representation of radian and sterdian

## GAIN

It is defined as the ability of the antenna to concentrate the radiated power in a given direction. Gain is a relative term in which actual antenna is compared with a reference antenna. It is defined as the ratio of maximum radiation intensity in a given antenna to the maximum radiation intensity from a reference antenna produced in the same direction.

G = max radiation intensity from test antenna ⁄ Max radiation intensity from a reference antenna

## DIRECTIVE GAIN

The extend to which a practical antenna concentrates its radiated energy relative to that of standard antenna is called as directive gain.

It is defined as the ratio of the radiation intensity in that direction to the average radiated power. It is a function of angles Θ and Φ.

$G_d = U(\Theta , \Phi) / U_{avg}$

Where $U_{avg} = W_r / 4\pi$

$W_r$ is the radiated power.

$G_d = U(\Theta , \Phi) / W_r / 4\pi$

$G_d = 4\pi\, U(\Theta , \Phi) / W_r$

$G_d = 4\pi\, U(\Theta , \Phi) \,/\, \int U \, d\Omega$

It is a qualitive measure of the extent to which the total radiated power is concentrated in one direction. It depends only on the distribution of radiated power in space.

## POWER GAIN

It compares the radiated power density of the actual antenna and that of the isotropic antenna on the basis that both are fed with the same input power.

$G_p$= Radiation intensity in a given direction / Average total input power.

$G_p = U(\Theta , \Phi) \,/\, W_T / 4\pi$

Where $W_{T=} W_r + W_l$

$W_r$ =Radiated power

$W_l$ =Ohmic losses in antenna

$G_p = 4\pi\, U(\Theta , \Phi) \,/\, W_T$

Power gain is also defined as the ratio of the power input supplied to test antenna to the power supplied to the reference antenna.

Factors depends on power gain are:

i) Sharpness of lobe

ii) Volume of the radiation pattern

## DIRECTIVITY

It is defined as the ratio of maximum radiation intensity $U(\Theta , \Phi)$ to the avg radiation intensity, $U_{avg.}$ The directivity is a dimensionless quantity. The maximum directivity is always less than 1

$$D = \frac{U(\theta,\phi)_{\max}}{U(\theta,\phi)_{average}}$$

## BANDWIDTH

This is the range of frequencies, within which the antenna characteristics(input impedance, pattern) conform to certain specifications . generally the range of frequencies over which the antenna system‘s SWR remains below a maximum value, typically 2.0 Antenna characteristics, which should conform to certain requirements, might be: input impedance, radiation pattern, beamwidth, polarization, side-lobe level, gain, beam direction and width, radiation efficiency. Separate bandwidths may be introduced: impedance bandwidth, pattern bandwidth, etc.

## BEAMWIDTH

Beamwidth is associated with the lobes in the antenna pattern. It is defined as the angular separation between two identical points on the opposite sides of the main lobe. The most common type of beamwidth is the half-power (3dB) beamwidth (HPBW). To find HPBW, in the equation, defining the radiation pattern, we set power equal to 0.5 and solve it for angles. Another frequently used measure of beamwidth is the first-null beamwidth (FNBW),which is the angular separation between the first nulls on either sides of the main lobe. Beamwidth defines the resolution capability of the antenna: i.e., the ability of the system to separate two adjacent targets

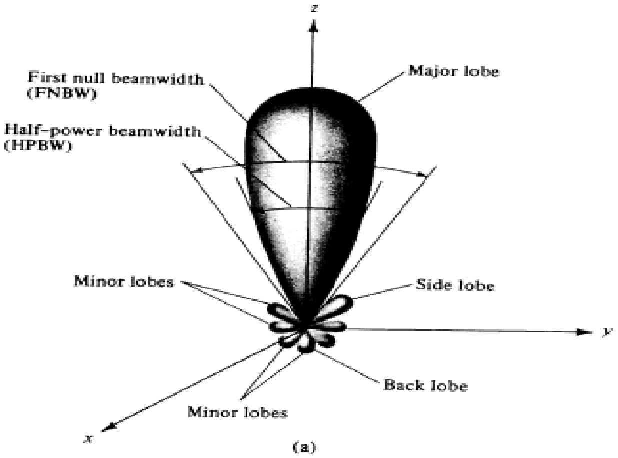


**Fig 1.5: Pattern in spherical co-ordinate system**

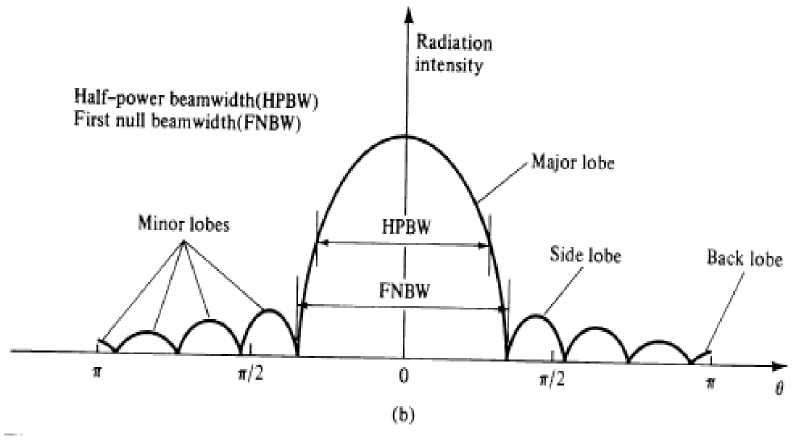


**Fig 1.6: Pattern in Cartesian co-ordinate system**

## INPUT IMPEDANCE:

This is the impedance measured at the antenna input terminals. In general it is complex and has two real parts and one imaginary part:

Radiation resistance: - represents conversion of power into RF waves (real)

Loss resistance – represents conductor losses, ground losses, etc. (real)

reactance – represents power stored in the near field (imaginary)

$$Z_A = R_A + jX_A$$

$R_A$ = input resistance
$X_A$ = input reactance

## RADIATION RESISTANCE

The power flowing through a circuit is , where V is the voltage (defined as energy per unit charge) and I is the current (defined as charge flow per unit time), so P has dimensions of energy per unit time. The physicist George Simon Ohm observed that the current flowing through most materials is proportional to the applied voltage, so many (but not all) objects have a well-defined resistance defined by R =v⁄ I (Ohm's law).

.

From Ohm's law for time-varying currents,

$$\langle P \rangle = \langle I^2 \rangle R$$

If

$$I = I_0 \cos(\omega t), \qquad \langle P \rangle = \frac{I_0^2 R}{2}$$

The radiation resistance of an antenna is defined by

$$R \equiv \frac{2\langle P \rangle}{I_0^2}$$

## POLARIZATION

The polarization of an antenna in a given direction is defined as the polarization of the plane wave transmitted by the antenna in that direction. The polarization of a plane wave is the figure the tip of the instantaneous electric-field vector E traces out with time at a fixed observation point. There are three types of typical antenna polarizations: the linear, circular, and elliptical polarizations, corresponding to the same three types of typical plane wave polarizations.

## LINEAR POLARIZATION

A plane wave is linearly polarized at a fixed observation point if the tip of the electric-field vector at that point moves along the same straight line at every instant of time.

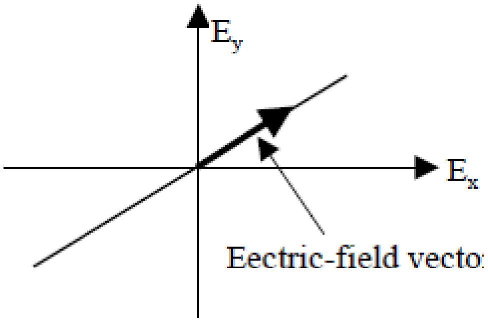


Fig 1.7: Linear polarisation

## CIRCULAR POLARIZATION

A plane wave is circularly polarized at a a fixed observation point if the tip of the electric-field vector at that point traces out a circle as a function of time. Circular polarization can be either right-handed or left-handed corresponding to the electric-field vector rotating clockwise (right-handed) or anticlockwise(left-handed).

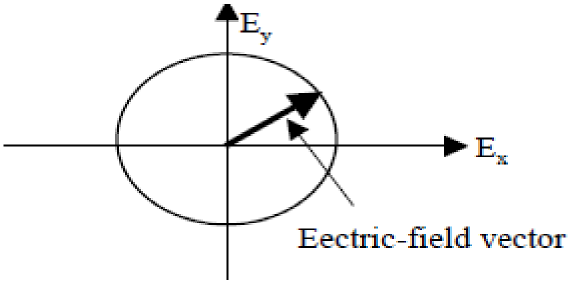


Fig3.8: circular polarisation

## ELLIPTICAL POLARIZATION

A plane wave is elliptically polarized at a a fixed observation point if the tip of the electric-field vector at that point traces out an ellipse as a function of time. Elliptically polarization can be either right-handed or left-handed corresponding to the electric-field vector rotating clockwise(right-handed) or anti-clockwise (left-handed).

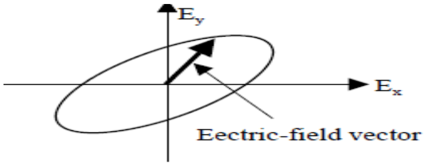
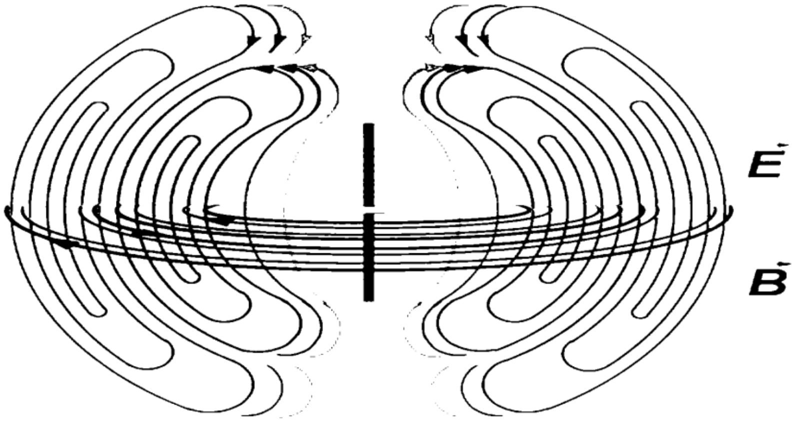


Fig1.9: Elliptical polarisation

## ANTENNA FIELD ZONES

The space surrounding the antenna is divided into three regions according to the predominant field behaviour. The boundaries between the regions are not distinct and the field behaviour changes gradually as these boundaries are crossed.



Radiation from a dipole

1.**Reactive near-field region**: This is the region immediately surrounding the antenna, where the reactive field dominates. For most antennas, it is assumed that this region is a sphere with the antenna at its centre.

2. **Radiating near-field (Fresnel) region** :This is an intermediate region between the reactive near-field region and the far-field region, where the radiation field is more significant but the angular field distribution is still dependent on the distance from the antenna.

3. **Far-field (Fraunhofer) region** :Here r >> D and r >> λ. The angular field distribution does not depend on the distance from the source any more, i.e., the far-field pattern is already well established.

## RADIATION FROM A CURRENT ELEMENT

To find the fields radiated by the current element, the two-step procedure is used. It will be required to determine first A and F and then find the E and H. An infinitesimal linear wire *(l <<λ)* is positioned symmetrically at the origin of the coordinate system and oriented along the *z* axis. The wire, in addition to being very small *(l <<λ)*, is very thin *(a <<λ)*. The spatial variation of the current is assumed to be constant and given by

$$\mathbf{I}(z') = \hat{\mathbf{a}}_z I_0$$

Since the source only carries an electric current I*e*, I*m* and the potential function Fare zero. To find A we write

$$\mathbf{A}(x, y, z) = \frac{\mu}{4\pi} \int_C \mathbf{I}_e(x', y', z') \frac{e^{-jkR}}{R} \, dl'$$

where (*x, y, z* ) represent the observation point coordinates, *(x',y', z')* represent the coordinates of the source, *R* is the distance from any point on the source to the observation point, and path *C* is along the length of the source.

$$\mathbf{I}_e(x', y', z') = \hat{\mathbf{a}}_z I_0$$

$$x' = y' = z' = 0 \text{ (infinitesimal dipole)}$$

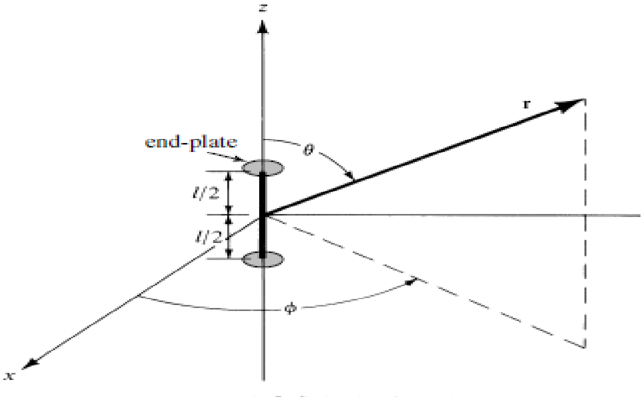
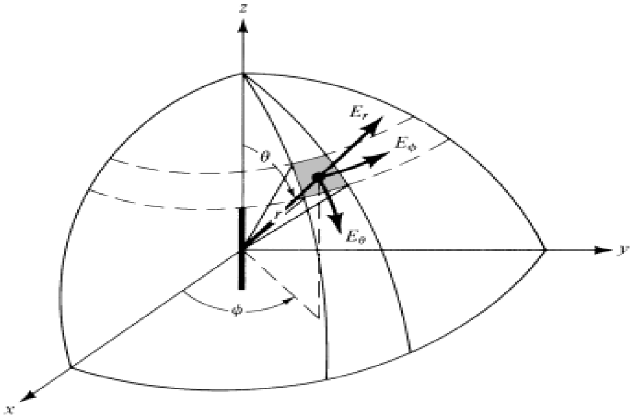


Fig 3.11:Infinitsmall dipole



ELECTRIC FIELD DISTRIBUTION

$$R = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2} = \sqrt{x^2 + y^2 + z^2}$$

$$= r = \text{constant}$$

$$dl' = dz'$$

The next step of the procedure is to find H*A* using (3-2a) and then E*A* with J = 0. To do this, it is often much simpler to transform from rectangular to spherical components

$$\begin{bmatrix} A_r \\ A_\theta \\ A_\phi \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix}$$

$$\boxed{\mathbf{A}(x, y, z) = \hat{\mathbf{a}}_z \frac{\mu I_0}{4\pi r} e^{-jkr} \int_{-l/2}^{+l/2} dz' = \hat{\mathbf{a}}_z \frac{\mu I_0 l}{4\pi r} e^{-jkr}}$$

$$A_x = A_y = 0$$

$$A_r = A_z \cos\theta = \frac{\mu I_0 l e^{-jkr}}{4\pi r} \cos\theta$$

$$A_\theta = -A_z \sin\theta = -\frac{\mu I_0 l e^{-jkr}}{4\pi r} \sin\theta$$

$$A_\phi = 0$$

Using the symmetry of the problem (no $\varphi$ variations), can be expanded in spherical coordinates and written in simplified form as

$$\mathbf{H} = \hat{\mathbf{a}}_\phi \frac{1}{\mu r} \left[ \frac{\partial}{\partial r}(r A_\theta) - \frac{\partial A_r}{\partial \theta} \right]$$

$$\mathbf{J} = 0.$$

$$\boxed{\begin{gathered} H_r = H_\theta = 0 \\ H_\phi = j \frac{k I_0 l \sin\theta}{4\pi r} \left[ 1 + \frac{1}{jkr} \right] e^{-jkr} \end{gathered}}$$

$$\boxed{\begin{aligned} E_r &= \eta \frac{I_0 l \cos\theta}{2\pi r^2}\left[1 + \frac{1}{jkr}\right] e^{-jkr} \\ E_\theta &= j\eta \frac{k I_0 l \sin\theta}{4\pi r}\left[1 + \frac{1}{jkr} - \frac{1}{(kr)^2}\right] e^{-jkr} \\ E_\phi &= 0 \end{aligned}}$$

$$\mathbf{E} = \mathbf{E}_A = -j\omega\mathbf{A} - j\frac{1}{\omega\mu\epsilon}\nabla(\nabla \cdot \mathbf{A}) = \frac{1}{j\omega\epsilon}\nabla \times \mathbf{H}$$

## RADIATION FROM A HALF WAVE DIPOLE

Half wave dipole is the simplest antenna and is used in complex systems like antenna arrays. It is a fundamental antenna of metal rod which has a physical length of half wavelength in free space. It is also called Hertz antenna. Dipole antenna is a symmetrical antenna whose two ends are at equal potential. Dipole is fed at the centre having maximum current at the centre.

$$I = I_m \sin\beta\,(h - z) \text{ for } z > 0$$
$$I = I_m \sin\beta\,(h + z) \text{ for } z < 0$$

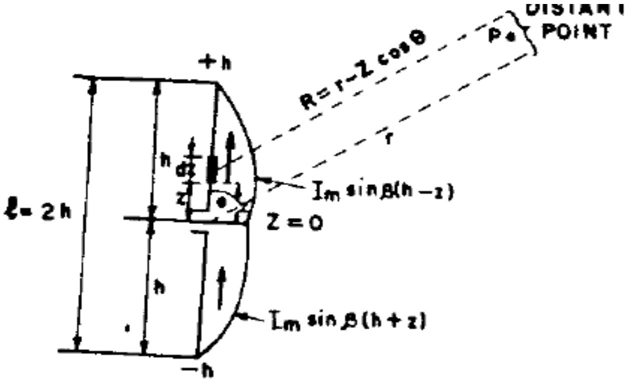


Fig 3.13:Half Wave Dipole

where $I_m$ = current maximum at the current loop.

Now vector potential at a distant point $P$ due to current element $I\,dz$ is given by

$$d A_z = \frac{\mu I\, dz . e^{-j\beta R}}{4\pi R} \quad \ldots$$

where $R$ = Distance between $I\,dz$ to distant point. The total vector potential due to all such current elements at distant point $P$ is given by

$$\int d A_z = \int_{-h}^{0} \frac{\mu I\, dz\, e^{-j\beta R}}{4\pi R} + \int_{0}^{+h} \frac{\mu I dz\ .\ e^{-j\beta R}}{4\pi R}$$

$$A_z = \frac{\mu I_m e^{-j\beta r}}{4\pi\beta r}\left[\frac{(1-\cos\theta)\cos(\pi/2\cos\theta)\ (1+\cos\theta)\cos(\pi/2\cos\theta)}{\sin^2\theta}\right]$$

$$= \frac{\mu I_m\ e^{-j\beta r}}{4\pi\beta r}\left[\frac{\cos(\pi/2\cos\theta)\ (1-\cos\theta+1+\cos\theta)}{\sin^2\theta}\right]$$

$$= \frac{\mu I_m e^{-j\beta r}}{4\pi r}\left[\frac{\cos(\pi/2\cos\theta)\ (2)}{\sin^2\theta}\right]$$

$$\because \left[\sin\beta z\,(1+\cos\theta)\right]_0^{\lambda/4} = \sin\frac{2\pi}{\lambda}\cdot\frac{\lambda}{4}(1+\cos\theta)$$

$$= \sin\frac{\pi}{2}(1+\cos\theta) = \cos(\pi/2\cos\theta)$$

$$\boxed{A_z = \frac{\mu I_m\ e^{-j\beta r}}{2\pi\beta r}\left[\frac{\cos(\pi/2\cos\theta)}{\sin^2\theta}\right]}$$

$$\mu H_\varphi = (\nabla\times H)_\varphi = \frac{1}{r}\left[\frac{\partial}{\partial r}(A_\theta\,.\,r)\right] = +\frac{1}{r}\left[\frac{\partial}{\partial r}(-A_z\sin\theta\,.\,r)\right]$$

$$\boxed{\mu H_\varphi = -\sin\theta\,\frac{\partial A_z}{\partial r}}$$

$$\mu H_\varphi = -\frac{\partial}{\partial r}\left[\frac{\mu I_m\ e^{-j\beta r}}{2\pi\beta r}\left\{\frac{\cos(\pi/2\cos\theta)}{\sin^2\theta}\right\}\right]\sin\theta$$

$$\mu H_\varphi = -\frac{\mu I_m\ e^{-j\beta r}\ (-j\beta)}{2\pi\beta r}\left\{\frac{\cos(\pi/2\cos\theta)}{\sin\theta}\right\}$$

$$H_\varphi = +\frac{j I_m\ e^{-j\beta r}}{2\pi r}\left\{\frac{\cos(\pi/2\cos\theta)}{\sin\theta}\right\}$$

$$|H_\varphi| = \left|\frac{I_m e^{-j\beta r}}{2\pi r}\ \frac{\cos(\pi/2\cos\theta)}{\sin\theta}\right|$$

$$\frac{E_\theta}{H_\varphi} = \eta = 120\pi$$

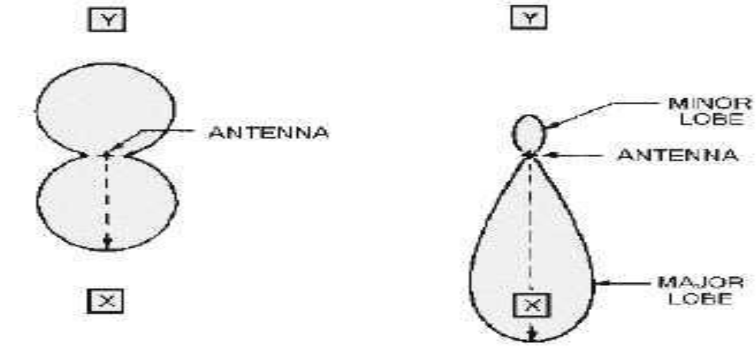
$$|E_\theta| = 120\pi\,|H_\varphi|$$

$$|E_\theta| = 120\pi \cdot \frac{I_m}{2\pi r}\left\{\frac{\cos(\pi/2\cos\theta)}{\sin\theta}\right\}$$

$$\boxed{|E_\theta| = \frac{60 I_m}{r}\left\{\frac{\cos(\pi/2\cos\theta)}{\sin\theta}\right\} \text{ volt/metre}}$$

## ARRAY ANTENNAS

An ARRAY is a combination of half-wave elements operating together as a single antenna. An array antenna is made up of more than one ELEMENT, but the basic element is generally the dipole. It provides more gain and greater directivity than single element antennas. Sometimes the basic element is made longer or shorter than a half-wave, but the deviation usually is not great. A DRIVEN element is similar to the dipole you have been studying and is connected directly to the transmission line. A DRIVEN ARRAY derives its power directly from the source. It obtains its power directly from the transmitter or, as a receiving antenna, it delivers the received energy directly to the receiver.



A. SINGLE ANTENNA PATTERN B. ARRAY PATTERN

Fig 3.14:Radiation pattern

Advantages of using antenna arrays

1. They can provide the capability of a steerable beam (radiation direction change) as in smart antennas.
2. They can provide a high gain (array gain) by using simple antenna elements.

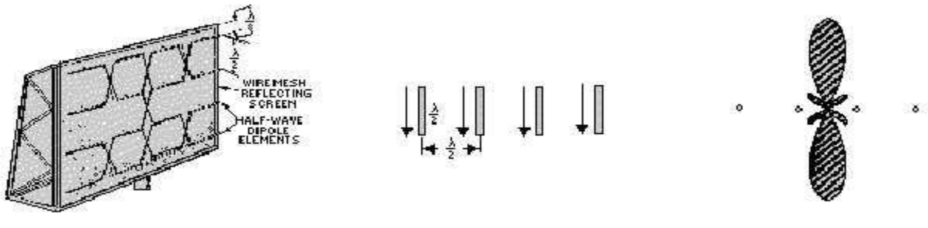
3. They provide a diversity gain in multipath signal

reception.

4. They enable array signal processing.

## BROADSIDE ARRAY

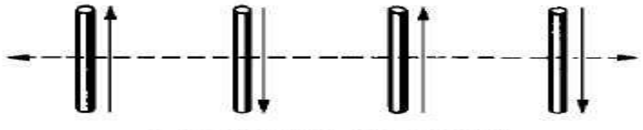
Physically, it looks somewhat like a ladder. When the array and the elements in it are polarized horizontally, it looks like an upright ladder. When the array is polarized vertically, it looks like a ladder lying on one side (view B). View C is an illustration of the radiation pattern of a broadside array. Horizontally polarized arrays using more than two elements are not common. This is because the requirement that the bottom of the array be a significant distance above the earth presents construction problems. Compared with collinear arrays, broadside arrays tune sharply, but lose efficiency rapidly when not operated on the frequencies for which they are designed.



A. ARRAY B. SIDE VIEW OF ARRAY C. TOP VIEW OF ARRAY

## END-FIRE ARRAYS

An end-fire array looks similar to a broadside array. The ladder-like appearance is characteristic of both (view A). The currents in the elements of the end-fire array, however, are usually 180 degrees out of phase with each other as indicated by the arrows. The construction of the end-fire array is like that of a ladder lying on its side (elements horizontal). The dipoles in an end-fire array are closer together (1/8-wavelength to 1/4 -wavelength spacing) than they are for a broadside array.



A. TOP VIEW OF ARRAY

B. SIDE VIEW OF ARRAY

Fig 3.16:Endfire array

Fig 3.15:Broadside array

Closer spacing between elements permits compactness of construction. For this reason an end-fire array is preferred to other arrays when high gain or sharp directivity is desired in a confined space. However, the close coupling creates certain disadvantages. Radiation resistance is extremely low, sometimes as low as 10 ohms, making antenna losses greater. The end-fire array is confined to a single frequency. With changes in climatic or atmospheric conditions, the danger of detuning exists.

## ARRAY OF TWO POINT SOURCES

### Case1:

2 isotropic point sources of same amplitude and phase

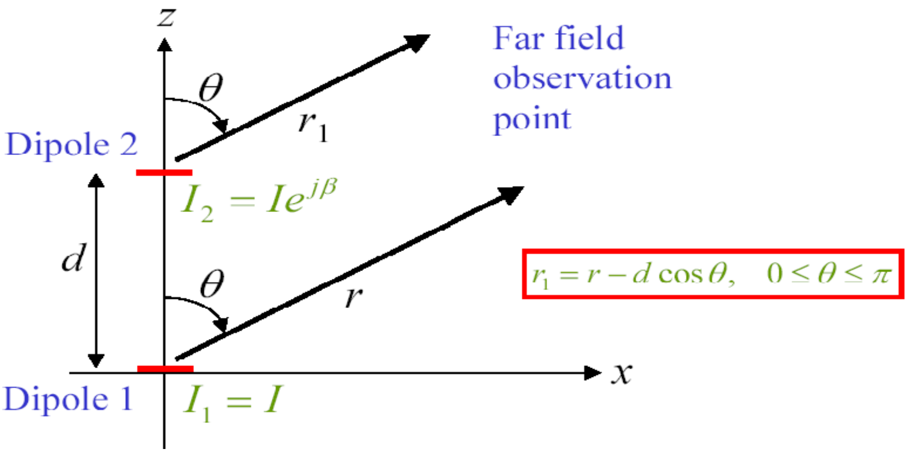


**Fig 3.17:Array of 2 point sources of same amplitude and phase**

Phase difference $=\beta d/2*\cos\theta=2\pi/\lambda*d/2*\cos\theta$

$E_1 = E_0 \exp(-j*\Psi/2)$

$E_2 = E_0 \exp(j*\Psi/2)$

The total field strength at a large distance r in the direction θ is :

$E = E_1 + E_2 = E_0[\exp(j*\Psi/2 + \exp(-j*\Psi/2)]$

Therefore: E = 2E0cos θ /2 ............ (1)

phase difference between E1 & E2 & ψ /2= dr/2*cos θ

E0 =amplitude of the field at a distance by single isotropic antenna

Substituting for ψ / in (1) & normalizing

**CASE 2:**

2 isotropic Point sources. The total field strength at a large distance r in the direction θ is :

$E = E_1 + E_2 = E_0[\exp(j*\Psi/2 - \exp(-j*\Psi/2)]$ Therefore: E = 2jE0SIN(ψ /2)

ψ = phase difference between E1&E2

oint sources of same a ψ /2=dr/2*cos θ

E0 =amplitude of the field at a distance by single isotropic antenna

mplitude but opposite phase

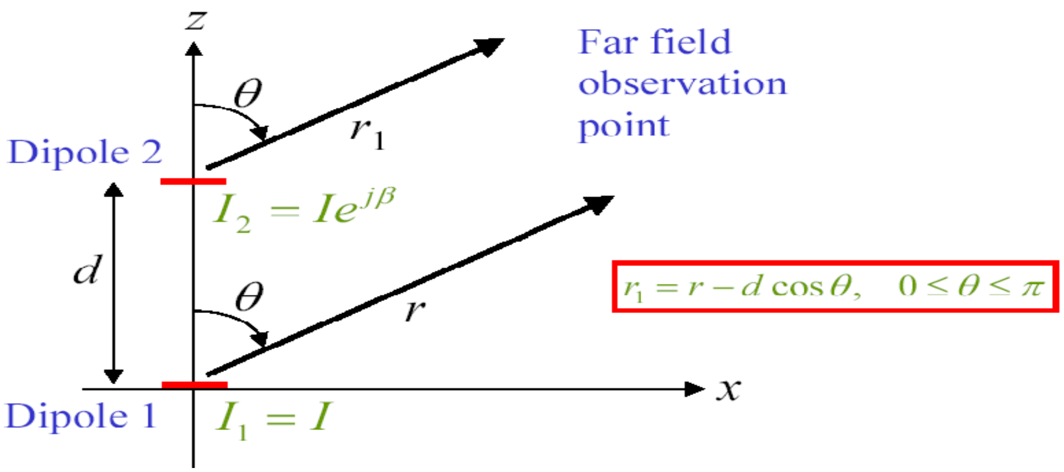


**Fig 3.18:Array of 2 point sources of same amplitude and opposite phase**

$E = 2E_0 \, COS(2\pi/\lambda * d/2 * \cos\theta)$

$E_{nor} = COS(dr/2 * \cos\theta)$

for $d = \lambda/2$

$E = COS(\pi/2 * \cos\theta)$

| | | |
|---|---|---|
| At $\theta = \pi/2$ | E=1... | Point of maxima= π/2(or) 3π/2 |
| At $\theta = 0$ | E=0... | Point of minima= 0 (or) π |
| At $\theta = \pm\pi/3$ | $E = 1/\sqrt{2}$ | 3db bandwidth point= ±π/3 |

**PATTERN MULTIPLICATION:**

The total far-field radiation pattern |E| of array (array pattern) consists of the original radiation pattern of a single array element multiplying with the magnitude of the array factor |AF|. This is a general property of antenna arrays and is called the principle of pattern multiplication.

The pattern multiplication principle states that the radiation patterns of an array of *N* identical antennas is equal to the product of the element pattern *Fe( )* (pattern of one of the antennas) and the array pattern *Fa( )*, where *Fa( )* is the pattern obtained upon replacing all of the actual antennas with isotropic sources.

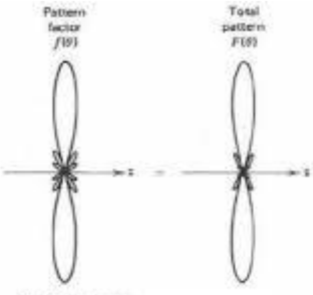
Given an antenna array of identical elements, the radiation pattern of the antenna array may be found according to the pattern multiplication theorem.

Pattern multiplication theorem

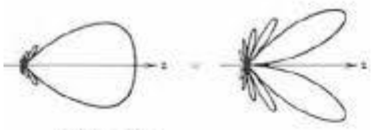
Array pattern=Array element pattern × Array factor

Array element pattern - the pattern of the individual array element. Array factor - a function dependent only on the geometry of the array and the excitation (amplitude, phase) of the elements.

Example (Pattern multiplication - infinitesimal dipole over ground)



Broadside case
($\theta_e = 90^\circ$, $\beta_e = 0$)



Endfire case
($\theta_e = 0^\circ$, $\beta_e L/2 = -3\pi$)

**ADAPTIVE ARRAYS**

Array Processing: Signal Processing is a wide area of research that extends from the simplest form of 1-D signal processing to the complex form of M-D and array signal processing. This article presents a short survey of the concepts, principles and applications of Array Processing.

Array structure can be defined as a set of sensors that are spatially separated, e.g. antennas. The basic problem that we attend to solve by using array processing technique(s) is to:

- Determine number and locations of energy-radiating sources (emitters).
- Enhance the signal to noise ratio SNR "signal-to-interference-plus-noise ratio (SINR)".
- Track multiple moving sources.

## Adaptive Antennas Defined as:

Systems comprising " multiple antenna elements (antenna arrays) " coherent processing " signal processing strategies (algorithms) that vary the way in which those elements are used as a function of operational scenario ! Providing " gain and interference mitigation " leading to improved signal quality and spectral efficiency.

srivenkateshwaraa
College of Engineering & Technology
ASPIRE TO EXCEL
Ariyur, Puducherry-605102.

ISO
21001

ELECTRONICS AND COMMUNICATION ENGINEERING

# ECT64-Antennas and Wave Propagation

YEAR-III SEMESTAR-VI

# QUESTION BANK (11 MARKS)

## UNIT-I

1.Derive the expression for radiation of half wave dipole, center-fed dipole and near, far field?

2. Explain the principle and concept of pattern multiplication?

3.Discuss in detail about broad side and end fire, binomial array antenna.

4.Draw the block diagram of adaptive array and explain briefly.

5.Discuss the antenna fundamentals for the following terms

(i) Power density (ii)Directivity (iii)Gain, Power gain (iv)Radiation resistance (v)Input Impedance. (vi)Radiation pattern (vii)Beamwidth, bandwidth (viii)Polarization

6. Derive the expression for far field components of a small loop antenna.

7. Explain linear array antenna and adaptive array in detail.

8.Derive an expression for array of two point sources (i) with equal amplitude and equal phase (ii) equal amplitude and opposite phase (iii) unequal amplitude and any phase.

9.Derive an expression of radiation from a current element and monopole element.

## UNIT-II

1.Discuss in detail about microstrip antenna, radiation mechanism and its applications.

2.Explain radiation from rectangulat aperture, uniform and tapered aperture.

3.Explain horn antenna, reflector antenna in detail.

4.Discuss the operation of slot antenna , lens antenna.

5.State babinet's principle and explain briefly.

6.Explain the feeding structures of parabolic antenna and cassegrain reflector.

## UNIT-III

1.Explain travelling wave wire antenna, v and rhombic antenna.

2.Discuss in detail about log periodic antenna & Biconical antenna.

3.(a)Define yagi-uda antenna and write the elements of yagi-uda antenna explain it.

(b) Explain the structure of spiral antenna.

4.Discuss briefly for the following antennas (i)loop antenna (ii) folded dipole (iii) helical antenna

## UNIT-IV

1.Discuss in detail about patch antenna and smart antenna.

2. List out the antenna measurements in detail.

3.(a) Define polarization and explain its types.

(b) Explain electromagnetic compatibility antenna and caliberation.

4.Draw the structure of electronic band gap and its applications.

5.Explain reconfigurable antenna, active antenna and dielectric antenna with neat diagram.

## UNIT-V

1.Explain the various factors involved in the propagation of radio waves.

2.Discuss in detail about

(i)Skip distance

(ii)Critical frequency

(iii)Maximun usable frequency

3.Explain the considerations in space wave propagation and its atmospheric effects.

4.Define Ionosphere? Write the ionospheric effects and its mechanism of ionospheric propagation.

5.Explain

(a) fading of signal and its types.

(b) Diversity reception

6.Discuss briefly about refection of radio waves by the surface of the earth in ground waves.

## UNIT-3 ANTENNA AND WAVE PROPAGATION

### Long Wire Antenna

* Long wire or travelling wave or non resonant or aperiodic antenna are those antennas in which there is no reflected wave i.e , standing waves does not travelling in such antennas .
*A long wire when excited by an RF source , the signal travels from the sending end to the other end and is reflected back towards the input .
* A standing wave current occurs due to impedance match at the ends , a pure travelling wave is produced with a progressive phase pattern .
* If the wire length is very long , travelling wave is still produced without impedance matching since a very small reflected signal is obtained compared to the sending signal due to loss of power over the long wire travelling to the other end.
*The simplest travelling wave wire antenna is a straight wire carrying travelling wave .
* The length of the wire is (l>>λ /2) and it is terminated with a matched load RL to make reflection nearly 0.

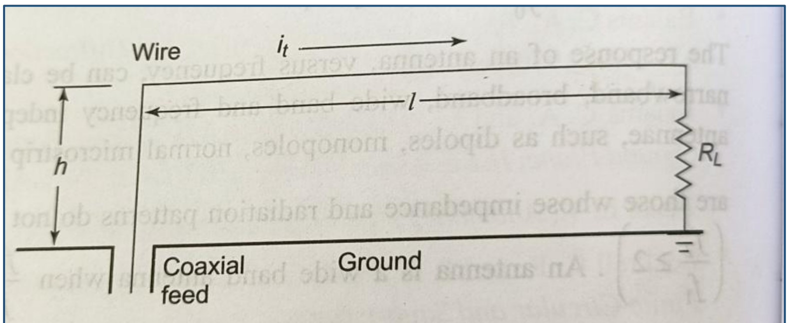
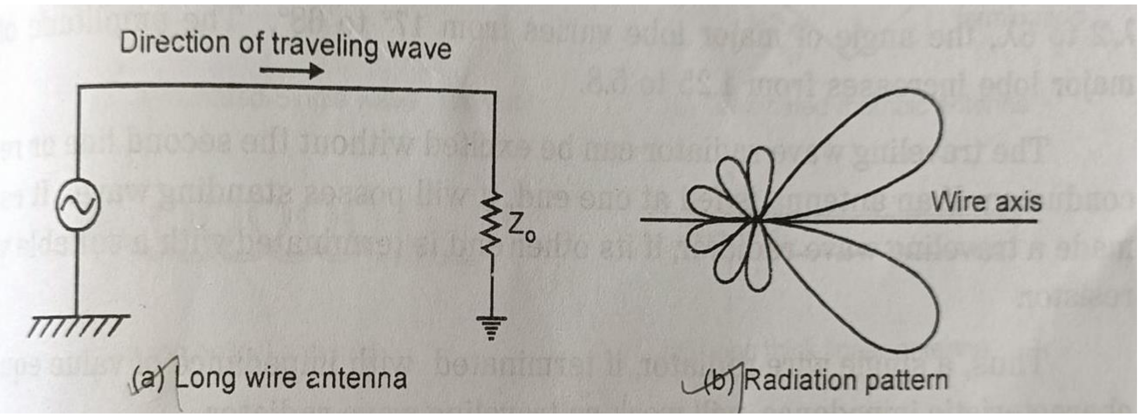


**Fig. 9.3** *Travelling wave wire antenna*

### FOR THE ANALYSIS OF ANTENNA THE FOLOWING ASSUMPTIONS ARE MADE:

1.Length of the wire is much longer than λ/2.
2.The wire is parallel to the ground.
3.Height of the wire above the ground is much lesser than length of the wire.

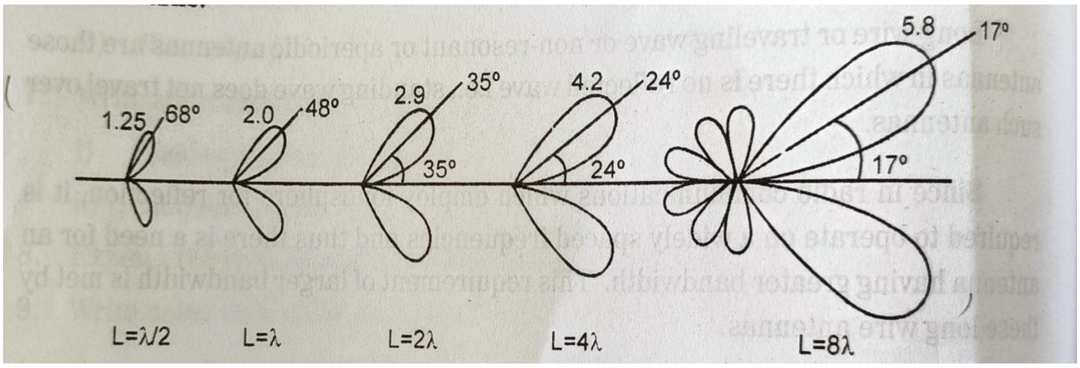
### RADIATION PATTERN OF TRAVELLING WAVE WIRE ANTENNA



(a) Long wire antenna
(b) Radiation pattern

## RADIATION PATTERN FOR TRAVELLING WAVE ANTENNA OF DIFFERENT LENGTHS

-->if radiation pattern for various lengths are plotted as shown in fig,it would be seen that as length of wire increases,the major lobes get closer and narrower to the wire axis.



It is further seen that for a variation of length of traveling wave radiator from 1/2 to 83, the angle of major lobe varies from 17 to 68. The amplitude of the major lobe increases from 1.25 to 5.8volt.

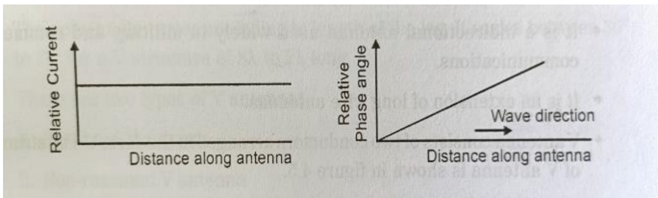
The traveling wave radiator can be excited without the second line or return conductor. If an antenna is fed at one end, it will posses standing wave. It can be made a traveling wave radiator, if its other end is terminated with a suitable value resistor.

Thus, a single wire radiator, if terminated with impedance of value equal to characteristic impedance, will work as traveling wave radiator.

A sinusoidal current distribution may be considered as the standing wave produced by two uniform traveling waves of equal amplitude moving in opposite directions along the antenna.

If only one such wave is present on the antenna, the current distribution is uniform. The amplitude is a constant along the antenna and the phase changes linearly with distance as shown in fig 4.3

**Current amplitude and phase relations along an antenna carrying a single uniform travelling wave:**



gure 4.3: Current amplitude and phase relations along an antenna carrying a single uniform traveling wave

## V ANTENNA

- An antenna having a V-shaped arrangement of conductors fed by a balanced line at the apex is called V-antenna.

- It is used for operations in VHF (Very High Frequency) band. This type of antenna is generally used on light aircraft.

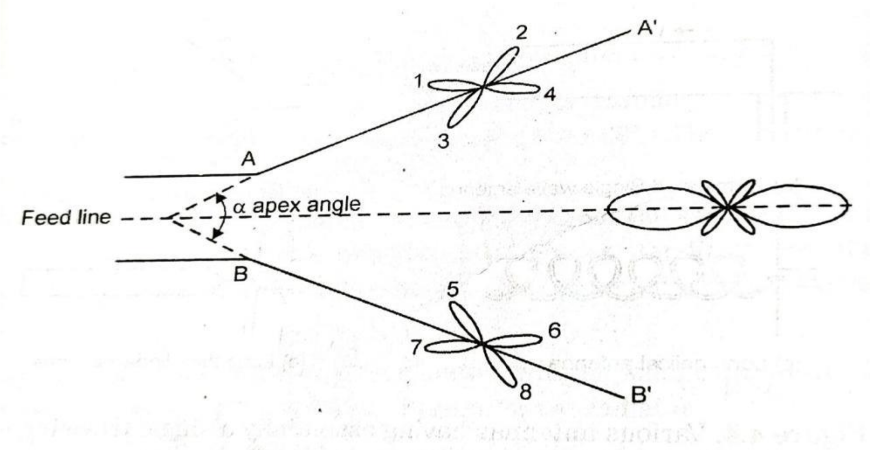


**Figure 4.5:** V antenna and its radiation pattern

*It is a bidirectional antenna used widely in military and commercial communications. It is an extension of long wire antennas.

*V antenna consists of two conductors arranged in the form V. The structure of V antenna is shown in figure 4.5. Each conductor is fed with currents of opposite polarity (180° out of phase)

*The included angle, length and elevation of the conductors are proportioned to give the desired directivity.

*Connecting the two wire feed line to the apex of the V and exciting the two sides of the V by 180° degree out of phase current cause the lobes to add along the line of the bisector and to cancel in other directions.

*The resultant is a bidirectional pattern which is sharper than the same length single long wire.

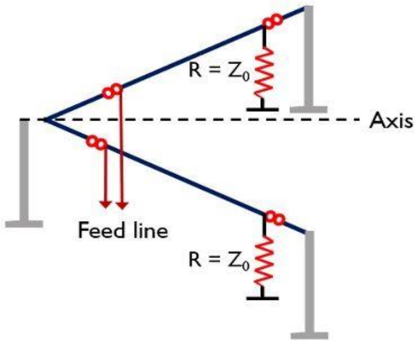
*The lobes are denoted as 1,2,3 and 4 on leg AA' and 5,6,7 and 8 on leg BB.

**There are two types of V antennas:**

1. Resonant V antenna

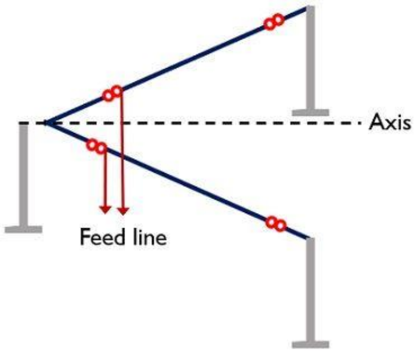
2. Non-resonant V antenna-
Resonant V antenna is shown in figure 4.5. Its radiation pattern is bidirectional.

The lobes B and A' combine and A cancels B' due to opposite direction.



Non-resonant V antenna

Electronics Desk



Resonant V antenna

Electronics Desk

**Advantages of V antenna:**

1. It is easy to construct.

2. It is cheap.

3. End tire and broadside antennas are easily constructed using V antenna.

**Disadvantages**

1. It provides strong minor lobes.

## RHOMBIC ANTENNA

The rhombic antenna is based on the principle of traveling wave radiator. It is also called as double V-type antenna or diamond shape antenna. It is used for HF transmission and reception, commercial point to point communication. A rhombic antenna consists of 4 straight wires arranged in the shape of diamond suspended horizontally above the surface of the earth as shown in figure 4.7. Rhombus is s square but the angles are not right angles. Four major lobes A,B,C and D combine together to give additional gain while lobes A'B'C' and D' get cancelled being in the opposite direction.

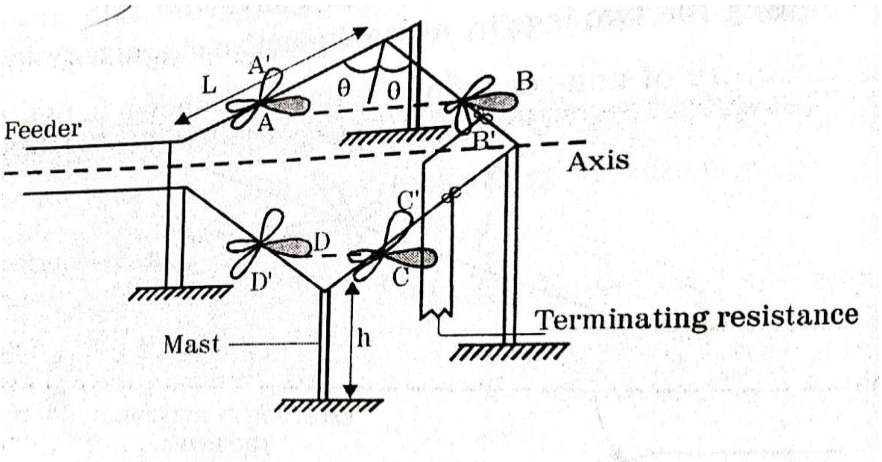


Figure 4.7: Rhombic antenna

*The remote end wires are in close, a terminating resistance 600 to 800 ohm can be conveniently connected at this location so that there is single outgoing traveling wave on the wires. The length of each leg is L and half of the included side angle is 0. The tilt angle θ is approximately equal to 90° minus the major lobe angle ie.. θ=90- beta.

*A rhombus is an equilateral parallelogram, generally with two opposite acute angles. Also called as diamond antenna due to its shape or traveling wave antenna as it is based on the principle of traveling wave radiator.

*It is horizontally installed over the ground at a height h. When rhombic antenna is used for transmission, the input is fed through a balanced (BALUN) line and the terminating resistor is adjusted so that traveling waves are set up in the four legs (sides) of the rhombic as shown in figure 4.8.

The maximum gain is along the direction of main axis, which passes through feed point to termination. Horizontally polarized waves are obtained. The presence of earth brings an elevation in the upward direction as in figure.

*The portion of the radiation cones which do not combine with the main lobe result in considerable side lobes having vertical and horizontal polarization both. This is one of the disadvantage of rhombic antenna. The two sides of rhombus are considered as the conductors of a two wire transmission line. The radiation pattern is unidirectional and is formed due to reinforcement of four lobes of four legs one on each. This unidirectional pattern can be converted into bidirectional simply by removing the terminating resistance.

The length of each leg varies from 2λ to 8λ. The directivity varies from 20 to 90. Power gain is of order of 50 to 60. The value of ß ranges from 17° to 24°.

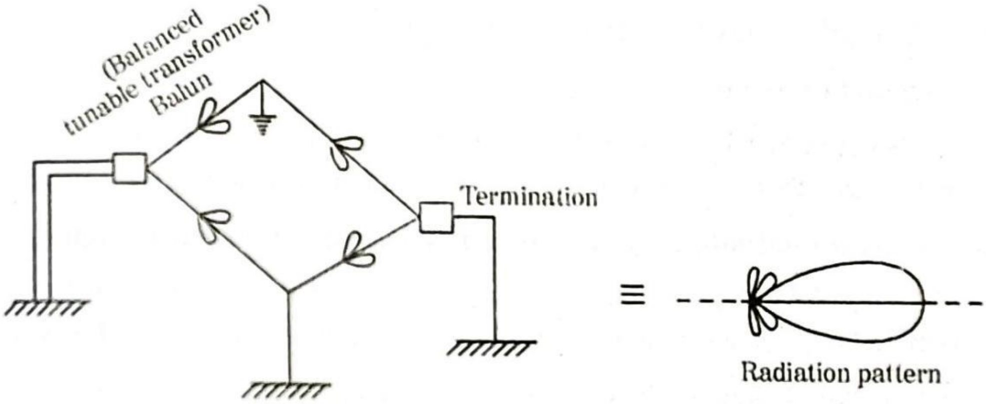


Figure 4.8: Rhombic antenna with Balun and termination

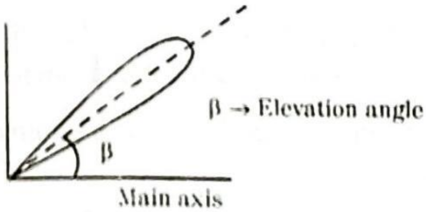


Figure 4.9: Radiation pattern in the presence of earth

**Advantages of Rhombic antenna:**

1. The input impedance and radiation pattern of rhombic antenna do not change rapidly over a considerable frequency range.

2. It is a highly directional broad band antenna with greatest radiated. n where enough space necessary for its installation is no problem.

3. Simple and cheap to erect.

**Disadvantages:**

1. It needs a larger space for installation.

2. Due to minor lobes, transmission efficiency is low.

## FOLDED DIPOLE ANTENNA

WHY FOLDED DIPOLE?

*A simple λ/2 dipole has a terminal resistance of about 73 ohm but matches with a 300 ohm to 600 ohm characteristic impedance and it requires an impedance transformer. In order to provide a good matching characteristics, variations of the single dipole element must be used.

*The folded dipole is an important modification of the conventional half wave dipole in which the two half wave dipole have been folded and joined together at the outer ends. One of the half wave dipoles is continuous while the other is split at the centre. Now the terminal resistance is nearly 300 ohm so that it can be directly connected to a 2-wire line having a characteristic impedance of the same value.

*If the conductors of the folded dipoles are of same diameter, then the currents with equal in magnitude and phase flows through the two dipoles. A folded dipole antenna can be designed with the length more than λ/2. Typically an input impedance of 2 conductor or 4 conductor  folded dipole of length 3/8* λ is 225 ohm, while that of 2 conductor folded dipole of length 3/4* λ is 450 ohm.

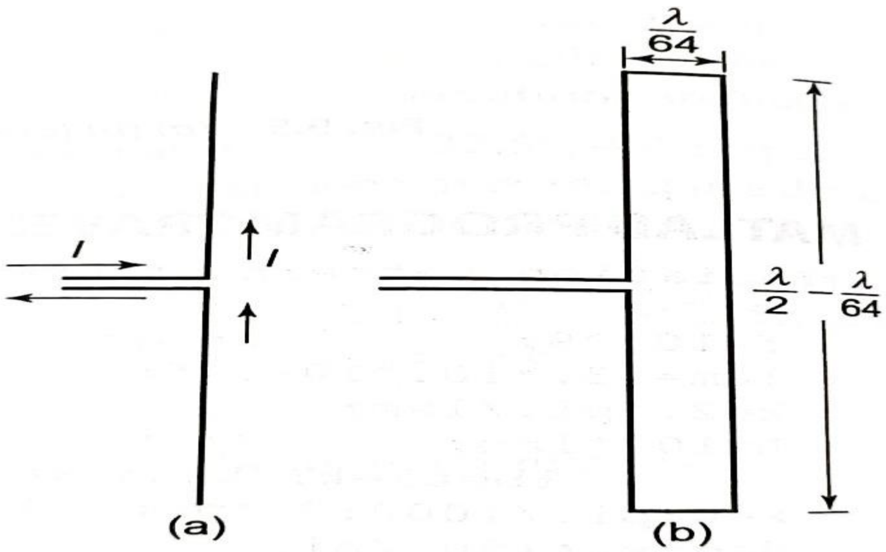


Fig. 9.6 Folded dipole antenna: (a) Ordinary dipole, (b) Folded dipole

**Different types of folded dipole antenna:**

1.Equation of input impedance.
2.Unequal radii.
3.Impedance transformation ratio(ITR).

**MODIFICATIONS OF FOLDED DIPOLES**

**1. T-match antennas**

Consider a 2-wire folded dipole and its terminal resistance is approximately 300 ohm. By doing the modification as shown in Fig.3.16, and is providing the terminal resistance approximately 600 ohm which is mainly depends on the value of D.

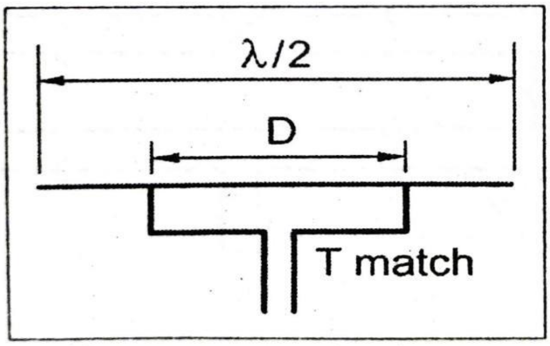


*Fig.3.16. T-match antenna*

### 2. Single-turn loop antenna

A 2-wire folded λ/2 dipole antenna is modified as a single-turn loop antenna and the arrows indicate the instantaneous current direction and the small dots indicate the locations of current minimum.

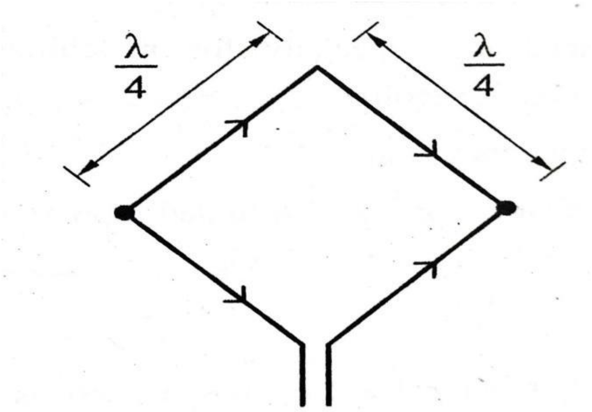


*Fig.3.17. Single-turn loop antenna*

### 3. Two-turn loop (or) quad antenna

A 4 wire folded λ/2 antenna is modified as 2-turn loop antenna.

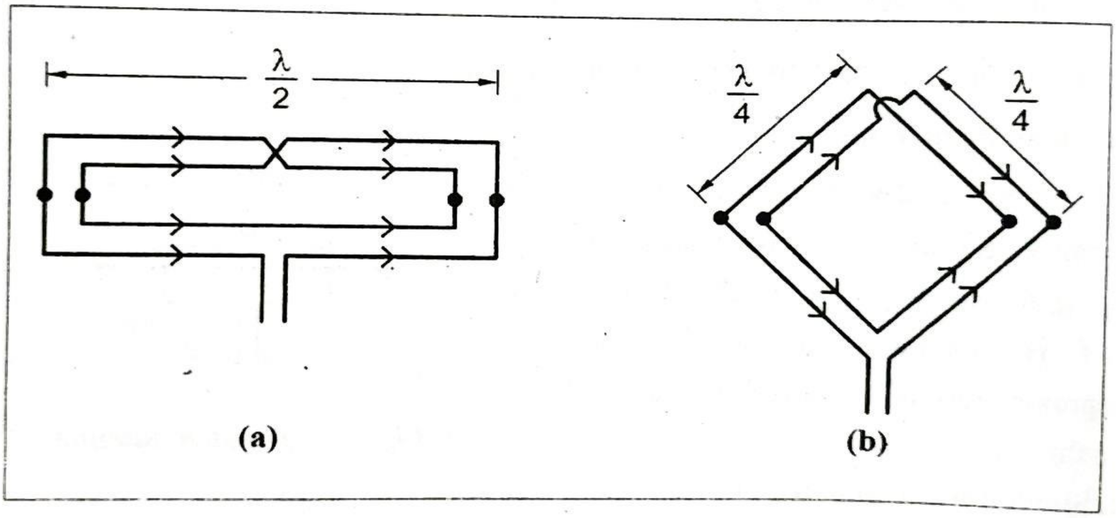


*Fig.3.18. (a) 4-wire folded antenna* *(b) Two- turn loop antenna*

## ADVANTAGES AND APPLICATION

### Advantages:

(1) High input impedance especially for matching network, so that impedance matching is required.

(ii) Wide range of frequency.

(iii) The bandwidth characteristics of a folded dipole antenna is far better that that of a single dipole of the same size.

### Application:

Folded dipole is used in wide band operation such as television as Yagi-uda antenna.

## YAGI-UDA ANTENNA

YAGI-UDA or simply yagi antennas are the most high gain antennas. The antenna was first invented by a Japanese Prof. S. Uda in early 1940's. Afterwards, it was described in English by Prof. H. Yagi.
It was read worldwide and the antenna became popular. Hence it was named as YAGI-UDA. This is the most common antenna used for TV reception. The gain of the antenna is around 7 dB and its radiation pattern is very much directive in one direction (normally receiving direction).

## CONSTRUCTION

A basic Yagi-Uda antenna consists of a driven element, a reflector and one or more directors. The driven element (Dr) is a resonant half-wave dipole (folded) made up of metallic rod at the frequency of operation. It is also called as active element, where the power from the transmitter is fed or which feeds received power to the receiver.

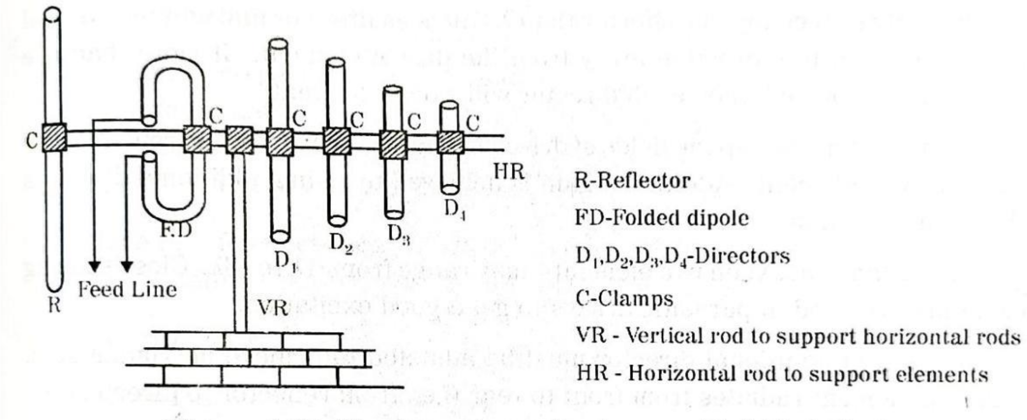


**Figure 4.12:** Six elements yagi antenna with folded dipole

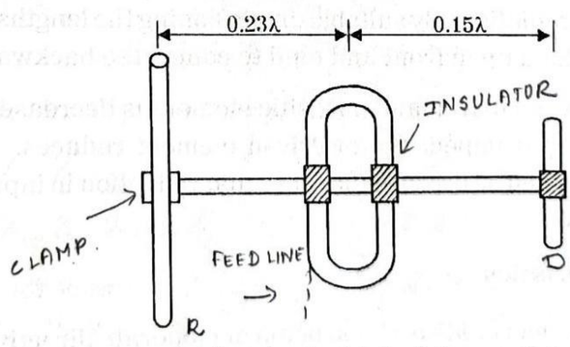


**Figure 4.13:** A typical television yagi antenna

- The parasitic elements in front of the driven element is known as director (D) and its number may be more than one, whereas the element in back of it is known as reflector (R), which derives power by radiation from the nearby driven element.

- The phase and amplitude of the currents through the parasitic elements mainly depends on the length of the elements and spacing between the elements.

- The length of the reflector is 5% more and the director is 5% less than the driven element which is λ/2 at the resonant frequency. In practice, the 3-element yagi array can be designed using the following expressions:

$$\text{Reflector (R) length} = \frac{500}{f(MHz)} \text{ feet (or) } \frac{152}{f(MHz)} \text{ meters}$$

$$\text{Driven } (D_R) \text{ element length} = \frac{475}{f(MHz)} \text{ feet (or) } \frac{143}{f(MHz)} \text{ meters}$$

$$\text{Director (D) length} = \frac{455}{f(MHz)} \text{ feet (or) } \frac{137}{f(MHz)} \text{ meters}$$

*Practically, the spacing between the driven element and the parasitic elements varies from 0.1 λ to 0.15 λ. The parasitic elements and the driven element could be clamped on a metallic support rod. The clamping over the support rod provides a rigid mechanical structure.

*The driven element is fed by a 2 wire balanced transmission line. But the reflector and director are not connected directly with the transmission line but they are coupled electrically with driven element. Increasing the number of directors will increase the power gain but decreases the
antenna bandwidth. A 3 element yagi antenna suitable for TV reception of moderate field strength.

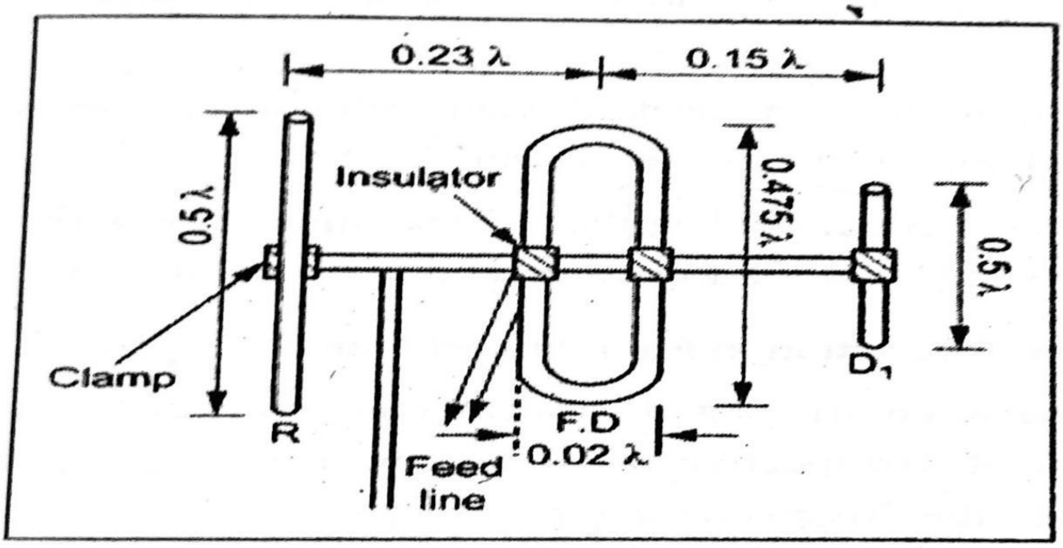


*Fig.3.21. A typical TV yagi antenna*

## WORKING PRINCIPLE

The length of the reflector is more than the folded dipole (driven element). Therefore it offers an inductive reactance (current lag the induced voltage) to the incoming signal.
The length of the director is shorter than the dipole. Hence it offers capacitive reactance (current leads the induced voltage).

### 1. Function of Reflector

- The radiation coming from the front at the reflector is absorbed and it retransmits the radiation towards the dipole in such a way that it adds with the incoming signal.

For any radiations coming from the back side, the reflector retransmits the radiation in such a way that it is out of phase with the direct radiation from back side at dipole and hence they can cancel each other.

### 2. Function of Director

*For the radiation coming from the front, the director generates its own radiation in such a way that it adds with this radiation at dipole and hence increases the signal strength.

*For radiation coming from the back, director generates its own radiation such that it cancels the radiation from back at the dipole.

*By suitable dimensioning, the lengths and spacing between the two elements, the radiated energy is added up in front and tend to cancel the backward radiation.

### 3. Compensation for Reduction in Input Impedance

*If the distance between driven and parasitic element is decreased, then it will load the driven element, irrespective of its length. Therefore the impedance at the input terminals of the driven element reduces.

*Folded dipole which has high impedance compared to the conventional half wave is used as driven element so that reduction in input impedance is compensated.

*Input impedance of folded dipole = $n^2$x Impedance of conventional half wave dipole

where, n-Number of elements in the folded antenna.

## GENERAL CHARACTERISTICS

The following are the general characteristics of the Yagi-Uda antenna:

(i) If Yagi-Uda antenna with three elements including one reflector, one driven element and one director are used, then it is commonly referred to as Beam Antenna.

(ii) This antenna gives unidirectional beam of moderate directivity with light weight, low cost and simplicity in feed system design.
(iii) With spacing of 0.1 λ to 0.15 λ., a frequency bandwidth of the order of 2% to 3% can be easily achieved.

(iv) It provides gain of about 7 to 8 dB and front to back ratio of about 20 dB.

(v) It is also known as super directive or super gain antenna due to its high gain and beam-width per unit area of the array.

## ADVANTAGES, DISADVANTAGES AND APPLICATIONS

### 1. Advantages

(i) Unidirectional radiation
(ii) Increased directivity
(iii) Simple construction.
(iv) Low cost.
(v) Light weight

### 2. Disadvantages

(i) It is sensitive to the frequency.
(ii) Bandwidth is reduced, if the array is constructed with more number of directors.

## 3. Applications

(i) Used in television reception.

(ii) Used as a transmitter in low frequency applications.

## LOG PERIODIC ANTENNA

## INTRODUCTION

*A log periodic antenna is a broadband narrow beam antenna. It is a frequency independent antenna.

### Frequency-Independent Concept

*If the structure of the antenna is defined in terms of angles only, then it comes under the category of frequency independent antenna. e.g., Log periodic antenna.

### Log-Periodic Concept

*Here, the geometry of the antenna structure is adjusted such that all the electrical properties of the antenna must repeat periodically with the logarithm of the frequency. For every repetition, the structure size changes by a constant scale factor by which the structure can either expanded or contracted. The log periodic principle can be understood with the help of the array of log periodic antenna known as Log Periodic Dipole Array (LPDA).

### Construction of LPDA:

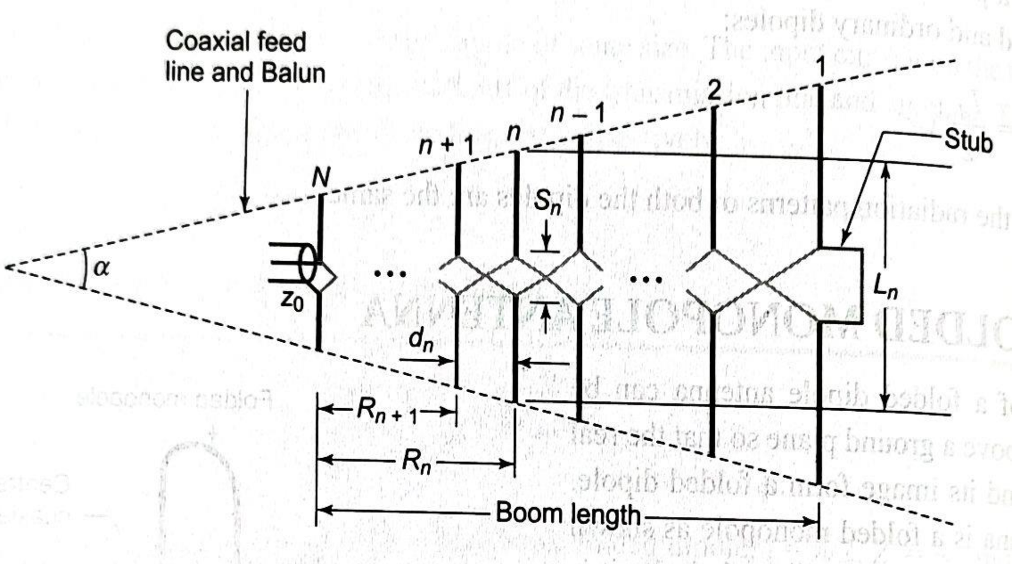


Fig. 9.10 *Log periodic dipole antenna*

*It consists of a number of dipoles of different lengths and spacings. The array is fed using a balanced transmission line which is connected at narrow end or apex of the array. Also the transmission line is transposed between each adjacent pairs of terminals of dipoles.

*The length of the dipoles increases from feed point towards other end such that the included angle a remains constant. The dipole lengths and the spacings between two adjacent dipoles are related through parameter called design ratio scale factor denoted by t.

## WORKING PRINCIPLE OF LPDA:

The Analysis of a log periodic dipole array can be done by considering t region of the antenna which is classified according to the length of the dipoles. The are three

(i) Inactive transmission-line region ( L< λ/2)

(ii) Active region (L ~ λ/2)

(iii) Inactive reflective region (L>λ/2)

### 1. Inactive Transmission Line Region (L< λ/2)

*It is the region in which the length of the dipoles is less than the resonant length λ/2 .therefore the elements present relatively high capacitance impedance.the spacing between the elements is comparatively smaller.

*The current in the region will be very small and hence it is considered as inactive region.These currents lead the voltage supplied by the transmission line.

*Transposition of transmission introduces 180° phase shift between adjacent dipoles.

* Hence currents in elements of these regions are small and hence small radiation in backward direction (towards left).

### 2. Active Region (L =λ /2)

*In this region, the dipole lengths are approximately equal to the resonant length (λ/2). Therefore the dipoles in this region offer resistive impedance. (Thus the element currents are of large value and in phase with the base voltage/Hence there is strong radiation towards left in backward direction and a little radiation towards right.

### 3. Inactive Reflective Region (L> λ/2)

The element (dipoles) lengths are longer than the resonant length (ie., L>λ/2) Hence the dipoles offer inductive impedance. The currents will be smaller in this region and also lag at the base voltage. Thus, any small amount of incident wave from the active region is reflected back towards the backward direction.

## USES OF LOG PERIODIC ANTENNA

It is mainly used in the field of HF communication where the multiband steerable (rotatable) and fixed antennas are generally used. It has the advantage that no power is wasted in terminating resistance.

(i) It is used for TV reception. Only one log periodic design will suffice for all the channels even upto UHF band.

(ii) It is best suited for all round monitoring. i.e., a single log periodic antenna will cover all the higher frequency bands, when there is no problem with the cost of the installation.

## SPIRAL ANTENNAS

*The spiral antennas are a frequency independent antenna) (It radiates a bi-directional main lobe perpendicular to the plane of the antenna, It produces circularly polarized waves within the band of operation and the radiation is elliptically polarized outside the band of operation.

* The surface of the equi angular spiral shape can be described completely by angles.
* It fulfills all the necessary conditions that are employed to design frequency independent antennas.

When the total arm length is comparable with the wavelength, the frequency the operation will be the lowest cut-off frequency and for all other frequencies above this, the pattern and impedance characteristics are frequency independent.

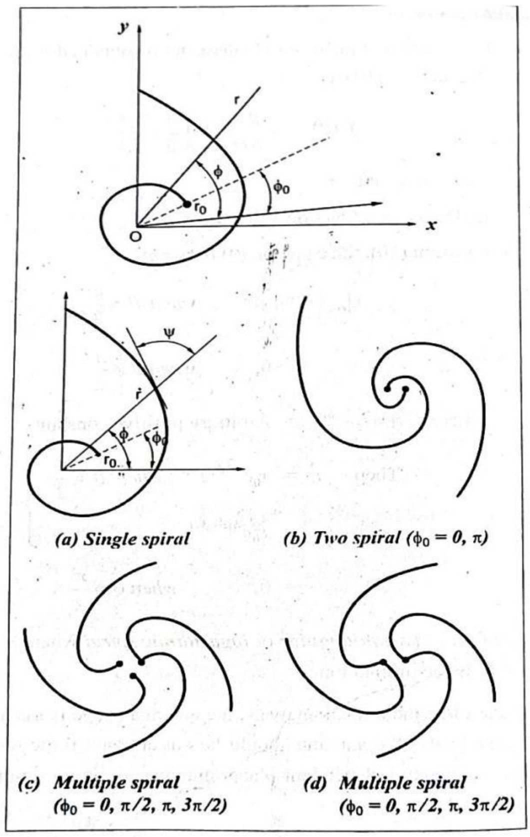


*Fig.9.2. Some simple shapes of frequency independent planar antennas*

## Types of spiral antenna

1. Planar spiral antenna
2. Conical spiral antenna

## Conical - Spiral Antenna

A tapered helix is a conical-spiral antenna and these were described and investigated extensively in the years following 1947.

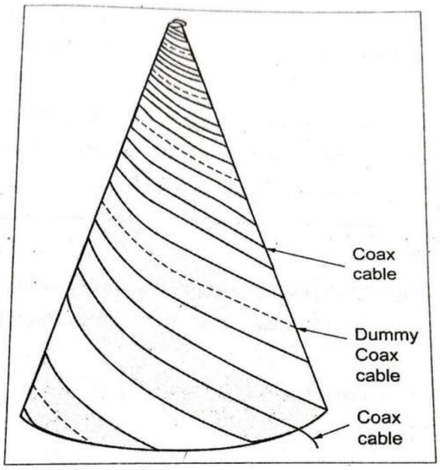


*Fig.9.4. Two arm balanced conical spiral antenna*

- The conducting conical spiral surface can be constructed conveniently using printed circuit technique, the conical arms on the dielectric cone which is also used as a support. The feed cable can be bounded to the metal arms which are wrapped around the cone.

- The conical equi-angular spiral antenna is fed at the apex by a means of a balanced transmission line carried up inside the cone along the axis of the cone.

- The main difference between the conical spiral and planar spiral antenna is that the conical spiral antenna provides unidirectional radiation in a single lobe towards the apex of the cone and with a maximum radiation along the axis.

- In conical antennas, the circular polarization and relatively constant impedances are preserved over large bandwidths.

## HELICAL ANTENNA

### INTRODUCTION

- Helical antenna is a simplest type of antenna (radiator) which provides circularly polarized waves; it is used in extra terrestrial communications where satellite relays are involved.

- The helical antenna is a broadband VHF and UHF antenna to provide circular polarization characteristics.

### Construction

- Helical antenna consists of a helix of thick copper wire or tubing wound in the shape of a screw thread and used with a flat metal called a ground plane or ground plate. The structure of the helical antenna is shown in fig.

- The helix is fed by a co-axial cable and it is connected between helix and ground plane. One end of the helix is connected to the centre conductor of the cable and the outer conductor is connected to the ground plane.

The ground plane is simply made of sheet or of screen or of radial and concentric conductors. The mode of radiation depends on the diameter of the helix "D" and tum spacing "S" (turn spacing is a measure between two centres of the turns)

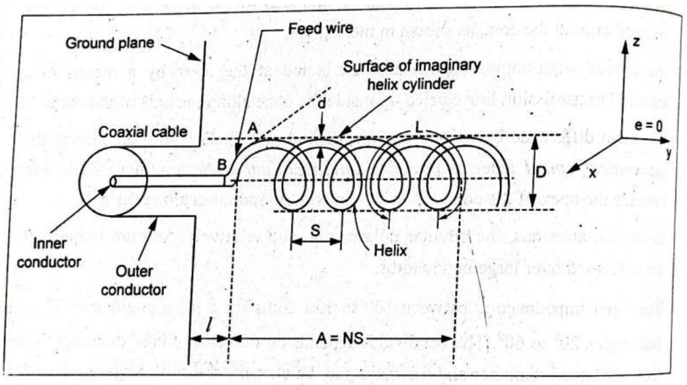


*Fig. 9.5. Helical Antenna* ✓

Radiation pattern of helical antenna (axial mode)

The following symbols are used to describe a helix

C = Circumference of helix = πD

d = Diameter of helix conductor

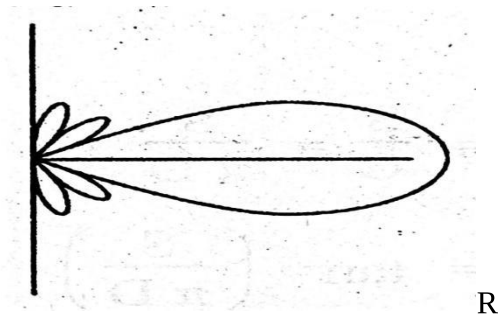
A = Axial length = NS

N = Number of turns

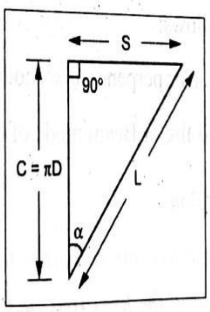
L = Length of one turn

l = Spacing of helix from ground plane

Alpha = Pitch angle



Radiation pattern of Helical Antenna



9.7. Inter-relation between circumference, spacing, turn length and pitch angle

**MODES OF RADIATION**

In general, a helical antenna can radiate in many modes. But the most important modes of radiation are as follows:

(i) Normal mode or perpendicular mode.

(ii) Axial or End fire or Beam mode of radiation.

**Normal Mode of Radiation**

In this normal mode of radiation, the radiation field is maximum in broadway (...) in the direction normal to the helix axis and is circularly polarized wave Here the dimensions of the helix is small compared with wavelength (ie) NL << $\alpha$.

Here, the radiation pattern is a combination of the equivalent radiation from a short dipole positioned on the same helix axis and a small loop which is also coaxial with helix axis.

When a = 0° helix corresponds to a loop and a 90° the helix becomes a linear dipole as shown in Figure.

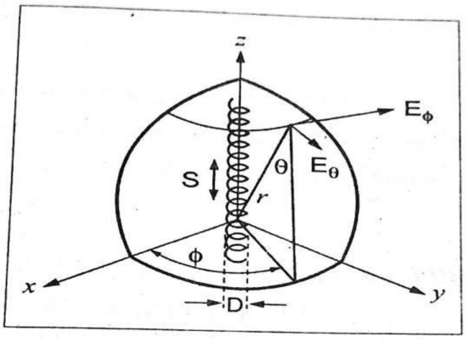
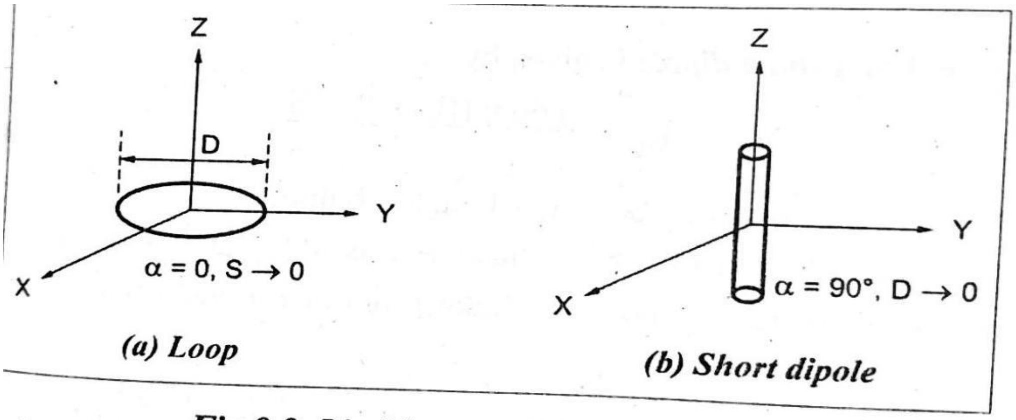


*Fig.9.8. Helix in 3-dimensional spherical coordinate*

If S= 0, helix collapse to a loop and if S = constant and D= 0, the helix straightens into a linear conductor (short dipole).



*(a) Loop*

*(b) Short dipole*

*Fig.9.9. Limiting conditions on helix*

Axial (OR) Beam Mode of Radiation

•The radiation field is maximum in the end fire direction that is along the helix axis and the polarization is circular or nearly circular. This mode occurs w the helix circumference (D) and spacing (S) are appreciable of the order of a wavelength.

• This mode produces a broad and fairly directional beam in the axial direction minor lobes at oblique angles. Because of this features, the helical antenna is for many practical applications.

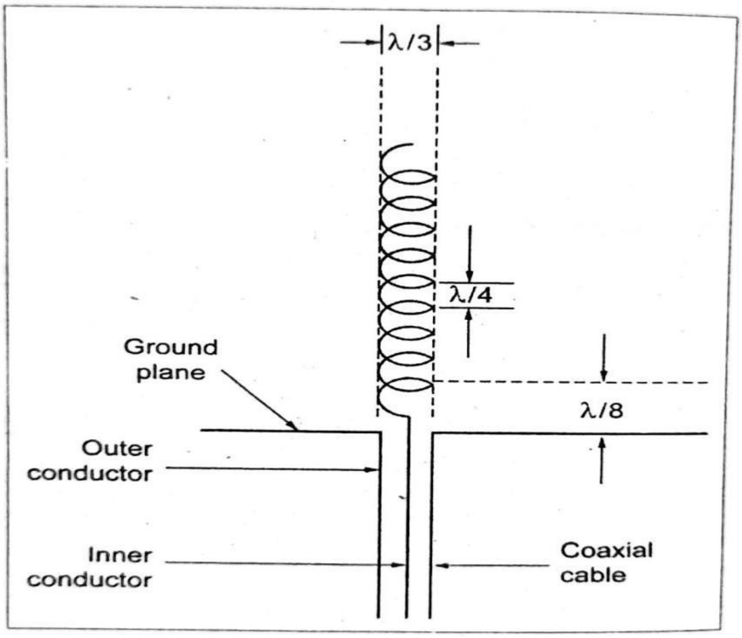


*Fig.9.10. Arrangement for generating axial mode*

The advantages of helical antenna are:

(i) N-turn helix is an end fire array of "n" sources.

(ii) The helix not only have a nearly uniform resistance input over a wide bandwidth, but it also operates as a super gain end fire array over the same bandwidth.

## APPLICATIONS OF HELICAL ANTENNA

The applications of helical antenna are,

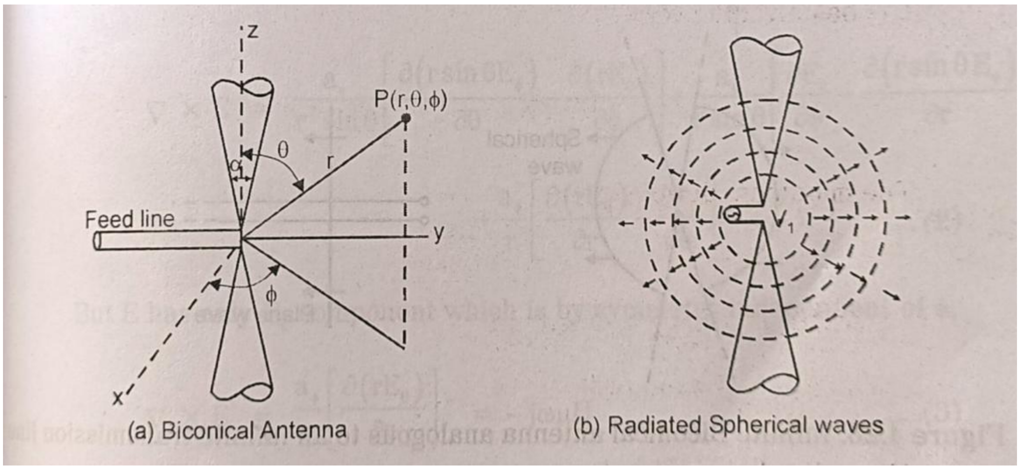
(i) Wide bandwidth, simplicity, highest directivity and circular polarization of helical beam antenna have made it indispensable for space communication application like telemetry, radio astronomy, satellite and space communications.

(ii) A single or an array of helical antenna is used to receive or transmit the VHF signals through ionosphere.

## BICONICAL ANTENNA

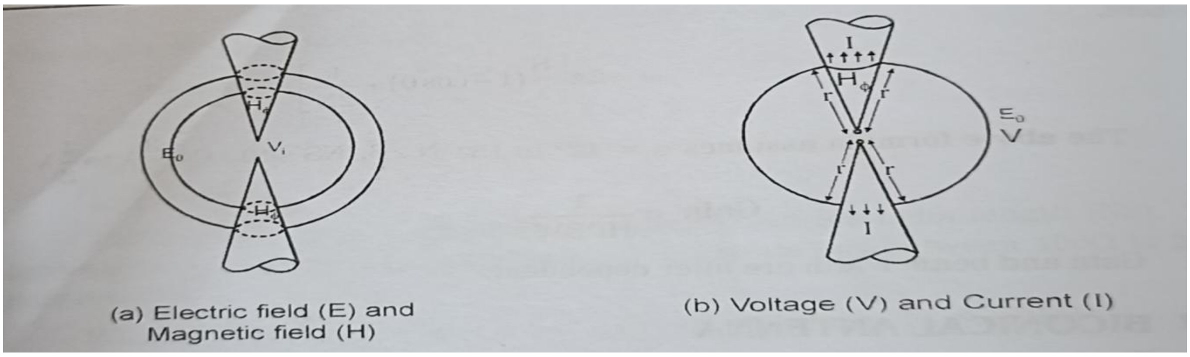
The biconical antenna is the name given to the double cone antenna as shown in figure 4.24. A biconical antenna consists of an arrangement of two conical conductors which is driven by potential, charge or an alternating magnetic field (and the associated alternating electric current) at the vertex.

The conductors have a common axis and vertex. The two cones face in opposite directions. Biconical antennas are broad band dipole antennas transceiving signals from 30MHz to 300MHz.



(a) Biconical Antenna (b) Radiated Spherical waves

The application of a voltage (V) at the input terminals will produce outgoing spherical waves as shown in figure 4.24 (b)which in turn produce a current I at any point (r, 0, 0) along the surface of the cone and voltage V between the points on the upper and lower cones at a distance r from the terminals as shown in figure.



(a) Electric field (E) and Magnetic field (H) (b) Voltage (V) and Current (I)

These can be used to calculate the characteristic impedance of the transmission line which is also equal to the input impedance of an infinite geometry The infinite biconical antenna is analogous to an infinite transmission line. The biconical antenna acts as a guide for a spherical wave in the same way that a uniform transmission line acts as a guide for a plane wave. This is shown in figure 4.26.

## SMALL LOOP ANTENNAS

The loop antenna is a radiating coil of any convenient cross-section of one or more turns carrying RF current. It may assume any shape as in figure 2.4. (Example: Rectangular, Square, triangular, Hexagonal and circular).

(a) Rectangular

(b) Triangular

(c) Square

(d) Circular

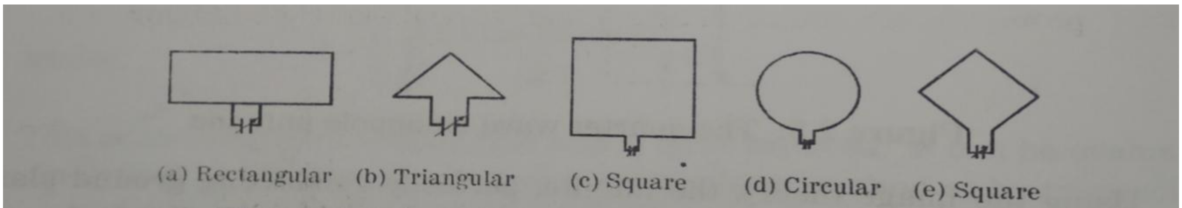

(a) Rectangular (b) Triangular (c) Square (d) Circular (e) Square

Figure 2.4: Loop antennas of different shapes

**Figure 2.4: Loop antennas of different shapes**

A loop antenna of more than one turn is called as frame. It is used in radio receiver, aircraft receiver, direction finding and UHF transmitters. Currents are of the same magnitude and phase throughout the loop if dimensions are small in comparison to wave length (a<<λ).

The radiation efficiency of closed loop antenna is low for transmission purposes.

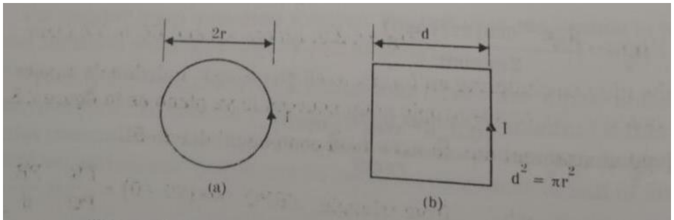
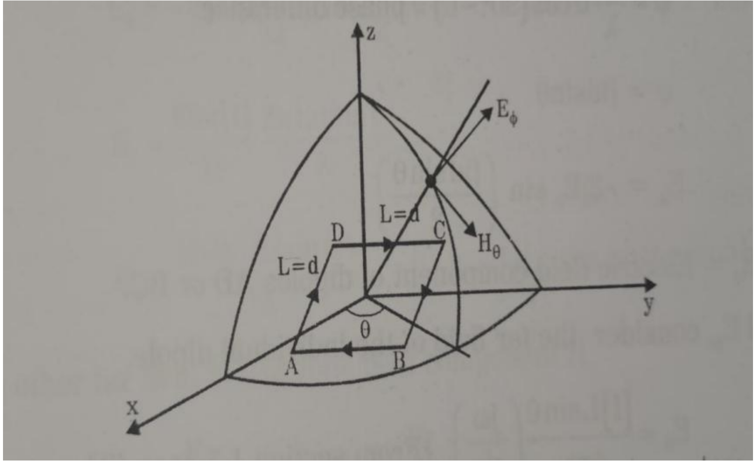


Figure 2.5: Circular loop and square loop of equal area

Radiated fields of small loop antenna:



Let the circular loop of radius r be represented by square loop of side length d such that areas of both are same. The loop is placed at the centre of the co- ordinate system as in figure 2.6 and its far field will have only E, component.

The sides AD and BC of the loop are being treated as short dipoles, their radiation pattern in horizontal plane x-y and vertical plane y-z in figure 2.7.

Both the dipoles radiating uniformly in all directions. Individuals dipoles AD and BC will behave like two isotropic point sources in yz plane as in figure 2.8.
E= field component due to AD+ field component due to BC.

srivenkateshwaraa
College of Engineering & Technology
ASPIRE TO EXCEL
Ariyur, Puducherry-605102.

ISO
21001

ELECTRONICS AND COMMUNICATION ENGINEERING

# ECT64- Antennas and Wave Propagation

YEAR- III SEMESTAR- VI

# UNIT- V

# PROPAGATION

**The three basic types of propagation:** Ground wave, space wave and sky wave propagation.

## Factors involved in the propagation of Radio Waves

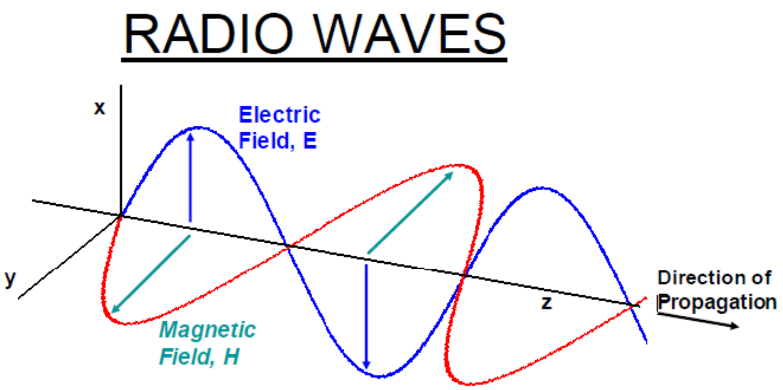
**Sky Wave Propagation:** Structure of the ionosphere – Effective dielectric constant of ionized region – Mechanism of refraction – Refractive index – Critical frequency – Skip distance – Effect of earth's magnetic field – Energy loss in the ionosphere due to collisions – Maximum usable frequency – Fading and diversity reception.

**Space Wave Propagation:** Reflection from ground for vertically and horizontally polarized waves – Reflection characteristics of earth – Resultant of direct and reflected ray at the receiver – Duct propagation.

**Ground Wave Propagation:** Attenuation characteristics for ground wave propagation – Calculation of field strength at a distance

## Propagation of Waves

The process of communication involves the transmission of information from one location to another. As we have seen, modulation is used to encode the information onto a carrier wave, and may involve analog or digital methods. It is only the characteristics of the carrier wave which determine how the signal will propagate over any significant distance. This chapter describes the different ways that electromagnetic waves propagate.

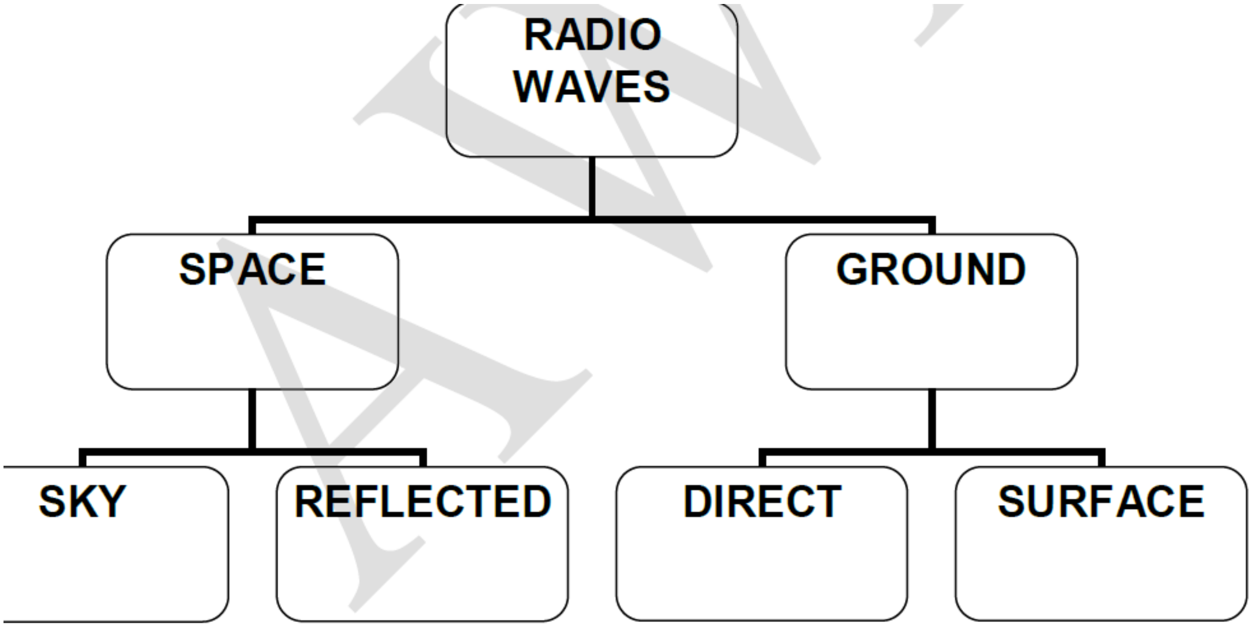


- Electromagnetic radiation comprises both an Electric and a Magnetic Field.
- The two fields are at right-angles to each other and the direction of propagation is at right-angles to both fields
- The Plane of the Electric Field defines the Polarisation of the wave.

**POLARIZATION**

The polarization of an antenna is the orientation of the electric field with respect to the Earth's surface and is determined by the physical structure of the antenna and by its orientation .

- Radio waves from a vertical antenna will usually be vertically polarized.
- Radio waves from a horizontal antenna are usually horizontally polarized.



**LINE OF SIGHT, GROUND WAVE, SKY WAVE**

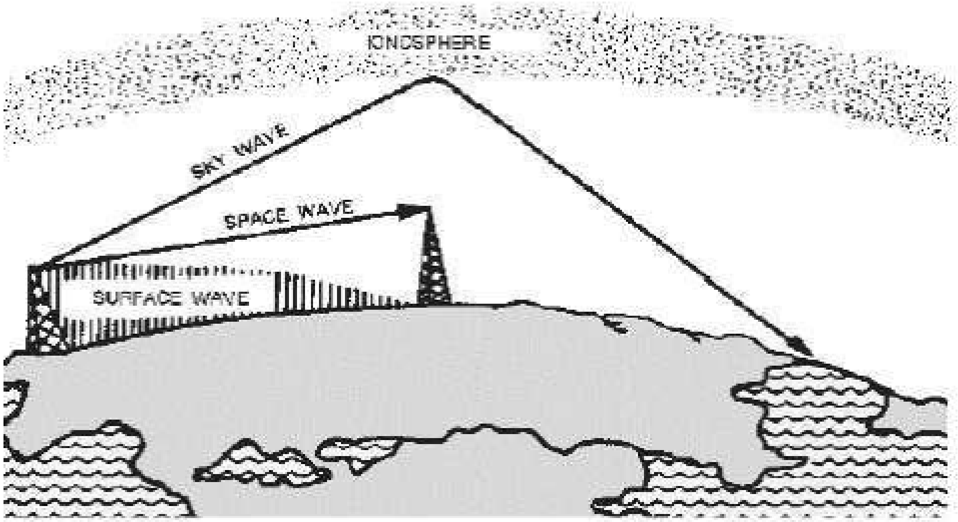
**Ground Wave** is a Surface Wave that propagates or travels close to the surface of the Earth.

**Line of Sight** (Ground Wave or Direct Wave) is propagation of waves travelling in a straight line. These waves are deviated (reflected) by obstructions and cannot travel over the horizon or behind obstacles. Most common direct wave occurs with VHF modes and higher frequencies. At higher frequencies and in lower levels of the atmosphere, any obstruction between the transmitting antenna and the receiving antenna will block the signal, just like the light that the eye senses.

• Space Waves: travel directly from an antenna to another without reflection on the ground. Occurs when both antennas are within line of sight of each another, distance is longer that line of sight because most space waves bend near the ground and follow practically a curved path. Antennas must display a very low angle of emission in order that all the power is radiated in direction of the horizon instead of escaping in the sky. A high gain and horizontally polarized antenna is thus highly recommended.

• Sky Wave (Skip/ Hop/ Ionospheric Wave) is the propagation of radio waves bent (refracted) back to the Earth's surface by the ionosphere. HF radio communication (3 and 30 MHz) is a result of sky wave propagation.

**LINE OF SIGHT, GROUND WAVE, SKY WAVE**



**Ground-Wave Propagation**

Radio waves follow the Earth’s surface

- AM broadcasts during the day
- Works best at lower frequencies (40, 80, and 160 meters)
- Relatively short-range communications
- Amateur priv’s are higher than broadcast frequencies, thus less ground-wave range

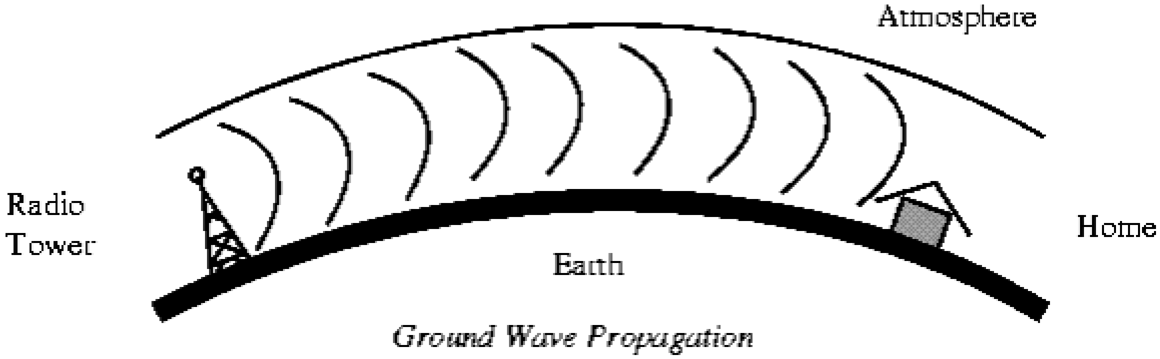
## REFLECTION OF RADIO WAVES BY THE SURFACE OF THE EARTH

### RF Propagation

There are three types of RF (radio frequency) propagation:

- Ground Wave
- Ionospheric
- Line of Sight (LOS)

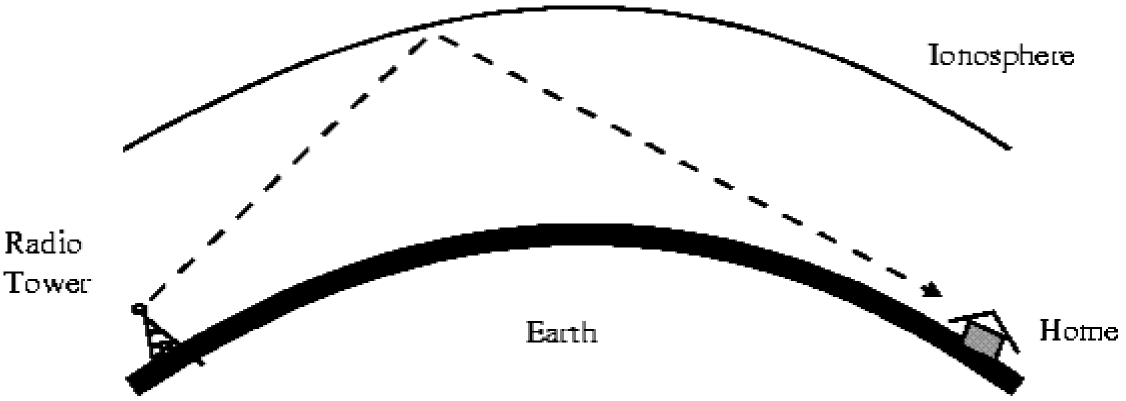
Ground wave propagation follows the curvature of the Earth. Ground waves have carrier frequencies up to 2 MHz. AM radio is an example of ground wave propagation.



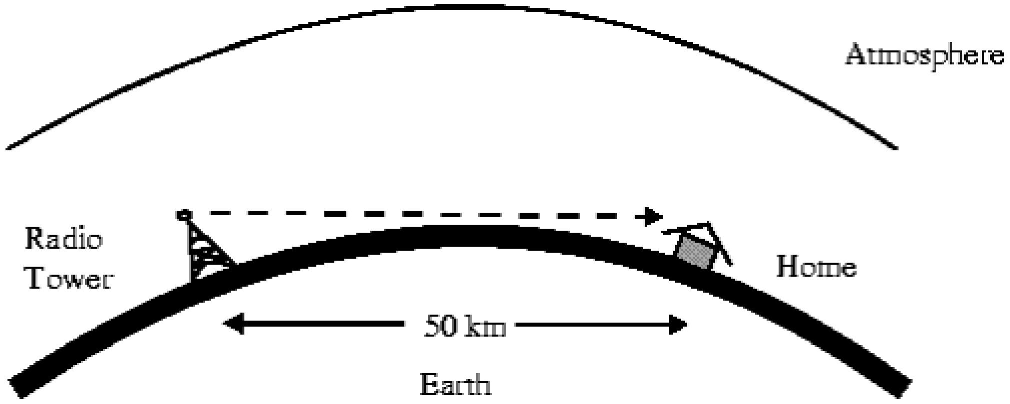
*Ground Wave Propagation*

Ionospheric propagation bounces off of the Earth's ionospheric layer in the upper atmosphere. It is sometimes called double hop propagation. It operates in the frequency range of 30 - 85 MHz. Because it depends on the Earth's ionosphere, it changes with the weather and time of day. The signal bounces off of the ionosphere and back to earth. Ham radios operate in this range.

## IONOSPHERE AND ITS EFFECT ON RADIO WAVES



Line of sight propagation transmits exactly in the line of sight. The receive station must be in the view of the transmit station. It is sometimes called space waves or tropospheric propagation. It is limited by the curvature of the Earth for ground-based stations (100 km, from horizon to horizon). Reflected waves can cause problems. Examples of line of sight propagation are: FM radio, microwave and satellite



The ionosphere is the region of the atmosphere that extends from about 30 miles above the surface of the earth to about 250 miles. it is appropriately named the ionosphere because it consists of several layers of electrically charged gas atoms called ions. the ions are formed by a process called ionization.

## MECHANISM OF IONOSPHERIC PROPAGATION

### IONIZATION

Ionization occurs when high energy ultraviolet light waves from the sun enter the ionospheric region of the atmosphere, strike a gas atom, and literally knock an electron free from its parent atom. a normal atom is electrically neutral since it contains both a positive proton in its nucleus and a negative orbiting electron. when the negative electron is knocked free from the atom, the atom becomes positively charged (called a positive ion) and remains in space along with the free electron, which is negatively charged. this process of upsetting electrical neutrality is known as ionization. the free negative electrons subsequently absorb part of the ultraviolet energy, which initially freed them from their atoms. as the ultraviolet light wave continues to produce positive ions and negative electrons, its intensity decreases because of the absorption of energy by the free electrons, and an ionized layer is formed. the rate at which ionization occurs depends on the density of atoms in the atmosphere and the intensity of the ultraviolet light wave, which varies with the activity of the sun. since the atmosphere is bombarded by ultraviolet light waves of different frequencies, several ionized layers are formed at different altitudes. lower frequency ultraviolet waves penetrate the atmosphere the least; therefore, they produce ionized layers at the higher altitudes. conversely, ultraviolet waves of higher frequencies penetrate deeper and produce layers at the lower altitudes. an important factor in determining the density of ionized layers is the elevation angle of the sun, which changes frequently. for this reason, the height and thickness of the ionized layers vary, depending on the time of day and even the season of the year.

### Recombination

Recall that the process of ionization involves ultraviolet light waves knocking electrons free from their atoms. a reverse process called recombination occurs when the free electrons and positive ions collide with each other. since these collisions are inevitable, the positive ions return to their original neutral atom state. the recombination process also depends on the time of day. between the hours of early morning and late afternoon, the rate of ionization exceeds the rate of recombination. during this period, the ionized layers reach their greatest density and exert maximum influence on radio waves. during the late afternoon and early evening hours, however, the rate of recombination exceeds the rate of ionization, and the density of the ionized layers begins to decrease. throughout the night, density continues to decrease, reaching a low point just before sunrise.

### Four distinct layers

The ionosphere is composed of three layers designated d, e, and f, from lowest level to highest level as shown in figure 4-10. the f layer is further divided into two layers designated f1 (the lower layer) and f2 (the higher layer). the presence or absence of these layers in the ionosphere and their height above the earth varies with the position of the sun. at high noon, radiation in the ionosphere directly above a given point is greatest. at night it is minimum. when the radiation is removed, many of the particles that were ionized recombine. the time interval between these conditions finds the position and number of the ionized layers within the ionosphere changing. since the position of the sun varies daily, monthly, and yearly, with respect to a specified point on earth, the exact position and number of layers present are extremely difficult to determine. however, the following general statements can be made:

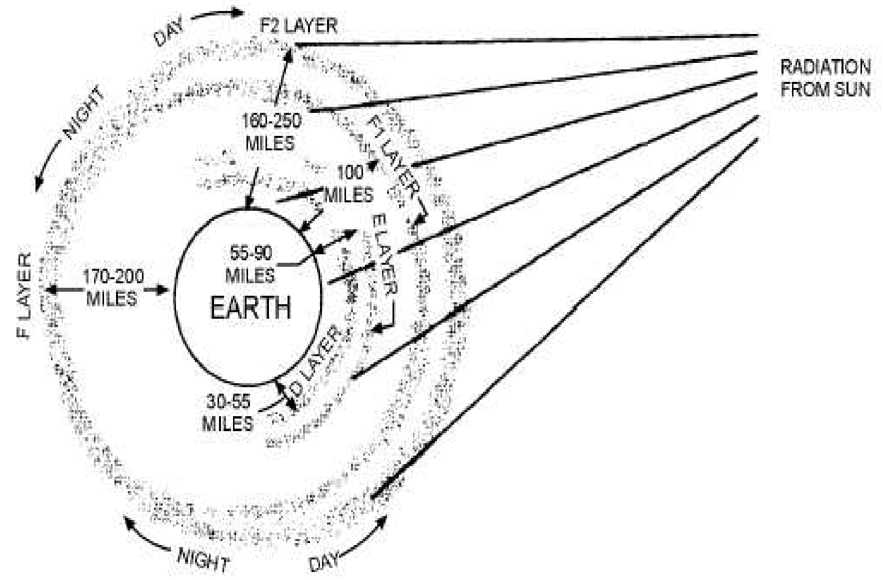


**Fig 5.1: layers of the ionosphere.**

The d layer ranges from about 30 to 55 miles. ionization in the d layer is low because it is the lowest region of the ionosphere. this layer has the ability to refract signals of low frequencies. high frequencies pass right through it and are attenuated. after sunset, the d layer disappears because of the rapid recombination of ions. b. the e layer limits are from about 55 to 90 miles. this layer is also known as the kennelly- heaviside layer, because these two men were the first to propose its existence. the rate of ionic recombination in this layer is rather rapid after sunset and the layer is almost gone by midnight. this layer has the ability to refract signals as high as 20 megahertz. for this reason, it is valuable for communications in ranges up to about 1500 miles. c. the f layer exists from about 90 to 240 miles. during the daylight hours, the f layer separates into two layers, the f1 and f2 layers. the ionization level in these layers is quite high and varies widely during the day. at noon, this portion of the atmosphere is closest to the sun and the degree of ionization is maximum. since the atmosphere is rarefied at these heights, recombination occurs slowly after sunset. therefore, a fairly constant ionized layer is always present. the f layers are responsible for high-frequency, long distance transmission

## Refraction in the ionosphere

when a radio wave is transmitted into an ionized layer, refraction, or bending of the wave, occurs. as we discussed earlier, refraction is caused by an abrupt change in the velocity of the upper part of a radio wave as it strikes or enters a new medium. the amount of refraction that occurs depends on three main factors: (1) the density of ionization of the layer, (2) the frequency of the radio wave, and (3) the angle at which the wave enters the layer.

## DENSITY OF LAYER

The relationship between radio waves and ionization density. each ionized layer has a central region of relatively dense ionization, which tapers off in intensity both above and below the maximum region. as a radio wave enters a region of increasing ionization, the increase in velocity of the upper part of the wave causes it to be bent back toward the earth. while the wave is in the highly dense center portion of the layer, however, refraction occurs more slowly because the density of ionization is almost uniform. as the wave enters into the upper part of the layer of decreasing ionization, the velocity of the upper part of the wave decreases, and the wave is bent away from the earth.

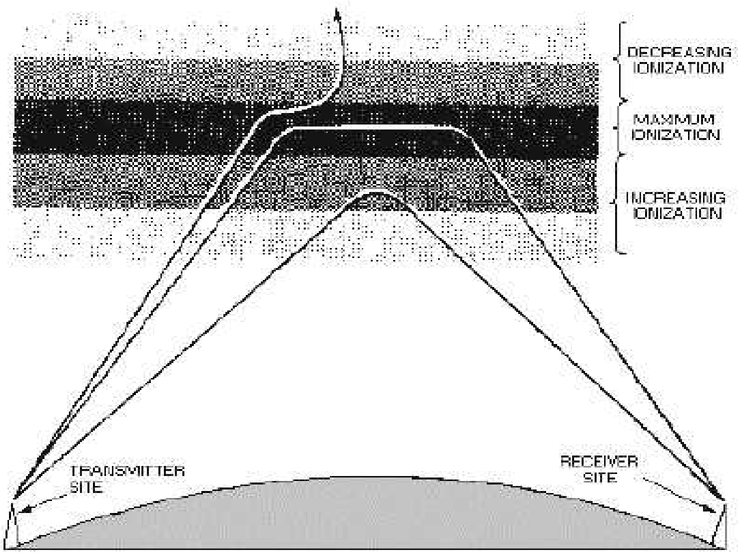


**Fig 5.2: effects of ionospheric density on radio waves.**

if a wave strikes a thin, very highly ionized layer, the wave may be bent back so rapidly that it will appear to have been reflected instead of refracted back to earth. to reflect a radio wave, the highly ionized layer must be approximately no thicker than one wavelength of the radio wave. since the ionized layers are often several miles thick, ionospheric reflection is more likely to occur at long wavelengths (low frequencies).

## Ground Wave Signal Propagation

The ground wave used for radio communications signal propagation on the long, and medium wave bands for local radio communications

Ground wave propagation is particularly important on the LF and MF portion of the radio spectrum. Ground wave radio propagation is used to provide relatively local radio communications coverage, especially by radio broadcast stations that require to cover a particular locality.

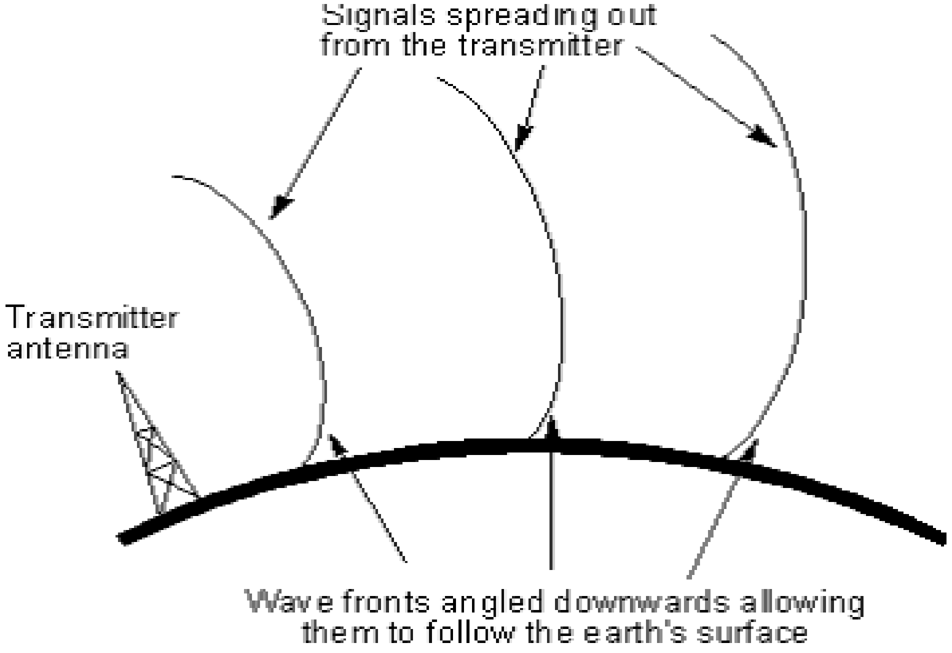
Ground wave radio signal propagation is ideal for relatively short distance propagation on these frequencies during the daytime. Sky-wave ionospheric propagation is not possible during the day because of the attenuation of the signals on these frequencies caused by the D region in the ionosphere. In view of this, radio communications stations need to rely on the ground-wave propagation to achieve their coverage.

A ground wave radio signal is made up from a number of constituents. If the antennas are in the line of sight then there will be a direct wave as well as a reflected signal. As the names suggest the direct signal is one that travels directly between the two antenna and is not affected by the locality. There will also be a reflected signal as the transmission will be reflected by a number of objects including the earth's surface and any hills, or large buildings. That may be present.
In addition to this there is surface wave. This tends to follow the curvature of the Earth and enables coverage to be achieved beyond the horizon. It is the sum of all these components that is known as the ground wave.

Beyond the horizon the direct and reflected waves are blocked by the curvature of the Earth, and the signal is purely made up from the diffracted surface wave. It is for this reason that surface wave is commonly called ground wave propagation.

## Surface wave

The radio signal spreads out from the transmitter along the surface of the Earth. Instead of just travelling in a straight line the radio signals tend to follow the curvature of the Earth. This is because currents are induced in the surface of the earth and this action slows down the wave-front in this region, causing the wave-front of the radio communications signal to tilt downwards towards the Earth. With the wave-front tilted in this direction it is able to curve around the Earth and be received well beyond the horizon.



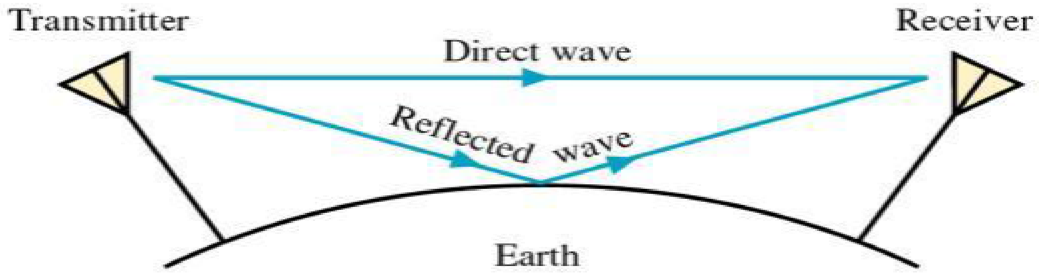
## EFFECT OF POLARISATION

The type of antenna has a major effect. Vertical polarisation is subject to considerably less attenuation than horizontally polarised signals. In some cases the difference can amount to several tens of decibels. It is for this reason that medium wave broadcast stations use vertical antennas, even if they have to be made physically short by adding inductive loading. Ships making use of the MF marine bands often use inverted L antennas as these are able to radiate a significant proportion of the signal that is vertically polarised.

At distances that are typically towards the edge of the ground wave coverage area, some sky-wave signal may also be present, especially at night when the D layer attenuation is reduced. This may serve to reinforce or cancel the overall signal resulting in figures that will differ from those that may be expected.

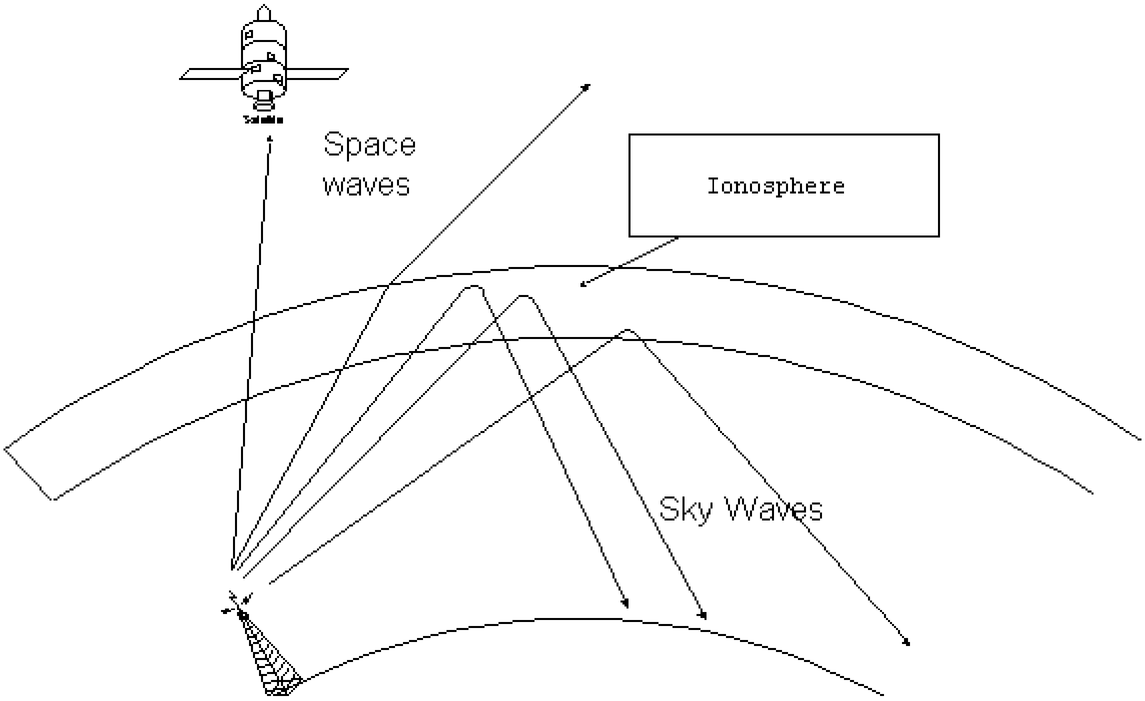
## SPACE (DIRECT) WAVE PROPAGATION

Space Waves, also known as direct waves, are radio waves that travel directly from the transmitting antenna to the receiving antenna. In order for this to occur, the two antennas must be able to ―see‖ each other; that is there must be a line of sight path between them. The diagram on the next page shows a typical line of sight. The maximum line of sight distance between two antennas depends on the height of each antenna.



## SKY WAVES

Radio waves in the LF and MF ranges may also propagate as ground waves, but suffer significant losses, or are attenuated, particularly at higher frequencies. But as the ground wave mode fades out, a new mode develops: the sky wave. Sky waves are reflections from the ionosphere. While the wave is in the ionosphere, it is strongly bent, or refracted, ultimately back to the ground. From a long distance away this appears as a reflection. Long ranges are possible in this mode also, up to hundreds of miles. Sky waves in this frequency band are usually only possible at night, when the concentration of ions is not too great since the ionosphere also tends to attenuate the signal. However, at night, there are just enough ions to reflect the wave but not reduce its power too much.



## Atmospheric Propagation

Within the atmosphere, radio waves can be reflected, refracted, and diffracted like light and heat waves.

### Reflection

Radio waves may be reflected from various substances or objects they meet during travel between the transmitting and receiving sites. The amount of reflection depends on the reflecting material. Smooth metal surfaces of good electrical conductivity are efficient reflectors of radio waves. The surface of the Earth itself is a fairly good reflector. The radio wave is not reflected from a single point on the reflector but rather from an area on its surface. The size of the area required for reflection to take place depends on the wavelength of the radio wave and the angle at which the wave strikes the reflecting substance. When radio waves are reflected from flat surfaces, a phase shift in the alternations of the wave occurs..

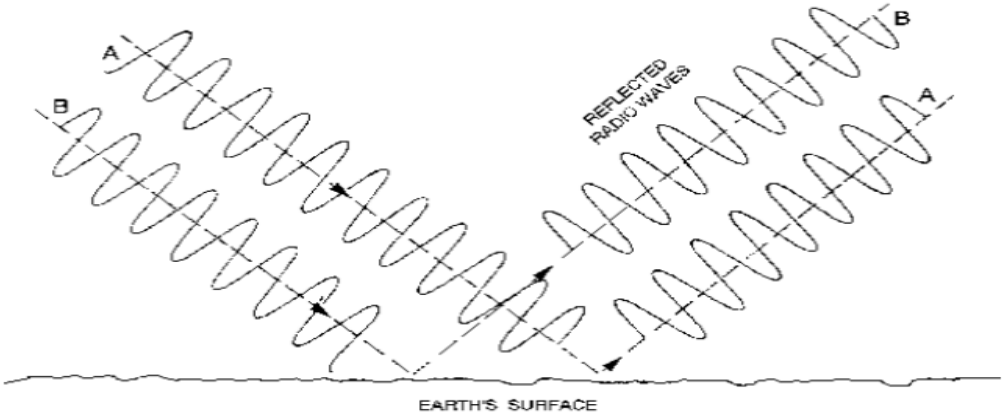
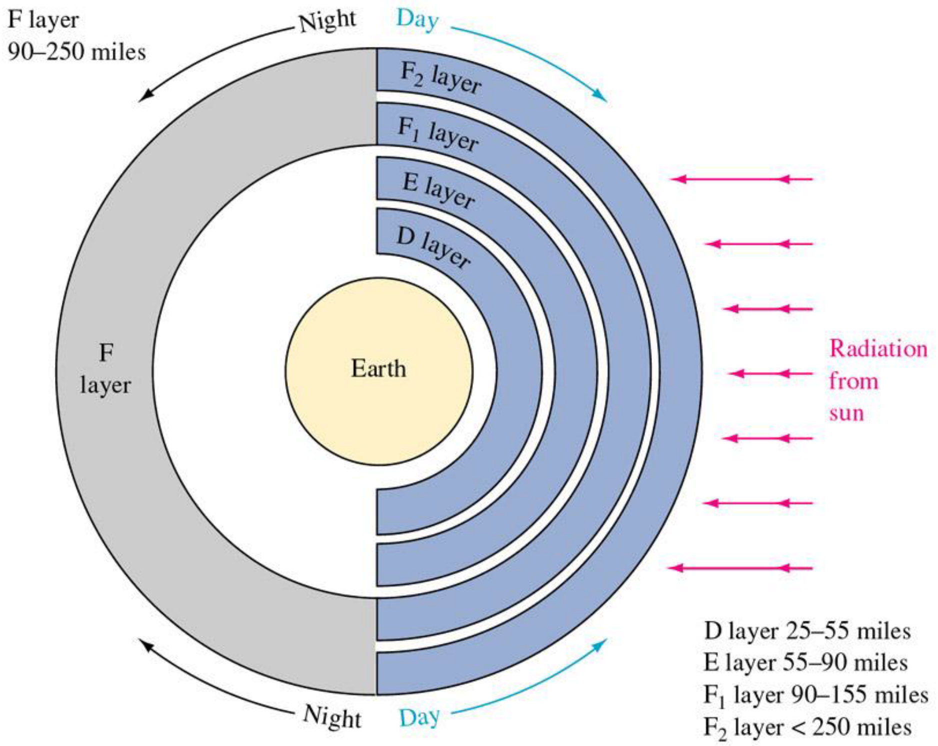


Fig 5.3: Reflection

Notice that the positive and negative alternations of radio waves (A) and (B) are in phase with each other in their paths toward the Earth's surface. After reflection takes place, however, the waves are approximately 180 degrees out of phase from their initial relationship. The amount of phase shift that occurs is not constant. It depends on the polarization of the wave and the angle at which the wave strikes the reflecting surface. Radio waves that keep their phase relationships after reflection normally produce a stronger signal at the receiving site. Those that are received out of phase produce a weak or fading signal. The shifting in the phase relationships of reflected radio waves is one of the major reasons for fading.

**THE IONOSPHERIC LAYERS**



**Ionospheric Storms**: Solar activity such as flares and coronal mass ejections produce large electromagnetic radiation incidents upon the ear disturbances of the ionosphere; changes the density distribution, electron content, and the ionospheric current system. These storms can also disrupt satellite communications and cause a loss of radio frequencies which would otherwise reflect off the ionosphere. Ionospheric storms can last typically for a day or so.

**D layer Absorption**: Occurs when the ionosphere is strongly charged (daytime, summer, heavy solar activity) longer waves will be absorbed and never return to earth. You don't hear distant AM broadcast stations during the day. Shorter waves will be reflected and travel further. Absorption occurs in the D layer which is the lowest layer in the ionosphere. The intensity of this layer is increased as the sun climbs above the horizon and is greatest at noon. Radio waves below 3 or 4 MHz are absorbed by the D layer when it is present.

When the ionosphere is weakly charged (night time, winter, low solar activity) longer waves will travel a considerable distance but shorter waves may pass through the ionosphere and escape into space. VHF waves pull this trick all the time, hence their short range and usefulness for communicating with satellites.

Faraday rotation: EM waves passing through the ionosphere may have their polarizations changed to random directions (refraction) and propagate at different speeds. Since most radio waves are either vertically or horizonally polarized, it is difficult to predict what the polarization of the waves will be when they arrive at a receiver after reflection in the ionosphere.

- Solar radiation, acting on the different compositions of the atmosphere generates layers of ionization
- Studies of the ionosphere have determined that there are at least four distinct layers of D, E, F1, and F2 layers.
- The F layer is a single layer during the night and other periods of low ionization, during the day and periods of higher ionization it splits into two distinct layers, the F1 and F2.
- There are no clearly defined boundaries between layers. These layers vary in density depending on the time of day, time of year, and the amount of solar (sun) activity.
- The top-most layer (F and F1/F2) is always the most densely ionized because it is least protected from the Sun.

## SOLAR CYCLE

Every 11 years the sun undergoes a period of activity called the "solar maximum", followed by a period of quiet called the "solar minimum". During the solar
th and leads to

## Refraction

Another phenomenon common to most radio waves is the bending of the waves as they move from one medium into another in which the velocity of propagation is different. This bending of the waves is called refraction. For example, suppose you are driving down a smoothly paved road at a constant speed and suddenly one wheel goes off onto the soft shoulder. The car tends to veer off to one side. The change of medium, from hard surface to soft shoulder, causes a change in speed or velocity. The tendency is for the car to change direction. This same principle applies to radio waves as changes occur in the medium through which they are passing. As the wave enters the dense layer of electrically charged ions, the part of the wave that enters the new medium first travels faster than the parts of the wave that have not yet entered the new medium. This abrupt increase in velocity of the upper part of the wave causes the wave to bend back toward the Earth. This bending, or change of direction, is always toward the medium that has the lower velocity of propagation.

## Diffraction

A radio wave that meets an obstacle has a natural tendency to bend around the obstacle as illustrated in figure. The bending, called diffraction, results in a change of direction of part of the wave energy from the normal line-of-sight path. This change makes it possible to receive energy around the edges of an obstacle as shown in view A or at some distances below the highest point of an obstruction, as shown in view B.

Although diffracted rf energy usually is weak, it can still be detected by a suitable receiver. The principal effect of diffraction extends the radio range beyond the visible horizon. In certain cases, by using high power and very low frequencies, radio waves can be made to encircle the Earth by diffraction.

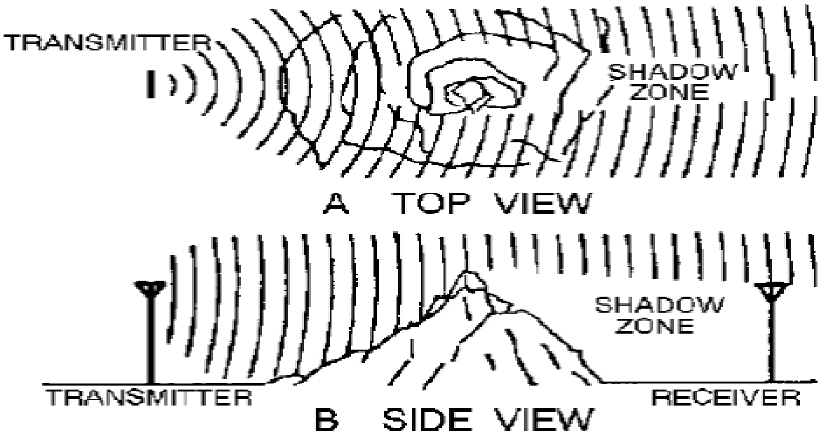


**Fig 5.4:Diffraction**

There are two principal ways in which electromagnetic (radio) energy travels from a transmitting antenna to a receiving antenna. One way is by GROUND WAVES and the other is by SKY WAVES. Ground waves are radio waves that travel near the surface of the Earth (surface and space waves). Sky waves are radio waves that are reflected back to Earth from the ionosphere.

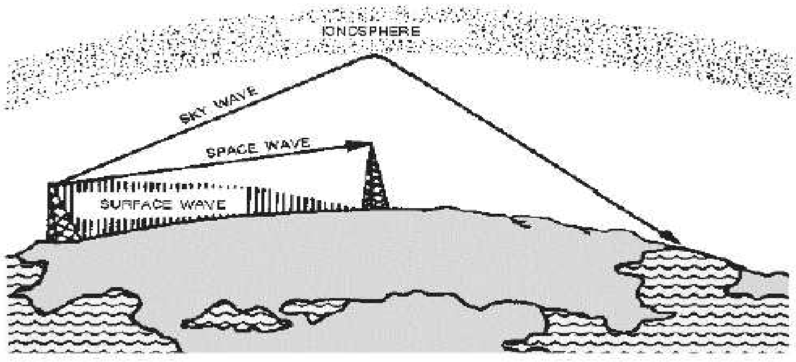


**Fig 5.4:Diagram showing different waves**

## Ground Waves

The ground wave is actually composed of two separate component waves. These are known as the SURFACE WAVE and the SPACE .The determining factor in whether a ground wave component is classified as a space wave or a surface wave is simple. A surface wave travels along the surface of the Earth. A space wave travels over the surface.

**Surface Wave**.

The surface wave reaches the receiving site by traveling along the surface of the ground as shown in figure . A surface wave can follow the contours of the Earth because of the process of diffraction. When a surface wave meets an object and the dimensions of the object do not exceed its wavelength, the wave tends to curve or bend around the object. The smaller the object, the more pronounced the diffractive action will be.

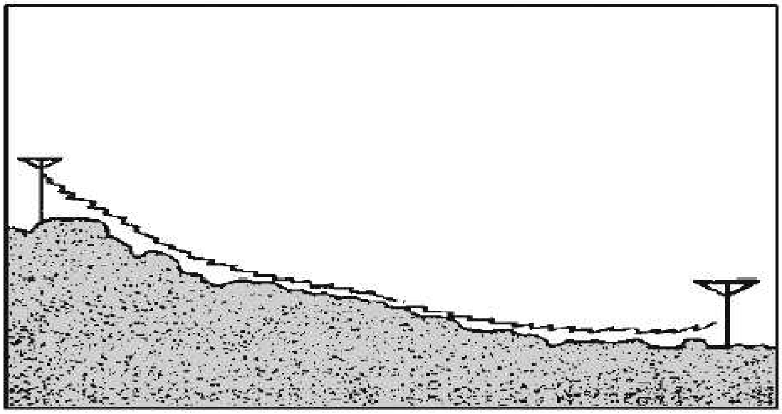


**Figure 5.5: Surface wave propagation**.

As a surface wave passes over the ground, the wave induces a voltage in the Earth. The induced voltage takes energy away from the surface wave, thereby weakening, or attenuating, the wave as it moves away from the transmitting antenna. To reduce the attenuation, the amount of induced voltage must be reduced. This is done by using vertically polarized waves that minimize the extent to which the electric field of the wave is in contact with the Earth. When a surface wave is horizontally polarized, the electric field of the wave is parallel with the surface of the Earth and, therefore, is constantly in contact with it. The wave is then completely attenuated within a short distance from the transmitting site. On the other hand, when the surface wave is vertically polarized, the electric field is vertical to the Earth and merely dips into and out of the Earth's surface. For this reason, vertical polarization is vastly superior to horizontal polarization for surface wave propagation. The attenuation that a surface wave undergoes because of induced voltage also depends on the electrical properties of the terrain over which the wave travels. The best type of surface is one that has good electrical conductivity. The better the conductivity, the less the attenuation.

Another major factor in the attenuation of surface waves is frequency. Recall from earlier discussions on wavelength that the higher the frequency of a radio wave, the shorter its wavelength will be. These high frequencies, with their shorter wavelengths, are not normally diffracted but are absorbed by the Earth at points relatively close to the transmitting site. You can assume, therefore, that as the frequency of a surface wave is increased, the more rapidly the surface wave will be absorbed, or attenuated, by the Earth. Because of this loss by attenuation, the surface wave is impractical for long-distance transmissions at frequencies above 2 megahertz. On the other hand, when the frequency of a surface wave is low enough to have a very long wavelength, the Earth appears to be very small, and diffraction is sufficient for propagation well beyond the horizon. In fact, by lowering the transmitting frequency into the very low frequency (vlf) range and using very high-powered transmitters, the surface wave can be propagated great distances. The Navy's extremely high-powered vlf transmitters are actually capable of transmitting surface wave signals around the Earth and can provide coverage to naval units operating anywhere at sea.

### Space Wave

The space wave follows two distinct paths from the transmitting antenna to the receiving antenna—one through the air directly to the receiving antenna, the other reflected from the ground to the receiving antenna. This is illustrated in figure 4-9. The primary path of the space wave is directly from the transmitting antenna to the receiving antenna. So, the receiving antenna must be located within the radio horizon of the transmitting antenna. Because space waves are refracted slightly, even when propagated through the troposphere, the radio horizon is actually about one-third farther than the line-of-sight or natural horizon.

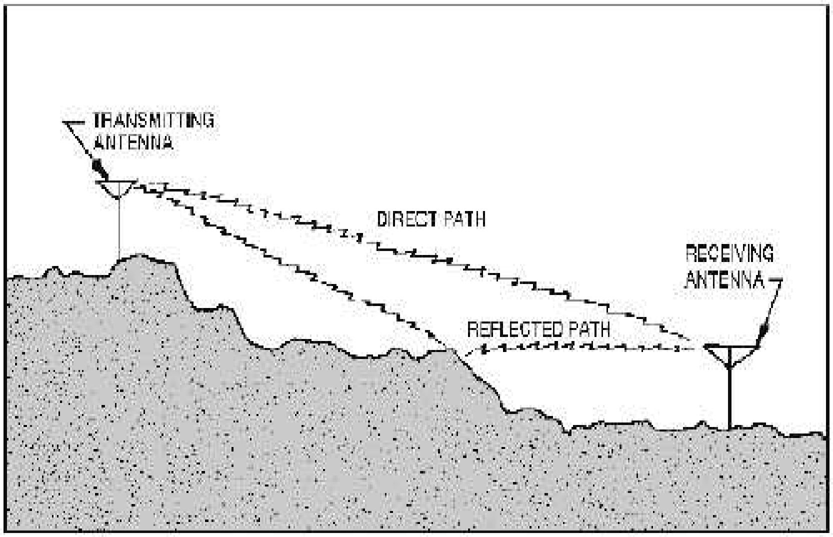


**Fig5.6:Space wave propagation**

Although space waves suffer little ground attenuation, they nevertheless are susceptible to fading. This is because space waves actually follow two paths of different lengths (direct path and ground reflected path) to the receiving site and, therefore, may arrive in or out of phase. If these two component waves are received in phase, the result is a reinforced or stronger signal. Likewise, if they are received out of phase, they tend to cancel one another, which results in a weak or fading signal.

### Sky Wave

The sky wave, often called the ionospheric wave, is radiated in an upward direction and returned to Earth at some distant location because of refraction from the ionosphere. This form of propagation is relatively unaffected by the Earth's surface and can propagate signals over great distances. Usually the high frequency (hf) band is used for sky wave propagation.

## FADING

The most troublesome and frustrating problem in receiving radio signals is variations in signal strength, most commonly known as FADING. There are several conditions that can produce fading. When a radio wave is refracted by the ionosphere or reflected from the Earth's surface, random changes in the polarization of the wave may occur. Vertically and horizontally mounted receiving antennas are designed to receive vertically and horizontally polarized waves, respectively. Therefore, changes in polarization cause changes in the received signal level because of the inability of the antenna to receive polarization changes. Fading also results from absorption of the rf energy in the ionosphere. Absorption fading occurs for a longer period than other types of fading, since absorption takes place slowly. Usually, however, fading on ionospheric circuits is mainly a result of multipath propagation.

## MULTIPATH FADING

MULTIPATH is simply a term used to describe the multiple paths a radio wave may follow between transmitter and receiver. Such propagation paths include the ground wave, ionospheric refraction, reradiation by the ionospheric layers, reflection from the Earth's surface or from more than one ionospheric layer, etc. Figure 2-21 shows a few of the paths that a signal can travel between two sites in a typical circuit. One path, XYZ, is the basic ground wave. Another path, XEA, refracts the wave at the E layer and passes it on to the receiver at A. Still another path, XFZFA, results from a greater angle of incidence and two refractions from the F layer. At point Z, the received signal is a combination of the ground wave and the sky wave. These two signals having travelled different paths arrive at point Z at different times. Thus, the arriving waves may or may not be in phase with each other. Radio waves that are received in phase reinforce each other and produce a stronger signal at the receiving site. Conversely, those that are received out of phase produce a weak or fading signal. Small alternations in the transmission path may change the phase relationship of the two signals, causing periodic fading. This condition occurs at point A. At this point, the double-hop F layer signal may be in or out of phase with the signal arriving from the E layer.

Multipath fading may be minimized by practices called SPACE DIVERSITY and FREQUENCY DIVERSITY. In space diversity, two or more receiving antennas are spaced some distance apart. Fading does not occur simultaneously at both antennas; therefore, enough output is almost always available from one of the antennas to provide a useful signal. In frequency diversity, two transmitters and two receivers are used, each pair tuned to a different frequency, with the same information being transmitted simultaneously over both frequencies. One of the two receivers will almost always provide a useful signal.

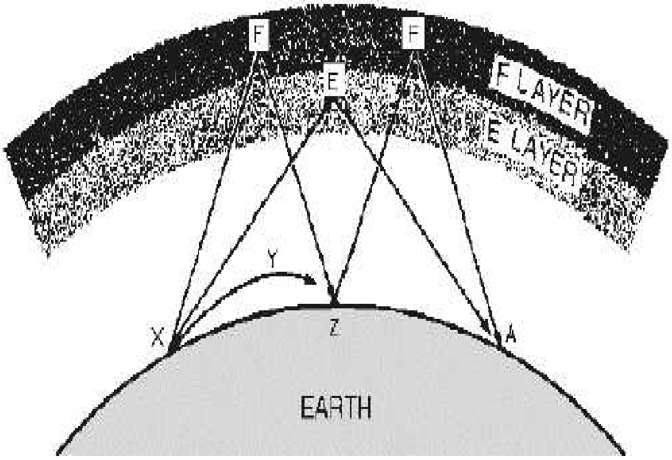


Fig 5.7:Multipath fading

### Selective Fading

Fading resulting from multipath propagation is variable with frequency since each frequency arrives at the receiving point via a different radio path. When a wide band of frequencies is transmitted simultaneously, each frequency will vary in the amount of fading. This variation is called SELECTIVE FADING. When selective fading occurs, all frequencies of the transmitted signal do not retain their original phases and relative amplitudes. This fading causes severe distortion of the signal and limits the total signal transmitted.

### Skip Distance/Skip Zone

The SKIP DISTANCE is the distance from the transmitter to the point where the sky wave is first returned to Earth. The size of the skip distance depends on the frequency of the wave, the angle of incidence, and the degree of ionization present.

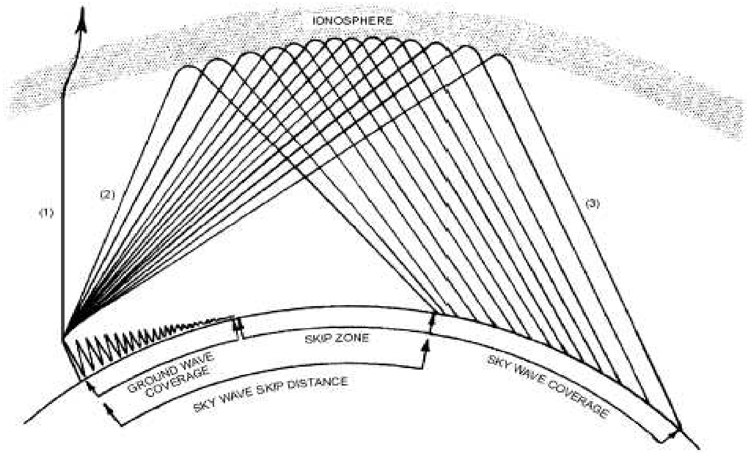


Fig 5.8: Relationship between skip zone, skip distance, and ground wave.

The SKIP ZONE is a zone of silence between the point where the ground wave becomes too weak for reception and the point where the sky wave is first returned to Earth. The size of the skip zone depends on the extent of the ground wave coverage and the skip distance. When the ground wave coverage is great enough or the skip distance is short enough that no zone of silence occurs, there is no skip zone. Occasionally, the first sky wave will return to Earth within the range of the ground wave. If the sky wave and ground wave are nearly of equal intensity, the sky wave alternately reinforces and cancels the ground wave, causing severe fading. This is caused by the phase difference between the two waves, a result of the longer path traveled by the sky wave.

## Maximum Usable Frequency

The higher the frequency of a radio wave, the lower the rate of refraction by an ionized layer. Therefore, for a given angle of incidence and time of day, there is a maximum frequency that can be used for communications between two given locations. This frequency is known as the MAXIMUM USABLE FREQUENCY (muf). Waves at frequencies above the muf are normally refracted so slowly that they return to Earth beyond the desired location, or pass on through the ionosphere and are lost. You should understand, however, that use of an established muf certainly does not guarantee successful communications between a transmitting site and a receiving site. Variations in the ionosphere may occur at any time and consequently raise or lower the predetermined muf. This is particularly true for radio waves being refracted by the highly variable F2 layer. The muf is highest around noon when ultraviolet light waves from the sun are the most intense. It then drops rather sharply as recombination begins to take place.

## CRITICAL FREQUENCY

For any given time, each ionospheric layer has a maximum frequency at which radio waves can be transmitted vertically and refracted back to Earth. This frequency is known as the CRITICAL FREQUENCY. It is a term that you will hear frequently in any discussion of radio wave propagation. Radio waves transmitted at frequencies higher than the critical frequency of a given layer will pass through the layer and be lost in space; but if these same waves enter an upper layer with a higher critical frequency, they will be refracted back to Earth. Radio waves of frequencies lower than the critical frequency will also be refracted back to Earth unless they are absorbed or have been refracted from a lower layer. The lower the frequency of a radio wave, the more rapidly the wave is refracted by a given degree of ionization.. Notice that the 5-megahertz wave is refracted quite sharply. The 20-megahertz wave is refracted less sharply and returned to Earth at a greater distance. The 100-megahertz wave is obviously greater than the critical frequency for that ionized layer and, therefore, is not refracted but is passed into space.

**Ionospheric Storms** - large scale changes in the chemical composition of the ionosphere resulting in changes to the MUF. Decreased MUFs restrict the frequencies available for use over a given distance. Ionospheric storms normally last for one to two day

Critical Frequency:

The highest frequency that will be returned to the earth when transmitted vertically under given ionospheric conditions

Critical Angle:

The highest angle with respect to a vertical line at which a radio wave of a specified frequency can be propagated and still be returned to the earth from the ionosphere

**Maximum usable frequency (MUF)**

The highest frequency that is returned to the earth from the ionosphere between two specific points on earth

Optimum Working frequency:

The frequency that provides for the most consistent communication path via sky waves

**Tropospheric Scattering**

Signals are aimed at the troposphere rather than the ionosphere

350 Mhz to 10GHz for paths up to 400 mi

Received signal = 10-6 th of the transmitted power

Fading a problem

## DIVERSITY TRANSMISSION AND RECEPTION

In telecommunications, a **diversity scheme** refers to a method for improving the reliability of a message signal by using two or more communication channels with different characteristics. Diversity is mainly used in radio communication and is a common technique for combatting fading and co-channel interference and avoiding error bursts. It is based on the fact that individual channels experience different levels of fading and interference. Multiple versions of the same signal may be transmitted and/or received and combined in the receiver. Alternatively, a redundant forward error correction code may be added and different parts of the message transmitted over different channels. Diversity techniques may exploit the multipath propagation, resulting in a diversity gain, often measured in decibels.

The following classes of diversity schemes can be identified:

- **Time diversity**: Multiple versions of the same signal are transmitted at different time instants. Alternatively, a redundant forward error correction codeis added and the message is spread in time by means of bit-interleaving before it is transmitted. Thus, error bursts are avoided, which simplifies the error correction.

- **Frequency diversity**: The signal is transmitted using several frequency channels or spread over a wide spectrum that is affected by frequency-selective fading. Middle-late 20th century microwave radio relay lines often used several regular wideband radio channels, and one protection channel for automatic use by any faded channel. Later examples include:

  - OFDM modulation in combination with subcarrier interleaving and forward error correction

  - Spread spectrum, for example frequency hopping or DS-CDMA.

- **Space diversity**: The signal is transmitted over several different propagation paths. In the case of wired transmission, this can be achieved by transmitting via multiple wires. In the case of wireless transmission, it can be achieved by antenna diversity using multiple transmitter antennas (transmit diversity) and/or multiple receiving antennas (reception diversity). In the latter case, a diversity combining technique is applied before further signal processing takes place. If the antennas are far apart, for example at different cellular base station sites or WLAN access points, this is calledmacrodiversity or site diversity. If the antennas are at a distance in the order of one wavelength, this is called microdiversity. A special case is phasedantenna arrays, which also can be used for beamforming, MIMO channels and space–time coding (STC).

- **Polarization diversity**: Multiple versions of a signal are transmitted and received via antennas with different polarization. A diversity combiningtechnique is applied on the receiver side.

- **Multiuser diversity**: Multiuser diversity is obtained by opportunistic user scheduling at either the transmitter or the receiver. Opportunistic user scheduling is as follows: at any given time, the transmitter selects the best user among candidate receivers according to the qualities of each channel between the transmitter and each receiver. A receiver must feed back the channel quality information to the transmitter using limited levels of resolution, in order for the transmitter to implement Multiuser diversity.

- **Cooperative diversity**: Achieves antenna diversity gain by using the cooperation of distributed antennas belonging to each node.

## FADING

In wireless communications, **fading** is deviation of the attenuation affecting a signal over certain propagation media. The fading may vary with time, geographical position or radio frequency, and is often modeled as a random process. A **fading channel** is a communication channel that experiences fading. In wireless systems, fading may either be due to multipath propagation, referred to as **multipath induced fading**, or due to shadowing from obstacles affecting the wave propagation, sometimes referred to as **shadow fading**.

The presence of reflectors in the environment surrounding a transmitter and receiver create multiple paths that a transmitted signal can traverse. As a result, the receiver sees the superposition of multiple copies of the transmitted signal, each traversing a different path. Each signal copy will experience differences in attenuation, delay and phase shiftwhile travelling from the source to the receiver. This can result in either constructive or destructive interference, amplifying or attenuating the signal power seen at the receiver. Strong destructive interference is frequently referred to as a **deep fade** and may result in temporary failure of communication due to a severe drop in the channel signal-to-noise ratio.

A common example of deep fade is the experience of stopping at a traffic light and hearing an FM broadcast degenerate into static, while the signal is re-acquired if the vehicle moves only a fraction of a meter. The loss of the broadcast is caused by the vehicle stopping at a point where the signal experienced severe destructive interference. Cellular phones can also exhibit similar momentary fades.

Fading channel models are often used to model the effects of electromagnetic transmission of information over the air in cellular networks and broadcast communication. Fading channel models are also used in underwater acoustic communications to model the distortion caused by the water.

## FADING TYPES

### SLOW VERSUS FAST FADING

The terms *slow* and *fast* fading refer to the rate at which the magnitude and phase change imposed by the channel on the signal changes. The coherence time is a measure of the minimum time required for the magnitude change or phase change of the channel to become uncorrelated from its previous value.

- **Slow fading** arises when the coherence time of the channel is large relative to the delay constraint of the channel. In this regime, the amplitude and phase change imposed by the channel can be considered roughly constant over the period of use. Slow fading can be caused by events such as **shadowing**, where a large obstruction such as a hill or large building obscures the main signal path between the transmitter and the receiver. The received power change caused by shadowing is often modeled using a log-normal distribution with a standard deviation according to the log-distance path loss model.

- **Fast fading** occurs when the coherence time of the channel is small relative to the delay constraint of the channel. In this case, the amplitude and phase change imposed by the channel varies considerably over the period of use.

In a fast-fading channel, the transmitter may take advantage of the variations in the channel conditions using time diversity to help increase robustness of the communication to a temporary deep fade. Although a deep fade may temporarily erase some of the information transmitted, use of an error-correcting code coupled with successfully transmitted bits during other time instances (interleaving) can allow for the erased bits to be recovered. In a slow-fading channel, it is not possible to use time diversity because the transmitter sees only a single realization of the channel within its delay constraint. A deep fade therefore lasts the entire duration of transmission and cannot be mitigated using coding.

The coherence time of the channel is related to a quantity known as the **Doppler spread** of the channel. When a user (or reflectors in its environment) is moving, the user's velocity causes a shift in the frequency of the signal transmitted along each signal path. This phenomenon is known as the Doppler shift. Signals traveling along different paths can have different Doppler shifts, corresponding to different rates of change in phase. The difference in Doppler shifts between different signal components contributing to a signal fading channel tap is known as the Doppler spread. Channels with a large Doppler spread have signal components that are each changing independently in phase over time. Since fading depends on whether signal components add constructively or destructively, such channels have a very short coherence time.

In general, coherence time is inversely related to Doppler spread, typically expressed as

$$T_c \approx \frac{1}{D_s}$$

where $T_c$ is the coherence time, $D_s$ is the Doppler spread. This equation is just an approximation,[1] to be exact, see Coherence time.

## BLOCK FADING

**Block fading** is where the fading process is approximately constant for a number of symbol intervals. A channel can be 'doubly block-fading' when it is block fading in both the time and frequency domains

## SELECTIVE FADING OR FREQUENCY SELECTIVE

**Selective fading** or **frequency selective fading** is a radio propagation anomaly caused by partial cancellation of a radio signal by itself — the signal arrives at the receiver bytwo different paths, and at least one of the paths is changing (lengthening or shortening). This typically happens in the early evening or early morning as the various layers in theionosphere move, separate, and combine. The two paths can both be skywave or one be groundwave.

Selective fading manifests as a slow, cyclic disturbance; the cancellation effect, or "null", is deepest at one particular frequency, which changes constantly, sweeping through the received audio.

As the carrier frequency of a signal is varied, the magnitude of the change in amplitude will vary. The coherence bandwidth measures the separation in frequency after which two signals will experience uncorrelated fading.

- In **flat fading**, the coherence bandwidth of the channel is larger than the bandwidth of the signal. Therefore, all frequency components of the signal will experience the same magnitude of fading.

- In **frequency-selective fading**, the coherence bandwidth of the channel is smaller than the bandwidth of the signal. Different frequency components of the signal therefore experience uncorrelated fading.

Fading can cause poor performance in a communication system because it can result in a loss of signal power without reducing the power of the noise. This signal loss can be over some or all of the signal bandwidth. Fading can also be a problem as it changes over time: communication systems are often designed to adapt to such impairments, but the fading can change faster than the adaptations can be made. In such cases, the probability of experiencing a fade (and associated bit errors as the signal-to-noise ratio drops) on the channel becomes the limiting factor in the link's performance.

## MITIGATION

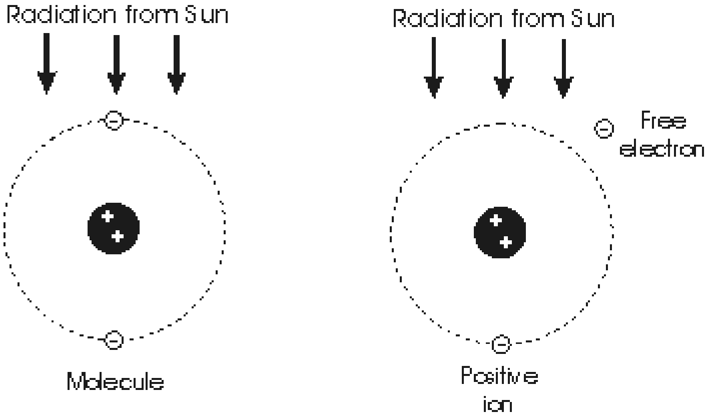
The effects of fading can be combated by using diversity to transmit the signal over multiple channels that experience independent fading and coherently combining them at the receiver. The probability of experiencing a fade in this composite channel is then proportional to the probability that all the component channels simultaneously experience a fade, a much more unlikely event.

Diversity can be achieved in time, frequency, or space. Common techniques used to overcome signal fading include:

- Diversity reception and transmission
- MIMO
- OFDM
- Rake receivers
- Space–time codes

## MECHANISM OF IONOSPHERIC PROPAGATION

The Sun emits vast quantities of radiation of all wavelengths and this travels towards the Earth, first reaching the outer areas of the atmosphere. In creating the ionisation it is found that when radiation of sufficient intensity strikes an atom or a molecule, energy may be removed from the radiation and an electron removed, producing a free electron and a positive ion. In the example given below, the simple example of a helium atom is give, although other gases including oxygen and nitrogen are far more common.



**Ionisation of molecules by solar radiation**

The radiation from the Sun covers a vast spectrum of wavelengths. However in terms of the effect it has on the atoms of molecules it can be considered as photons. The electrons in the atoms or molecules can be considered as orbiting the central nucleus consisting of protons and neutrons. Electrons are tied or bound to their orbit around the nucleus by electro-static forces, the electron is negatively charged and the nucleus is positively charged. There are equal numbers of electrons and protons in any molecule and as a result it is electro-statically neutral.

When a photon strikes the atom, or molecule, the photon transfers its energy to the electron as excess kinetic energy. Under some circumstances this excess energy may exceed the binding energy in the atom or molecule and the electron escapes the influence of the positive charge of the nucleus. This leaves a positively charged nucleus or ions and a negatively charged electron, although as there are the same number of positive ions and negative electrons the whole gas still remains with an overall neutral charge.

Most of the ionisation in the ionosphere results from ultraviolet light, although this does not mean that other wavelengths do not have some effect. Additionally, each time an atom or molecule is ionised a small amount of energy is used. This means that as the radiation passes further into the atmosphere, its intensity reduces. It is for this reason that the ultraviolet radiation causes most of the ionisation in the upper reaches of the ionosphere, but at lower altitudes the radiation that is able to penetrate further cause more of the ionisation. Accordingly, extreme ultra-violet and X-Rays give rise to most of the ionisation at lower altitudes. This reduction in these forms of radiation protects us on the surface of the Earth from the harmful effects of these rays.

The level of ionisation varies over the extent of the ionosphere, being far from constant. One reason is that the level of radiation reduces with decreasing altitude. Also the density of the gases varies. In addition to this there is a variation in the proportions of monatomic and molecular forms of the gases, the monatomic forms of gases being far greater at higher altitudes. These and a variety of other phenomena mean that there are variations in the level of ionisation with altitude.

The level of ionisation in the ionosphere also changes with time. It varies with the time of day, time of year, and according to many other external influences. One of the main reasons why the electron density varies is that the Sun, which gives rise to the ionisation is only visible during the day. While the radiation from the Sun causes the atoms and molecules to split into free electrons and positive ions. The reverse effect also occurs. When a negative electron meets a positive ion, the fact that dissimilar charges attract means that they will be pulled towards one another and they may combine. This means that two opposite effects of splitting and recombination are taking place. This is known as a state of dynamic equilibrium. Accordingly the level of ionisation is dependent upon the rate of ionisation and recombination. This has a significant effect on radio communications.

## CRITICAL FREQUENCY

The critical frequency is an important figure that gives an indication of the state of the ionosphere and the resulting HF propagation. It is obtained by sending a signal pulse directly upwards. This is reflected back and can be received by a receiver on the same site as the transmitter. The pulse may be reflected back to earth, and the time measured to give an indication of the height of the layer. As the frequency is increased a point is reached where the signal will pass right through the layer, and on to the next one, or into outer space. The frequency at which this occurs is called the critical frequency.

The equipment used to measure the critical frequency is called an ionosonde. In many respects it resembles a small radar set, but for the HF bands. Using these sets a plot of the reflections against frequency can be generated. This will give an indication of the state of the ionosphere for that area of the world

## MAXIMUM USABLE FREQUENCY, MUF

When a signal is transmitted using HF propagation, over a given path there is a maximum frequency that can be used. This results from the fact that as the signal frequency increases it will pass through more layers and eventually travelling into outer space. As it passes through one layer it may be that communication is lost because the signal then propagates over a greater distance than is required. Also when the signal passes through all the layers communication will be lost.

The frequency at which radio communications just starts to fail is known as the Maximum Usable Frequency (MUF). As a rule of thumb it is generally three (for the F region) to five (for the E region) times the critical and it is true for low angles of incidence, although more exact methods are available for determining this figure.

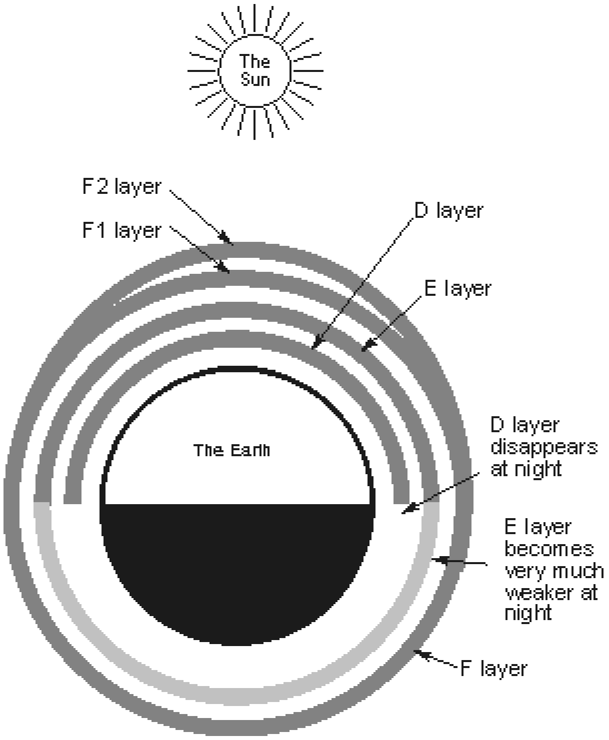
It is possible to calculate the relationship more exactly:

$$\mathbf{MUF = F / (\sec\theta)}$$

Where:

MUF = Maximum Usable Frequency
F = frequency
$\theta$ = the angle the incident ray makes with a vertical line through the point of incidence.

The factor sec $\theta$ is called the MUF factor and it is a function of the path length if the height layer is known. By using typical figures for the heights of the different ionospheric regions the factors may be determined.



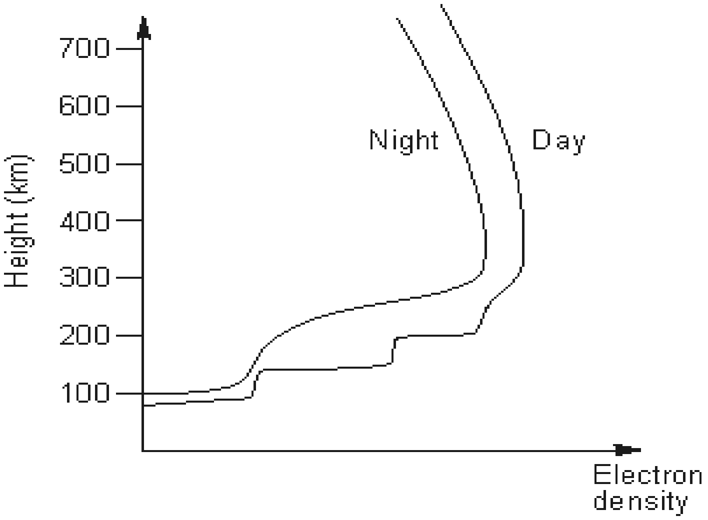
**A simplified view of the layers in the ionosphere over the period of a day**

Other effects like the season and the state of the Sun also have a major effect. Sunspots and solar disturbances have a major impact on the level of radiation received, and these effects are covered in other articles on this website on Sunspots and Solar Disturbances. The season also has an effect. Again this is covered in other articles on the Radio-Electronics.Com website. However very briefly, the radiation received from the Sun varies in the same way that heat from the Sun varies according to the season, and accordingly the level of ionisation and free electrons changes. However this is a very simplified view as other facts also come into play.

## IONOSPHERIC LAYERS

The traditional view of the ionosphere indicates a number of distinct layers, each affecting radio communications in slightly different ways. Indeed, the early discoveries of the ionosphere indicated that a number of layers were present. While this is a convenient way of picturing the structure of the ionosphere it is not exactly correct.

Ionisation exists over the whole of the ionosphere, its level varying with altitude. The peaks in level may be considered as the different layers or possibly more correctly, regions. These regions are given letter designations: D, E, and F regions. There is also a C region below the others, but the level of ionisation is so low that it does not have any effect radio signals and radio communications, and it is rarely mentioned.



**The typical electron distribution in the ionosphere**

The different layers or regions in the ionosphere have different characteristics and affect radio communications in different ways. There are also differences in the exact way they are created and sustained. In view of this it is worth taking a closer look at each one in detail and the way they vary over the complete day during light and darkness.

### D Region

The D region is the lowest of the regions within the ionosphere that affects radio communications signals to any degree. It is present at altitudes between about 60 and 90 kilometres and the radiation within it is only present during the day to an extent that affects radio waves noticeably. It is sustained by the radiation from the Sun and levels of ionisation fall rapidly at dusk when the source of radiation is removed. It mainly has the affect of absorbing or attenuating radio communications signals particularly in the LF and MF portions of the radio spectrum, its affect reducing with frequency. At night it has little effect on most radio communications signals although there is still a sufficient level of ionisation for it to refract VLF signals.

The layer is chiefly generated by the action of a form of radiation known as Lyman radiation which has a wavelength of 1215 Angstroms and ionises nitric oxide gas present in the atmosphere. Hard X-Rays also contribute to the ionisation, especially towards the peak of the solar cycle.

### E Region

The region above the D region is the E region. It exists at altitudes between about 100 and 125 kilometres. Instead of attenuating radio communications signals this layer chiefly refracts them, often to a degree where they are returned to earth. As such they appear to have been reflected by this layer. However this layer still acts as an attenuator to a certain degree.

Like the D region, the level of ionisation falls relatively quickly after dark as the electrons and ions re-combine and it virtually disappears at night. However the residual night time ionisation in the lower part of the E region causes some attenuation of signals in the lower portions of the HF part of the radio communications spectrum.

The ionisation in this region results from a number of types of radiation. Soft X-Rays produce much of the ionisation, although extreme ultra-violet (EUV) rays (very short wavelength ultra-violet light) also contribute. Broadly the radiation that produces ionisation in this region has wavelengths between about 10 and 100 Angstroms. The degree to which all of the constituents contribute depends upon the state of the Sun and the latitude at which the observations are made.

### F Region

The most important region in the ionosphere for long distance HF radio communications is the F region. During the daytime when radiation is being received from the Sun, it often splits into two, the lower one being the F1 region and the higher one, the F2 region. Of these the F1 region is more of an inflection point in the electron density curve (seen above) and it generally only exists in the summer.

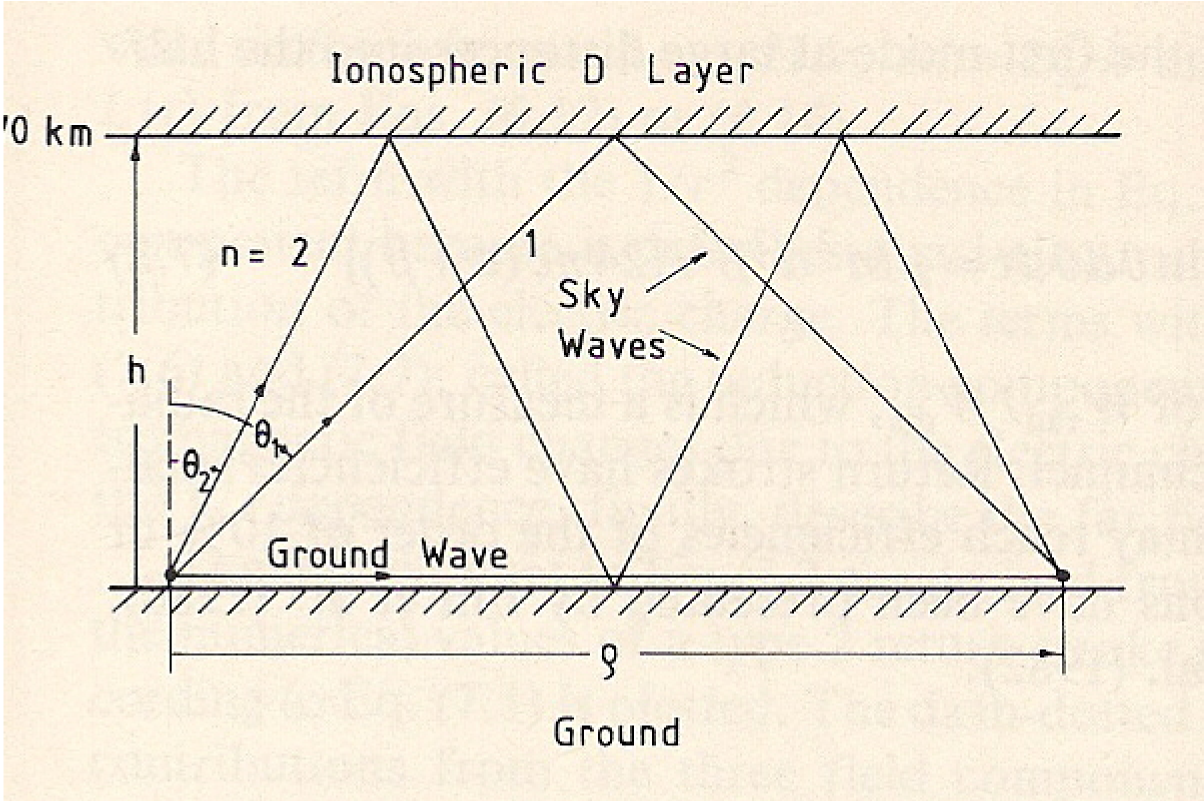
Typically the F1 layer is found at around an altitude of 300 kilometres with the F2 layer above it at around 400 kilometres. The combined F layer may then be centred around 250 to 300 kilometres. The altitude of the all the layers in the ionosphere layers varies considerably and the F layer varies the most. As a result the figures given should only be taken as a rough guide. Being the highest of the ionospheric regions it is greatly affected by the state of the Sun as well as other factors including the time of day, the year and so forth.

The F layer acts as a "reflector" of signals in the HF portion of the radio spectrum enabling world wide radio communications to be established. It is the main region associated with HF signal propagation.

Like the D and E layers the level of ionisation of the F region varies over the course of the day, falling at night as the radiation from the Sun disappears. However the level of ionisation remains much higher. The density of the gases is much lower and as a result the recombination of the ions and electrons takes place more slowly, at about a quarter of the rate that it occurs in the E region. As a result of this it still has an affect on radio signals at night being able to return many to Earth, although it has a reduced effect in some aspects.

The F region is at the highest region in the ionosphere and as such it experiences the most solar radiation. Much of the ionisation results from ultra-violet light in the middle of the spectrum as well as those portions of the spectrum with very short wavelengths. Typically the radiation that causes the ionisation is between the wavelengths of 100 and 1000 Angstroms, although extreme ultra-violet light is responsible for some ionisation in the lower areas of the F region.

**RAY PATH**



**Geometry of ray propagation within the Earth-ionosphere waveguide. The ground wave and two sky waves are displayed**

In the VLF(very low frequency) range, the transfer function is the sum of a ground wave which arrives directly at the receiver and multihop sky waves reflected at the ionospheric D-layer showed in the above figure.

For the real Earth's surface, the ground wave becomes dissipated and depends of the orography along the ray path. For VLF waves at shorter distances, this effect is, however, of minor importance, and the reflection factor of the Earth is $R_e = 1$, in a first approximation.

At shorter distances, only the first hop sky wave is of importance. The D-layer can be simulated by a magnetic wall ($R_i = -1$) with a fixed boundary at a virtual height h, which means a phase jump of 180° at the reflection point.In reality, the electron density of the D-layer increases with altitude, and the wave is bounded as shown in Figure 2.

The sum of ground wave and first hop wave displays an interference pattern with interference minima if the difference between the ray paths of ground and first sky wave is half a wavelength (or a phase difference of 180°). The last interference minimum on the ground ($z = 0$) between the ground wave and the first sky wave is at a horizontal distance of

$$\rho_1 \approx 2\ f\ h^2/c$$

with c the velocity of light.