Chapter 4

The Ionosphere

We can define the ionosphere as the height region of the earth's atmosphere where the concentration of free electrons is so large that it affects radio waves. The ionosphere was discovered when it was observed that radio waves can propagate over large distances, and one therefore had to assume the existence of an electrical conductive layer in the upper atmosphere which could reflect the waves.

The electrically conductive region stretches from about 50km to 500km above the ground (see figure 4.3), and the concentration of electrons n_e varies from 10^7 particles per m^3 at 50km to a maximum of 10^{12} particles per m^3 at 250-300km.

The ionosphere is formed when energetic electromagnetic-and particle radiation from the sun and space ionize air molecules, creating plasma in the upper atmosphere. This plasma is weakly ionized; the ratio between electron density and density of neutral air never becomes larger that 10^{-7}, even at the altitude when n_e reaches its maximum.

The regular ionospheric layers we will describe in this chapter are formed by extreme ultraviolet (EUV) and X-ray radiation from the sun, and have a characteristic variation with the time of day and latitude. In polar regions, i.e., north of 65°, energetic electrons and protons precipitate along the magnetic field lines and give rise to particle impact ionization. Irregular ionospheric layers are formed, which are associated with the northern light phenomena. These layers can cause strong perturbations in radio-wave propagation and cause problems for communication and navigation systems.

4.1 Introduction

The Scottish physicist, Balfour Stewart, understood as early as 1882 that there had to be an ionized region in the atmosphere. Compass measurements of the earth's magnetic field showed variations, which Steward thought could only be due to electric currents in the upper atmosphere. He concluded that the upper atmosphere was more ionized in the daytime than at night, and more in summertime than in wintertime, and more at sunspot maxima than minima (time of day, season, solar cycle dependence).

In December 1901, the Italian Marconi sent radio waves from Cornwall, England to Newfoundland, Canada. British scientists Heaviside and Kenelly concluded that the waves had to follow the curvature of the earth along electrically conductive layers in the upper atmosphere. There had to be an "ionosphere" that acted like a mirror for radio waves with wave length $\lambda >\approx 20\ m$. Together with other scientists they decided to measure the electric properties of the upper atmosphere. The Briton Appleton was among the first who studied reflections from the upper atmosphere by means of interference. He used continuous radio waves and detected displacements by the Doppler principle.

Soon, several ionized layers where discovered, and Appleton suggested a subdivision ordered alphabetically starting with the *E*-layer (Heaviside and Kenelly) at the bottom, and with an *F*-layer above it. Measurements showed that the *F*-layer was divided in two parts, each having its peak. The layers were named F_1 and F_2, respectively. Later, a *D*-layer below the *E*-layer was discovered (see figure 4.3).
The atmosphere above the *F*-layers (>500km) is called the magnetosphere, since the magnetic field has a dominant impact on the movement of the electrically charged particles in this region.

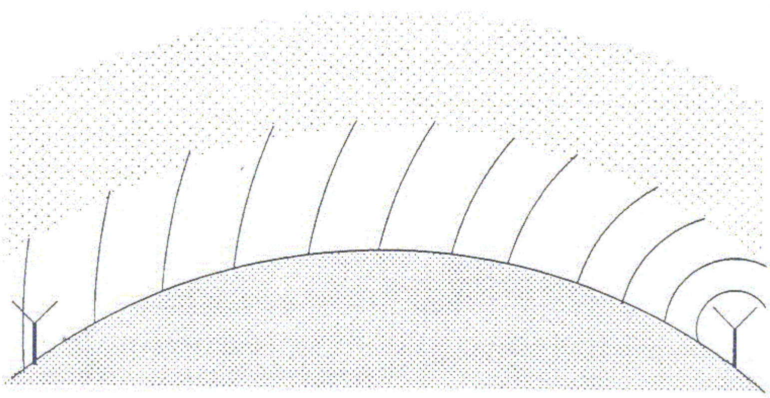

Figure 4.1: Marconi's radio sending from England to the USA on December 12th, 1901 established that there had to exist an electric conductive layer in the atmosphere. Radio waves have ever since been the main tool in the exploration of the ionosphere.

4.2 Formation of the layers of the ionosphere

In this section we are going to describe briefly how the ionosphere is formed. The grade of ionization depends on the intensity and the wavelength of the incoming radiation, as well as the composition of the atmosphere. Refer also to the illustration of the composition of the atmosphere in figure 3.11 page 51.

4.2.1 Photo ionization

The ionospheric layers are formed by photo ionization of atoms (*X*) and molecules (*XY*). Ionization is mainly owed to EUV-radiation from the sun.

$$\begin{aligned} h\nu + X &\rightarrow X^+ + e \\ h\nu + XY &\rightarrow XY^+ + e \end{aligned} \tag{4.1}$$

Ions are lost by recombination through several possible processes:

$$X^+ + e \rightarrow X + h\nu \quad \text{(Slow)} \tag{4.2}$$

$$XY^+ + e \rightarrow X^* + Y^* \quad \text{(Fast)} \tag{4.3}$$

The latter is called dissociative recombination, and leads to the splitting of a molecule into two atoms in an excited state. The process is effective because impulse and energy can easily be distributed among X^* and Y^*. In addition, free electrons can form negative ions by attachment:

$$XY + e \rightarrow XY^-$$

And the negative ions can be lost by photo detachment:

$$XY^- + h\nu \rightarrow XY + e \tag{4.4}$$

We get a continuity equation of the form

$$\frac{dn_e}{dt} = \underset{production}{q} \quad \underset{-}{-} \quad \underset{loss}{\alpha \left[XY^+\right] n_e}$$

With the notation $[XY^+]$ we mean the concentration of the molecule XY^+. Electric neutrality requires that

$$n_e + \left[XY^-\right] = \left[XY^+\right]$$

With the exception of the lowest part of the ionosphere, where $[XY^-] \approx 0$ and $n_e = [XY^+]$. We get

$$\frac{dn_e}{dt} = q - \alpha\, n_e^2 \tag{4.5}$$

The time derivate

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \vec{u} \cdot \nabla$$

in the more general case where $\vec{u}$ is the velocity of the air through the volume element we look at. The simplest models assume $\vec{u} \cdot \nabla n_e = 0$ so that the equation becomes

$$\frac{dn_e}{dt} = q - \alpha\, n_e^2 \tag{4.6}$$

The recombination coefficient α depends of what kind of ion species are present. In the equation, α may be replaced by an effective recombination coefficient

$$\alpha_{eff} = \frac{1}{N^+} \sum_{i=1}^{n} \left(\alpha_i \left[XY^+\right]_i \right) \tag{4.7}$$

Where α_i refers to a certain ion type. Typical ions are O_2^+, NO^+, O^+ in the *E*- and *F*-layer, and composite ions of the type $NO^+ (H_2O)_n$ in the lower ionosphere. For NO^+ and O_2^+, $\alpha \approx 5 \cdot 10^{-7}$ cm^3/s.
When negative ions are present (typical below 75km) we can equate the density of the negative ions with N^- and define $\lambda = \frac{N^-}{n_e}$. By assuming electrical neutrality, the continuity equation becomes:

$$\frac{d}{dt}\left(n_e + N_-\right) \quad = \quad q - \alpha_{eff}\left(n_e + N_-\right) n_e \tag{4.8}$$

$$\frac{d}{dt}\left(n_e(1+\lambda)\right) \quad = \quad q - \alpha_{eff}\left(1+\lambda\right) n_e^2 \tag{4.9}$$

If we assume λ time independent we get

$$\frac{dn_e}{dt} = \frac{q}{(1+\lambda)} - \alpha_{eff}\ n_e^2 \tag{4.10}$$

This equation is similar to Eq. 4.6 except for the fact that the ion production q is replaced by an "effective" production $\frac{q}{(1+\lambda)}$ which is always less than q.

4.2.2 Chapman layers

Sydney Chapman presented in 1931 a simple mathematical model for the formation of ionized layers, which was based on the fact that energetic photons from the sun split air molecules into electrons and positive ions. The model describes the major characteristics of the observed variations in the different layers of the ionosphere.

We will now outline the fundamental theory behind the formation of ionized layers in the atmosphere. The goal is to develop a simple model for how the plasma density varies with height and the sun's zenith angle. We start out with the following assumptions:

- There is only one type of gas present
- The atmosphere is horizontally stratified.
- Radiation is monochromatic and parallel.
- The atmosphere is isotherm (scale height $H = \frac{kT}{mg} = \text{constant}$).

In chapter 3 we found that energy absorbed along the radiation path can be given as (Eq. 3.30)

$$\frac{1}{\sec(\chi)} \cdot \frac{dI}{dz} = I_\infty \sigma n \cdot e^{-\tau \sec(\chi)} \tag{4.11}$$

where σ is the absorption cross section and n is the number of absorbing molecules/atoms per unit volume. The rate of ion production should be proportional to the rate at which radiation is absorbed. If η electron-ion pairs are produced per unit of energy absorbed, the ion production rate becomes

$$q(\chi, z) = I_\infty \sigma\ \eta\ n(z) \cdot e^{-\tau \sec(\chi)} \tag{4.12}$$

Since $\tau = \sigma\ n\ H$ we get

$$q(\chi, z) = \frac{I_\infty \eta}{H} \tau \cdot e^{-\tau \sec(\chi)} \tag{4.13}$$

The maximum ion production q_m is found by calculating τ when $\frac{dq}{dz} = 0$. We find that

$$\tau = \frac{1}{\sec(\chi)} = \cos(\chi)$$

and

$$q_m(\chi, z_m) = \frac{I_\infty \eta}{eH \sec(\chi)} \tag{4.14}$$

For $\chi = 0$ (sun in zenith),

$$q_{m,o} = q_m(0, z_{m_0}) = \frac{I_\infty \eta}{eH} \tag{4.15}$$

Notice that the altitude for maximum ion production z_{m0} in this case is the altitude where the optical depth $\tau = 1$.

4.2.3 Chapman variations

We found a simple expression for how the ion production varies with height z and the sun's zenith angle χ. We also found an expression for the maximum ion production $q(0, z_{m0})$ at $\chi = 0$ and at height z_{m0}. It is convenient to introduce a normalized height parameter z', which measures the height in unit of scale height, and with z_{m0} as reference height.

$$z' = \frac{(z - z_{m_0})}{H} \tag{4.16}$$

We have then chosen a reference altitude $z' = 0$ at where vertically incoming radiation reaches an optical depth of $\tau = 1$. We introduce z' by looking at the height variation of $n(z)$:

$$n = n_0 \cdot e^{-\frac{z}{H}} = n_0 \cdot e^{(-\frac{z_{m0}}{H})} \cdot e^{-\frac{(z - z_{m0})}{H}} = n_0 \cdot e^{-\frac{z_{m0}}{H}} \cdot e^{-z'} \tag{4.17}$$

We put this in equation 4.13

$$q(\chi, \mathrm{z}) = \frac{I_\infty \eta}{H} \sigma\, n\, H \cdot e^{-\sigma \,\mathrm{n}\, \mathrm{H} \sec(\chi)}$$

and remember that $\tau = \sigma\, n\, H$ ($\tau = 1$ where $z = z_{m0}$).

$$\sigma\, n_0 H \cdot e^{\left(-\frac{z_{m0}}{H}\right)} = 1$$

After some calculation we get

$$q(\chi, \mathrm{z'}) = \frac{I_\infty \eta}{eH} \cdot e^{\left(1 - z' - \sec(\chi)\, e^{-z'}\right)} \tag{4.18}$$

or

$$q(\chi, z') = q_{m_0} \cdot e^{\left(1 - z' - \sec(\chi) \cdot e^{-z'}\right)} \tag{4.19}$$

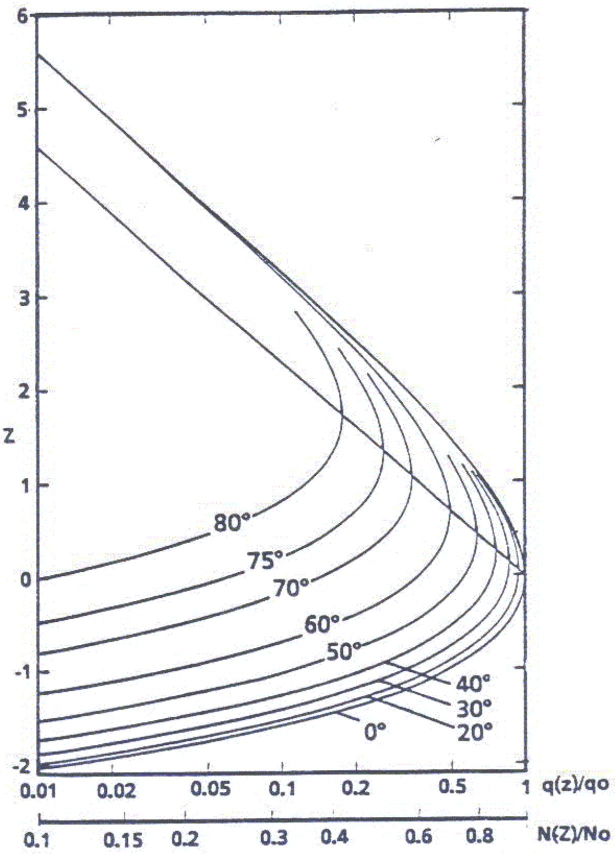

Figure 4.2.: The normalized photo ionization rate $q(z)/q_0$ and the electron density $N(z)/N_0$ according to Chapman's theory show the Chapman layer variations with altitude z and zenith angle χ. Notice that the horizontal axis is logarithmic.

Where $q_{m0} = q_m(0,0)$ is the maximum ion production at $\chi = 0$. Figure 4.2 shows how ion production in such a Chapman-layer varies with z' and χ.

4.2.4 Electron density in a Chapman layer

The number of ion pairs per unit volume consists of production q and loss L, and can be described by the continuity equation

$$\frac{dN^+}{dt} = q - L \tag{4.20}$$

where N^+ is the number of positive ions per volume unit. In the case of charge neutrality, and if no negatively charged ions are present, the electron density is $n_e = N^+$. The loss term then has to be $L = \alpha\, n_e N^+$ where α is a constant (α is the recombination coefficient). The continuity equation can then be written as

$$\frac{dn_e}{dt} = q - \alpha n_e^2 \tag{4.21}$$

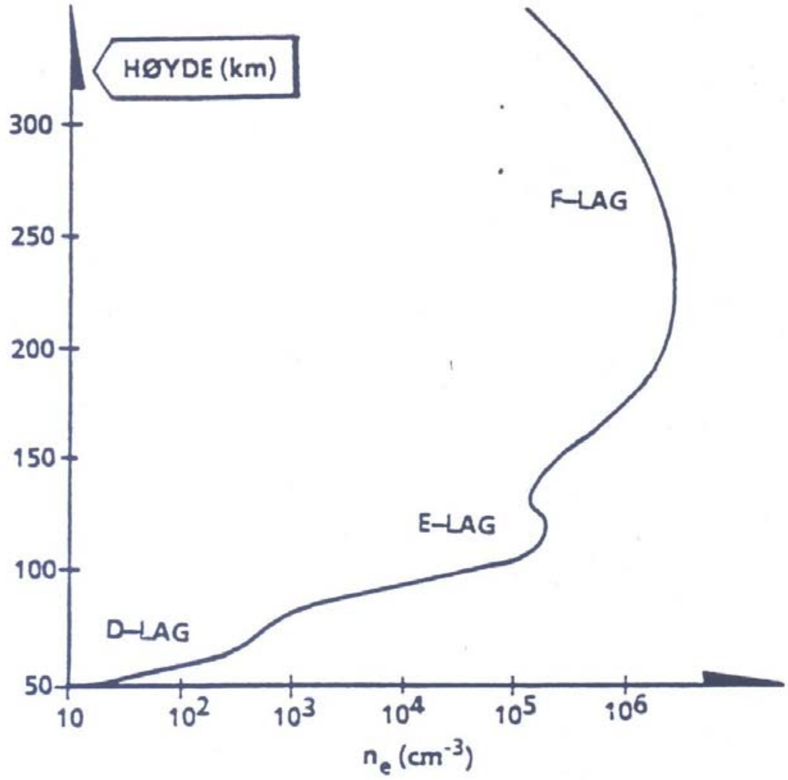

Figure 4.3.: Typical electron density profile in the normal ionosphere

At equilibrium, $\frac{dn_e}{dt} = 0$ and $q = \alpha\, n_e^2$. This results in a peak electron density n_{emo}

$$n_e(\chi, z') = n_{em_0} \cdot e^{\frac{1}{2}\left(1 - z' - \sec(\chi) \cdot e^{-z'}\right)} \tag{4.22}$$

where

$$n_{em_0}^2 = \frac{q_m(0,0)}{\alpha} \quad \text{and} \quad n_{em} = \sqrt{\frac{q_m(0,0)}{\alpha} \cos(\chi)}$$

These equations show that the electron density in a Chapman layer varies with height and sun angle as square root of the ion production q. We get a daily and seasonal variation of the layer. A Chapman layer has its highest peak electron density, and lowest height of this peak, when the sun is at its highest in the sky. The layer disappears at night.

4.3 The layers of the ionosphere

Figure 4.3 shows a typical electron density profile for a normal ionosphere at daytime. As we can see, it is divided into three different layers:

- D-layer (50-95km)
- E-layer (95-150km)
- F-layer (150-500km)

The electron concentration in these layers varies with the exposure to different types of radiation, different types of recombination and various transport processes, and none of the layers behave exactly like the ideal Chapman layer with respect to variations in altitude, time of day, and latitude. In addition, none of the layers disappear totally when the sun is below the horizon, because of scattered radiation, and transport mechanisms, which can transport plasma from a sunlit region to a dark region of the atmosphere. Let us now look more closely at each of these layers.

4.3.1 The *F*-layer

In the altitude above ca. 150km, ions and electrons are formed when the atmosphere's major components, O and N_2, absorb EUV (Extreme Ultraviolet) radiation with wavelength $10nm < \lambda < 90nm$. The primary ions are O^+ and N_2^+, but these react quickly with neutral atoms and molecules

$$\begin{aligned} O + hv &\rightarrow O^+ + e \\ N_2 + hv &\rightarrow N_2^+ + e \end{aligned}$$

$$\begin{aligned} O^+ + O_2 &\rightarrow O_2^+ + O \\ O^+ + N_2 &\rightarrow NO^+ + N \\ N_2^+ + O &\rightarrow NO^+ + N \end{aligned}$$

The most important ions are therefore O^+, NO^+ and O_2^+. These recombine to

$$O^+ + XY \rightarrow XO^+ + Y \tag{4.23}$$

$$XO^+ + e \rightarrow X^* + O^* \tag{4.24}$$

Where X and Y can be O or N. The continuity equation for O^+ is given by

$$\frac{d}{dt}\left[O^+\right] = q - \gamma \left[O^+\right]\left[XY\right] = q - \beta \left[O^+\right] \tag{4.25}$$

where $\beta = \gamma[XY]$. The factor β will in a certain height be approximately constant because the neutral density $[XY]$ does not change significantly during the ionization process (the ionization rate is small). The continuity equation becomes:

$$\frac{d}{dt}\left[XO^+\right] = \beta \left[O^+\right] - \alpha \left[XO^+\right] n_e \tag{4.26}$$

We can write $[O^+] + [XO^+] = n_e$. By combining these two equations we get

$$\frac{dn_e}{dt} = q - \alpha \left[XO^+\right] n_e \tag{4.27}$$

At high altitudes (>200km) where O^+ is the dominant ion, and the loss rate becomes proportional to the electron density.

$$\left[O^+\right] >> \left[XO^+\right] \rightarrow \left[O^+\right] \approx n_e \qquad => \frac{dn_e}{dt} = q - \beta\, n_e$$

In the lowest part of the *F*-layer, XO^+ ions dominate, and we get a loss proportional to the square of the electron density.

$$\left[O^+\right] << \left[XO^+\right] \rightarrow \left[XO^+\right] \approx n_e \qquad => \frac{dn_e}{dt} = q - \alpha\, n_e^2$$

As mentioned earlier, the *F*-layer deviates from an ideal Chapman layer. This deviation results from complicated recombination processes, and also from the divergence term $\vec{u} \cdot \nabla n_e$, which is usually not negligible. The continuity equation should therefore be

$$\frac{\partial n_e}{\partial t} = q - \alpha \left[XO^+\right] n_e - \vec{u} \cdot \nabla n_e \tag{4.28}$$

4.3.2 The *E*-layer

The *E*-layer stretches from ~95km to 150km above the ground and is the layer of the ionosphere that is in closest agreement with the Chapman description. The ion production in the normal *E*-layer is caused by X-rays (*1nm < λ< 10nm*) and ultraviolet radiation (*100nm < λ < 150nm*) dissociating O_2 and N_2 to O_2^+ and N_2^+. N_2^+ disappears quickly by charge exchange.

$$N_2^+ + O_2 \quad \rightarrow \quad O_2^+ + N_2 \tag{4.29}$$

$$N_2^+ + O \quad \rightarrow \quad NO^+ + N \tag{4.30}$$

so that O_2^+ and NO^+ are the dominating ions. Recombination is therefore dissociative.

$$O_2^+ + e \quad \rightarrow \quad O^* + O^* \tag{4.31}$$

$$NO^+ + e \quad \rightarrow \quad N^* + O^* \tag{4.32}$$

and the continuity equation gets into the standard form

$$\frac{dn_e}{dt} = q - \alpha_{eff} n_e^2$$

Thin sporadic layers appear in the *E*-layer due to, among other things, the incidence of metal ions (i.e., Na^+, Mg^+,and Fe^+) with a long lifetime, which are affected by dynamic processes. Sporadic *E*, denoted as (*Es*), can be extremely thin and dense layers. They are not associated with the normal *E*-layer, and occur more occasionally both day and night.

4.3.3 The *D*-layer

Electron density in the *D*-layer normally does not have a distinct peak (see figure 4.3). Both the ion production and the recombination processes are highly complicated between 50km and 95km. *H*-Lymann – α radiation ($\lambda = 121.5nm$), which is a very strong spectral line, penetrates down into the *D*-layer and has enough energy to ionize NO, which is found in small amounts (see figure 3.11 Check)). This production mechanism dominates from about 70-95km, but the contribution from solar X-rays can be important as well. Below about 70km, normal ionization from high-energy cosmic radiation dominates. In the *D*-layer, a large number of complicated and heavy positive and negative ions is formed, and the recombination processes are therefore both height- and temperature-dependent.

4.4 The disturbed ionosphere

Energetic particles like electrons, protons and α - particles from the sun and the magnetosphere can penetrate the atmosphere and contribute to an extraordinary production of ions and electrons in the ionosphere. Such particle precipitation is closely related to the northern light phenomenon, and the layers of the ionosphere vary, therefore, often very irregularly, in the auroral zone. Especially important is the ion production that takes place in the *E*-and *D*-layer. Here, the increase in electron density can give rise to great disturbances in radio-wave propagation conditions (reflection, absorption, and scatter).
Disturbances in the polar ionosphere can cause a total breakdown in shortwave communication ("radio blackouts"), and the electric currents that are induced in the ionosphere can influence power supply, corrosion in oil pipelines, etc.

Updated 1 SEPT2009