2.1. Introduction

The Ionosphere occupies a key position in solar-terrestrial physics, because extending up to the magnetosphere and down to the domain of meteorology, it impinges on both (Rawer and Suchy, 1967). Ionospheric physics is a big research field in its own right and contributes to other fields because ionization acts as a very good tracer for neutral atmosphere, being chemically reactive, sensitive to physical conditions and relatively easy to observe by both ground-based and space techniques.

The ionosphere is defined as that part of the earth’s atmosphere where ionization exists in sufficient quantities to play a role in affecting various processes taking place in that region. These processes may be physical, chemical as well as dynamical.

Conventionally, the ionosphere is divided into three regions: D, E and F (Fig. 2.1). The F region is further divided into $F_1$ and $F_2$ during day time. $F_1$ layer is absent at night. The region above the $F_2$ layer is called topside ionosphere.

2.2 The ionizing radiations.

The ionosphere has its origin in the interaction of ionizing radiations with the neutral atmosphere. The ionizing radiations responsible for the formation of the ionosphere are solar extreme ultraviolet (EUV) and X-rays, and corpuscular radiation i.e. galactic cosmic rays and energetic particles having their origin within the sun earth system.
Fig. 2.1
2.2.1 Solar EUV radiation and X-rays

Rocket and satellite observations have led to a detailed identification of the solar emission line spectrum, while the intensities of these radiations are still plagued by some uncertainties (Hinteregger, 1965). Although radiation at wavelengths below 1000 Å are the predominant ionizing agents, wavelengths as high as 2400Å contribute to the ionization of trace constituents in the atmosphere. In addition to ionization, solar UV radiation in the Schumann-Runge continuum (1350-1750 Å) is responsible for the dissociation of molecular oxygen into atomic oxygen.

The rate of production of ion pairs (electrons and ions), q, as a function of altitude can be expressed by

\[ q(z) = \sigma I n(z)e^{-\tau} \]

where \( \sigma \) is the ionization cross section, \( n(z) \) is the number density of the ionizable atmospheric constituent, \( I \) is the photon-flux incident on the atmosphere and \( \tau \) is the optical depth which is given by

\[ \tau(z_0) = \int_{z_0}^{\infty} \sigma_a n(z)dz = \sigma_a n(z_0)H \sec \gamma \]

where \( \sigma_a \) is the absorption cross section, \( n(z) \) is the number density of the absorbing constituent and \( H \) is its scale height, and \( \gamma \) is the solar zenith angle.

The concept of unit optical depth is generally used to describe the penetration depth of radiation into the atmosphere. At unit optical depth (\( \tau = 1 \)), there occurs the maximum of ion production for radiation incident on an isothermal atmosphere. Thus, for vertical incidence \( \gamma = 1 \) occurs at a level \( z_0 \) above which the content of absorbing constituent \( n(z_0)H \) equals the reciprocal of the absorption cross section \( (\sigma_a) \).

D-region

Between 70 km and 85 km, Lyman \( \alpha \) at 1216 Å acting on the trace
constituent nitric oxide (NO) is the most important ionizing agent for the quiet D-region, whereas galactic cosmic rays are responsible for the formation of the lower most part of the D-region, below 70 km. The flux of solar Lyman α (1216Å) has been found to lie between 3 and 6 ergs cm⁻² sec⁻¹ (1.8×10¹¹ and 3.6×10¹¹ photons cm⁻² sec⁻¹), the uncertainty possibly being due to instrument calibration. With increasing solar activity X-ray emission in the wavelength range from 2-8 Å becomes increasingly important as ionizing source of the D-region. The X-ray intensities in this wavelength region increase by 5 orders of magnitude (10⁵) from completely quiet sun to a class 3 flare; in the absence of flares the solar cycle variation is a factor of 300 at 8Å. The relationship between SID's and X-ray flare emission has now also been confirmed by satellite observations (Chubb et al., 1964).

E-region

The important solar radiations for the production of ionization in the region from 85-140 km are X-rays (10-100Å), Lyman β (1026 Å), the C III (977 Å) line and the Lyman continuum (800-910 Å).

The X-ray range can be divided into two parts having comparable absorption cross sections, 10-31 Å and 31-100 Å, the dividing point being the K absorption edge of molecular nitrogen. The range from 10-31 Å, the intensity of which is highly variable, contains much less energy than the more stable range from 31-100 Å, so that the latter contributes most of the E-region ionization. The flux at sunspot maximum for the range from 10 to 100 Å has been found to be ~0.9 erg cm⁻² sec⁻¹ with a variation of a factor of 7 from solar maximum to minimum (Friedman, 1962). From considerations of the occurrence of unit optical depth, Lyman β and C III, which ionize O₂ exclusively, and the Lyman continuum are the ultraviolet ionizing radiations of potential importance to the formation of the E-region. The total flux of
these UV radiations is about 0.5 erg cm\(^{-2}\) sec\(^{-1}\) at solar maximum. The question of the X-ray or UV origin of the E-region is still not completely resolved. Watanabe and Hinteregger (1962) have put forward some argument in favour of the ultraviolet, while observations of solar X-ray and UV emission during the IGY eclipse expedition (Chubb et al., 1961) tend to support the theory of X-ray control of the E-region. It was found that at totality 10 to 13% of the x-ray flux still remained, whereas the Lyman \(\alpha\) intensity dropped by a factor of 2000; Lyman \(\beta\) and the continuum must have decreased correspondingly since they originate even lower in the solar atmosphere. This is also in agreement with the requirement of a residual ionizing flux of 10 to 15% at totality inferred from a number of ionospheric eclipse observations (Ratcliffe, 1956).

A strong correlation between E-layer critical frequency and solar decimeter wave emission, especially in the range from 3-30 cm has been established (Kundu, 1960) which indicates that the X-rays responsible for E-region ionization originate primarily in the 10 000 – 20 000 km height range of the corona, which is also the source of the radio emissions.

F-region

The main source of ionization in the lower F-region (F\(_1\)) is the Lyman continuum (910-800 Å) and the region 350-200 Å containing about 1 erg cm\(^{-2}\) sec\(^{-1}\) including the intense He II (304Å) resonance line. At higher altitudes there are additional contributions from wavelengths in the range from 700 to 500 Å which includes the emission lines of He I (584 Å), Mg X (625 Å) and Si XII (520, 500 Å) with a total flux of about 0.4 erg cm\(^{-2}\) sec\(^{-1}\). The maximum rate of absorption takes place between 160 and 180 km and less than 10% penetrates below 145 km.
2.3 General Behaviour

An insight into the behaviour of the ionosphere can be had from the response of the electron concentration $N$ to various processes occurring in the ionosphere (Yeh and Liu, 1976). The basic principle that governs it is the conservation of charge which mathematically is written as:

$$\frac{dN}{dt} = q - \frac{1}{V} \nabla \cdot (N \mathbf{V})$$  \hspace{1cm} (1)

Where the left hand side is the rate of change of electron concentration. The terms on the right hand side are the rate of production per unit volume denoted by $q$, the rate of loss per unit volume through chemical processes denoted by $1$, and the divergence of ionization away from the point of interest through motion with a velocity $V$, denoted by $\nabla \cdot (N \mathbf{V})$. Even though the velocity appears explicitly only in the last term, the effect of motion is also felt implicitly in the other two terms. As a matter of fact, some of the major anomalous ionospheric phenomena are caused by changes in the first two terms in response to motions.

The major contribution to ionization production is solar radiation. Sometimes cosmic rays may ionize the bottom of D region to form temporarily a C region. In high latitudes, especially auroral zones, energetic electrons and protons may enter to ionize through collisions. But at middle and low latitudes almost all stable ionizations in the ionosphere are created by the photo-ionization processes. This has been discussed in section 2.2.

The loss of ions in the ionosphere depends on their chemical reactions with other constituents. The ion chemistry of the D region is fairly complex, involving both binary and tertiary reactions as well as both positive and negative ion reactions. Some of these reactions are controlled by minor constituents whose concentrations in the D region are poorly known. There is thus a great deal of uncertainty about these reaction rates in the D-region. The ion chemistry of the E and F regions involves binary positive
ion reactions and their reaction rates are known fairly accurately (Ferguson, 1974). Because of these chemical reactions, the electron concentration is in general dependent on the concentration of ions and neutrals. This dependence implies that the equations are now coupled i.e. one must simultaneously consider continuity equations for ions and possibly neutrals in addition to the one for electrons given by equation (1). Fortunately, in the region where the ion chemistry is complex the reaction rates are reasonably rapid so that only production and loss terms need be considered. When this happens the ionosphere is said to be under chemical equilibrium except during rapid changes such as solar flares and solar eclipses. The presence of chemical reactions also implies the dependence of electron concentration on the composition and neutral concentrations. If, for example, the ionosphere is lifted up to a higher altitude where chemical losses are taking place at a slower rate, the electron concentration will be larger than the case in the absence of this uplift. Conversely when there is downward motion the electron concentration is expected to decrease.

The third term on the right of equation (1) is the motion term. This term is unimportant in D and E regions, but very important in the F-region. In general, three processes need be considered in this motion term. They are (i) diffusion, (ii) electromagnetic drift, and (iii) ion drag. Through these three processes, movements may be induced in the neutral atmosphere. Any force induced by these processes in the vertical direction is small as compared with gravity. Therefore the neutral atmosphere is regarded as unmovable in the vertical direction by these processes.

In the F-region the gyro-radius is small when compared with the mean free path. Consequently the charged particles can move freely along the magnetic field lines but not across them. This implies that the diffusion is primarily along the magnetic field lines. Also because of the strong coulomb
force between electrons and ions all the charged particles must diffuse with the same speed to maintain charge neutrality. This kind of diffusion is called ambipolar diffusion. This ambipolar diffusion takes place in the ionosphere because of concentration gradient and temperature gradient.

The ionospheric electric field responsible for producing electromagnetic drift may come about either because of currents in the dynamo region (altitude 120 km) or because of motions in the magnetosphere. This electric field induces an ionospheric motion which in the F-region, is perpendicular to both electric field and the earth’s magnetic field and hence is popularly referred to as the EXB drift. The process of ion drag is responsible for coupling ionospheric motions to the neutral atmospheric motions and vice-versa. At any time when there is a relative motion between the charged particles and neutrals there is also a drag force. The horizontal drag force on the neutrals has to work against the viscous force. In the height range of 200-300 km the drag force is strong in daytime so that the neutral air is more fully accelerated; on the otherhand, the viscosity is strong at night so that the acceleration on the neutral air by ion drag is slow and less effective. The reason for this difference is the large day to night variation in the ratio $N_e/N_n$ at heights 200-300 km, (where $N_e$, $N_n$ are electron and neutral particle densities respectively).

Other factors which also influence the electron concentration in the ionosphere are thermospheric winds, acoustic gravity waves, atmospheric tides and planetary waves etc.

Atmospheric gravity waves play an important role in transferring energy from the troposphere and mesosphere to ionospheric heights. While propagating upwards they interact with the ionosphere to produce Travelling Ionospheric Disturbances (TIDs). A knowledge about the various characteristics of TIDs can throw light on the generation mechanism,
propagation characteristics and the nature of sources of the gravity waves.

2.4 Response of the Ionosphere to the Propagation of Electromagnetic Waves

When an electromagnetic wave passes through the ionosphere the charged particles in a macroscopic volume acquire a drift motion in addition to the random thermal motion. The equation of drift motion for charged particles is (Booker, 1984):

\[ \text{Mass} \times \text{acceleration} = \text{Electric forces} + \text{Magnetic forces} + \text{gravitational forces} + \text{pressure gradient forces} + \text{collisional forces} \]

or\[ Nm \frac{Dv}{Dt} = NqE + NqvxB + Nmg - \nabla p \]

where \( N \) is the number density of particles of mass \( m \). Collisions have been neglected for the time being. Assuming that there is a steady drift velocity \( v_0 \) and a steady magnetic field \( B_0 \), we can write

Total velocity = \( v_0 + v \); Total field = \( B_0 + B \)

Substituting these in (3) and neglecting the higher order terms,

\[ Nm(\frac{Dv}{Dt} + v_0v) = NqE + Nq (v_0xB_0 + vxB_0) + Nmg - \nabla p. \]

The steady background velocity can be reduced to zero by choosing the appropriate frame of reference, so that

\[ Nm \frac{Dv}{Dt} = NqE + NqvxB_0 + Nmg - \nabla p. \]

Hence the equations for the drift motion for electrons, ions and neutral particles, become

\[ N_{e}\frac{\partial v_e}{\partial t} = N_{e}q_{e}E + N_{e}\frac{v_e}{x}B_0 + N_{e}m_{e}g - \nabla P_e \]

\[ N_{i}\frac{\partial v_i}{\partial t} = N_{i}q_{i}E + N_{i}q_{i}\frac{v_i}{x}B_0 + N_{i}m_{i}g - \nabla P_i \]

\[ Nm\frac{\partial v}{\partial t} = Nmg - \nabla p \]

Where \( e \) and \( i \) stand for electron & ion respectively. These equations
imply acoustic waves, whose velocities are of the order of \((kT/m)^{1/2}\) (\(k\) boltzmann constant, \(T\) temperature). These velocities are much smaller than the plasma wave velocities which are of electromagnetic nature and therefore have a velocity of propagation of the order of \(C\), the velocity of light. By comparison the velocity of acoustic waves can be neglected. This corresponds to taking the temperature as negligible and ionosphere can thus be taken as a cold plasma as far as electromagnetic waves are concerned.

Now as \(p = NkT\), the concept of cold plasma implies that gas pressure for electrons, ions and neutral particles is unimportant for the propagation of electromagnetic waves through the plasma. If we also neglect the effect of gravity, the equations of drift motion become

\[
\begin{align*}
N_e m_e \frac{\partial v_e}{\partial t} &= N_e q_e (E + v_e \times B_0) \\
N_i m_i \frac{\partial v_i}{\partial t} &= N_i q_i (E + v_i \times B_0) \\
N m \frac{\partial v}{\partial t} &= 0
\end{align*}
\]

These can also be written as (for electrons and ions):

\[
\begin{align*}
m_e \frac{\partial^2 r_e}{\partial t^2} &= q_e (E + \frac{\partial v_e}{\partial t} \times B_0) \\
m_i \frac{\partial^2 r_i}{\partial t^2} &= q_i (E + \frac{\partial v_i}{\partial t} \times B_0)
\end{align*}
\]

Where \(r_e, r_i\) are drift displacements for electrons and ions respectively.

In a homogeneous medium, a wave is represented in cartesian coordinates \((x, y, z)\) by

\[
\exp[j(wt - k_x x - k_y y - k_z z)]
\]

where \(k = (k_x, k_y, k_z)\) is called the propagation vector and \(w\) the angular frequency. If the position vector \((x, y, z)\) is denoted by \(r\), then it becomes

\[
\exp[j(wt - k \cdot r)]
\]

Further if \(n = kC/w\) is a refractive index vector, then it is

\[
\exp[jw(t - c^{-1}n \cdot r)]
\]
So eqs. 10, 11 become

\[
(jw)^2 \text{m}_e \text{e} = q_e (E + jw r x B_0) \\
(jw)^2 \text{m}_i \text{i} = q_i (E + jw r x B_0)
\]

The complex free electric moment per unit volume \((P_f)\) is

\[P_f = N_e q_e \text{e} + N_i q_i \text{i}\]

Maxwell’s equations in a plasma are

\[
B = \mu_0 \text{E} \\
D = \varepsilon_0 \text{E} + P_f
\]

where \(D = \varepsilon_0 \text{E} + P_f\)

is the complex electric flux density vector, and

\[B = \mu_0 \text{E}\]

Substituting the value of \(H\) from (18) after using (21) and that of \(D\) from (20) in (19) we get

\[P_f = -\varepsilon_0 [E + n \times (n \times E)]\]

Concept of Dispersion relation:- In a homogeneous medium, waves can exist which can be represented by the exponential wave function \(\exp[j(w t - k_x x - k_y y - k_z z)]\), but arbitrary values cannot be assigned to \(w, k_x, k_y\) and \(k_z\). The relation which connects these four quantities is called the dispersion relation. Arbitrary values can be assigned to any three of them. The dispersion relation then determines the fourth.

The dispersion relation for the magnetoplasma is obtained by equating the above expression for \(P_f\) to that derived in equation (17) from the equation of drift motion of the particles.

Let us choose the axes such that the imposed magnetic field is in the z-direction. Further let the direction of propagaton be in the y-z plane, making an angle \(\theta_p\) with the direction of the imposed magnetic field. Then refractive index vector is

15
\[ n = n(0, \sin \theta_p, \cos \theta_p) \]

where \( n \) is the refractive index for a wave of angular frequency \( \omega \) for which the propagation vector is in the direction \((0, \sin \theta_p, \cos \theta_p)\). We get

\[ P_{fx} = \epsilon_0 (n^2-1) E_x \]  
\[ P_{fy} = \epsilon_0 \left\{ (n^2 \cos^2 \theta_p - 1) E_y - n^2 \sin^2 \theta_p \cos \theta_p E_z \right\} \]  
\[ P_{fz} = \epsilon_0 \left\{ -n^2 \sin \theta_p \cos \theta_p E_y + (n^2 \sin^2 \theta_p - 1) E_z \right\} \]

Writing eq. (17) in terms of susceptibilities (Booker, 1984)

\[ P_{fx} = \epsilon_0 (k_T E_x + k_H E_y) \]  
\[ P_{fy} = \epsilon_0 (-k_H E_x + k_T E_y) \]  
\[ P_{fz} = \epsilon_0 k_L E_z \]

Where \( k_L, k_T \) are susceptibilities along and transverse to the direction of magnetic field respectively when electric field is applied in the direction of magnetic field. \( k_H \) is susceptibility in the Hall direction when electric field is applied transverse to the magnetic field. The susceptibilities are given by (Booker, 1984)

\[ k_L = \frac{\omega^2 N_e}{w^2 + \omega^2 N_i} \]  
\[ k_T = \frac{\omega^2 N_e/(w^2 M_e - w^2)}{w^2 N_i/(w^2 M_i - w^2)} \]  
\[ k_H = -\frac{1}{\omega^2} \left( \frac{\omega^2 N_e}{w^2 N_e - w^2} + \frac{\omega^2 N_i}{w^2 M_i - w^2} \right) \]

Eliminating \( P_{fx}, P_{fy}, P_{fz} \) between these two sets of equations

\[ (n^2-1-k_T)E_x - k_H E_y = 0 \]  
\[ k_H E_x + (n^2 \cos^2 \theta_p - 1-k_T) E_y - n^2 \sin \theta_p \cos \theta_p E_z = 0 \]  
\[ -n^2 \sin \theta_p \cos \theta_p E_y + (n^2 \sin^2 \theta_p - 1-k_L) E_z = 0 \]

Eliminating \( E_x, E_y, E_z \) we get the dispersion relation of the form

\[ A n^4 + B n^2 + C = 0 \]

where

\[ A = (1+k_L) - (k_L-k_T) \sin^2 \theta_p \]  
\[ B = -2(1+k_L)(1+k_T)+\{(1+k_T)(k_L-k_T) - k^2_H \} \sin^2 \theta_p \]  
\[ C = (1+k_L)(1+k_T)^2 + k^2_H \]
Solving (35) we get
\[ n^2 = (1+k_L)[(1+k_T) - k_H \left\{ \frac{1}{2} \sin^2 \phi_p + \frac{1}{4} \sin^4 \phi_p - \frac{1}{2} \cos^2 \phi_p \right\}^{1/2}] / [(1+k_L) - (k_L-k_T)\sin^2 \phi_p] \] (39)

where \[ S = [(1+k_T)(k_L-k_T) - k^2 H] / [k_H (1+k_L)] \] (40)

Because \( k_L, k_T \) and \( k_H \) depend on the angular frequency \( \omega \), we can know from the above equation how the refractive index \( n \) depends upon frequency. We can also know from the above equation how \( n \) depends upon the angle \( \phi_p \) between the direction of propagation of phase and the direction of the imposed magnetic field.

The alternative signs in the dispersion relation mean that, for a given frequency \( \omega \) and a given direction \( \phi_p \) of phase propagation, two waves can propagate in a magnetoplasma. They have different refractive indices. When \( n \) is real, they have different phase velocities.

\[ V = C/n \]

They are known as characteristic waves.

Behaviour of wave polarization:- In order to understand the behaviour of the polarization of the incoming wave the two values of \( n \) are substituted into (32) or (33) and (34) to get

\[ \frac{E_y}{E_x} = \frac{n^2 - 1-k_T}{k_H} \] (41)

\[ \frac{E_z}{E_y} = \frac{n^2 \sin \phi_p \cos \phi_p}{(n^2 \sin^2 \phi_p - 1-k_L)} \] (42)

These complex ratios give the relative amplitudes and phases of the rotating components of the electric vector along \( x, y \) and \( z \) axis. The tip of the resultant electric vector executes an ellipse whose plane does not in general coincide with the coordinate plane. Neither is the plane of the electric ellipse, in general perpendicular to the direction of phase propagation. Because the ratios of the cartesian components of complex electromagnetic vectors describe the state of the electric polarization of
the wave, they are known as polarization ratios.

Radio approximation of dispersion relation: As the frequencies used for studying the ionosphere lie in the VHF band, the dispersion relation can be simplified. The radio approximation corresponds to dropping all ionic terms and is applicable when \( w >> (w_{me}/w_{ni})^{1/2} \). Then

\[
n^2 = (1 - w^2 N/w^2) \left\{ \frac{(1 - w^2 N/w^2) - 1/2 \sin^2 \theta_p w^2 Me/w^2}{\left[ (1 - w^2 Me/w^2)^{-1/2} \sin 2\theta \frac{w^2 Me}{w^2} \right] + \left\{ \frac{1/4 \sin^4 \theta_p w^4 Me/w^4 + \cos^2 \theta_p w^2 Me/w^2 (1 - w^2 N/w^2)^2} \right\}^{1/2}} \right\} (43)
\]

By moving the radical in eq. (43) to the denominator

\[
n^2 = 1 - \frac{(w^2 N/w^2)(1 - w^2 N/w^2) - 1/2 \sin^2 \theta_p w^2 Me/w^2}{\left[ (1 - w^2 Me/w^2)^{-1/2} \sin 2\theta \frac{w^2 Me}{w^2} \right] + \left\{ \frac{1/4 \sin^4 \theta_p w^4 Me/w^4 + \cos^2 \theta_p w^2 Me/w^2 (1 - w^2 N/w^2)^2} \right\}^{1/2}} (44)
\]

The upper sign refers to the O-wave and the lower sign to the X-wave.

Effect of collisions on dispersion relation: Now collisions can be introduced by replacing \( w_N^2 \) by \( w_N^2 /(1 - j \nu_e/w) \) and \( w_{Me} \) by \( w_{Me}/(1 - j \nu_e/w) \) (Booker, 1984), where \( \nu \) represents the collision frequency. We get

\[
n^2 = 1 - w^2 N/w^2 \left\{ w^2 N/w^2 - j \nu_e/w \right\}/\Delta \}
\]

where

\[
\Delta = (1 - j \nu_e/w)(1 - w^2 N/w^2 - j \nu_e/w) - 1/2 \sin^2 \theta_p w^2 Me/w^2
\]

\[
+ \left\{ 1/4 \sin^4 \theta_p w^4 Me/w^4 + \cos^2 \theta_p w^2 Me/w^2 (1 - w^2 N/w^2)^2 \right\}^{1/2}
\]

Now making the substitutions

\[
X = w^2 N/w^2 \quad \quad Y = w_{Me}/w \quad \quad Z = \nu_e/w
\]

\[
Y = Y \cos \theta_p \quad Y_T = Y \sin \theta_p
\]

\[
n^2 = 1 - X(1 - jZ)/(1 - jZ)(1 - X - jZ) - 1/2 Y^2 T
\]

\[
+ \left\{ 1/4 Y_T^4 Y^2 (1 - X - jZ)^2 \right\}^{1/2}
\]

As is generally the case in satellite radio transmissions, \( w >> \nu \), then
\[ n^2 = 1 - 2x(1-x) / [2(1-x) - y^2 T_1 (y^4 T + 4(1-x)^2 y^2 L)^{1/2}] \]  

(47)

Depending upon whether the direction of propagation is parallel to the earth's magnetic field (\( \theta_p = 0^\circ \)) or, perpendicular to it (\( \theta_p = 90^\circ \)), the propagation is said to be longitudinal or transverse respectively. In some cases, even if the direction of propagation is not exactly parallel or transverse to the magnetic field, the propagation characteristics can be approximated to that of longitudinal or transverse propagation. Propagation is quasi-longitudinal if, under the radical in eq. (47) \( Y_L^2 \) term dominates i.e.

\[ Y_L^2 \gg \frac{Y_T^4}{4(1-x)^2} \]

It is quasi-transverse if \( Y_T^4 \) term dominates i.e.

\[ Y_T^4 \gg 4(1-x)^2 Y_L^2 \]

2.5 Exploration of the Ionosphere

Experimental methods of two fundamentally different kinds have led to the knowledge of the ionosphere (Ratcliffe, 1972). One, called in situ measurements, makes use of equipment carried aloft by rockets or artificial satellites to study the ionosphere in their neighbourhood. Experimental methods of the other type make use of radio waves produced either by natural processes or by man made transmitters.

The most important ways in which man-made radio waves have been used to study the ionosphere can be classified as follows:

(i) Waves are reflected from the ionosphere by a process similar to the total reflection of light waves, and the time interval between the emission of a wave and its return to the sending point is measured. The radio frequency of the emitted pulse is altered smoothly and the echo time is recorded as a function of frequency; the record is called ionogram. The
apparatus is called ionosonde, when it is on the ground, and is used to explore the lower part of the ionosphere; when it is in a space vehicle it is used to explore the upper part or the topside, it is then called a topside sounder. The emissions from a topside sounder sometimes produce disturbances in the local ionosphere that last much longer than the emitted pulse, they have been called resonances and have proved useful in ionospheric investigations.

These measurements are used to deduce electron density as a function of height. However it is difficult to make observations of ionization in D layer heights in same detail as for greater heights. Further it is difficult to derive information about the valley region (120-140km) between E and F layer owing to the fact that this region is shielded by the underlying ionization.

(ii) Partial Reflection: The totally reflected waves recorded by ionosondes are usually so strong that transmitters of comparatively small power can be used at places where electrical interference is not particularly large. If a high power sounder is used in a place where there is little interference it is possible to investigate much weaker echoes returned after partial reflection from the ionosphere. In one method, the partial reflection occurs because the electrons have a distribution that is irregular on a scale much greater than the distance between them and much less than a radio wavelength; it is called the method of partial reflection. In another method the partial reflection represents the wave energy returned from individual electrons, each scattering independently, in the way first described by J.J. Thomson. It is called the method of incoherent scatter or Thomson scatter.

(iii) Wave interaction. The temperature of the ionospheric electrons in a small defined region of the atmosphere is increased by the absorption of
energy from a short pulse of radio wave: the effect of this heated region on another wave traversing it is then observed.

(iv) Doppler Shifts: A wave emitted from a space vehicle is observed at the ground and the Doppler change of frequency resulting from the movement of the vehicle in the line of sight, or from changes in the intervening ionosphere, is measured.

(v) Faraday Rotation: A linearly polarized wave emitted from a satellite has its plane of polarization rotated as it traverses the anisotropic ionosphere; the changes in the polarization of the wave received at the ground are observed as the satellite moves or as the ionosphere changes. A detailed theory for determining the ionospheric electron content using the Faraday rotation technique is given in section 2.6.

The most important naturally occurring waves that have been used in ionospheric research can be grouped as follows:

(vi) Electron whistlers are audio frequency electromagnetic waves radiated impulsively from lightning flashes near earth; they are produced when different Fourier components of the impulse travel through the ionosphere at different speeds so as to be received as a note of decreasing frequency, known as whistler. Reception can be either on the earth, or in a space vehicle. Most whistlers provide evidence about the electron content of the regions they traverse. Certain types receivable also provide an estimate of the relative concentration of ions of different masses in the neighbourhood of the satellite.

(vii) Ion whistlers occur at infrasonic frequencies in the range 0.65 to 5 Hz. They have their origin in the action of the solar wind, or of a stream of energetic particles, on the magnetosphere and they travel to the earth as hydromagnetic waves. Their spectrum at the receiver is determined by dispersion of these waves much as the spectrum of audio frequency whistlers
is determined by dispersion of electromagnetic waves.

(viii) Very Low Frequency (VLF) noise (sometimes called hiss) occurs naturally in the ionosphere. Although there is no accepted theory to account for its origin it is often found that its spectrum is terminated abruptly at a low frequency. By measuring this low frequency deductions have been made about the electron and ion content of the magnetosphere.

2.6 Theory of Faraday Rotation

The phase of an electromagnetic wave radiated by the transmitter of a geostationary satellite at the location S and received at a ground station located at R, using geometrical optics, is given by

\[ \phi_{\text{radians}} = \omega t + \phi_0 = \omega \cdot \frac{L}{c} \int_{R}^{S} ds + \phi_0 = 2\pi f/c \int_{R}^{S} nds + \phi_0 \]  

(48)

where \( \phi_0 \) = initial phase at the satellite.

Expanding (47) by Taylor series and retaining only linear terms

\[ n = 1 - x (1 - x) / (2(1 - x) - Y_T^2 + \{Y_T^4 + 4(1 - x)^2 Y_L^2\}^{1/2}) \]  

(49)

In the case of quasi longitudinal propagation (\( Y_T \to 0 \)) it reduces to

\[ n = 1 - x / [2(1 + Y_L)] \]  

(50)

With a further approximation (Ebel et al, 1969) we can write

\[ n = n_{Ox} = 1 - (1 + Y_L) x / 2 = 1 - x / 2 + xy_L / 2 \]

\[ = 1 - 40.305 \frac{N}{f^2} + 11.285 \times 10^{11} \frac{NB\cos\theta}{f^3} \]  

(51)

where \( n_{Ox} \) = refractive index of the ordinary/extra ordinary component of the electromagnetic wave. Substituting the values of \( n \) from (51) in (48)

\[ \phi_{Ox} = 2\pi f/c \int_{R}^{S} ds + 8.447 \times 10^{-7} f \int_{R}^{S} Nds + 2.365 \times 10^4 / \int_{R}^{S} f^2 N\cos\theta ds + \phi_0 \]  

(51)
Now Faraday rotation is half the value of the phase difference between the ordinary and extra ordinary components. Therefore
\[
\mathcal{N} = \frac{1}{2} (\phi_o - \phi_x) = 2.365 \times 10^4 / f^2 \int_{S}^{R} N B \cos \theta ds \tag{52}
\]

Assuming the ionospheric electron density contours to be spherically stratified and thickness of the ionosphere to be small compared to the radius of the earth, the integral along the ray path in eq. (52) can be replaced by an equivalent integral in the vertical direction. Then
\[
\mathcal{N} = 2.365 \times 10^4 / f^2 \int_{0}^{h} B \cos \theta \sec \gamma dh \tag{53}
\]
where \(\gamma\) = zenith angle of the ray path, \(h\) = height of the satellite.

\(B \cos \theta \sec \gamma\) is a geometrical factor involving the earth's magnetic field and is known as the magnetic field factor (M). Since M varies slowly with height (Yeh and Gonzales, 1960) it is possible to assign a weighted mean value (\(\overline{M}\)) to it and express eq. (53) as
\[
\mathcal{N} = 2.365 \times 10^4 / f^2 \int_{0}^{h} N \overline{M} dh = 2.365 \times 10^4 / f^2 \overline{M} N_T \tag{54}
\]
Where \(N_T = \int_{0}^{h} N dh\), the total electron content of the ionosphere in a vertical column of unit cross-section up to the satellite altitude.

If \(\mathcal{N}\) is in degrees and B in ampereturns/metre then eq. (54) becomes
The analytical form of the weighted Magnetic field factor ($\bar{M}$) is given by:

$$\bar{M} = \int_{0}^{h} \frac{NB \cos \theta \sec f \, dh}{N \, dh}$$

Equations 54, 55 give the first-order relationship between the amount of Faraday rotation suffered by a plane polarized wave while traversing the distance between a satellite and a receiver at ground and the total electron content of the ionosphere in a vertical column of unit cross-section. Therefore, the total electron content of the ionosphere can be obtained if $\bar{M}$ is obtained from beacon satellite data and $\bar{M}$ for a station is known.