CHAPTER - III

RADIO COMMUNICATION SYSTEM AND RADIO WAVES IN THE IONOSPHERE

3.1 Radio communication system.

The radiocommunication systems in use today may be divided into one of two classes : firstly, Broadcasting - both radio and television - and, secondly, radio links.

Table 3.1 gives the Classification of various Frequencies

Frequency	Classification	Abbereviation
10 - 30 KHz	very low frequencies	v.l.f.
30 - 300 KHz	low frequencies	l.f.
300 - 3000 KHz	medium frequencies	m.f.
3 - 30 MHz	high frequencies	h.f.
30 - 300 MHz	very high frequencies	v.h.f.
300 - 3000 KHz	ultra high frequencies	u.h.f.
3 to 30 GHz	super high frequencies	s.h.f.
30 to 300 GHz	extra high frequencies	e.h.f.

The B.B.C. broadcast Radio 1, 2, 3 and 4 programmes at a number of differet frequencies in the medium and v.h.f. bands and at one frequency in the low frequency band. In the medium - frequency band, amplitude - modulated d.s.b. transmissions are used with carrier frequencies in the range 647 - 1546 KHz and having a bandwidth of approximately 9000 Hz.

For reasonably high - quality reception of music, an audio bandwidth of at least 15 KHz is required. This occupies an r.f. bandwidth of 30 KHz which cannot be accommodated in the congested medium frequency band. High quality broadcast transmissions are therefore provided in the v.h.f. band using carrier frequencies in the range of 88.1 - 96.8 MHz. V.h.f. signals have a limited range of a hundred kilometers or so and so a number of stations are required within a fairly small area. The carrier frequencies allotted to the stations in a given area are about 200 KHz apart, to minimize inter-station interference. (In the B.B.C. v.h.f. broadcast system a r.f. badwidth of 180 KHz is necessary to provide the 15 KHz audio bandwidth).

IT MAY BE NOTED THAT FOR STANDARD TIME SIGNALS FREQUENCY STABILITY IS HIGHER.

There are some common basic terms and building blocks in communication system (Fig. 3.1)

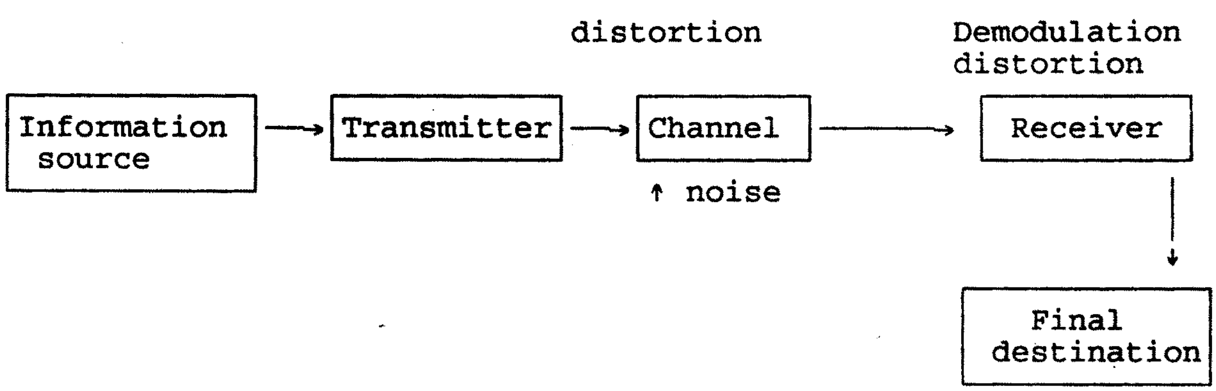


Fig. 3.1 Communication system Block diagram.

1. Information source :- The origin of message is from some information source. The amount of information conveyed in any message is measured in "bits".

2. Transmitter :- The message to be communicated has to be first converted to an electric signal by the help of a suitable transducer. The electric signal so obtained has to be suitably

processed and amplified before being sent on the channel by transmitter.

3. Channel :- This refers to the medium over which the transmitted signal travels. Some unwanted (random) energy termed as noise is added to the signal at the transmitter, receiver as well at the channel stage.

4. Receiver :- There are wide range of communications receivers availabe, each being for specific purpose governed by specific factors such as frequency of operation, bandwidth, noise figure, output level, type of modulation etc.

5. Destination :- The destination of the received signal also governs the type of receivers.

Modulation is necessary for transmission because an unmodulated carrier cannot be used to transmit intelligence as it has fixed amplitude, frequency and phase. For information transmission, any one of the three variables has to be deviated from its unmodulated value. Variation can occur in a more gentle fashion, as in the amplitude modulation used in ordinary broadcasting. It is also possible to vary the frequency (frequency modulation or f.m.) or the phase (phase modulation). Consider fucntion $f(t) = A \exp(jwt)$. To achieve amplitude modulation, the amplitude of carrier wave is made proportional to the instantaneous amplitude of the modulating signal. It is well known that a complicated amplitude function which is repeated as a function of time can be considered to be made up of a number of continuous waves at discrete frequency intervals added with different amplitudes and phases. A nonrepeated amplitude function can be composed of a similar sum of component waves, but in this case, all frequencies are, in general, represented. The details of such consideration will be found in Fourier Transforms. The theory of Fourier integral can be used to show that the spectrum ϕ (w) contains a single frequency only if A is independent of time. However, in any practical case, the amplitude of the signal must at some time have been zero. Therefore one can say that, in principle, there is no such thing as a pure sine or cosine wave, and we always have more than one frequency involved in any physical problem. However, most soluble problems are linear in frequency. In such a case it is sufficient to decompose the initial wave in to its component waves, solve the problem for the components, and then reassemble the resulting waves in to single solution. In contrast, the detection (or demodulation) of a signal in a radio receiver is necessarily a nolinear process for which the calculations cannot be made on the individual frequencies seperately.

The general expression for a sinusidal carrier wave is

$$v = V \sin (wt + \theta). \qquad 3.1$$

Where v is instantaneous voltage of the wave. V is peak value or amplitude of wave. w is angular velocity of the wave in radians per second; w is related to the frequency of the wave by expressin $w = 2\pi f$ where f is frequency in Hertz, θ is the phase of the wave at the instant when t = 0.

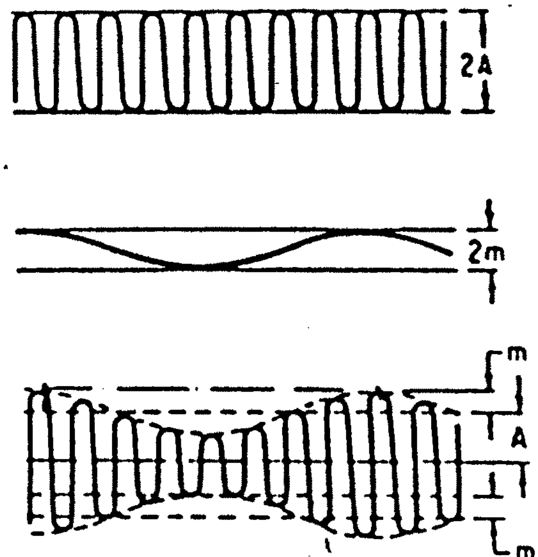


Fig.3.2
Amplitude modulated wave

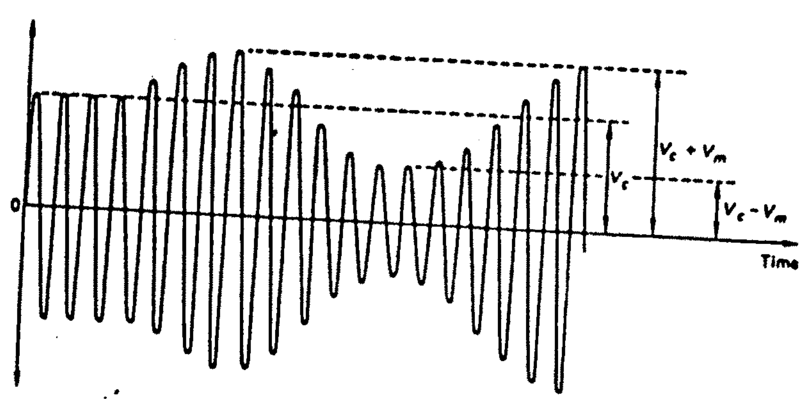


Fig.3.3

sinusolidally modulated wave

For amplitude modulation the amplitude V of the wave may be varied with frequency of wave remaining constant.

If a carrier wave is amplitude modulated, the amplitude of the carrier is caused to vary in accordance with instananeous value of the modulated signal. The difference in frequency between the carrier wave and modulating signal will normally be much higher, and may often be of the order of thousands of Hertz.

Amplitude of modulating wave varying with time is written as $A[1+ m \cos(w2t-k2z)]\cos(w1t-k1z)$. Where m is called modulation index. A be amplitude of carrier wave. Then resultant wave takes the form , Fig (3.2).

$$E = A \cos(w1t-k1z) + (Am/2)\cos[(w1+w2)t - (k1+k2)z] + (Am/2)\cos[(w1-w2)t - (k1-k2)z] \quad 3.2$$

Thus modulated wave can be expressed as the sum of three component waves; the first, $A \cos(w1t -k1z)$,is called the carrier wave, while two others , which are waves whose frequencies are the sum and difference of two given frequencies and whose propagation constants are given in like manner, are called the side-bands of the modulated wave.

Now suppose that the modulating wave is function of f(w2t), periodic over the interval $-\pi <= w2t <= \pi$ Then f(w2t) can be written as a Fourier series, with the addition of the phase term -k2z and the modulated wave E(t) can be written as ,

$$E(t) = Eo[1 + mAo + m \sum_{\nu=1}^{\alpha} A\nu\cos(\nu w2t-k2z) + m \sum_{\nu=1}^{\alpha} Bn\sin(\nu w2t - k2z)] \cos(w1t - k1z).$$

Exapanding,

$$
\begin{aligned}
E(t) = Eo[(1 + mAo)\cos(w1t - k1z) \\
&+ (m/2) \quad Ay \cos[(w1 + y\, w2)t - (k1+k2)z] \\
&+ (m/2) \quad Ay \cos[w1 - y\, w2)t - (k1-k2)z] \\
&+ (m/2) \quad By \sin[w1 + y\, w2)t - (k1+k2)z] \\
&+ (m/2) \quad By \sin[w1 - y\, w2)t - (k1-k2)z]] \qquad \text{eq } 3.3
\end{aligned}
$$

The modulated wave therefore consists of an infinite set of waves at the discrete frequencies w1, w1(+/-)w2, w1(+/-)2w2, ---- added together with proper amplitudes and phases.

To illustrate the demodulation process , let us consider the case of the carrier frequecy w1 modulated by a pure sine wave of frequency w2. The expanded form of the wave is given in eq 3.2, which shows that we have the carrier at frequency w1 plus two sidebands at (w1+w2) and (w1-w2). A common form of detector, is the square law detector , for which the output is proportional to square of input . Thus squaring eq 3.3 and reducing the squares and products of cosines, we get.

$$
\begin{aligned}
E^2 (t) = A^2 \{(½ + m^2/4) + m \cos (w2\, t - k2z) \\
&+ (½ + m^2/4) \cos (2w1\, t - 2k1\, z) + (m^2/4) \cos (2w2t + 2k2\, z) \\
&+ (m/2) \cos [(2w1 + w2)t - (2k1 + k2)\, z) \\
&+ (m/2) \cos [\, (2w1 - w2)t - (2k1 - k2)z] \\
&+ (m^2/8) \cos [2(w1 + w2)t - 2(k1 + k2)\, z] \\
&+ (m^2/8) \cos [2\, (w1 - w2)t - 2(k1 - k2)\, z]\} . \qquad \text{eq } 3.4
\end{aligned}
$$

The output spectrum contains the frequencies o, (d-c), w2, 2w1, 2w2, 2w1 + w2, 2w1 - w2, 2w1 + 2w2 and 2w1 - 2w2.

We note that these frequencies are those which would result from forming sums and the two sideband frequencies.

In the radio applications the carrier frequency is very much greater than any of sideband frequencies w2, w3, etc. Therefore it is possible to follow the detector by a filter whose upper cutoff is only slightly greater then wk, where wk is the highest frequency of modulating wave. If filter does not pass direct current, the output will consists of the modulating frequencies w2,...,wn, plus those sum and differences terms wn ± wm, where n, m = 2,3,..... k, which fall in frequency range $0 \leq w < wk$.

Now suppose the modulating wave has the form

$$1 + \sum_{\nu = 2}^{k} \epsilon_{\nu} \cos (W\nu\, t - K\nu\, z) \qquad 3.5$$

Then the amplitude of the output frequencies Wn will be proportional to ϵ_n, while the amplitudes of the sum and differnce terms will be proportional to $(\epsilon_n \epsilon_m) / 4$. If all ϵn are small, the amplitudes of the sum and difference terms will be negligible, so that the detected output is proportional to the original modulating wave with an appropriate phase shift.

3.2 (a) Modulation depth.

The envelope of an amplitude - modulated carrier wave varies in accordance with the waveform of the modulating signal and hence there must be a relationship between maximum and minimum values of the modulated wave and the amplitude of the modulating signal. This relationship is expressed in terms of the modulation factor of the modulated wave. The modulation factor M of an amplitude modulated wave is defined by the expression,

$$M = \frac{\text{maximum amplitude} - \text{minimum amplitude}}{\text{maximum amplitude} + \text{minimum amplitude}}$$

When expressed as a percentage M is known as the modulation depth, or the depth of modulation or the percentage mpdulation. Consider sinusoidally modulated wave. Fig 3.3

Since the envelope of the modulated carrier wave must vary in accordance with the modulating signal, its maximum amplitude must be equal to the amplitude of the carrier wave plus the amplitude of modulating signal, i.e. (Vc + Vm). Similarly minimum amplitude of the modulated wave must be equal to (Vc - Vm). Where Vc is the amplitude of the carrier wave and Vm is the amplitude of the modulating signal.

For sinusoidal modulation, therefore, modulation factor becomes

$$m = \frac{(Vc + Vm) - (Vc - Vm)}{(Vc + Vm) + (Vc - Vm)} = \frac{Vm}{Vc} \qquad 3.6$$

The modulation factor is equal to the ratio of the amplitude of the modulating signal to the amplitude of the carrier wave.

3.2 (b) Ionospheric modulated wave :-

We propose to investigate the modulation factor of an ionospheric modulated wave, by registering an encoded received time signals, with a PC-XT and PC - 20386 computer interface systems using a PC - 207 ADD on CARD supplied by Mircodynolog firm.

3.3 Radio waves in the ionosphere.

The ionosphere consists of a number of ionized regions above earth's surface, which play a most important part in the propagation of radio waves. Our knowledge of it is derived almost entirely from radio measurements, and it is therefore important to understand the process by which the waves are reflected. The ionosphere is beleived to influence radio waves mainly bacause of the pressure of free electorns. The early experiments showed that the electrons must be arranged aproximately in horizontal stratified layers, so that the number density is a function only of the height above the earth's surface. The terminology which is now generally accepted is based upon the vertical distribution of temperature and composition. The ionosphere must be almost electrically neutral, for if there were any appreciable space charge, it would give rise to large electric forces which would prevent formation of stable layers. There must therefore be at least as many positive ions as electrons, per unit volume. Besides negative electrons there may also be heavy ions formed by the attachment of electrons to air molecules. Heavy ions of both signs might play a part in the propagation of radio waves, and heavy ions, whether positive or negative, has a mass approximately 60,000 times that of an electron, and it is shown that, for all frequencies above a few hundred cycles, ions must be about 60,000 times more numerous than electrons if they are to have a dectectable effect. If this could happen at all it would only be in the very lowest region of the ionosphere, but there seems to be no evidence that heavy ions give any observable effect. It is therefore assumed, that only the free electrons can affect radio propagation.

The terminology to decribe the various regions of the upper atmosphere will be that based upon the temperature distribution of neutral atmosphere and the following terminology is used. The mesosphere, which lies in the height range of 50 to 85 Kilometers, is a region of decreasing temperature with height. The thermosphere, above 85 Kilometers, is a region in which the temperature increases with height. these "regions" are not well defined and the transision regions are called "pauses".

In addition to the terminology based on temperature, others have been deviced based on alternative physical quantities and processes. Two of these are illustrated in Figs. 3.4 (a) & 3.4 (b) For example, one terminology is based on the fact that turbulence predominates below about 100 Kilometers, where as diffusive seperation sets in above about 110 Kilometers. Above about 500 Kilometers is region called the exosphere. The term ionosphere was first applied by Sir. Robert Watson- Watt to that part of the atmosphere in which free ions exists in sufficient quantities to affect propagation of radio waves. The ionosphere therefore can be, considered as lying between about 40 to 50 Kilometers and several earth radii. This definition is essentially that adopted by the Institute of Radio Engineers.

It is convenient to define a region as a section of atmosphere within which there can exist ion distributions called layers. A division of the ionosphere into regions is given in Table 3.2 together with layers which may exist within these regions. The electron distribution within region may not be a peak of electron density.

Table 3.2

Height range (Km.)	Regions	Layers
50 - 90	D	D
90 - (120 - 140)	E	E1, E2, E3
Above (120 - 140)	F	F1, F1½, F2

3.4 Pressure and density variations

The relationship between pressure P and density ϱ at any height h is given by the "borometric" equation, which is derived as follows :

Consider an elementary cylinder of height dh and unit cross section. The net pressure difference dp between the top and bottom surface must equal the downward force due to the weight of the fluid in the cylinder, and is given by,

$$dp = - \varrho \, gdh \qquad 3.7$$

$$dp = - N \bar{m} \, gdh,$$

Where g is the acceleration due to gravity, N is the number density of molecules and $\bar{m}$ is the mean molecular mass. It should be remembered that N, $\bar{m}$ and g are all functions of height ; example, g varies inversely as the square of the radius vector.

The perfect gas law is

$$P = NKT \tag{3.8}$$

Where K is Boltzmann's constant = 1.372 x 10 $^{-16}$ erg / deg. and T is absolute temperature.

From 3.7 and 3.8 equations

$$\frac{dp}{P} = - \frac{\bar{m}g}{KT} dh = - \frac{dh}{H} \tag{3.9}$$

Here $H = \frac{KT}{mg}$ is known as the "scale height" of the atmosphere, which again is height dependent. At the earth's surface the value H is about 8 Km.

In isothermal atmosphere, integration of 3.9 equation gives

$$P = Po \exp \left(- \frac{h - ho}{H}\right) \tag{3.10 (a)}$$

$$\rho = \rho_o \exp \left(- \frac{h - ho}{H}\right) \tag{3.10 (b)}$$

Where P is the pressure at a height h corresponding to a pressure Po at height ho. Similary for the density ρ. The quantity (h - h0/ H), which is the height differnece in scale height units, is often denoted by Z. Below about 100 Kilometers .

it is possible to measure T (h) and $\bar{m}$(h) directly, by means of rockets and hence to determine H. Above about 100 Km, however, it is much easier to measure ϱ (h) and determine H from (3.9) equation expressed in the form.

$$1/H = -d/dh(\log p) = (\varrho g) / {}_h\int^{\alpha} \varrho g dh, \quad 3.11$$

The following formulae are useful in calculating the scale height at an any altitude.

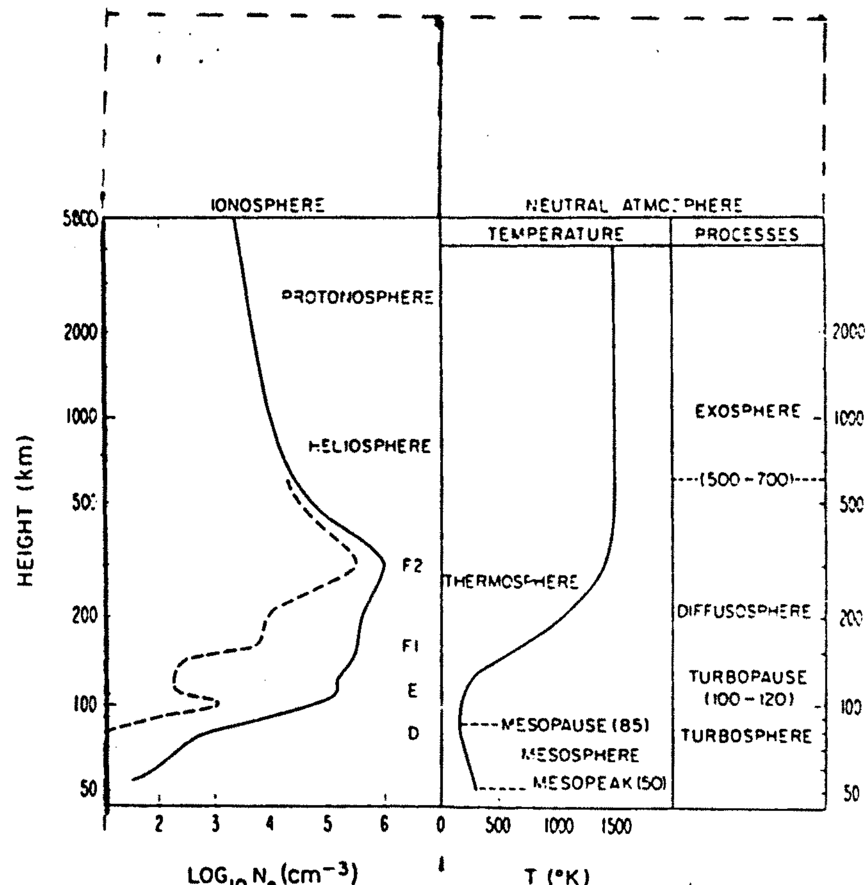
$$H(Km) = 0.848\ (1+ h/a)^2\ T^{o}K / m\ (gm/mole), \quad 3.12$$

Where a is the earth's radius (6370 km approximately)

Eq 3.12 can be approximated by

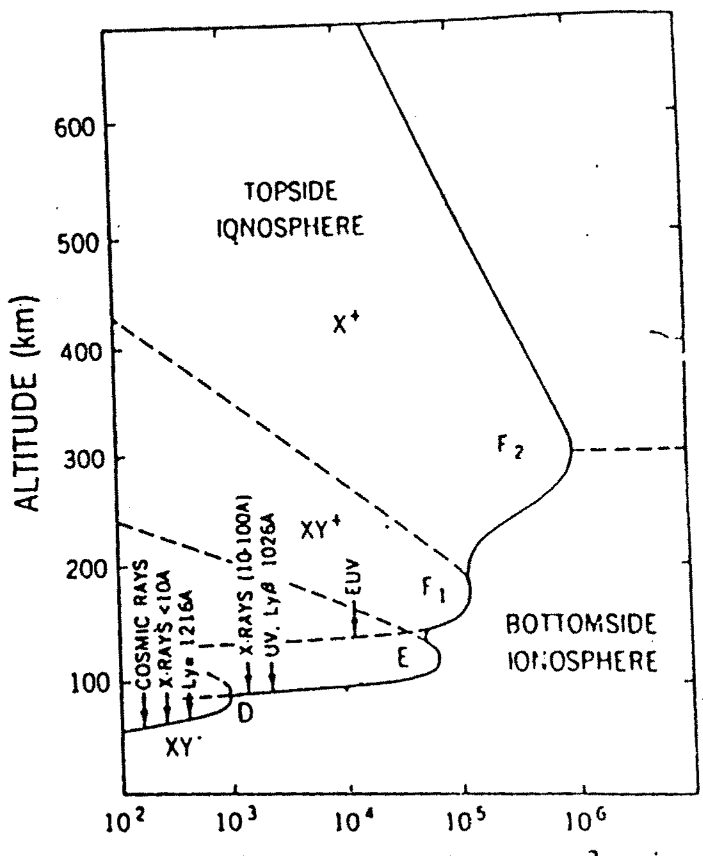
$$H = 0.93 \frac{T}{M}, \quad 3.13$$

Which is correct, to within 7.5 percent, between 50 km and 500 km. It is exact at 300 km. The temperature or scale height structure gives rise to the terminology indicated in the column labelled "temperature" in Fig. 3.4. Although the terminology is self explanatory, the causes of the observed temperature structure deserve comment. The mesosphere is heated by the absorption by Ozone of ultraviolet light, in the upper atmosphere, they tend to reunite with positive ions (recombination) and to attach themselves to neutral molecules to form negative ions such as $O2^-$ (attachment). The $O2^-$ ion is latter neutralized in further



Atmospheric nomenclature.

Fig.3.4(a)



TOPSIDE
IQNOSPHERE
X+
F2
XY+
F1
E
D
XY-
BOTTOMSIDE
IONOSPHERE
COSMIC RAYS
X-RAYS <10A
Ly α 1216A
X-RAYS (10-100A)
UV, Ly β 1026A
EUV
ALTITUDE (km)
600
500
400
300
200
100
10²
10³
10⁴
10⁵
10⁶
ELECTRON DENSITY (cm⁻³)

Fig.3.4(b)

reactions. Electrons can also leave a given volume by moving out of it (diffusion and / or drift). Nitrogen ions also play a role in breakdown and reformation of Ozone molecules.

For our present purposes we shall consider recombination and attachment only. Let the number densities of electrons, of positive ions, and of molecules to which attachment is possible be denoted by N(e), $N(A^{+})$ and N(A), respectively. The rate at which electorns are lost by recombining with positive ions is given by

$$dN(e) / dt = - \alpha N(e) \cdot N(A+) \qquad 3.14$$

Where α is recombination coefficient. If we assume that there are few negative ions compared with the electron concentration, we have $N(e) \approx N(A+)$ so that,

$$d(N(e)) / dt = - \alpha \{N(e)\}^2 \qquad 3.15$$

Consider the process of attachment : the rate of disappearance will be proportional to both electron concentration and neutral atom concentration, that is,

$$d N(e) / dt = - b N(e) N(A) \qquad 3.16$$

Where b is constant of proportionality. If we assume that the number of neutral molecules is enormously greater than number of electrons, so that it does not change appreciably when a few of the atoms are converted in to negative ions, we may write

$$d(N(e)) / dt = - \beta N(e) \qquad 3.17$$

Where β is known as the attachment coefficient. The coefficients α and β may vary with height because the reactions usually involves three bodies rather than two.

The rate of change of electron density N is, therefore, given by the following continuity equations for the cases of recombination and attachment, respectively

$dN / dt = q - \alpha N^2$ 3.18

$dN / dt = q - \beta N$ 3.19

Where q is the rate of electron production, i.e., number of electorns produced per second.

It should be noted here that if electrons are produced at a rate of q and are lost by attachment, respectively to neutral atoms to form negative ions, and if these ions are subsequently lost by ionic recombination, it can be shown that the continuity equation assumes the form

$dN / dt = q_{eff} - \alpha_{eff} N^2$ 3.20

Where q_{eff} and α_{eff} are effective rates of production and disappearence. Thus the net effect is to make the process appear as simple recombination. The effective recombination coefficient α_{eff} becomes relatively large in the lower (D and E) regions of the atmosphere.

Another set of reactions which is thought to be of importance in the upper atmosphere is electorn disappearance first by atom - ion exchange between a neutral molecule (XY) and the positive ion (A^+), followed by recombination between the electron and the XA^+ ion :

Production A + photon ⟶ $A^+ + e$, eq 3.21 (a)

Loss $A^+ + XY$ ⟶ $XA^+ + Y$ eq 3.21 (b)

$XY^+ + e$ ⟶ $X' + A'$ eq 3.21 (c)

The primes indicate that the atoms X and A may be left in excited states.

With these reactions it can be shown that, if the number density of XY is high in the F -region and low in E and D regions(as is thought to be the case if A represents atomic oxygen and XY molecular oxygen), the continuity equation takes the form

$$dN / dt = q - \beta_{eff} \cdot N, \qquad 3.22$$

Where β eff is an effective attachment coefficient. We now see that the forms of the continuity equation may not be those expected from simple recombination and attachment.

3.5 Formation of a Chapman Layer

The simplest type of ionized layer that can be deduced from theoretical consideration is known as a Chapman Layer. The derivation is based on the following assumptions :

1) An atmosphere with only one type of gas.

2) Plane stratification.

3) A parallel beam of monochromatic ionizing radiation from the sun.

4) An Isothermal atmosphere.

To start with, let us invoke assumptions 1, 2, 3 only. Let the

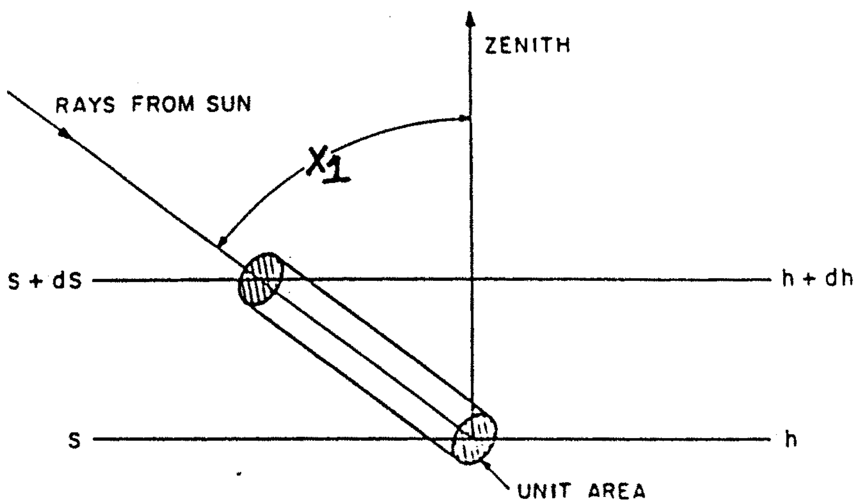


Fig.3.5 *Absorption of radiation in a slab of gas.*

inoizing radiation of intensity $S\alpha$ be incident at a Zenith angle X1.

Absorption of radiation in a slab of gas on the top of atmosphere. As the radiation is absorbed, as it penetrates in to the atmosphere, its intensity diminishes. Let S be the energy flux at a height h and S + ds the flux at a height h + dh as sketched in Fig. 3.5 Let a be absorption corss section of the atoms of the gas and N their number density. The energy absorbed ds in a cylinder of unit cross section and axis parallel to the direction of the incident beam is given by

$$ds = s\,\sigma N\,dh \sec X1 \qquad 3.23$$

Eq 3.23 merely states that the energy absorbed is proportional to the total corss section and to intensity of the incident radiation.

Upon integration we obtain :

$$\int ds / s = \sec X1 \int N\,\sigma\,dh = -\tau \sec X1 \qquad 3.24$$

Where, $\tau = -\int N\,\sigma\,dh$ 3.25

is optical depth of the atmosphere down to the height h. The minus sign arises because τ increases in the opposite direction to h.

At great heights $S \longrightarrow S\infty$ as $\tau \longrightarrow 0$, eq.3.24 gives

$$S = S\infty \exp(-\tau \sec X1) \qquad 3.26$$

The energy absorbed per unit volume is given by

$$ds / dh \sec X_1 = N\sigma S = N\sigma S\infty \exp(-\tau \sec X1) \qquad 3.27$$

Let η be the number of ion pairs produced per unit quantity of energy absorbed, i.e. the ionization efficiency. The number of ions pairs produced per unit volume per second is

$q\,(X1, h) = N \sigma S\infty\, n \exp(-\tau \sec X1)$ 3.28

Now τ is a function of X1 and h, using 3.7, 3.8 and 3.9 we obtain

$$\tau = \sigma \int N dh = (\sigma / mg) \int dp = \sigma N KT / mg = \sigma NH \qquad 3.29$$

Substitution of (3.29) into (3.28) yields

$q\,(X1, h) = (\tau S\infty / H \epsilon)\, \eta \exp(1 - \tau \sec X)$ 3.30

Where $\epsilon = 2.718$ .... Note that we have neglected the variation of `g' with height. Notice that 3.30 holds for any temperature distribution. To obtain Chapman's formula. We invoke assumption of, as isothermal atmosphere, where H is independent of height. We can define a quantity Z by the relation

$$Z = -\log \tau \qquad 3.31\ (a)$$

$$\tau = \exp(-z) \qquad 3.31\ (b)$$

Substitution of 3.31 (b) in the 3.30 gives

$q(X1, Z) = ((S\infty) n / \epsilon H) \exp\{1 - z - \sec X1 \exp(-z)\}$ 3.32(a)

$q(X1, Z) = qo \exp\{1 - z - \sec X1 \exp(-z)\}$ 3.32(b)

Where,

$qo = S\infty \cdot \eta / \epsilon H$ 3.33

is the rate of production of ion pairs at the level Z = 0 when the sun is overhead, i.e. when To = 1, and where To is the value of T_1 at the level Z = 0,

From 3.29, 3.31 (b) and 3.32 (a) we get,

$$\exp(-z) = T_1 / To = [P / Po] \exp(-(h - ho) / H)$$

i.e. $Z = (h - ho) / H$ 3.34

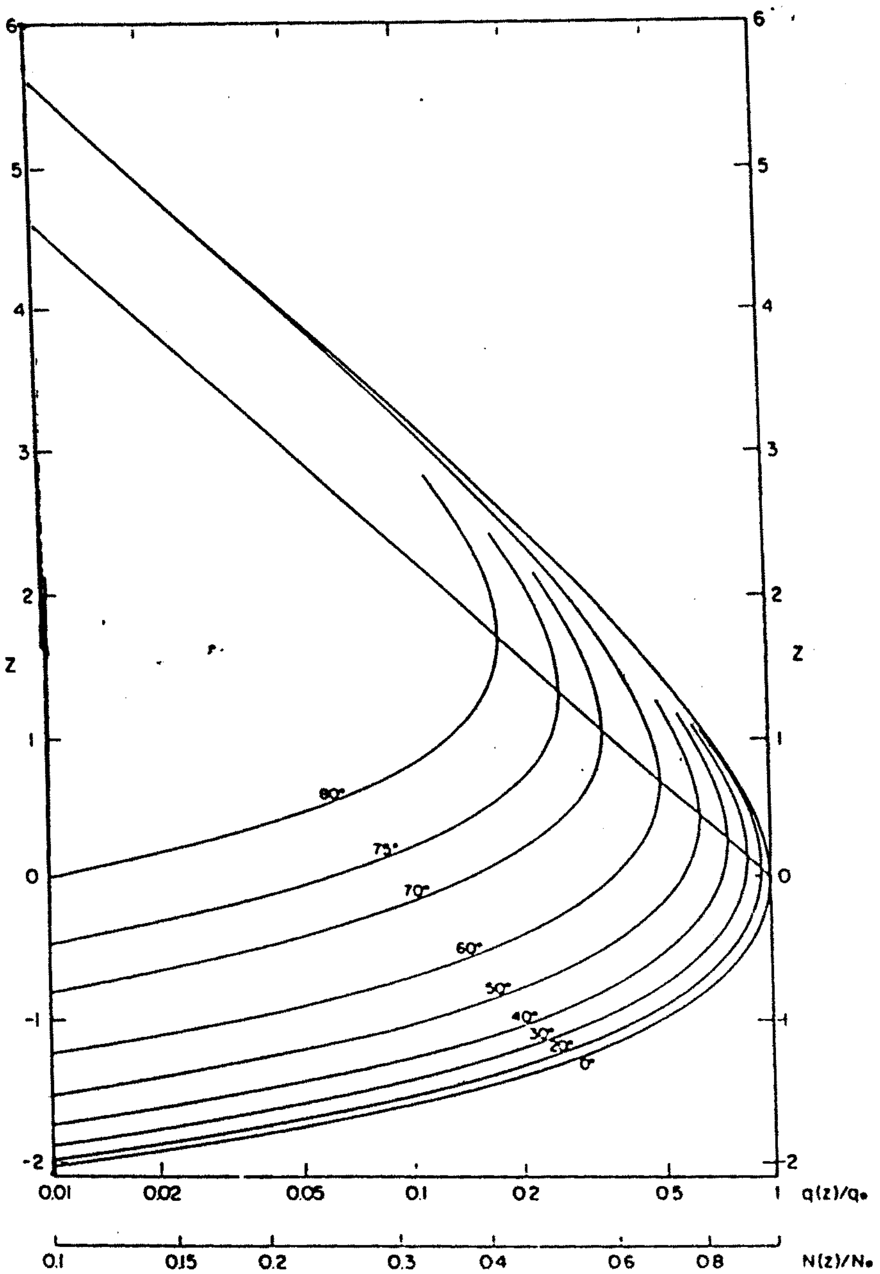


Fig.3.6 *Normalized rate of photoionization $\dot{q}(z)/q_0$ and electron density $N(z)/N_0$ according to Chapman theory.*

The reference height ho is therefore, the height of maximum ion production when the sun is overhead.

If in 3.32 (b), we replace Z by Z - log secX1 and put $X_1 = 0$, we obtain

$$q\ (Zo,\ Z - \log \sec X1) = \sec X1\ q(X1,\ Z) \qquad 3.35$$

Eq 3.35 gives us an important scaling rule. That is, the curve q(X1, Z) has the same shape as q(0, Z), but is moved upwards by log secX and is diminished by cos X_1. The height of peak of ion production Zm or hm is given by

$$Zm = \log \sec X1 \qquad 3.36\ (a)$$

$$hm = ho + H \log \sec X1 \qquad 3.36\ (b)$$

The variation of q(X1) for various values of Z is shown in Fig 3.6. It can be seen that the use of a logarithmic scale for q makes the shapes of the q curves identical. The peak rate production is given by

$$qm = qo \cos X1 \qquad 3.37$$

The intensity of the ion production is determined by the flux of ionizing radiation and the efficiency n changing the flux does not alter the height of maximum production. Thus if there are several wavelengths present in the incident radiation for which ionization coefficients are markedly different, several distinct layers will result. The most strongly absorbed radiation produces the uppermost layer and vice versa. Likewise, different ionized layers would result with different gases in the atmosphere.

3.6 Plane waves and Spherical waves :

The curvature of the earth

Radio waves travelling from a transmitter to a receiver near the earth's surface may take one of several paths. A wave may travel over the earth's surface, and it is then known as the ground wave. The earth is an imperfectly conducting curved surface, and the theory of the propagation of the ground wave involves many problems of the greatest mathematical interest. Another wave may travel upto the ionosphere, be reflected there, and return to the receiver. It is with a single reflection of this kind, the wave originates at a source of small dimensions so that wave-front is approximately spherical, but by the time it reaches the ionosphere the radius of curvature is so large that the wave can be treated as plane. This involves an approximation, that the error is negligible, except in certain special cases rarely met with in practice. Similarly, The ionospheric layers are curved because of the earth's curvature, but in most problems this curvature can be neglected.

3.7 Effect of collisions and of the earth's magnetic field.

The motion of an electron in the ionosphere is affected by the earth's magnetic field and by the collisions which the electron makes with other particles. It was shown by Lorentz that the collisions have the same effect as a retarding force proportional to velocity. Which arises at high frequencies and can be treated by `ray-theory' methods. The retarding force is most important at low frequencies and here the wavelength is so long that ray theory methods are inapplicable and a`full-wave' treatment must be used.

This effect of earth's magnetic field is to make the ionosphere a doubly refracting medium. This is a great complication, and often leads to a differential equation of the fourth order. It is convenient to neglect the earth's magnetic field for many purposes, so that the differential equation to be solved is only of the second order.

One reason for this is that only in this way, can the differential equation be reduced to a form whose solution can be studied by mathematics, but by neglecting the earth's magnetic field, principles can be established which can then be extended to more general cases, where magnetic field is included.

In seismolgy some study has been made of the propagation of elastic waves in media whose properties vary gradually from place to place. For example, some parts of the oceanbed are horizontally stratified and for sound waves, behave very like an inverted ionosphere.

3.8 Theory of wave propagation :

The purpose of this chapter is to present the basic ideas needed for an understanding of the propagation of radio waves in the ionosphere. In the presence of earth's magnetic field, the ionosphere behaves as a birefringent medium and so it is necessary to discuss the properties of such a medium.

These properties include the decomposition of a plane polarized wave into Ordinary and Extraordinary waves and the Deviation of the direction of energy flow away from the direction of Phase propagation.

We shall derive the Appleton formula in the Chapter IV, for the complex refrective index of a magneto - ionic medium and consider its properties.

The weakness of the Appleton theory will be considered there at and more generalized formulas will be discussed. Finally, we shall also discuss the propagation of wave packets in a magneto - ionic medium.

3.9 Electromagnetic waves

The principles underlying the propagation of radio waves are based on the relationship between electricity and magnetism which were discovered in the eighteenth and nineteenth centuries. It is now generally known that the attractive and repulsive properties of electrified bodies and of lodestone were first recorded by the ancient civilizations. Little further progress was made until the eighteenth century when attempts were made to establish laws of electric and magnetic forces similar to the gravitational law of Newton. This search bore fruit when, in 1785, the French scientist Coulomb, using a torsion balance, showed that the mechanical force between two small electrified spheres varies inversly as the square of the distance r between them. Later work showed that the force was proportional also to the amounts of charge q1, q2 on the two bodies and so we obtain the relationship.

$$F = (1 / 4\pi \epsilon_o) \; (q_1 q_2 / r^2), \qquad 3.38$$

Where ϵ_o is permittivity of free space (8.85×10^{-12} F/M).

It is useful to introduce the concept of electric field E at this point. Let us assume that the charge q2 produces an electric field E at the site of q1 and that this field exerts the force. F on q1 where

$$F = q1\ E \qquad 3.39$$

Thus the electric field E at a point may be defined as the force experienced by a small unit test charge. It is assumed that the testcahrge itself does not alter the field.

The field at a vector distance r from a point charge q is, therefore given by,

$$E = (1 / 4\pi \epsilon_o) \; (q / r^3) \; (\bar{r}) \qquad 3.40$$

so that the

vector E is in the direction of r. If the charge is distributed throughout a volume of space, we have,

$$E = (Ar / 4\pi \epsilon_o) \int dq / r^2 , \qquad 3.41$$

Where dq is an element of charge and Ar is a unit vector in the direction from dq. Many electrical and magnetic problems, practicularly static field problems are best tackled through the idea of potential difference between two points A and B can be defined as the work done by an external force on a unit test charge in moving it from A to B as shown in Fig. 3.7

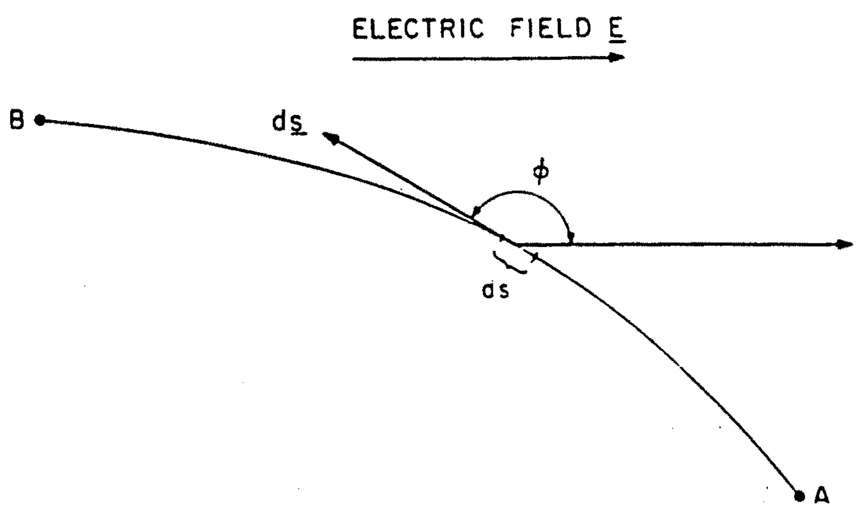


Fig.3.7 *Potential difference.*

The zero of potential therefore, arbitrary; and it is sometimes convenient to define the zero at infinity. The work done on the test charge in moving it from A to B is

$$V_B - V_A = - {}_A\int^B E \cos \phi \, ds = - {}_A\int^B E \cdot ds \qquad 3.42$$

and hence

$$E = - dv / ds = - \text{grad } V. \qquad 3.43$$

Gauss showed that the integral of the eletric field over a surface A enclosing a medium of permitivity ϵ is given by

$$\epsilon \int_A E \cdot d_A \int \rho \, dv = Q \qquad 3.44$$

Where ρ is charge density, V is the volume enclosed, and Q is the total charge enclosed.

Similarly, for a magnetic field of induction B Wb / m² (Webers per square meter),

$$\int_A B \cdot dA = 0 \qquad 3.45$$

This follows since there are no free magnetic poles. Before proceeding to a consideration of current electricity let us consider some elementary properties of dielectrics. Consider a parallel plate capacitor with a large area of cross section (so that edge effects are negligible), and distance d between plates, connected to a voltage source v, and let σ be the charge density on One of the plates, From 3.44 we can show that the field E between the plates is given by

$$\epsilon o \; E = \sigma \qquad 3.46$$

and by definition,

$$E = V / d \qquad 3.47$$

Now on the introduction of a slab of dielectric of thickness d between the plates, the positive nuclei in the atoms of the dielectric will tend to be displaced in the direction of smaller field while the negative charges (electrons) will tend to be displaced up the field, i.e. the medium becomes polarized. Thus a backelectromotive force is set up which will tend to decrease the field between the plates.

Now the total field E, as given by (eq 3.47) is independent of the dielectric. Hence, the charge density on the plates must increase by an amount σi, for instance. We can regrad the polarization of the dielectric as conferring on it an effective permitivity ϵ so that from (3.44) we obtain,

$$\epsilon E = \sigma + \sigma i \qquad 3.48$$

Imagine a cylinder of dielectric of unit cross section with its axis perpendicular to the plates. At the end surfaces will be charges + σi and - σi at a distance d apart. Hence we have the equivalent of a dipole of moment o-id. The electric moment per unit volume is, therefore, σi which is called the polarization P. The total quantity $\sigma + \sigma i$ is known as the displacement D. Hence we have the relationship.

$$D = \epsilon o E + P \qquad 3.49$$

Now in certain media (anisotropic) the application of an electric field in one direction gives rise to a polarization in some other direction. Under these circumstances (eq.3.49) has to be expressed as vector equation :

$$D = \epsilon_0 E + \bar{p}$$ 3.50

This relation is of great importance in ionospheric radio propogation.

Maxwell's maintained that as far as the electromagnetic field was concerned, the displacement current was just as important as a flow of electricity and that it must be added to the current expression. This enables us to write
Maxwell's equation in integral form as follows

$$\oint_A D \cdot dA = \int \rho \, dv,$$ 3.51 (a)

$$\oint_A B \cdot dA = 0 ,$$ 3.51 (b)

$$\oint H \cdot ds = \int J \cdot dA + d / dt \int D \cdot dA,$$ 3.52 (a)

$$\oint E \cdot ds = - d / dt \int B \cdot dA,$$ 3.52 (b)

3.10 A solution of Maxwell's equation

In order to solve the set of equations (eqs 3.51 and 3.52) it is necessary to express them in their differential forms. This can be achieved by means of the following two vector theorems :

Green's Theorem :

Let F be any vector function and Va volume bounded by the closed surface A; then,

$$\int F \cdot dA = \int_V \operatorname{div} F \, dv$$ 3.53

Stoke's Theorem :

Let f be a vector function and A a surface (not closed) bounded by a simple closed curve S; then,

$$\S_S F \cdot dS - \S_A \text{curl } F \cdot dA \qquad 3.54$$

Applying these transformations to 3.21 we obtain

$$\text{div } D = \rho \qquad 3.55 \text{ (a)}$$

$$\text{div } B = 0 \qquad 3.55 \text{ (b)}$$

$$\text{curl } H = J + \dot{D} \qquad 3.55 \text{ (c)}$$

$$\text{curl } E = - \dot{B} = \mu \dot{H} \qquad 3.55 \text{ (d)}$$

Where the dot signifies the time derivative.

CHAPTER - 3

REFERENCES

[1] D.C. Green (2^{nd} Edition)
(P.No. (30-31), (53-54))

[2] Umesh Sinha.
Principles of Communication Engineering.
Published by Smt.Sumitra Handa, 16/7698,
New Market, New Rohtak Road,
New Delhi (1^{st} Edition 1981)
P.No.

[3] John M.Kelso
Radio ray propagation.
(P.No. (29-33)).
Mc Graw - Hill, Inc. (1964)

[4] K.G.Budden
Radio waves in the ionosphere.
(P.No. (1-7), (150), (175-180))

[5] Kenneth Davies
Ionospheric radio propagation
(P.No. (2-4), (8-17), (47), (53)).
Mc Graw - Hill Inc. (1965).