

SOME PROPERTIES OF A MODEL OF THE F1 LAYER
OF THE IONOSPHERE.

A Thesis submitted for the
Degree of Master of Science of
Rhodes University

by
GERHARD DE JAGER.

January, 1963.

Except where it is clear from the text that I am describing the work of others, or where it is obvious that I am making a survey of existing knowledge about the ionosphere, the work described in this thesis is my own.

Gerhard de Jager

CONTENTS.

	<u>Page No.</u>
ACKNOWLEDGEMENTS.	viii
INTRODUCTION.	x
CHAPTER 1. THE IONOSPHERE AND ITS EXPLORATION BY RADIO.	1
1.1 A conducting region in the upper atmosphere.	1
1.2 The structure of the ionosphere.	2
1.3 Exploring the ionosphere by radio.	2
1.4 Plan of the rest of the chapter.	3
1.5 The refractive index.	4
1.6 The condition for reflection.	6
1.7 Obtaining the group refractive index from μ .	8
1.8 The equivalent height.	9
1.9 The shape of the $h'(f)$ curve.	10
1.10 The deduction of an ionogram from an $N(h)$ curve.	11
CHAPTER 2. ORIGIN AND SHAPE OF THE ELECTRON LAYERS.	13
2.1 Introduction.	13
2.2 The production of ion-electron pairs.	13
2.2.1 Chapman theory of production.	14
2.2.2 Some extensions to the Chapman theory.	15
2.3 The loss of electrons.	15
2.3.1 Simple recombination.	16
2.3.2 Attachment.	17
2.3.3 A more complicated loss process.	17
2.3.4 More than one species of recombining ions.	19

2.4	Layer formation.	19
2.4.1	Chapman theory.	19
2.4.2	The Bradbury hypothesis.	20
2.5	The effect of movements.	21
2.5.1	The continuity equation.	21
2.5.2	Diffusion in the F2 region.	22
2.6	Layer formation with diffusion taken into account.	23
2.6.1	Attachment like loss operative in the F2 region.	23
2.6.2	The F2 peak formed by diffusion.	23
2.7	A transition region present in the F layer.	26
CHAPTER 3.	THE F1 LAYER.	27
3.1	Two models.	27
3.2	The model used by Mariani.	27
3.3	The model used by Hirsh.	28
3.3.1	Identification of the species of atoms and molecules involved.	28
3.3.2	Transition from a linear to a quadratic loss law.	29
3.3.3	The effective attachment coefficient a function of height.	31
3.3.4	Mixing and diffusive equilibrium.	32
3.3.5	The F1 layer on this model.	34
CHAPTER 4.	A NON-EQUILIBRIUM MODEL OF THE F1 LAYER.	36
4.1	Introduction.	36
4.2	The production function.	38

6.6	Enhancement of the F1 layer.	72
6.6.1	Introduction.	72
6.6.2	The f_oF1 cusp on an ionogram.	73
6.6.3	The production function.	74
6.6.4	Changes in the loss rate.	77
6.6.5	The nett rates of change of ion and electron densities.	81
6.6.6	Conclusion.	82
6.7	The upward movement of the base of the F2 layer.	82
6.8	The $f_oF1\frac{1}{2}$ critical frequency.	83
6.8.1	Introduction.	83
6.8.2	The point of inflection that gives rise to the $f_oF1\frac{1}{2}$ cusp.	84
6.8.3	Identification of the species of positive ions involved.	84
6.8.4	The explanation for an $F1\frac{1}{2}$ point of inflection proposed by Ratcliffe.	85
6.8.5	Application of this explanation to the present model.	88
6.8.6	An alternative viewpoint.	89
6.8.7	The electron density.	90
6.8.8	Discussion.	92
6.9	Variation in height of the transition level.	92
6.10	The $N_{min} - \delta t$ anomaly during an eclipse.	95
6.11	Discussion.	99
CHAPTER 7. DISCUSSION.		101

7.1	The problem.	101
7.2	The conditions under which the $F1\frac{1}{2}$ layer appears.	101
7.3	Middle latitudes.	102
7.4	The results of the model and McElhinny's observations.	102
7.5	The problem of assuming near complete diffusive separation of the ionizable constituent and the molecular species XY.	106
CHAPTER 8. SUGGESTIONS FOR FURTHER RESEARCH.		108
SUMMARY.		109
REFERENCES.		111

ACKNOWLEDGEMENTS.

I would like to thank Professor J.A. Gledhill, my research director, for his guidance and unfailing interest in my work, and for his willingness to discuss problems with me at all times. I am very grateful.

Also I would like to thank Drs. J.E.C. Gliddon and P.C. Kendall, who arranged for certain computations, in connection with this project, to be done, Mr. A.D.M. Walker for his advice on the deductions of ionograms, and Mrs. H. Locke for much help and advice in the preparation of the diagrams.

The assistance of Mr. and Mrs. V.C. Moran and Mr. G. Walters, in connection with the photographic work, is very much appreciated.

I am very grateful to Miss K. Longfield for her skill and patience in the typing of this thesis.

Finally I wish to thank my father for his encouragement over the years.

Mysterious, in the blaze of day,
Nature pursues her tranquil way:
The veil she wears, if hand profane
Should seek to raise, it seeks in vain,
Though from her spirit thine receives,
When hushed it listens and believes,
Secrets - revealed, else vainly sought,
Her free gift when man questions not, -
Think not with levers or with screws
To wring them out if she refuse.

GOETHE: Faust.

INTRODUCTION.i. The problem.

The present work was initially aimed at providing an explanation for some of the phenomena that occur in the ionosphere at sunrise. The approach that was taken was to determine the changes that take place on a theoretical model of the ionosphere and then to compare these with observations. A prerequisite for this approach was a theoretical model that would show, among other things, a bifurcation of the F layer during daytime without making unjustified arbitrary assumptions. The absence of such a model, particularly as far as non-equilibrium conditions are concerned, resulted in the present attempt to provide such a model for the F1 region.

ii. Presentation.

In Chapter 1 an account is given of the structure of the ionosphere and the usual way in which experimental observations of it are obtained. Chapters 2 and 3 describe the theory of layer formation and the specific case of the F1 layer.

In Chapter 4 a non-equilibrium model of the F1 region of the ionosphere is proposed. The behaviour of this model under normal day conditions and during an eclipse is then considered in Chapters 5 and 6 respectively. In Chapter 7 the results are discussed and some suggestions for further research follow in Chapter 8.

To facilitate easy comparison of the eclipse results Figures 6.1, 6.2, 6.3, 6.4, 6.6, 6.17 and 6.18 are appended in such a way they they can be placed next to each other.

iii. Nomenclature.

A few words on the terms used in this thesis are needed.

The word "peak" is used to indicate the maximum in any quantity when considered as a function of height, so that $\frac{dN}{dh} = 0$ at the peak.

Workers on the ionosphere at first used the term "F1 layer" to refer to the layer below the daytime F2 layer. It gradually became clear that this F1 layer does not exhibit a definite peak, but is sometimes marked only by a "ledge". In this thesis it is, nevertheless, convenient to refer to the F1 layer, except where special attention is to be drawn to the fact that no peak in the layer occurs. The term "F1 ledge" is then used.

Diffusion describes both ambi-polar or plasma diffusion which refers to the diffusion of a plasma through a neutral gas, and the diffusion of one neutral gas through another. Its meaning in any specific case should be clear from the context.

1.

CHAPTER 1.

THE IONOSPHERE AND ITS EXPLORATION BY RADIO.

1.1 A conducting region in the upper atmosphere.

In 1839 CARL FRIEDRICH GAUSS criticized the assumption that the seat of the currents responsible for variations in the terrestrial magnetic field was contained wholly within the surface of the earth. In his "General Theory of Terrestrial Magnetism" published in that year he comes to the conclusion that the existence of an electrically conducting region in the upper atmosphere could not be precluded on the basis of the knowledge then available. KAISER (1962) points out that this is the earliest known time at which the existence of such a conducting region was mentioned.

Twenty-one years later, in 1860, Lord KELVIN mentioned the idea of a conducting region in the upper atmosphere in a discussion of atmospheric electricity (CHALMERS, 1962). In the absence of experimental verification BALFOUR STEWART in 1878 repeated the suggestion made by Gauss in 1839.

The existence of such a conducting region was once more postulated by KENNELLY (1902) and independently by HEAVISIDE (1902). This time it was invoked in an attempt to explain Marconi's success in transmitting radio signals across the Atlantic. The radio waves would then be reflected back and forth between the ocean and the layer and in this way be propagated round the curved surface of the earth.

2.

Direct evidence for the existence of a conducting layer came in 1925 when APPLETON and BARNETT (1925) detected a radio wave coming down from the sky after being reflected. Many experiments have since been conducted in studying this layer, including measurements of the electron densities by rockets and satellites in situ.

1.2 The structure of the ionosphere.

The region in the earth's atmosphere in which free electrons exist in significant number is called the ionosphere, a term first proposed by WATSON WATT. This region extends even below 80 km. but the electron density exhibits maxima at about 120 km. and 300 km. Figure 1.1 illustrates this showing curves of the electron density N plotted against the height h . The peak in N at 120 km. is termed the E layer and the one at 300 km. the F_2 layer (terms first suggested by Appleton). A subsidiary maximum, the F_1 ledge, is sometimes present at about 200 km, but is often absent or present only as a point of inflection in the $N(h)$ curve.

1.3 Exploring the ionosphere by radio.

In 1925 a method similar to radar and sonar was employed for the first time in exploring the ionosphere (BREIT and TUVE, 1926). In it the time delay between transmission and reception of a pulse of radio waves after being reflected by the ionosphere is determined and from this the height of the reflecting layer is computed.

When a pulse of electro-magnetic radiation of radio frequency is incident vertically on and reflected from the ionosphere, the

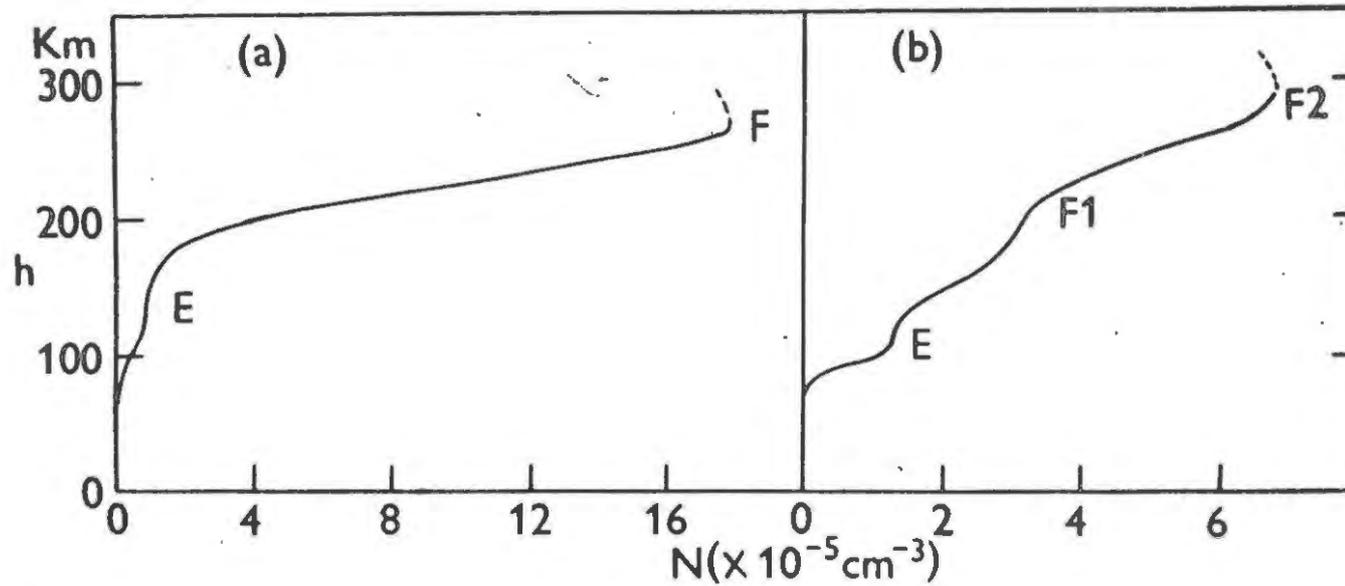


Figure 1.1/ Electron distributions showing the electron concentration N as a function of height h . The curves correspond to the ionograms of figure 1.2. The E and F2 layers, and the F1 ledge, correspond to the critical frequencies fE , $fF1$ and $fF2$.

time delay t is usually expressed in terms of the equivalent height h' . h' would be the height of the reflecting layer had the pulse travelled with the free space velocity c all the way, because the relation

$$2h' = ct \quad (1.1)$$

would then apply. By using a vertical incidence sounder of variable frequency f the variation of h' with f can be determined and so $h'(f)$ curves or ionograms can be obtained. Figure 1.2 shows typical traces of ionograms such as would result from $N(h)$ distributions like those in Figure 1.1. From Figure 1.2 it is evident that when the exploring frequency approaches the marked frequencies h' starts to increase rapidly. These marked frequencies are called critical frequencies or penetration frequencies and they will just penetrate the maxima in the electron density at the E and F2 layer peaks.

It should be noted in particular that the height h' is not the true height. As soon as the pulse enters the ionosphere its group velocity decreases, which means that a retardation of the pulse takes place. The height h' is therefore obviously always greater than the true height of the reflecting layer.

1.4 Plan of the rest of the chapter.

The final check on our theoretical model will be to compare some theoretically constructed ionograms with ones obtained experimentally. To be able to do this we shall now continue by describing how such a deduction can be done. It is, however, necessary to start at the basic magneto-ionic theory which describes the

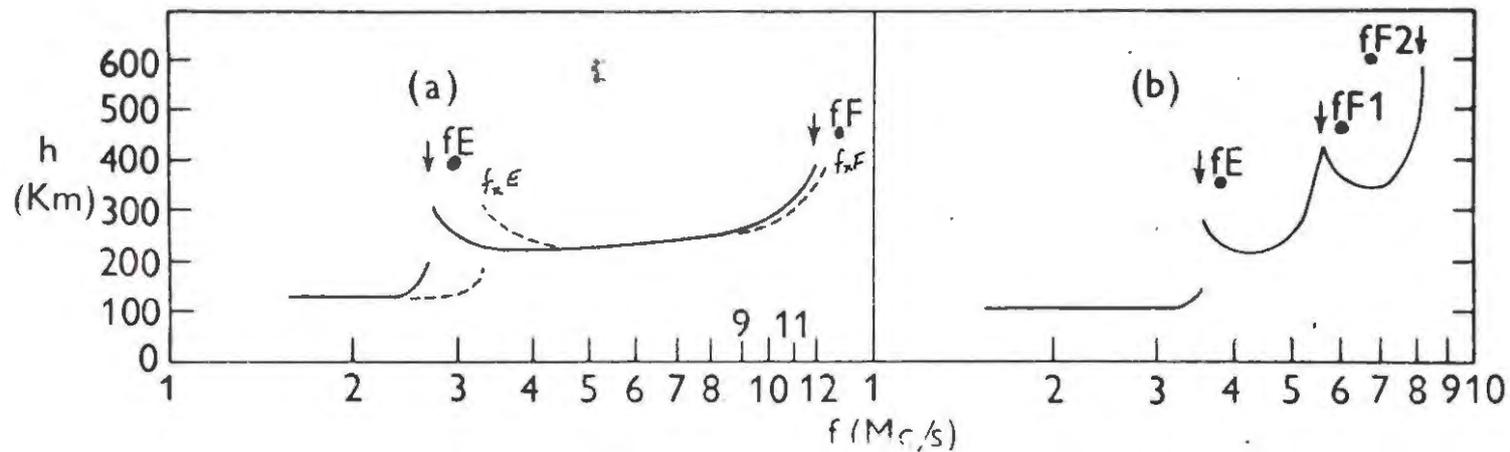


Figure 1.2 Observed 'ionograms' showing the equivalent height h' of reflection of a radio wave as a function of frequency f . The broken line represents the 'extraordinary wave' resulting from the birefringence of the ionosphere; it is not shown in (b). f_E , f_{F1} and f_{F2} mark the 'critical frequencies'.

4.

propagation of an electro-magnetic wave in a gas consisting of neutral molecules and free charges, the free charges being distributed uniformly throughout the gas. From this theory we shall quote an expression for the phase refractive index and show which conditions result in the reflection of a wave. After obtaining a relationship between the group and phase refractive indices we shall show how the equivalent height depends on the group refractive index and finally describe how an ionogram can be constructed from a given $N(h)$ distribution. As we proceed through the chapter we shall refer to the first two figures to point out features to be expected in an ionogram as a result of the theory under discussion.

1.5 The refractive index.

The magneto-ionic theory on which our expression for the complex refractive index n is based, considers the effect of electrons on the propagation of an electromagnetic wave, the effect of the more massive ions being very small. A formula for n which takes into account the effects of collisions of electrons with neutral particles and those due to the presence of a magnetic field was first derived by APPLETON (1927, 1932) and independently by HARTREE (1931). When the wave frequency is greater than about 1 Mc/s (as it is for the conditions we are interested in), the effects of collisions may be neglected. The expression for n when collisions are ignored can be found in standard textbooks on the subject e.g. BUDDEN (1961) and is quoted as

$$n^2 = 1 - \frac{X}{1 - \frac{Y_T^2}{2(1-X)} \pm \left\{ \frac{Y_T^4}{4(1-X)^2} + Y_L^2 \right\}^{\frac{1}{2}}} \quad (1.2)$$

where $n = \mu - j\chi'$, the complex refractive index

μ = phase refractive index = $\frac{c}{v}$

χ' = attenuation factor

c = velocity of electromagnetic waves in vacuo

v = phase velocity of electromagnetic radiation in the medium

$X = f_0^2/f^2$

$Y_T = Y \sin \theta$ $Y_L = Y \cos \theta$ $Y = f_H/f$

f = wave frequency

f_0 = plasma frequency = $(1/2\pi) \cdot \left(\frac{Ne'^2}{\epsilon_0 m} \right)^{\frac{1}{2}}$

f_H = gyro frequency = $Be'/2\pi m$

e' = charge on the electron

m = mass of the electron

ϵ_0 = electric permittivity of free space

μ_0 = magnetic permittivity of free space

B = magnitude of magnetic induction in the medium

= $\mu_0 H$ if the relative magnetic permeability of the medium is unity.

θ = angle made by the magnetic field with the direction of the radiation.

(Rationalised units are used in the above expressions).

A point to note about this equation is that it is double valued due to the plus and minus terms in the denominator. The presence

6.

of the earth's magnetic field thus renders the ionosphere birefringent for electromagnetic waves. The plus sign in the denominator will give us the refractive index for the ordinary wave, the minus sign n for the extraordinary wave except in the case where the wave normal is in the direction of the imposed field.

1.6 The condition for reflection.

Assuming the ionosphere to be horizontally stratified, we find that on applying Snell's law to a wave entering it from below at an angle of incidence i ,

$$\mu \sin \phi = \sin i \quad (1.3)$$

where ϕ is the angle contained between the wave normal and the vertical. It is clear that the wave will be refracted more and more if N is increasing with height. At a level where N is sufficient to reduce $\sin \phi$ to unity and therefore $\mu = \sin i$, the wave will travel horizontally and after this will return to the ground. For vertical incidence this will happen if

$$\mu = 0. \quad (1.4)$$

Although this condition is derived on the simple ray theory, a full wave treatment of the problem (BUDDEN, 1961) shows the same condition to hold even if the earth's magnetic field is not neglected.

In the case of no magnetic field

$$n = 1 - X \quad (1.5)$$

and reflection will take place when $n = 0$, i.e. when $f = f_o$, so that

$$N = \frac{4\pi^2 m^2 f_o^2}{e'^2} \quad (1.6)$$

$$= 1.24 \times 10^4 f_o^2 \quad (1.7)$$

if N is in electrons per cm^3 and f_o is in Megacycles. By using this relation one can plot any electron density distributions on a true height-plasma frequency diagram.

When a magnetic field is present and the wave frequency f is greater than the gyro-frequency (as it usually is), the conditions under which $n = 0$ are given by

$$X = 1 \quad \text{Ordinary wave} : f = f_o \quad (1.8)$$

$$X = 1 - Y \quad \text{Extraordinary wave} : f_o^2/f^2 = 1 - f_H/f \quad (1.9)$$

$$\text{or } f_o^2 = f^2 - f \cdot f_H$$

$X = 1 + Y$ also gives a zero, but in the ionosphere the wave cannot penetrate to the level at which this condition is satisfied because it gets reflected lower down where $X = 1 - Y$.

We notice that the condition for reflection for the ordinary wave is the same as for the case in which the earth's magnetic field is neglected. Obviously, if N is an increasing function of h , for a given probing frequency f the extraordinary wave will be reflected at a lower height than the ordinary wave. This fact results in the extraordinary ray trace on an ionogram being shifted

to the high frequency side as can be seen on referring to Figure 1.2.

1.7 Obtaining the group refractive index from μ .

From equation (1.2) we can calculate μ for any values of X and Y. To obtain a value of μ' from this, we follow an approach due to SHINN and WHALE (1952).

Standard theory on wave propagation shows that in a dispersive medium the velocity of propagation of a wave group is not the same as it is in a vacuum, but is related to the phase velocity and the frequency in the following way

$$\frac{1}{U} = \frac{d}{df} \left(\frac{f}{v} \right) \quad (1.10)$$

where U is the group velocity and v is the phase velocity.

Defining the group refractive index as $\mu' = \frac{c}{U}$ by analogy with the case of the phase velocity, and differentiating (1.10), yields

$$\mu' = \mu + f \frac{d\mu}{df} \quad (1.11)$$

on substitution of the refractive indices.

Shinn and Whales's method is based on the differentiation of equation (1.2) with respect to f and substitution into equation (1.11) to give

$$\mu' = \frac{1}{\mu(1-X) \left[\frac{1 - Y_I^2}{1 - X} \right] + \left[\frac{Y_I^4}{4(1-X)^2} + Y_L^2 \right]^{\frac{1}{2}}} \times$$

$$\left\{ 1 - \mu^2 - X^2 \pm \frac{(1 - \mu^2)(1 + X)}{2 \left[\frac{Y_I^4}{40 - X^2} + Y_L^2 \right]^{\frac{1}{2}}} \right\} + \mu \quad (1.12)$$

WALKER (1962) calculated values of μ' for Grahamstown by this method and these values were used in the present work.

1.8 The equivalent height.

Let us consider a slab of thickness dh at a height h in a horizontally stratified ionosphere. On passing through it a pulse of electromagnetic waves will travel with a group velocity U where

$$U = c/\mu'$$

If it passes through the slab in time dt , we have

$$dt = dh/U \quad (1.13)$$

In the case where the pulse travels from the ground, is reflected and returns to the ground, the total time elapsed will be given by

$$t = 2 \int_0^h dh/U$$

which, on substitution for U yields

$$ct = 2 \int_0^h \mu' dh \quad (1.14)$$

Equation (1.1) also gave an expression for ct . Equating the two values we obtain

$$h' = \int_0^h \mu' dh \quad (1.15)$$

This formula for h' is the basis for all the methods of obtain-

ing the true height h from the experimentally measured values of h' . To do this is, however, not a simple matter because, as we have seen, the expression for μ' is in general rather complicated, with the result that numerical methods must be resorted to.

1.9 The shape of the $h'(f)$ curve.

We are now in a position to discuss the appearance of an ionogram such as in Figure 1.2 in the light of some theoretical considerations. The fact that the ionosphere is rendered birefringent by the presence of the earth's magnetic field leads us to expect two traces. In Figure 1.2 the ordinary wave trace is indicated by the solid line, the extraordinary by the dashed line.

Most of the retardation of the wave takes place in a region where the wave frequency is near the plasma frequency. In regions where the gradient of $\frac{dN}{dh}$ is small, i.e. near the E and F layer peaks, the wave travels for an appreciable distance in a region where the wave frequency is not far removed from the plasma frequency. Very much retardation is therefore to be expected at the critical frequency where the wave just penetrates the electron density peak. The critical frequencies for the ordinary wave are marked f_oE and f_oF and for the extraordinary wave f_xE and f_xF .

The conditions for reflection outlined in section 1.6 result in the shifting of the extraordinary wave trace to lower levels and thus to the high frequency side. From the resulting difference in the critical frequencies between the ordinary and extraordinary waves attempts have been made to determine the magnetic field

strength in the ionosphere (SINGH and TOLPADI, 1962).

If the wave is reflected from a region in which the gradient $\frac{dN}{dh}$ is large, the wave frequency will be well removed from the plasma frequency for the greatest part of the path traversed, so that little retardation is again expected. We see that a point of inflection on the $N(h)$ curve will also give rise to a cusp on an ionogram such as appears in Figure 1.2 at the point marked f_oF_1 which is due to the F_1 ledge that appears in Figure 1.1.

1.10 The deduction of an ionogram from an $N(h)$ curve.

The equivalent height of reflection of a radio wave incident vertically on the ionosphere is given by

$$h' = \int_0^h \mu' dh \quad (1.16)$$

where the group refractive index is a function of N , the wave frequency f and the direction of the earth's magnetic field.

Following TITHERIDGE (1959), the $N(h)$ curve may be divided up into a series of heights h_r at a given series of frequencies which we shall call the Titheridge frequencies. Corresponding to every h_r is an equivalent height h'_r . Let the mean value of μ' for a wave frequency f_n over the lamination h_{r-1} to h_r be $\overline{\mu'_{rn}}$ so that we can write

$$\overline{\mu'_{rn}} = \frac{1}{h_r - h_{r-1}} \int_{h_{r-1}}^{h_r} \mu' dh \quad (1.17)$$

Titheridge proposed a method whereby h' can be easily evaluated

from h in an $N(h)$ curve making use of the changes in height from one frequency to the next and of the changes in group retardation due to a given segment of ionization as the wave frequency is increased. The values of $\overline{\mu}'_{rn}$ for Grahamstown for a set of Titheridge frequencies had previously been calculated by WALKER (1962) and use was made of these in this thesis.

Equation (1.16) can be approximated by

$$h'_n = \sum \overline{\mu}'_{rn} \Delta h \quad (1.18)$$

when the electron density distribution is thought of as being divided up into slabs of thickness Δh by the Titheridge scaling frequencies. By a straightforward application of equation (1.18) h' can be evaluated for every Titheridge frequency so as to produce an ionogram.

CHAPTER 2.ORIGIN AND SHAPE OF THE ELECTRON LAYERS.2.1 Introduction.

Ionograms obtained at regular intervals during the diurnal, seasonal and sunspot cycles at different latitudes enable us to study the behaviour of the ionosphere in detail. From the experimental observations that have become available in this way it is evident that a theory to account for all the observed phenomena would be very complex. Since simplified models can account for many of the observed effects and give us an insight into the physical processes involved, they are still used extensively. In this thesis we shall investigate the properties of such a simple model. In the sections that follow we shall outline some of the theories advanced to account for the processes of electron production and loss and the resulting formation of the layers of ions and electrons.

2.2 The production of ion-electron pairs.

Experiment leads us to believe that electromagnetic radiation in the ultra-violet region, emanating from the sun, is the main agent responsible for the production of the ionized layers, (WATANABE and HINTEREGGER, 1962). According to the above-mentioned authors soft X-rays play some part in the formation of the E layer. The role of corpuscular radiation in the formation of the ionized layers is not very clear at present, but it is thought to be rather insignificant as far as the E and F1 layers are concerned. Any theory

based on the assumption that the ionizing agent is electromagnetic radiation should therefore be more acceptable than others.

2.2.1 Chapman theory of production.

CHAPMAN (1931a,b) first considered the problem of the absorption of monochromatic radiation from the sun in an atmosphere of uniform composition in which the density decreases exponentially with increase in altitude. In this model the atmosphere is assumed to be in equilibrium under gravity and the temperature T and the molecular weight m are assumed to be constant, so that we obtain the following expression for the density f at a height h with respect to the density f_0 at some reference level h_0

$$f = f_0 \exp(-z) \quad (2.1)$$

$$\text{where } z = \frac{h - h_0}{H}$$

$$\text{and } H = \frac{kT}{mg}, \text{ termed the scale height by Chapman}$$

$$k = 1.38 \times 10^{-16} \text{ ergs per degree, called Boltzmann's constant}$$

$$\text{and } g = \text{acceleration due to gravity.}$$

H will be a constant if the variation in the gravitational acceleration with altitude is neglected.

If q ion-electron pairs are produced at level h in a cm^3 per second by a beam of monochromatic radiation of intensity I

$$q = IBh \quad (2.2)$$

where n is the number density of and B' the ionization cross-section for the ionizable gas. I is the product of I_{∞} , the flux of ionizing photons incident at the top of the horizontally stratified ionosphere, and a term that depends on the total number of ionizable particles in a column of unit cross-section along the path traversed by the radiation in the atmosphere. Substituting appropriate expressions for I and n in equation (2.2) we obtain

$$q(z) = q_0 \exp \left[1 - z - \exp(-z) \text{Ch}(\chi) \right] \quad (2.3)$$

where q_0 is the peak rate of production at the reference height h_0 for overhead sun and $\text{Ch}(\chi)$ is a function defined by CHAPMAN (1931b) which depends on the solar zenith angle χ and the scale height. $\text{Ch}(\chi)$ approximates to $\sec \chi$ for χ less than 75° .

2.2.2 Some extensions of the Chapman theory.

CHAPMAN (1939) extended this theory to the case in which solar radiation was absorbed in the atmosphere in absorption bands. Another important extension to the simple theory was to abandon the restrictions imposed by the assumption of a constant scale height. GLEDHILL and SZENDREI (1950) and NICOLET (1951) considered the case in which the scale height increased with altitude. The theory has further been adapted (RISHBETH and SETTY, 1961), to account for two ionizable constituents having different scale heights and ionizing cross-sections but absorbing the same radiation.

2.3 The loss of electrons.

Electron loss from a given volume can take place through recombination with positive ions, attachment to neutral atoms or

molecules, and by moving outside the given volume. In this section we shall consider the former two processes leaving the transport process for a later one.

2.3.1 Simple recombination.

Let the number densities of electrons and positive ions be N and $n(A^+)$ respectively. We can then write

$$\frac{dN}{dt} = -\alpha N.n(A^+) \quad (2.4)$$

for the loss rate of electrons by this process. In the above equation α is the recombination coefficient. If this is the only process (or, in the actual case, if the number of negative ions formed is negligible in comparison with the number of positive ions), N will be approximately equal to $n(A^+)$ and we can write

$$\frac{dN}{dt} = -\alpha N^2 \quad (2.5)$$

in which the loss rate is proportional to N^2 . Radiative recombination



and dielectronic recombination



(BATES, 1956) would both fall under this scheme, but for the conditions prevailing in the upper atmosphere the value of α for both these processes would only be of the order of $10^{-12} \text{ cm}^3 \text{ sec}^{-1}$. (In the above expressions $h\nu$ represents a photon and e an electron).

2.3.2 Attachment.

Since the effect of ions on the propagation of a radio wave is negligible in a region where the number densities of ions and electrons are comparable, an electron loses its effectiveness on combining with an atom or molecule to form an ion. If $n(a)$ is the number density of neutral atoms or molecules to which the electrons can attach themselves, the rate at which electrons become ineffective by this process is given by

$$\frac{dN}{dt} = -b.N.n(a) \quad (2.6)$$

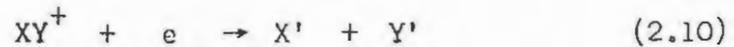
where b is the rate constant for this process. If we take $n(a)$ to be so great that it is not affected by the process described here, we can write

$$\frac{dN}{dt} = -\beta N \quad (2.7)$$

in which the "loss" rate is proportional to the first power of the electron density. β is called the attachment coefficient.

2.3.3 A more complicated loss process.

From experimental observations it has become clear that the effective recombination coefficient in the E and F1 layers is of the order of $10^{-8} \text{ cm}^3 \text{ sec}^{-1}$ (RATCLIFFE, 1956) and that the loss rate is proportional to the square of the electron density. Since the simple recombination processes described in section 2.3.1 would therefore be inadequate to account for the recombination operative in these regions, BATES and MASSEY (1946) were led to propose the recombination scheme represented by the following expressions:



in which (2.8) represents the photoionization of atom A

(2.9) represents a charge transfer process between ion A^+ and molecule XY, and

(2.10) represents the recombination of an electron with the molecular ion XY^+ with the resulting dissociation of the molecule that may leave the atoms in excited states (denoted by primes).

The rate for process (2.9) is given by

$$\frac{dn(A^+)}{dt} = -\lambda_1 n(A^+) \cdot n(XY) \quad (2.11)$$

and for (2.10) by

$$\frac{dN}{dt} = -\lambda_2 n(XY^+) \cdot N \quad (2.12)$$

where λ_1 and λ_2 are the respective rate constants.

At first it was thought that (2.9), representing charge transfer (BATES, 1956), was operative in the ionosphere but at present an ion-atom interchange process seems more likely. BATES and NICOLET (1960) consider a number of such processes that fall under the general scheme:



Equations (2.11) and (2.12) will represent these reactions as well if the rate constants are taken to refer to the reactions (2.13)

and (2.14).

2.3.4 More than one species of recombining ions.

The experimental determinations of the effective recombination coefficients have yielded widely varying results. It has become clear that determinations of α from eclipse measurements, nocturnal decay of ionization and the "sluggishness" of the ionosphere do not agree. McELHINNY (1959) and MINNIS (1958) have discussed the two-ion hypothesis considered by BATES and McDOWELL (1957) in an attempt to explain eclipse results on the basis of the presence of two kinds of positive ions. BOWHILL (1961) has extended this approach to an investigation in detail of the form of the recombination when several species of positive ions are present. He claims that this model explains the discrepancies observed quite adequately, and that the results are in agreement with rocket measurements of ionospheric composition.

2.4 Layer formation.

2.4.1 Chapman theory.

CHAPMAN (1931a) assumed that the layers of electrons were formed as a result of equilibrium between the production of electrons by solar radiation and recombination loss proportional to N^2 and independent of height. Each observed layer (including the F1 ledge, which, at the time, was thought to be an independent layer) was supposed to be due to an independent production peak of ionizing radiation acting on a different ionizable constituent. The resulting equation of continuity takes the form

$$\frac{dN}{dt} = q - \alpha N^2 \quad (2.15)$$

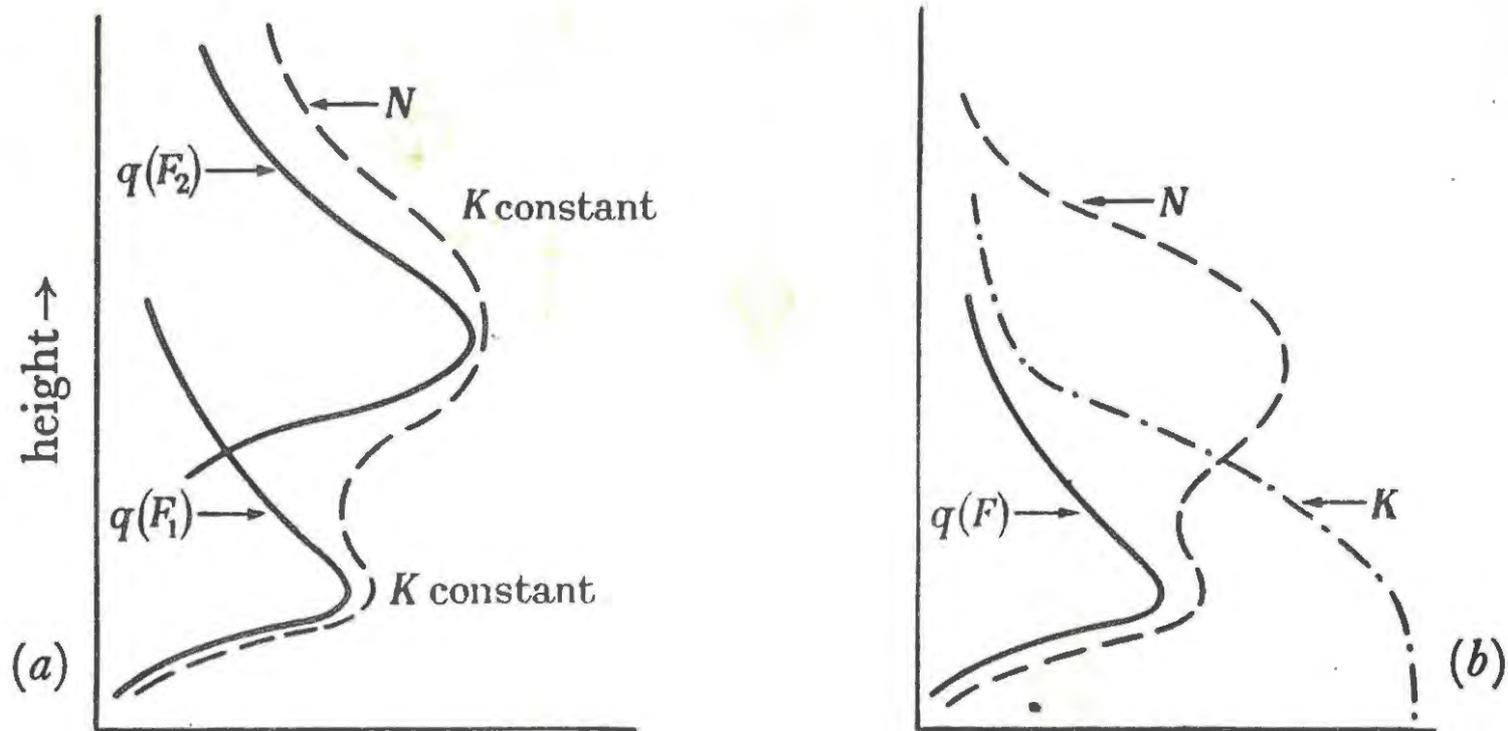
where α is independent of height and q is given by equation (2.3).

Although equation (2.15) does not admit of an explicit solution, a numerical integration showed that this theory was remarkably successful in explaining many of the observed features of the E layer and F1 ledge. Its success in explaining the observed variations in the F2 layer was much less encouraging.

The layers that result from the conditions described have a distinct shape and are called Chapman-alpha layers (RATCLIFFE and WEEKES, 1960). They are illustrated in Figure 2.1.(a).

2.4.2 The Bradbury hypothesis.

Another possibility for the formation of the two F layers during daytime was advanced by BRADBURY (1938). On this hypothesis the same incident radiation acting on the same ionizable constituent would be responsible for both the F1 ledge and the F2 layer. A separate F2 layer would then be formed through the rapid decrease in the effective recombination coefficient with increase in height. A small value of the effective recombination coefficient would require a large value of the electron density for the loss term to become equal to the production term, a condition required for equilibrium. Such a small value of α therefore results in an accumulation of electrons at high levels to form the F2 layer. This hypothesis is illustrated in Figure 2.1.(b).



electron density, N ; loss coefficient, K and rate of production, q

(a) Chapman (b) Bradbury

FIGURE 2.1 To illustrate the 'Chapman' and the 'Bradbury' hypotheses for the formation of the F_2 layer.

We note that if the effective recombination coefficient decreased indefinitely with height faster than the rate at which the production function decreased, the resulting electron density would increase indefinitely. This is a consequence of the fact that the theory ignores the possibility of the ionizable constituent being depleted.

However, another factor enters to prevent such an indefinite increase and this factor is considered in the next section.

2.5 The effect of movements.

2.5.1 The continuity equation.

On taking account of the mean drift velocity \bar{v} , the continuity equation for the electron density takes the form

$$\frac{dN}{dt} = q - L - \text{div}(N\bar{v}) \quad (2.16)$$

where L represents the loss term and the divergence term the effect of movements. If we consider only the vertical component of \bar{v} because vertical movements are usually thought to be the most important, equation (2.16) takes the form

$$\frac{dN}{dt} = q - L - \frac{d}{dh} (Nw) \quad (2.17)$$

where w is the vertical component of \bar{v} .

The vertical drift velocity could be due to temperature differences, diffusion or electromagnetic forces. It appears that movements due to electromagnetic forces are most important in the E region and that those due to diffusion are most important in the

F region. We shall concern ourselves solely with the effects to be expected as a result of diffusion.

2.5.2 Diffusion in the F2 region.

FERRARO (1945) considered what effects diffusion would have on the electron density distribution in the ionosphere. The electrostatic attraction between the electrons and positive ions prevents any appreciable separation between them and therefore no subsequent space charge results. The "ambi-polar" diffusion under gravity that ensues is accounted for in the following equation derived by Ferraro:

$$\frac{\partial N}{\partial t} = q - \alpha N^2 + D \left(\frac{\partial^2 N}{\partial h^2} + \frac{3}{2H} \frac{\partial N}{\partial h} + \frac{N}{2H^2} \right) \quad (2.18)$$

In this equation the diffusion coefficient D is inversely proportional to the concentration of the neutral molecules n and directly proportional to $\sin^2 I'$ where I' is the angle of inclination (the "dip" angle) of the earth's magnetic field at the point under consideration. The effects of diffusion would therefore be expected to be small near the equator where I' is small and to increase with height as n decreases. Ferraro solved equation (2.18) by quadratures and came to the conclusion that the effect of diffusion in the E and F1 regions of the ionosphere is negligible, but that it may be appreciable though small in the F2 region. At most the effect of diffusion would be to lower the F2 layer as a whole by a quarter of a scale height.

2.6 Layer formation with diffusion taken into account.

2.6.1 Attachment-like loss operative in the F2 region.

Recognizing the fact that movements might have important consequences on the observed distribution of electron densities in the F2 region, RATCLIFFE et al (1956) determined the most likely ways in which electrons are lost in this region. Their results could be explained more satisfactorily if a loss law of the attachment type was operative, the loss coefficient decreasing with height. They suggested the following expression for the way in which the loss coefficient K varies:

$$K = 10^{-4} \exp\left(\frac{300 - h}{50}\right) \text{ per sec.} \quad (2.19)$$

Their results thus confirm Bradbury's hypothesis for the formation of the F layers.

At the end of section 2.4.2 we noted that the electron density might increase indefinitely with increase in height if the loss rate decreases faster than the production rate. It turns out that diffusion under gravity prevents this from happening.

2.6.2 The F2 peak formed by diffusion.

As it became clear that the loss law operating in the F2 region is attachment-like, FERRARO and ÖZDOĞAN (1958) solved equation (2.18) with the term αN^2 replaced by KN where K is the loss coefficient, equal to 10^{-4} sec^{-1} and independent of height. The results obtained predicted changes in the height and peak density of the F2 layer which were at variance with the observed behaviour of the F2

layer. Their choice of a loss coefficient independent of height apparently vitiated the result to a certain extent. Later rocket results also showed that the neutral particle densities assumed by them were too low by a factor of 10. YONEZAWA (1956), however, showed that when electrons are continuously produced and lost as on the Bradbury hypothesis, the effect of ambi-polar diffusion is to produce a peak in the electron density near a level where q , the loss rate and D are all of the same order.

A more general equation than the one derived by Ferraro was obtained by SHIMAZAKI (1957). He showed that w_D , the vertical component of the drift velocity due to diffusion, is given by

$$-w_D = D(h) \left[\frac{1}{N} \frac{\partial N}{\partial h} + \frac{1}{T} \frac{\partial T}{\partial h} + \frac{\psi}{H} \right] \quad (2.20)$$

where ψ is the ratio of the mean molecular weight of the plasma to that of the neutral atmosphere, H the scale height of the atmosphere and T the temperature. The continuity equation (2.17) then becomes

$$\frac{\partial N}{\partial t} = q - KN + \frac{\partial}{\partial h} \cdot N \cdot D \left[\frac{1}{N} \frac{\partial N}{\partial h} + \frac{1}{T} \frac{\partial T}{\partial h} + \frac{\psi}{H} \right] \quad (2.21)$$

where K is given by equation (2.19) and is height dependent.

Three different types of solution to this equation will be considered here.

RISHBETH and BARRON (1960) obtained a solution for the special case of an isothermal atmosphere with $\psi = \frac{1}{2}$ under equilibrium conditions i.e. with $\frac{\partial N}{\partial t} = 0$. They came to the conclusion that the

maximum electron density, and the height at which it occurs are such that the magnitudes of the production, loss and diffusion processes are approximately equal, and that their results are not materially altered if the scale height of the atmosphere varies with height.

The second solution is due to GLIDDON (1959) who obtained an analytical solution to equation (2.21) for the special case of an isothermal atmosphere in which $\psi = \frac{1}{2}$. The solution was used by GLIDDON and KENDALL (1960) to obtain quantitative results for the variation of electron density in the F2 layer. Their model exhibited many of the features observed in the actual F2 layer behaviour. Thus their results give quantitative support to the Bradbury hypothesis about the F2 layer formation.

BRIGGS and RISHBETH (1961) proposed an analogue computer to solve the general equation (2.21). The analogue demonstrated many features observed in the ionosphere. In particular, it exhibited the same type of diurnal variation in the height and maximum electron density of the F2 layer and simulated the behaviour of the F2 region during eclipses and night-time reasonably faithfully.

From the above and from the work of others, it therefore becomes clear that diffusion is the mechanism responsible for the formation of the peak in the electron density in the F2 layer. Because this overcomes the difficulty of a possible indefinite increase in N with height, it gives still more support to the Bradbury hypothesis.

2.7 A transition region present in the F layer.

If we accept the fact that most of the F1 ledge effects can be explained on the assumption of a quadratic loss law, and the results of RATCLIFFE et al (1956) which imply that in the F2 region the loss process is better represented by a term that contains only the first power of the electron density, we must investigate the possibility of a gradual changeover from a quadratic to a linear loss law as we ascend through the F region. This will be our concern in the next chapter.

CHAPTER 3.THE F1 LAYER.3.1 Two models.

The experimental results concerning the F1 layer support the assumption that a recombination type of loss operates in this region (RATCLIFFE and WEEKES, 1960). Just above the F1 ledge an attachment type of loss law must take over if the observations (RATCLIFFE et al, 1956) on the F2 layer are to be adequately explained. MARIANI (1958), and HIRSH (1959), (YONEZAWA, TAKAHASHI and ARIMA (1959)) have analysed this problem.

3.2 The model used by Mariani.

As a basis of his model Mariani assumed an expression for the loss coefficient first suggested "very tentatively" by BURKARD (1956) and used by Mariani in the form

$$\alpha = \alpha_0 n_e (N.T)^{-1} \quad (3.1)$$

where α_0 is a constant, T denotes the temperature, and n the particle concentration.

The bifurcation of the F layer is supposed to arise as a consequence of discontinuity in the physical conditions in the upper atmosphere. The expression has some properties to be expected for a loss coefficient in the F region such as an inverse dependence on the electron density and a dependence on the particle concentration. If the concentration of the particles concerned decreases with altitude, the loss coefficient would also decrease, thus implying the acceptance

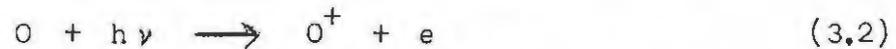
of the Bradbury hypothesis. The main disadvantage of this model, as admitted by Burkard, is that a knowledge of possible recombination processes to which this coefficient applies is not available.

3.3 The model used by Hirsh.

In his analysis of the possible origin of the F1 layer, Hirsh used the scheme for the loss of electrons first proposed by Bates and Massey and described in section 2.3.3. Let us consider the consequences of such a process in more detail.

3.3.1 Identification of the species of atoms and molecules involved.

It has been shown, mostly through the work of BATES and NICOLET (BATES, 1955; NICOLET, 1960; BATES and NICOLET, 1960) that the atomic species to be identified with atom A in reaction (2.8) is oxygen. We may therefore write with reasonable certainty



for the production process in the F region. NICOLET and MANGE (1954) have shown that molecular oxygen is not completely absent in the higher F region even though the radiation responsible for the dissociation of molecular oxygen penetrates to the E region. They attribute the presence of oxygen at these levels to the effects of complete mixing of the atmosphere, since upward diffusion alone cannot account for the observed concentrations. The presence of O_2 offers a possibility for the identification of the molecular species XY involved in the loss process described in (2.13) and (2.14) as





Another possible reaction scheme considered by Bates and Nicolet is the following:-



Measurements carried out by JOHNSON et al (1958) with mass spectrometers mounted on rockets have shown that all the ions involved in both these processes are present in the ionosphere. To keep a certain amount of generality we shall therefore consider the reaction scheme:-



with respective rate constants λ_1 and λ_2 .

3.3.2 Transition from a linear to a quadratic loss law.

On neglecting the effect of movements the continuity equation corresponding to (3.7) is

$$\frac{d}{dt} n(O^+) = q - \lambda_1 n(O^+) \cdot n(XY) \quad (3.9)$$

and that corresponding to (3.8)

$$\frac{d}{dt} N = q - \lambda_2 n(OY^+) \cdot N \quad (3.10)$$

We may also assume that the ionosphere is electrically neutral

since the electrostatic forces between ions and electrons will prevent any appreciable separation. We therefore have

$$N = n(OY^+) + n(O^+) \quad (3.11)$$

When equilibrium has set in

$$\frac{d}{dt} N = 0 = \frac{d}{dt} n(O^+) \quad (3.12)$$

Under these conditions

$$\begin{aligned} q &= \lambda_1 n(O^+) \cdot n(XY) \\ &= \lambda_2 n(OY^+) \cdot N \end{aligned}$$

On eliminating $n(O^+)$ and $n(OY^+)$ through use of (3.11) we obtain

$$q = \frac{\lambda_1 \lambda_2 n(XY) \cdot N^2}{\lambda_1 n(XY) + \lambda_2 N} \quad (3.13)$$

If $\lambda_1 n(XY) \gg \lambda_2 N$ as it is at low levels in the ionosphere where $n(XY)$ is thought to be large and N not so large, equation (3.13) reduces to

$$q = \lambda_2 \cdot N^2 \quad (3.14)$$

which would represent an equilibrium layer in which a quadratic loss law is operative with an effective recombination coefficient $\alpha = \lambda_2$.

If $\lambda_1 n(XY) \ll \lambda_2 N$ at higher levels then equation (3.13) becomes effectively

$$q = \lambda_1 N \cdot n(XY) \quad (3.15)$$

which would apply to an equilibrium layer in which the electron loss is proportional to the first power of the electron density.

In an ionosphere that changes in shape with time N obviously is a function of both height and time. Since λ_1 , λ_2 and $n(XY)$ are all taken to be time independent this means that the inequalities from which (3.14) and (3.15) result would also be functions of both height and time, a fact ignored by HIRSH (1959) and not mentioned by RATCLIFFE (1959). Some of the consequences to be expected when this fact is not overlooked will emerge in later chapters of this thesis.

3.3.3 The effective attachment coefficient a function of height.

The value of $n(XY)$ is taken to be very large in comparison with N at all levels so that any changes in it due to the ion-atom interchange process mentioned may be neglected. This means that we may assume $n(XY)$ to be a function of height only. Equation (3.15) may then be written

$$q = \gamma \cdot N \quad (3.16)$$

where γ is an effective attachment coefficient given by

$$\gamma = \lambda_1 \cdot n(XY) \quad (3.17)$$

and is a function of height. (λ_1 , it will be remembered, is the rate coefficient for the ion-atom interchange process).

The concentration of the molecular species XY is usually assumed to decrease exponentially with height, so that we may further write

$$n(XY) = n_0(XY) \cdot \exp\left(-\frac{h - h'_0}{H_Y}\right) \quad (3.18)$$

wherein $n_0(XY)$ is the value of $n(XY)$ at some reference level h'_0 and H_Y is the scale height of the XY molecules. Writing $k = H_1/H_Y$ for the ratio of the scale height H_1 of the ionizable constituent to the scale height of the XY molecules, we obtain by substitution for H_Y in equation (3.18) the following

$$n(XY) = n_0(XY) \cdot \exp(-kz) \quad (3.19)$$

if h'_0 is taken to refer to the same level as h_0 , the reference level of the production function (2.3). Substituting for $n(XY)$ from (3.19) into (3.17) we obtain

$$\gamma = \lambda_1 n_0(XY) \cdot \exp(-kz)$$

which may be written

$$\gamma = \gamma_0 \cdot \exp(-kz) \quad (3.20)$$

where $\gamma_0 = \lambda_1 n_0(XY)$ is the value of γ at $z = 0$ and is a constant.

3.3.4 Mixing and diffusive equilibrium.

In his presidential address to the Physical Society RATCLIFFE (1959) discussed some of the consequences to be expected if the loss process mentioned in section 3.3.1 is the one taking place in the F region of the ionosphere.

Substituting from equations (3.20) and (2.3) into (3.16) we obtain

$$N = \frac{q_0}{\gamma_0} \cdot \exp\left[1 - z(1 - k) - \exp(-z) \cdot CH(\chi)\right] \quad (3.21)$$

If the atmospheric gases are completely mixed, they will be distributed with the same scale height and k will be unity. Equation (3.21) shows that under equilibrium conditions the electron density will then approach an asymptotic value with increase in altitude. This result is due to the neglect of diffusion and of the depletion of the ionizable constituent by the ionizing radiation.

In the case where the gases are separated by diffusion and have come to diffusive equilibrium, each gas will be distributed with its own scale height. We can then obtain a value for k for an isothermal atmosphere from

$$k = \frac{H_i}{H_y} = \frac{m_y g}{kT} \cdot \frac{kT}{m_i g} = \frac{m_y}{m_i} \quad (3.22)$$

where m_y is the mean molecular weight of the XY molecules and m_i that of the ionizable constituent. If atomic oxygen is the ionizable constituent and molecular oxygen the substance XY, k will be 2. If the XY molecules are to be identified with N_2 , k will be 1.75.

By an examination of the observed electron density distributions Ratcliffe showed that many of the observed features of the F2 layer could be explained on the assumption that the degree of mixing is variable. Considering values of k between 1.1 and 1.5 he came to the conclusion that such a variable degree of mixing could account for the observed increase in the F2 electron density peak in middle latitudes during the winter when less mixing would cause k to increase.

3.3.5 The F1 layer on this model.

Hirsh based his model of the F1 ledge on the loss process employed by Ratcliffe in his discussion and described in section 3.3.1.

On substituting from equations (2.3) and (3.20) into equation (3.13) we obtain

$$N = \frac{q_0 Q(z, \chi)}{2 \gamma_0 \exp(-kz)} \left[1 + \sqrt{1 + C \frac{\exp(-2kz)}{Q(z, \chi)}} \right] \quad (3.23)$$

where $Q(z, \chi) = \exp [1 - z - \exp(-z) \sec \chi]$

$$\text{and } C = \frac{4\gamma^2}{q_0 \alpha} \quad (3.24)$$

This is the equation that was solved by Hirsh for different values of C and k . The resulting $N(h)$ profiles lead to the type of ionograms observed only if the values chosen for C lie above certain limits. In particular, if $k = 2$, C has to be greater than 600 when $\chi = 0$ for a peak to appear in the $N(h)$ curve at the "F1 critical frequency". Large values of C imply large values of the effective attachment coefficient through equation (3.24) and in fact the values of γ_0 deduced by Hirsh are

$$5.6 \times 10^{-2} \text{ sec.}^{-1} > \gamma_0 > 3.5 \times 10^{-3} \text{ sec.}^{-1}.$$

The expression

$$\gamma = 10^{-4} \cdot \exp\left(\frac{300 - h}{45}\right) \quad (3.25)$$

arrived at by RATCLIFFE et al. (1956) (hereafter referred to as

R.S.S.T.) yields a value of $\gamma_0 = 1.5 \times 10^{-3} \text{ sec.}^{-1}$ at 180 km (where $h = h_0$). This value is lower than any of the possibilities considered by Hirsh who could suggest no explanation for the discrepancy.

The disagreement may be due to:

(a) The extrapolation of equation (3.25) below 250 km, since it was derived from data taken between 250 and 350 km.

(b) The assumptions made by Hirsh in arriving at his estimate of γ_0 , viz.

$$(i) \quad \frac{dN}{dt} = 0$$

$$(ii) \quad \alpha = 5 \times 10^{-9} \text{ cm.}^3 \text{ sec.}^{-1}$$

and

$$(iii) \quad N_m = N_0 (\cos \chi)^{\frac{1}{2}}$$

where N_m is the peak electron density in the F1 ledge and N_0 is the value of N_m when $\chi = 0^\circ$. This last assumption applies strictly to a layer in which only a quadratic loss law is operative.

In view of the fact that the extrapolation and the assumptions made may be unjustifiable, the discrepancy cannot be considered serious. In fact, by abandoning the equilibrium assumption in Hirsh's model it is shown in this thesis that such a model can explain many features of the F1 layer. Such a non-equilibrium model is considered in the next chapter.

CHAPTER 4.A NON-EQUILIBRIUM MODEL OF THE F1 LAYER.4.1 Introduction.

The assumption of equilibrium

$$\frac{dN}{dt} = \frac{dn(O^+)}{dt} = 0$$

made by Hirsch in his model of the F1 layer is clearly not very realistic for times near sunrise and during eclipses. In our model of the F1 layer, to be presented here, we have therefore abandoned this assumption.

As our point of departure we take the reaction scheme proposed by Bates and Massey and subsequently modified by Bates and others. It is described in sections 2.3.3 and 3.3.1. The continuity equations which result are given in section 3.3.2 and restated here as

$$\frac{d}{dt} n(O^+) = q - \gamma n(O^+) \quad (4.1)$$

and

$$\frac{d}{dt} N = q - \alpha N \cdot n(OY^+) \quad (4.2)$$

where $\gamma = \lambda_1 \cdot n(XY)$ is the ion-atom interchange rate coefficient, also called the effective attachment coefficient and α is the effective recombination coefficient.

It is proposed to integrate equation (4.1) by numerical

methods at constant height levels with suitable values assumed for q and γ . From the values of $n(O^+)$ thus obtained figures for the electron density can then be found through a similar integration of equation (4.2) with the aid of the equation of charge neutrality

$$N = n(O^+) + n(OY^+) \quad (4.3)$$

The integrations are to be done at height intervals of half a scale height and time intervals of five minutes from before dawn, when we assume $n(O^+) = N = n(OY^+) = 0$, up to local noon. Below a level of about 300 km. the neglect of any residual ionization produced during the previous day is well justified (LONG, 1962) because at these levels there are only small concentrations of electrons present at sunrise. Above the 300 km. level the neglect is not so readily justifiable, but the error introduced is relatively small due to the rapid build-up of ionization after sunrise.

Because an investigation of the behaviour of the model under eclipse conditions was intended and the values of the obscuration factor at Grahamstown for the solar eclipse of 25 December 1954 were available, the equations were solved with values of q appropriate to that date.

We shall consider the values to be adopted for the factors q , γ and α and the solution of the equations of continuity in the rest of this chapter.

4.2 The production function.

Values of the production term were calculated from equation (2.3), which can be written in the form

$$q = \frac{q_0 \cdot \exp(1 - z)}{\exp[-\exp(-z) \cdot \text{Ch}(\chi)]} \quad (4.4)$$

q_0 was taken as 400 electrons per cm^3 per sec., a value appropriate to the year 1954 (SEITY, 1960). The height h_0 is taken as 170 km., (R.S.S.T. 1956; WATANABE and HINTEREGGER, 1962). Values of Chapman's grazing incidence integral $\text{Ch}(\chi)$ was obtained from a table published by WILKES (1953). Two quantities are needed for entry into this table.

- (a) The value of x , the distance, measured in terms of the scale height, from the centre of the earth to the level under consideration, and
- (b) χ , which is obtained from the equation

$$\cos \chi = \sin \delta \cdot \cos \theta + \cos \delta \cdot \sin \theta \cdot \cos \phi \quad (4.5)$$

where δ is the declination of the sun; θ is the latitude, and ϕ is the sun's hour angle.

A value of $x = 200$ was used. This corresponds to a height of 230 km. and a scale height of 33 km., values appropriate to the region just above the F1 ledge. This choice of x was convenient because $\text{Ch}(\chi)$ is tabled only at intervals of $x = 50$. Values of χ were obtained for Grahamstown (latitude $33^{\circ}18'S$, longitude $26^{\circ}32'E$) at five minute intervals of South African Standard Time

(U.T. + 2 hours) for 25 December.

4.3 The recombination coefficient.

No agreement has yet been reached on the value of the effective recombination coefficient α appropriate to the F1 region. Values from 10^{-9} to 10^{-7} cm.³ sec.⁻¹ have been suggested. As a more or less representative value we have assumed α to be 10^{-8} cm.³ sec.⁻¹ (KING and LAWDEN, 1962).

4.4 The ratio of the scale heights.

The ratio of the scale heights k was discussed in section 3.3.4. The arguments advanced by Ratcliffe for values of k between 1.1 and 1.5 are well substantiated and if we accept these, we have to conclude that nearly complete mixing takes place even in the F2 region of the ionosphere.

From the work done by Hirsh it is, however, evident that small values of k will necessitate large values of the constant $C = \frac{4\gamma_0^2}{q_0\alpha}$ for the F1 ledge to appear in the $N(h)$ curve. This in turn implies large values of γ_0 and so aggravates the discrepancy, mentioned in section 3.3.5., between the values for γ_0 obtained by Hirsh and those by R.S.S.T.

It should be kept in mind that Ratcliffe was mostly concerned with explaining the behaviour of the F2 layer in his arguments for k between 1.1 and 1.5. Hirsh, on the other hand, specifically investigated F1 ledge effects. He also showed that a large value

of k accentuates the F1 ledge, thus causing its behaviour to be more easily discernible.

In the present model we have taken $k = 1.8$. This value will lead to a distinct F1 ledge, but will also allow for a considerable degree of diffusive separation. It underestimates complete diffusion if the substance XY is O_2 but slightly overestimates it if XY is to be identified with N_2 .

4.5 The ion-atom interchange rate coefficient.

The most reliable value of γ so far determined is that by R.S.S.T. and is given by

$$\gamma = 10^{-4} \exp\left(\frac{300 - h}{50}\right) \text{ sec.}^{-1} \quad (4.6)$$

for the height interval 250 km. to 350 km. Arguments, based on other experimental observations, have been advanced for both higher (VANZANDI et al., 1960) and lower (YONEZAWA and TAKAHASHI, 1960) values of the coefficient. Most of these determinations are based on data from levels well above the F1 ledge. At these heights the R.S.S.T. figure of $10^{-4} \text{ sec.}^{-1}$ for γ at 300 km. appears to be a good average.

In our model we have normalised the function

$$\gamma = \gamma_0 \exp(-kz) \quad (4.7)$$

to assume a value close to $10^{-4} \text{ sec.}^{-1}$ at 300 km. With $k = 1.8$ this yields a value of

$$\gamma_0 = 10^{-2} \text{ sec.}^{-1}$$

at 170 km, which is larger than the value obtained by a direct extrapolation of (4.6) to 170 km. It does, however, lie in the range established by Hirsh. It also agrees with the experimental values of γ_0 of KING and LAWREN (1962) who obtained an average value of $\gamma_0 = 10^{-2} \text{ sec.}^{-1}$ in determinations that were carried out on the basis of data obtained from the F1 region.

With the values of q_0 , α and γ_0 chosen here the constant

$$C = \frac{4 \gamma_0^2}{q_0 \alpha} = 100.$$

4.6 The solution of equation (4.1).

The complicated nature of the function q makes the direct integration of equation (4.1) difficult. A step by step integration, assuming an initial value of $n(O^+)$ is, however, easily tractable. On this method the value of $n(O^+)$ at the end of an interval is used in the calculation of the loss term $\gamma n(O^+)$ during the next interval. This means that a small change in $n(O^+)$ during the interval is implicitly assumed. At low levels, where γ is large, this turned out to be an untenable assumption, because the solution started behaving in an oscillatory manner. As an illustration of this we show here a sample computation for the level $z = -0.5$ where $\gamma = 246 \times 10^{-4} \text{ sec.}^{-1}$. At 0525 S.A.S.T. the O^+ concentration is taken as zero. At the end of a five minute interval the value is $0.000001682 \text{ ions cm.}^{-3}$, which is the value of $n(O^+)$ at the start of the next interval.

TABLE 4.1 Sample computation of $n(O^+)$.

SAST	$n(O^+) \times 10^{-4}$	q	$\gamma n(O^+)$	$q - \gamma n(O^+)$	$\Delta n(O^+)$
	ions cm^{-3}	ions cm^{-3}	ions cm^{-3}	ions cm^{-3}	ions cm^{-3}
		sec^{-1}	sec^{-1}	sec^{-1}	sec^{-1}
0525	0.0	0.00005608	0.0	0.00005608	0.000001682
0530	0.000001682	0.0004446	0.0004138	0.0000308	0.000000924
0535	0.000002606	0.002338	0.000641	0.001697	0.00005091
0540	0.00005352	0.009329	0.013166	-0.003837	-0.00001151
0545	0.00004201	0.02926	0.01033	0.01893	0.0005678
0550	0.0006098	0.07748	0.15001	-0.07253	-0.002176
0555	-0.001566	..			

To overcome this difficulty equation (4.1) was solved for each time interval on the assumption that q varies linearly with time. A computed value of q was available for five minute intervals, so that an assumed linear change between calculated values would not deviate much from the correct ones. The change in q was assumed to be given by

$$q = D + B.t \quad (4.8)$$

during the n th interval with $D = q_{n-1}$ the value of q at the beginning of the interval so that

$$B = \frac{q_n - q_{n-1}}{T} \quad (4.9)$$

for intervals of T seconds, when q_n is the value of q at the end of the interval. On substituting into equation (4.1) for q we obtain

$$\frac{d}{dt} n(O^+) = D + Bt - \gamma n(O^+) \quad (4.10)$$

This equation can be written in the form

$$y' = D + Bt - by \quad (4.11)$$

where $y = n(O^+)$ and $b = \gamma$, and the prime denotes the first derivative with respect to time.

Equation (4.11) is an exact equation of the first order and first degree and may be solved by first multiplying through by an integrating factor $\exp(bt)$ and regrouping to obtain

$$\exp(bt).dy + \exp(bt).(by - D - Bt).dt = 0 \quad (4.12)$$

On integrating we obtain the solution

$$\exp(bt).(y - \frac{D}{b} - \frac{Bt}{b} + \frac{B}{b^2}) = \text{constant} \quad (4.13)$$

which may be written in the form

$$y - \frac{D}{b} - \frac{Bt}{b} + \frac{B}{b^2} = C'.\exp(-bt)$$

On substituting back for b and y we obtain

$$n(O^+) = \frac{a}{\gamma} - \frac{B}{\gamma^2} + C'.\exp(-\gamma t) \quad (4.14)$$

for the number density of atomic oxygen ions during the n th interval.

Before equation (4.14) can be used in a computation it is necessary to fix the value of the constant of integration C' . Seeing that the final solution of equation (4.1) is done over

5 minute intervals C' can be found in terms of the value of $n(O^+)$ at the beginning of the interval under consideration. Let $n(O^+)_p$ denote the value of $n(O^+)$ at the end of the interval and $n(O^+)_{p-1}$ its value at the beginning, so that we may write

At time $t = 0$

$$\begin{aligned} n(O^+)_{n-1} &= \frac{q_{n-1}}{\gamma} - \frac{B}{\gamma^2} + C' \cdot \exp(-\gamma \cdot 0) \\ &= \frac{q_{n-1}}{\gamma} - \frac{B}{\gamma^2} + C' \end{aligned}$$

and therefore $C' = n(O^+)_{n-1} - \frac{q_{n-1}}{\gamma} + \frac{B}{\gamma^2}$

At time $t = T$

$$\begin{aligned} n(O^+)_n &= \frac{q_n}{\gamma} - \frac{B}{\gamma^2} + C' \cdot \exp(-\gamma \cdot T) \\ &= \frac{q_n}{\gamma} - \frac{B}{\gamma^2} + \exp(-\gamma \cdot T) \left\{ n(O^+)_{n-1} - \frac{q_{n-1}}{\gamma} + \frac{B}{\gamma^2} \right\} \end{aligned}$$

On substituting for B and collecting terms we obtain

$$\begin{aligned} n(O^+)_n &= \frac{q_n}{\gamma} \left(1 - \frac{1}{\gamma T} \right) \cdot \left\{ 1 - \exp(-\gamma T) \right\} \\ &+ \frac{q_{n-1}}{\gamma^2 T} \cdot \left\{ 1 - \exp(-\gamma T) \right\} + n(O^+)_{n-1} \exp(-\gamma T). \quad (4.15) \end{aligned}$$

Equation (4.15) appears formidable, but since all the terms in the coefficients q_n and q_{n-1} reduce to single constants when constant time intervals are used and the integration takes place at a given height, it offers no great difficulty for computation. Assuming an initial value of zero for the oxygen ion density

before sunrise, the value of $n(O^+)$ was thus found at five minute intervals up to local noon.

4.7 The solution of equation (4.2).

The analytical solution of equation (4.2) is notoriously difficult. It is a form of the Ricatti differential equation which admits of solution only in particular cases, of which equation (3.2) unluckily does not appear to be one. We therefore have to fall back on a numerical procedure, the one selected being that due to Picard, a method described by REDISH (1961). Picard's method is essentially an iterative process of developing a series solution, and we shall merely outline it here.

$$\text{Let } y' = f(x, y)$$

$$\text{and } y = y_0 \text{ at } x = x_0$$

$$\text{Then } y = y_0 + \int_{x_0}^x f(x, y) dx$$

This equation is solved by iteration, with the successive approximations to y , $y^{(s+1)}$ given by

$$y^{(s+1)} = y_0 + \int_{x_0}^x f(x, y^{(s)}) dx.$$

Obtaining the first approximation to equation (4.2) by this method, we get

$$N^{(1)} = N_0 + \int_0^T \left[q - \alpha N_0 \cdot n(OY^+) \right] dt$$

which can be written as

$$N^{(1)} = N_0 + \left[\bar{q} - \alpha N_0 \cdot n(OY^+)_0 \right] \cdot T \quad (4.16)$$

for a period of T seconds.

In this equation $\bar{q} = \frac{q_{n-1} + q_n}{2}$ is the mean value of q for the interval and $n(OY^+)_0$ is the value of $n(OY^+)$ at the beginning of the interval, obtained from

$$n(OY^+)_c = N_0 - n(O^+)_c \quad (4.17)$$

where N_0 is the value of N and $n(O^+)_c$ is the value of $n(O^+)$ at that time.

This method assumes that N and $n(OY^+)$ during the interval can be adequately represented by N_0 and $n(OY^+)_0$, the constant initial values. The method should be well suited for solving the continuity equation near noon when the changes in electron and ion concentrations are small. Another time when the method should serve to obtain a reasonably accurate solution is just after sunrise when the loss term is very small compared with the production term. (We assume the production function is given accurately enough by the linear change in q between successive computed values). The largest error can therefore be expected when the loss term has become appreciable and the electron concentration is still changing rapidly.

As an example of the error to be expected we take $z = 0$ (where q changes most rapidly) and select the interval during

which the value of N changes most. This is between 0645 and 0650 S.A.S.T. when N changed from 4.905×10^4 to $5.774 \times 10^4 \text{ cm.}^{-3}$. During this interval the production of electrons amounted to $1.525 \times 10^4 \text{ cm.}^{-3}$ and the loss to $0.656 \times 10^4 \text{ cm.}^{-3}$. Had we taken N_0 as $5.774 \times 10^4 \text{ cm.}^{-3}$ the loss term would have amounted to 0.777 cm.^{-3} and would have resulted in a value of $5.648 \times 10^4 \text{ cm.}^{-3}$ for N at the end of the interval. The former value of N is an overestimate, the latter an underestimate. We can assume that the mean value would be near the correct one, so that the way in which the computation was performed leads to an overestimate of about 1%. It should also be noted that only the rate of increase in N is affected in this way, because the final value attained by N is determined by q and the loss coefficients which fix the equilibrium values of N , $n(O^+)$ and $n(OY^+)$ at levels that attain equilibrium.

The example shows that by using five minute intervals in the computation the error introduced is at most about 1% during the forenoon. During the late afternoon when the production term becomes small in comparison with the loss term, larger errors are to be expected. The method would then lead to a too rapid decrease in the electron density.

4.8 Procedure for computation.

The first step in the computation is to obtain successive values of q for a given height by the method described in section 4.2. With a value of γ appropriate to this height (section 4.5) equation (4.15) yields values of $n(O^+)$ at intervals of 5 minutes,

if an initial condition of

$$n(O^+) = N = n(OY^+) = 0$$

is assumed. Equation (4.18) then yields values of N at these intervals if use is made of the equation of charge neutrality (4.3).

Part of the computation for the level $z = 1.5$ is shown as an example in Table 4.2. Columns 2, 3 and 4 show electron and ion densities at the times given in column 1; column 5 shows the production rate. In columns 6, 7 and 8 $n(O^+)$ is computed; in 9, 10, 11 and 12 the electron density change is computed. By way of clarification a few further points may be noted:

- (a) The coefficients of q_n , q_{n-1} and $n(O^+)_0$ in columns 6, 7 and 8 are calculated from the expressions in equation (4.15).
- (b) The value of $n(O^+)$ at any time is found by adding the figures in column 6, 7 and 8 in the previous row.
- (c) $\bar{q} = 150q_n + 150q_{n-1}$ for a 300 second period.
- (d) N is the algebraic sum of columns 9, 10 and 11.
- (e) After 0510 S.A.S.T. the concentrations are given in numbers per cc $\times 10^{-4}$.

TABLE 4.2 Specimen calculation of electron and ion densities through the step by step solution of equations (4.1) and (4.2).

1	2	3	4	5	6
SAST	N	$n(O^+)$	$n(OY^+)$	q	$131.28 \times q_{n-1}$
	electrons cm^{-3} .	ions cm^{-3} .	ions cm^{-3} .	ions $cm^{-3} sec^{-1}$.	ions $cm^{-3}(5min)^{-1}$.
0445	0.0	0.0	0.0	0.0028	0.0
0450	0.424	0.397	0.027	0.0399	0.371
0455	6.838	6.302	0.536	0.2667	5.242
0500	52.83	47.84	4.99	1.0350	35.01
0505	248.1	219.4	28.7	2.830	135.9
0510	827.8	712.7	115.1	5.980	371.5
	10^{-4}	10^{-4}	10^{-4}		10^{-4}
0515	0.2149	0.1794	0.0355	10.60	0.07850
0520	0.4634	0.3739	0.0895	16.53	0.1392
0525	0.8692	0.6769	0.1923	23.42	0.2170
0530	1.5985	1.0991	0.4994	30.98	0.3074
0535	2.526	1.641	0.885	38.80	0.4067
0540	3.505	2.293	1.212	46.77	0.5094
0545	4.661	3.040	1.621	54.57	0.6140
0550	5.955	3.865	2.090	62.28	0.7164
0555	7.334	4.750	2.584	69.65	0.8176
0600	8.745	5.678	3.065	76.68	0.9144
0605
..

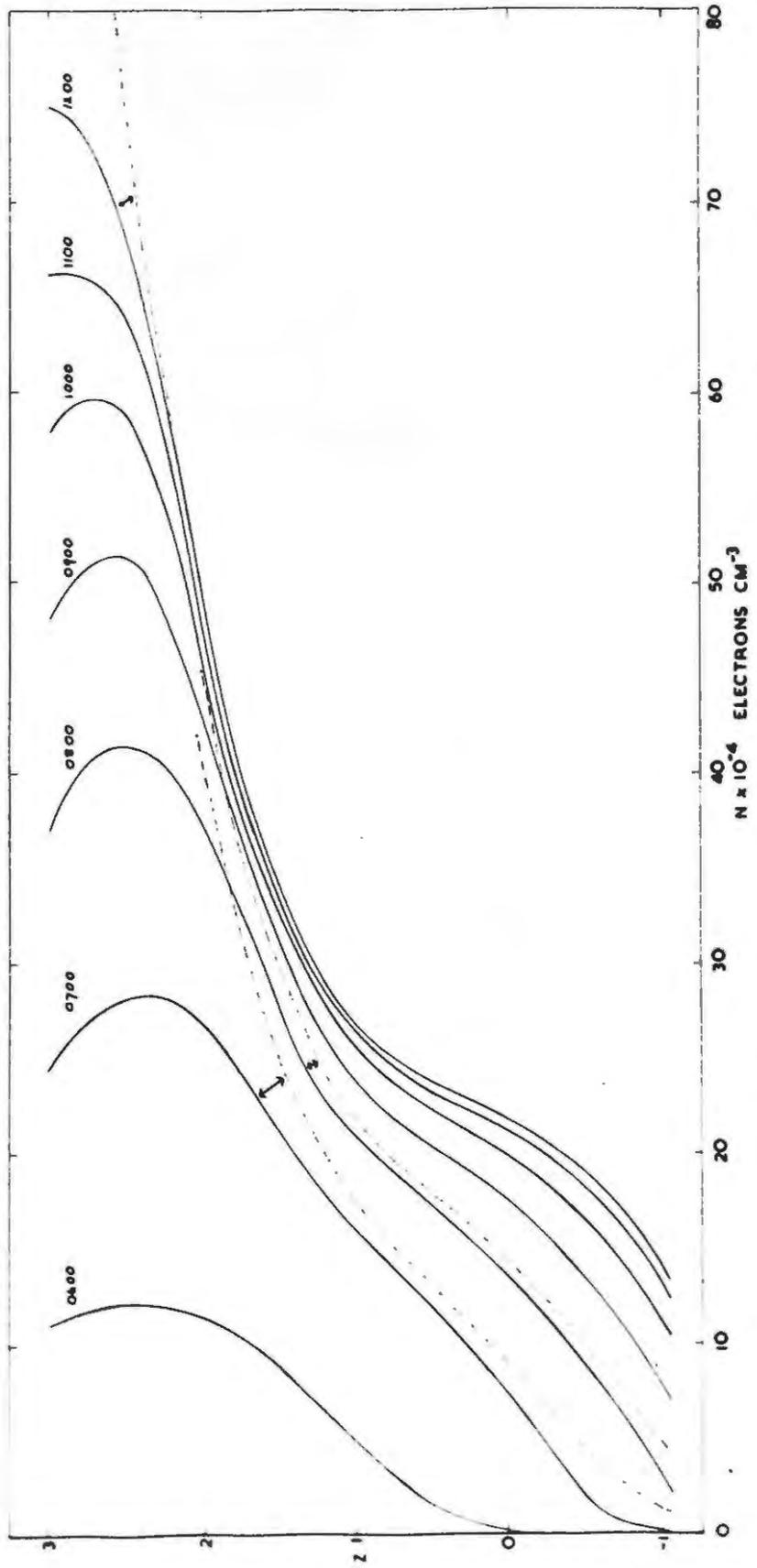


FIG 5.1 nA CURVES AT GIVEN TIMES (----EQUILIBRIUM CURVES).

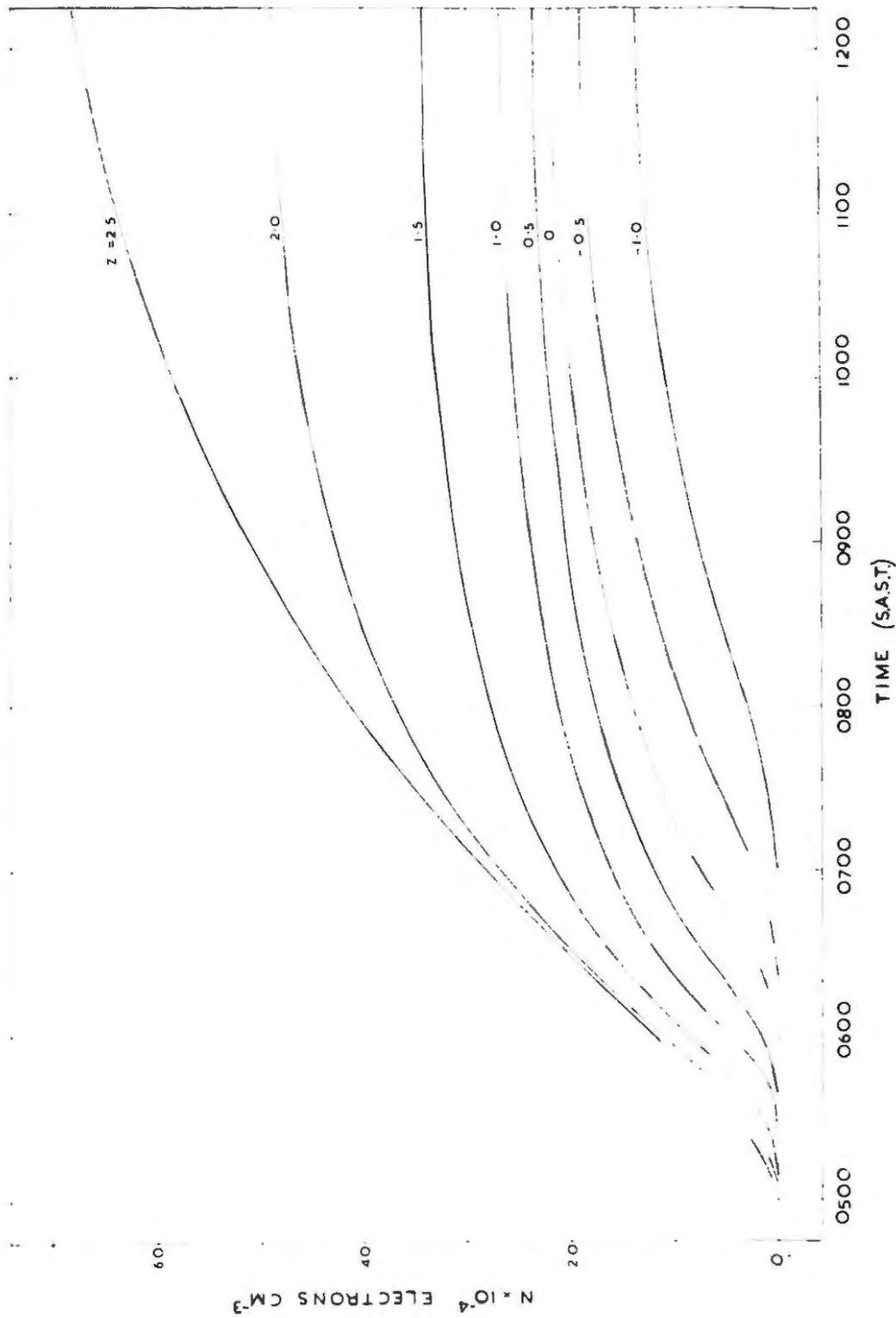


FIG52 N(t) CURVES AT CONSTANT HEIGHTS

intervals obtained by the methods described in Chapter 4. In Figure 5.2 the corresponding $N(t)$ curves are shown at height intervals of $z = \frac{1}{2}$ from $z = -1.0$ to $z = 2.5$. Also shown in Figure 5.1 are three equilibrium electron density distributions calculated by the method described by HIRSH (1959). These are shown by dashed lines for the times 0700, 0800 (not extended beyond the point) and 1200 hours and are connected by arrows to the corresponding non-equilibrium distributions at these times.

In these figures the bifurcation of the F layer during a normal morning is shown clearly, a fact to be expected of a model based on that proposed by Hirsh.

5.3 The F1 ledge.

Figure 5.1 shows that the F1 ledge on the present model is reasonably pronounced during the period near noon but not at all so early in the morning. Neither is the point to be identified with the F1 ledge maximum on the $N(h)$ curve clearly discernible. In investigating the variations in this maximum it is therefore desirable to construct ionograms to make the identification of f_oF1 possible. From this critical frequency the F1 maximum electron density and the height at which it occurs can then be determined. This was done by the method described in section 1.10. The ionograms for the ordinary wave in the presence of a magnetic field were thought to be sufficient to indicate the variations present in the model.

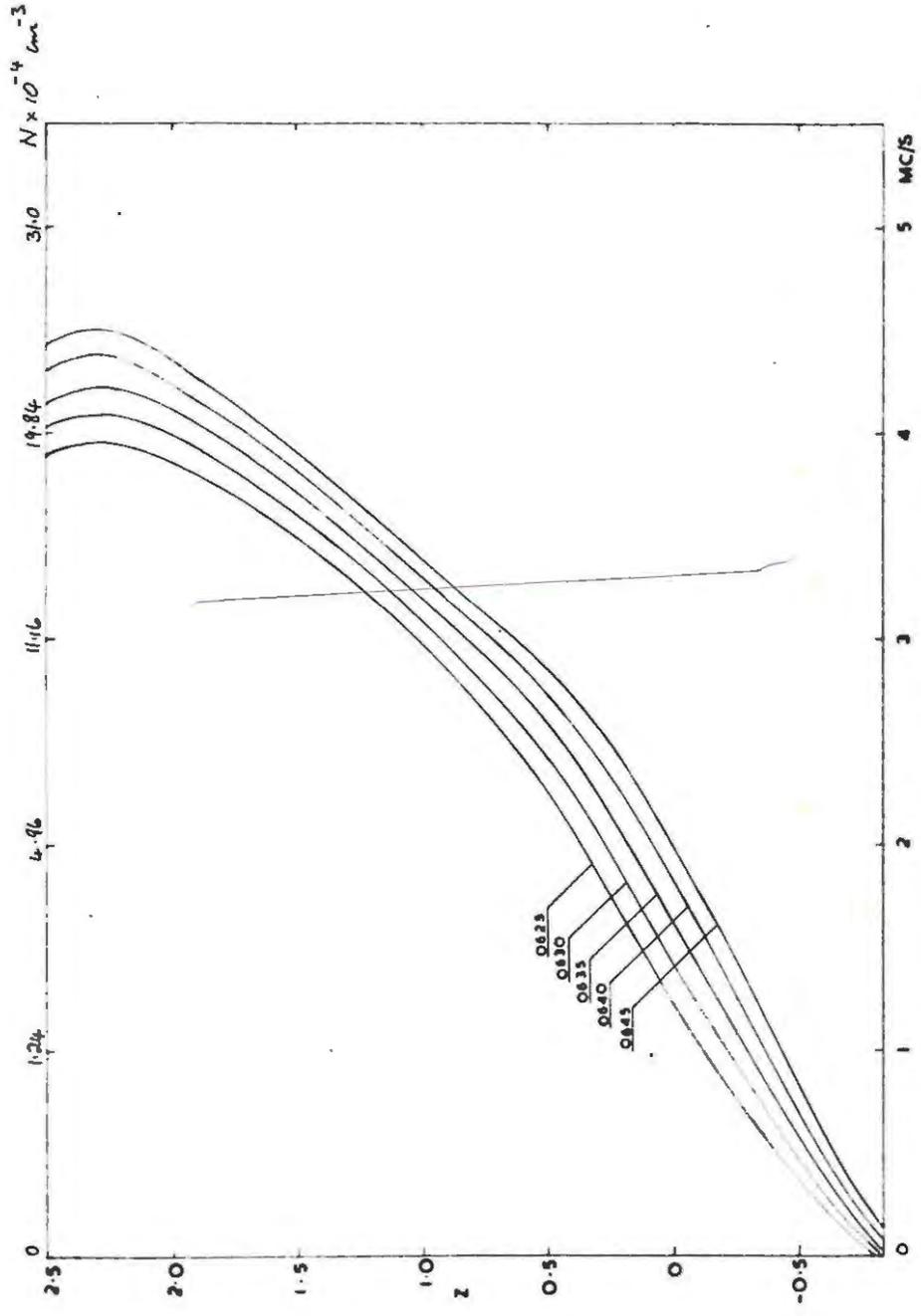


FIG. 3 $N(z)$ CURVES NEAR TIME OF INITIAL BIFURCATION

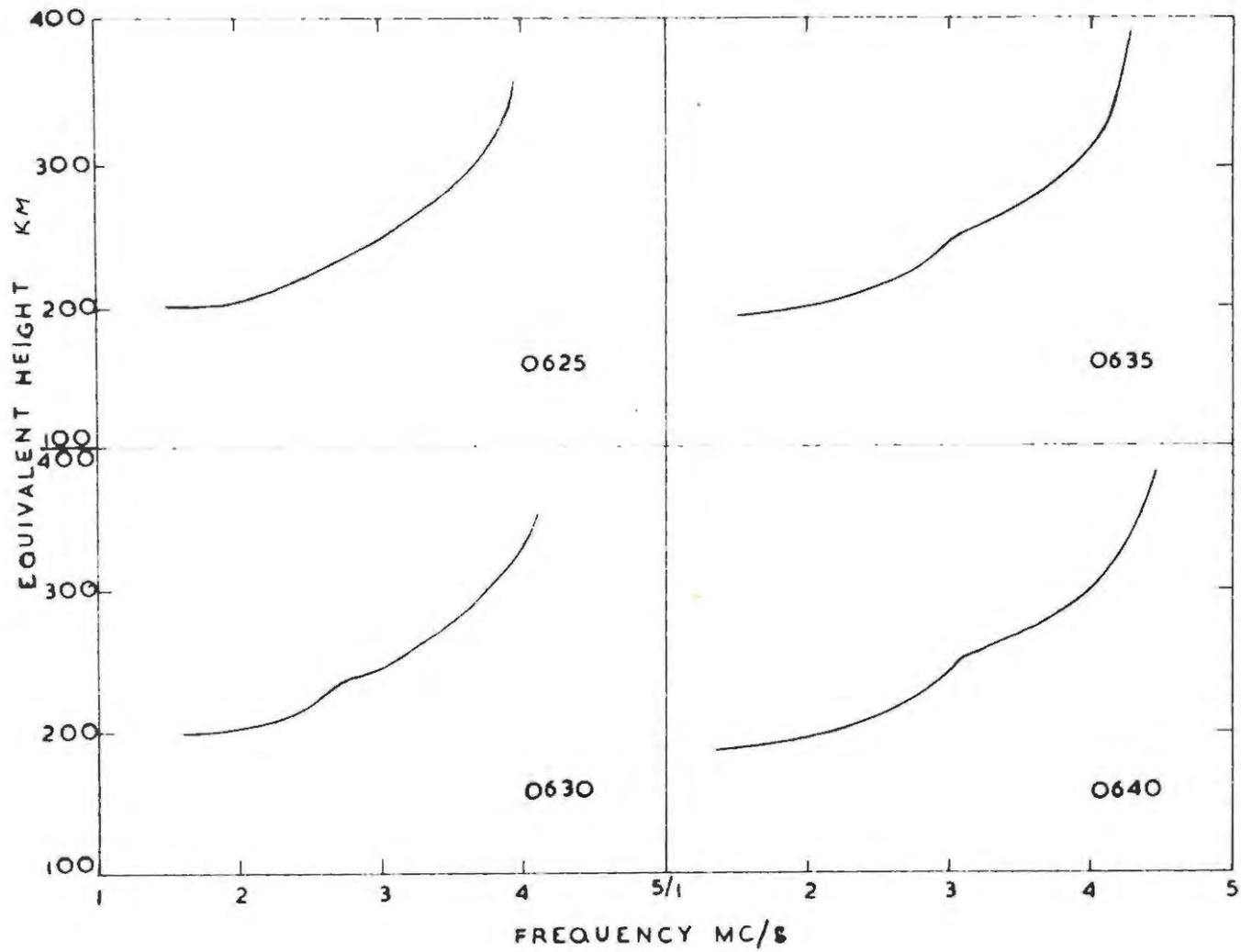


FIG 5.4 IONOGRAMS CONSTRUCTED FROM $N(z)$ CURVES IN FIG 5.3

in passing. SETTY (1960) states that at Cambridge bifurcation occurred at or near ground sunrise except during December (winter) when the F1 ledge was not very pronounced and first appeared about an hour and a half after ground sunrise. The second observation is that by SEN (1949) who reported that bifurcation of the F layer at Singapore (latitude $1^{\circ}19'09''$) occurred one and a half hours after sunrise (presumed to mean ground sunrise) during January and February.

Let us consider the following possible explanation for the discrepancy between the time delay observed and that predicted by the model:

The average control day noon $N(h)$ profile constructed by McElhinny from observations shows a more pronounced F1 layer than the noon $N(h)$ profile obtained from the present model (SZENDREI and McELHINNY, 1956, Fig. 1b). This discrepancy will be discussed in more detail in Chapter 7.

Let us for the moment accept that the observed F1 ledge was more pronounced. It is well known that the F1 ledge first appears as a point of inflection on the bottom of the F2 layer and develops as the solar zenith angle decreases. If the rate at which it develops on the model is the same as that in the ionosphere during the control period, the more pronounced F1 ledge observed would be expected to appear earlier than that on the model. The assumption that the rate of development is the same in the two cases can be partly justified by the fact that the observed be-

haviour of the F1 ledge is approximately like that of a Chapman layer and that a Chapman production function is assumed for the model. It is also clear from Fig. 5.6 (see section 5.3.2) that this rate is at most as great in the ionosphere as it is on the model.

5.3.2 The variation of f_oF1 with time.

On the simple Chapman theory of layer formation it can be shown (CHAPMAN, 1931a) that the electron production function varies in the following way:

$$q_m = q_o \cdot \cos \chi \quad (6.1)$$

where q_m is the value of the maximum in the $q(h)$ curve at any time and q_o is the value of q_m when $\chi = 0^\circ$.

We can therefore write

$$\frac{q_m}{q_o} = \cos \chi. \quad (6.2)$$

The assumption of equilibrium

$$\frac{dN}{dt} = 0 \quad (6.3)$$

in a region in which electrons are lost by a purely recombination process, proportional to N^2 , leads to

$$q = \alpha N^2. \quad (6.4)$$

On substituting from (6.4) into (6.2) with obvious subscripts, we

obtain

$$\frac{q_m}{q_o} = \frac{N_m^2}{N_o^2} = \cos^2 \chi \quad (6.5)$$

so that

$$N_m = N_o (\cos \chi)^{\frac{1}{2}}. \quad (6.6)$$

We have previously obtained the relationship

$$N = 1.24 \times 10^4 \cdot f^2 \quad (6.7)$$

between N and the frequency f of the exploring radio wave. Substitution for N in equation (6.6) from (6.7) yields

$$f_m = f_o \cdot (\cos \chi)^n \quad (6.8)$$

where f_m corresponds to N_m , and f_o to N_o with $n = 0.25$ for a purely quadratic loss law region.

To obtain $f_o F1$ accurately the same approach was followed as in the determination of the time at which bifurcation is first observed. h_o was again taken as 170 km and $H = 30$ km, values appropriate to the F1 region (NICOLET, 1960). Only partial ionograms were constructed showing the $h'(f)$ curve near $f_o F1$ because in the present case only the region near the F1 ledge inflection point was of interest. These partial ionograms are shown in Figure 5.5.

From the F1 critical frequencies so obtained the gradient of $\log \frac{f_m}{f_o}$ vs $\log \cos \chi$ gave $n = 0.19$. It should be mentioned

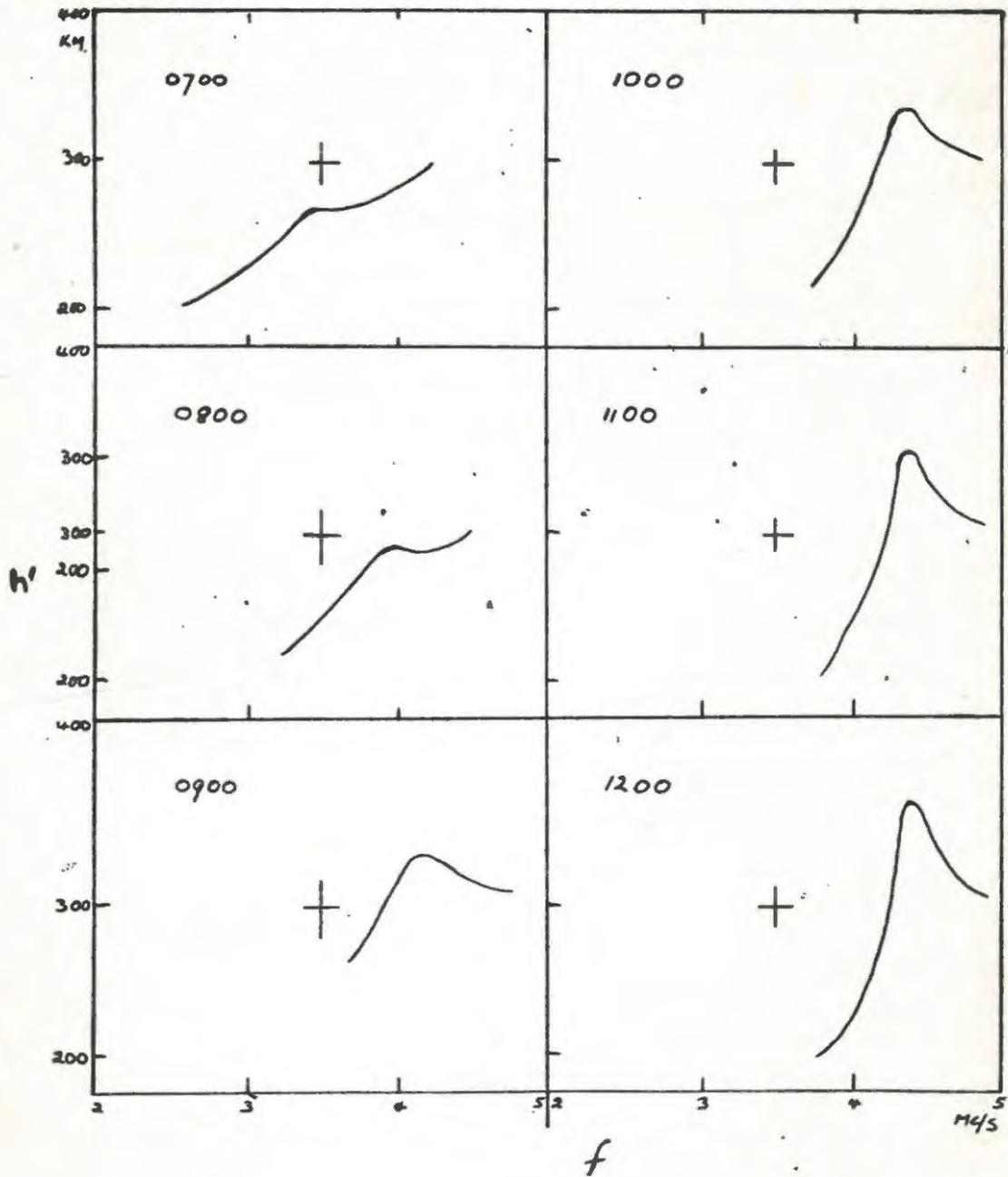


FIG 5.5 PARTIAL IONOGRAMS NEAR f_oF_2 (ORDINARY RAY).

that this gradient increases when χ becomes large due to the departure from equilibrium. In the determination of n the gradient near $\chi = 0$ was therefore used. The figure of 0.19 for n is lower than that for a purely quadratic region. It does, however, not differ significantly from $n = 0.20$ as determined by HIRSH (1959). This is to be expected since the equilibrium assumption is made in the method by which it is determined. Experimental determinations of n (ALLEN, 1948) yield values between 0.13 and 0.28 with a mean of 0.20.

In Figure 5.6 the variation of the electron density at the F1 ledge maximum with time is shown graphically, that on the model being represented by a solid line and the experimental observations by McElhinny represented by small circles with a broken line passing through them. The value of n for the latter observations is 0.17. In view of the assumptions made in the determination of n , no close agreement can be expected between theoretical determinations and the observed values, the reported difference being within the acceptable range.

5.3.3 The height of the F1 layer maximum electron density.

By using the f_oF1 critical frequencies obtained from the ionograms in Figure 5.5, the height of the point of inflection on the $N(z)$ curves in Figure 5.1 can be found. (It should be kept in mind that we have assumed $z = 0$ at 170 km and the scale height to be 30 km).

In Figure 5.7 the height of the F1 layer maximum on the present

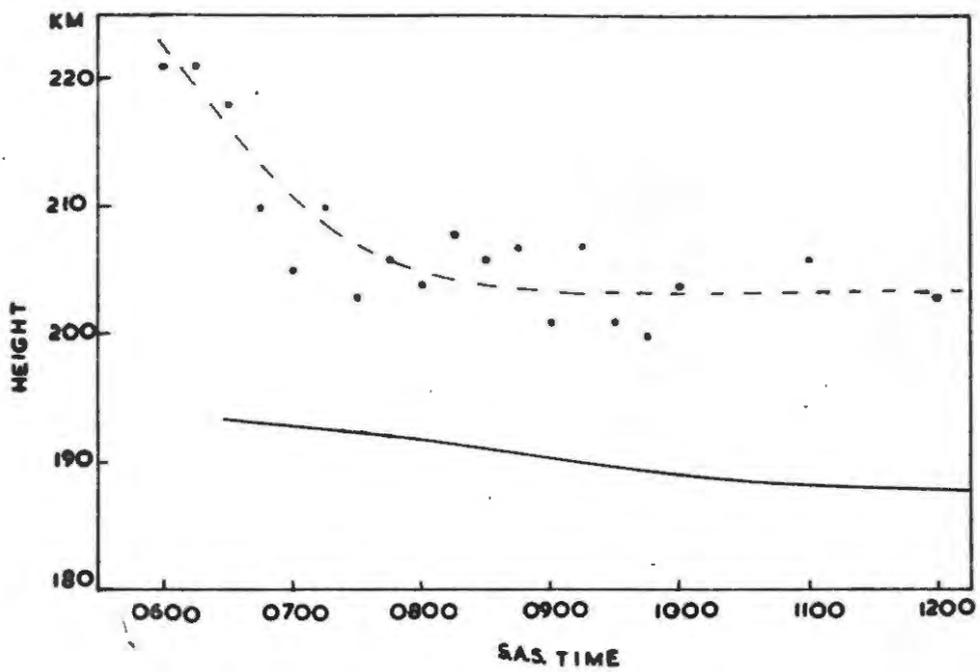


FIG 5.7 HEIGHT OF F1 LAYER KM (--- OBSERVED)

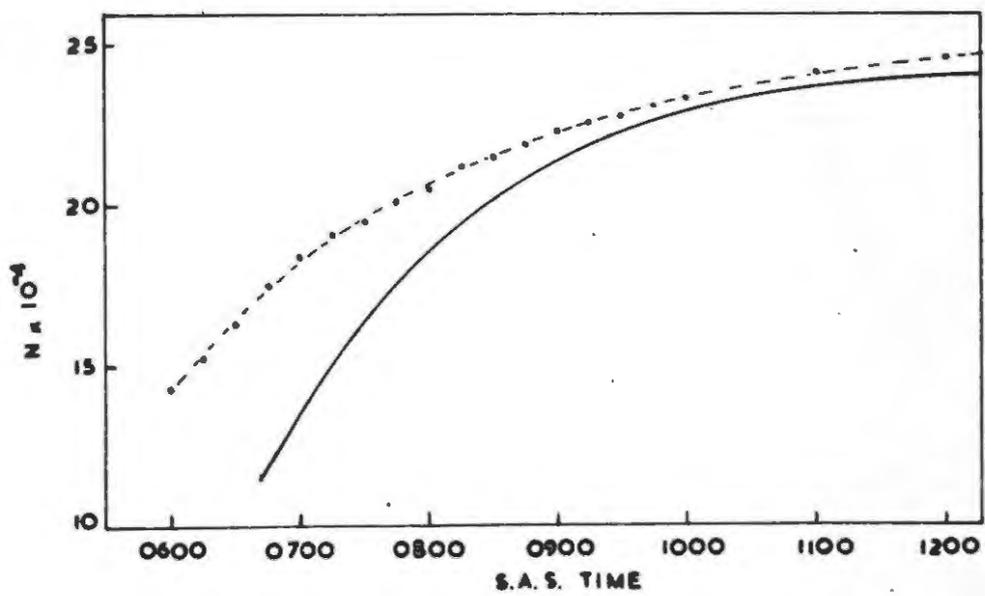


FIG 5.6 F1 MAXIMUM N_{max} CURVE

model is indicated by a solid line. Indicated by small circles are the heights as measured by McELHINNY (1958). The broken line was drawn to show the general tendency of the observations.

Two noticeable features emerge from this figure. The first is the rather rapid decrease in the height of the observed F1 ledge maximum immediately after sunrise. The other is the difference in heights at which the F1 maximum occurs on the model and in the observations. Apart from this the model shows the same slow decrease in the height of the F1 maximum as that observed.

We shall consider three different possible reasons for the observed discrepancies.

The first is a possible explanation for the rapid decrease in height observed just after sunrise. The neglect of night-time ionization is thought to be responsible for its absence on the model. The following argument, though very rough, does show how the decrease could be accounted for. It may be very approximately illustrated by a series of diagrams like those shown in Figure 5.8.

Consider the case of no night-time ionization first. This is illustrated in Figure 5.8(a), electron density being plotted along the abscissa and height on the ordinate. The distribution in (a) i is taken as that due to freshly formed ionization. With no night-time ionization added the level of the F1 ledge maximum must remain unaltered.

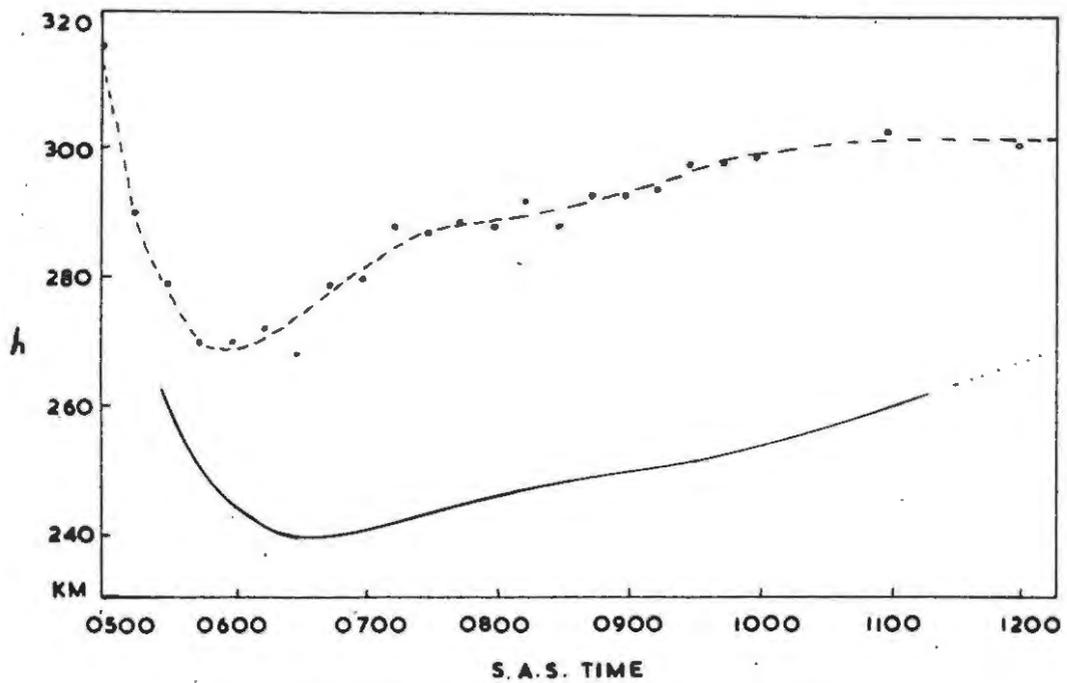


FIG 5.9 HEIGHT OF F2 MAXIMUM (---OBSERVED)

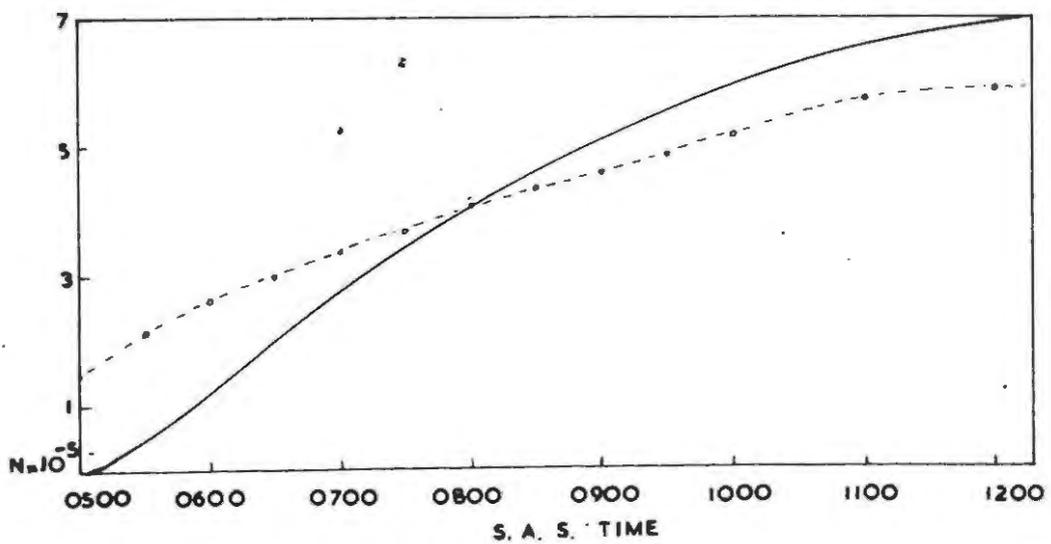


FIG 5.10 F2 ELECTRON DENSITY
MAX

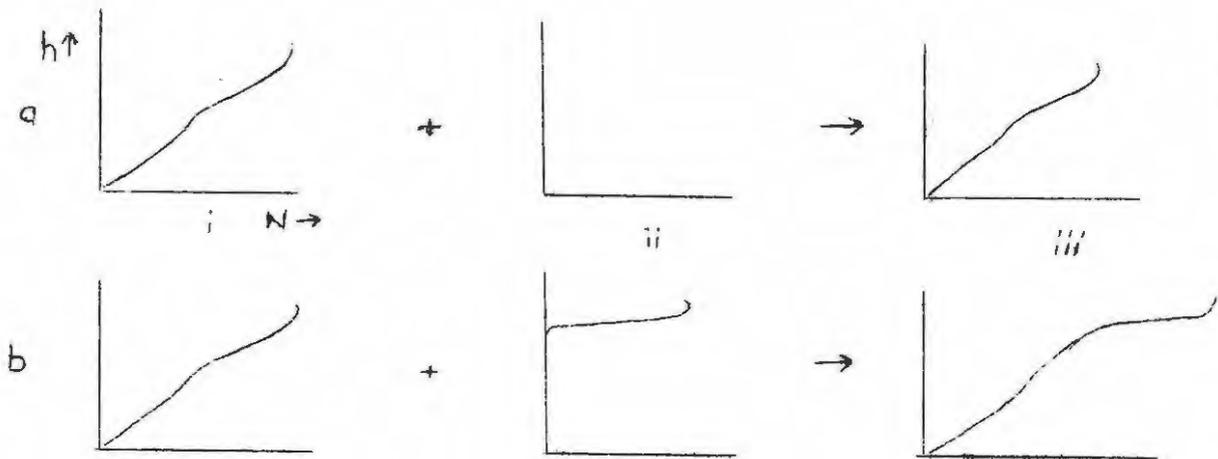


Figure 5.8. To show how the F1 ledge could be caused to appear at higher levels at sunrise.

In (b) we consider the effects to be expected if a significant amount of night-time ionization is present some distance above the level at which the F1 ledge is formed at the time that "fresh" ionization starts forming. The same distribution of "fresh" ionization as that considered in (a) is added to the residual night-time ionization distribution shown in (b) ii to yield the distribution in (b) iii. In this distribution two discontinuities are shown. The discontinuity at the higher level results from the addition of "fresh" ionization to the night-time distribution, the lower discontinuity from the "normal" process by which the F1 ledge is formed. In general, the F1 ledge maximum will probably not be identifiable with either of these because they may occur at nearly the same electron density and probably merge to form a single point of inflection that would occur at a higher

level than that at which "normal" F1 layer maximum appears. As the "fresh" ionization increases with a decrease in χ the effect of the residual ionization would be expected to disappear gradually, thus causing a lowering in the level at which the F1 ledge point of inflection occurs.

The same process offers another explanation for the difference in the times at which bifurcation is first observed. If the F1 ledge observed early in the morning is affected by the presence of night-time ionization, the discontinuity in the $N(h)$ curve associated with it will appear as soon as the "fresh" ionization at the base of the bank of night-time ionization becomes appreciable. This occurs earlier than the time at which the F1 ledge becomes visible for the first time under conditions where no night-time ionization is present.

The following reasons are thought to be responsible for the discrepancies between the height of the F1 maximum on the model and that observed by McElhinny and others.

In the first instance the experimental observations should be treated with caution because they were obtained through Kelso scaling of the ionograms neglecting the presence of a magnetic field, a method which also does not allow for the effects to be expected when a valley exists between electron density peaks or when significant low level ionization is present. This scaling method is known to yield "true" heights about 10 to 20 km too high.

In general the F1 ledge maximum is observed at levels of 170

to 180 km. The height at which it occurs on the model is about 190 km, i.e. about a third of a scale height above the generally observed level. A probable explanation for this is the assumption of near complete diffusive separation made in the model. The effect of more mixing (a smaller value of k) would be to decrease the gradient of the ion-atom interchange rate coefficient and therefore to lower the F1 maximum bringing it closer to the level at which the peak of the production function occurs.

5.4 The F2 layer peak.

The results of the non-equilibrium model presented in Figure 5.1 show an F2 layer electron density peak. No such peak occurs in the equilibrium model advanced by Hirsh. SETTY (1960) was the first to show that this peak arises from the solution of the continuity equations under non-equilibrium conditions. Because diffusion effects are neglected on the present model, the F2 peak it shows cannot be expected to exhibit the same behaviour as the observed F2 peak. Through the work of RISHBETH (1960) it has, however, become clear that diffusion and vertical drift due to electromagnetic forces have no very marked effect on the behaviour of the F layer near sunrise. There is therefore some merit in comparing the variations of the F2 peak on the model with observation.

In Figure 5.9 the height variation of the F2 peak is shown, the solid line representing the theoretical variation and the broken line the variation observed by McElhinny. Shortly

after sunrise a sharp decrease in height occurs in both cases. After this the height of the electron density at the peak increases steadily on the model until noon while the observations show the gradual attainment of an equilibrium value. This levelling off observed in the F2 peak height is presumably an effect of diffusion which decreases the height of the F2 peak but which becomes predominant only above a certain height.

The same mechanism is thought to be responsible for the electron density at the F2 peak in the ionosphere reaching a quasi-equilibrium value sooner than on the model. This effect is shown in Figure 5.10 in which the F2 electron density peak is plotted as a function of time, the solid line representing the theoretical variation and the broken line the observations. In this case diffusion would cause a greater loss rate through the lowering of the peak thus bringing about equilibrium between production and loss at an earlier time.

The observed height of the F2 peak is approximately 30 km greater than that shown by the model. Kelso scaling, neglecting magnetic field effects, would once again yield too great "true" heights, in this case even more so because of the greater electron densities involved than in the F1 layer. The model also neglects any increase in the scale height with increase in altitude, a neglect that causes the F2 peak to occur at a lower level than in the ionosphere in this particular case.

5.4 Discussion.

The model presented accounts reasonably well for the observed variations in the F1 region of the ionosphere and even gives a fair account of the F2 layer behaviour. Features apparently well represented are the bifurcation of the night-time F layer and the variation of f_oF_1 with change in the solar zenith angle. The neglect of night-time ionization, a scale height variation with altitude and ambi-polar diffusion is probably responsible for some discrepancies in the heights of layer peaks.

The behaviour of the model under more rapid variations in the production rate than that which occurs during the course of a normal day should show how well it reflects variations due to large departures from equilibrium. This we investigate in the next chapter.

CHAPTER 6.THE F1 REGION DURING AN ECLIPSE.6.1 Introduction.

In the ionosphere the largest departures from equilibrium conditions are to be expected during periods when the rate of change of the electron density is not approximately zero. Thus, during sunrise, when

$$\frac{dN}{dt} \approx q$$

and during an eclipse, near the time of maximum obscuration, when

$$\frac{dN}{dt} \approx \text{loss rate}$$

equilibrium conditions clearly do not apply at all levels. During the later part of an eclipse many interesting phenomena can occur, such as the formation of valleys between the electron density peaks (GLEDHILL, 1959) and of a point of inflection responsible for an additional critical frequency between f_oF1 and f_oF2 (RATCLIFFE, 1956). These phenomena cannot be explained on a model of the ionosphere in which equilibrium between production and loss rates is assumed at all times. The behaviour during an eclipse of the non-equilibrium model proposed in Chapter 4 should therefore prove interesting. In this chapter the results of such an investigation are presented. Special attention is paid to various eclipse phenomena that occur.

6.2 The eclipse.

We have mentioned that the model proposed should exhibit

approximately the same behaviour in the F1 region as the ionosphere showed over Grahamstown during the control period in McElhinny's investigation. To facilitate comparison with the eclipse observations by McElhinny on the 25th December, 1954, the obscuration factor at 200 km for this annular eclipse was used in the computation of the eclipse effects to be expected on the model. The particulars of the eclipse are given in Table 6.1. During annularity 85.5% of the solar disc was covered.

TABLE 6.1. Times of contact at various heights: South African Standard Time = U.T. + 2 hours.

Height km	1st contact	2nd contact	3rd contact	4th contact
100	0646	0800	0807	0935
200	0645	0758	0805	0934
300	0643	0757	0803	0932

6.3 Solution of the continuity equations.

Under eclipse conditions the production function q must be multiplied by the appropriate factor at each instant of time to obtain the production rates. We have used the production rates computed from $q_0 = 400 \text{ electrons cm}^{-3} \cdot \text{sec}^{-1}$ and for 25 December, 1954 in conjunction with the mentioned eclipse factor. In solving equations (4.2) and (4.3) q was multiplied by the value of f at the time for which q was calculated. The mean of q , \bar{q} , and of f then gave the production rate during any interval. Except for this change the procedure for computation was the same

as before. In order to obtain more accurate $N(z)$ profiles during the eclipse the equations were solved at intervals of $1/4H$ from $z = 2.75$ to $z = 0.75$ and at intervals of $1/2H$ from $z = 0.5$ to $z = -1.0$.

6.4 Results.

6.4.1 The electron density.

Some of the results of the computation are shown graphically in Figures 6.1, 6.2 and 6.3. (These can be found at the back of this thesis). In Figure 6.1 the variation in electron density with time is shown for the given constant levels of z . The corresponding $N(z)$ profiles at given times are shown in Figures 6.2 and 6.3 in which the electron density is plotted in terms of the plasma frequency.

Consider the $N(t)$ curves in Figure 6.1 first. The increasing delay that appears between the time of maximum obscuration and the time at which the minimum in N occurs at increasing heights between $z = 0.5$ and $z = 2.0$ is due to a decrease in the effective loss coefficient. It has been reported by GLEDHILL (1959) in his model of the ionosphere in which the loss rate was assumed to decrease with height. Above $z = 2.0$ the time lag is observed to decrease with an increase in height. This effect may be explained by the even slower recombination rates at these higher levels.

Observe the very slight decrease in the electron density at $z = 2.75$ during the first part of the eclipse. This is due to the very small loss rate at this level. The same reason causes the electron density to increase at a relatively great rate before

the eclipse starts. As the production rate starts increasing immediately after the maximum phase of the eclipse it very soon overtakes the loss rate with the result that the electron density minimum occurs near the time of maximum obscuration.

At a level of about $z = 2$ the loss rate is greater than at $z = 2.75$ but still not nearly as large as the production rate. Consequently the production rate overtakes it only some considerable time after maximum obscuration. At $z = 0$ the production and loss rates are much greater than at $z = 2.0$ and nearly equal in magnitude. A small time lag results causing the minimum in N once again to occur near the time of maximum obscuration.

Between $z = 0.5$ and $z = -0.5$ the time lag is observed to increase with decrease in height. This is due to a decrease in the loss rate with decrease in altitude. In this case, however, the decrease in the loss rate is not due to a decrease in the loss coefficient, which is assumed independent of height in this region (see section 6.9) but due to a decrease in electron and ion densities because at the time under consideration these densities have not built up to anywhere near their noon equilibrium values. At higher levels the decrease in N follows the reduction in q more slowly and by the time that the obscuration is at a maximum the F2 peak occurs at or above the level $z = 2.75$, where N has decreased least.

Between 0640 and 0800 hours the lines for the levels $z = 2.0$ to $z = 2.75$ cross over. This is associated with the upward move-

ment of the F2 peak, caused by the decrease in the loss coefficient with increase in h . An upward movement of the base of the F2 layer is also apparent from the $N(z)$ curves in Figure 6.2. This is discussed in section 6.7.

During the period 0820 to 0930 a clearly discernible crowding together of the electron density constant level contours takes place. In the model discussed by Gledhill the lines actually crossed over during a period just after maximum obscuration, signifying the formation of a valley between the F1 and F2 peaks. It is clear from Figure 6.1 that no such valley formed on the present model. By referring to the $N(z)$ curves for the period mentioned in Figures 6.2 and 6.3 it can be seen that this crowding together of the electron density contours is associated with the reduction in the height gradient of N . This would be expected to lead to a more enhanced F1 ledge during this time.

More information could be gleaned from the direct study of these curves, but their interpretation is greatly facilitated by the construction of ionograms which show up phenomena in a way more usually met with in the description of eclipse effects.

Before we turn to the interpretation by way of the ionograms let us consider the results of the variation in the ion composition.

6.4.2 The variations in the concentration of O^+ and OY^+ ions.

The present model also yields the concentration of ions at different levels and times, information that is not obtained

through the usual methods of investigating the F1 layer. In Figure 6.4 (at the back of the thesis) the time variation of $n(O^+)$ is shown for the heights indicated and in Figures 6.5 and 6.6 (the latter at the back of this thesis) the height distributions of $n(O^+)$ at the times shown. Similarly the concentration-time curves for given heights and the concentration-height curves at given times for the OY^+ ions are shown in Figures 6.7 and 6.8 (a), (b) and (c) respectively.

As an illustration of the way in which the composition of the plasma varies during the course of the eclipse, the OY^+ ion and electron densities at 0700 and 0830 hours are plotted against z in Figure 6.8 (d) and both positive ion and electron concentrations at 0925 in Figure 6.9.

Several interesting features appear in these figures:-

- i. Figure 6.4 shows a similar crowding together of the O^+ concentration contours as the $N(t)$ curves exhibit, although to a considerably lesser extent.
- ii. Figure 6.6 shows a point of inflection in the $n(O^+) - z$ curve which moves towards greater concentrations and from about $z = 1$ to about $z = 2$ during the period 0840 to 0950 S.A.S.T.
- iii. Figure 6.7 shows a crossing over in the $n(OY^+)$ -time curves of $z = 0.5$, 0 and -0.5 after the maximum phase of the eclipse. This crossing over is due to the downward movement of the peak in the $n(OY^+)$ -height distribution during this period. It is more clearly illustrated in Figure 6.8 (c).

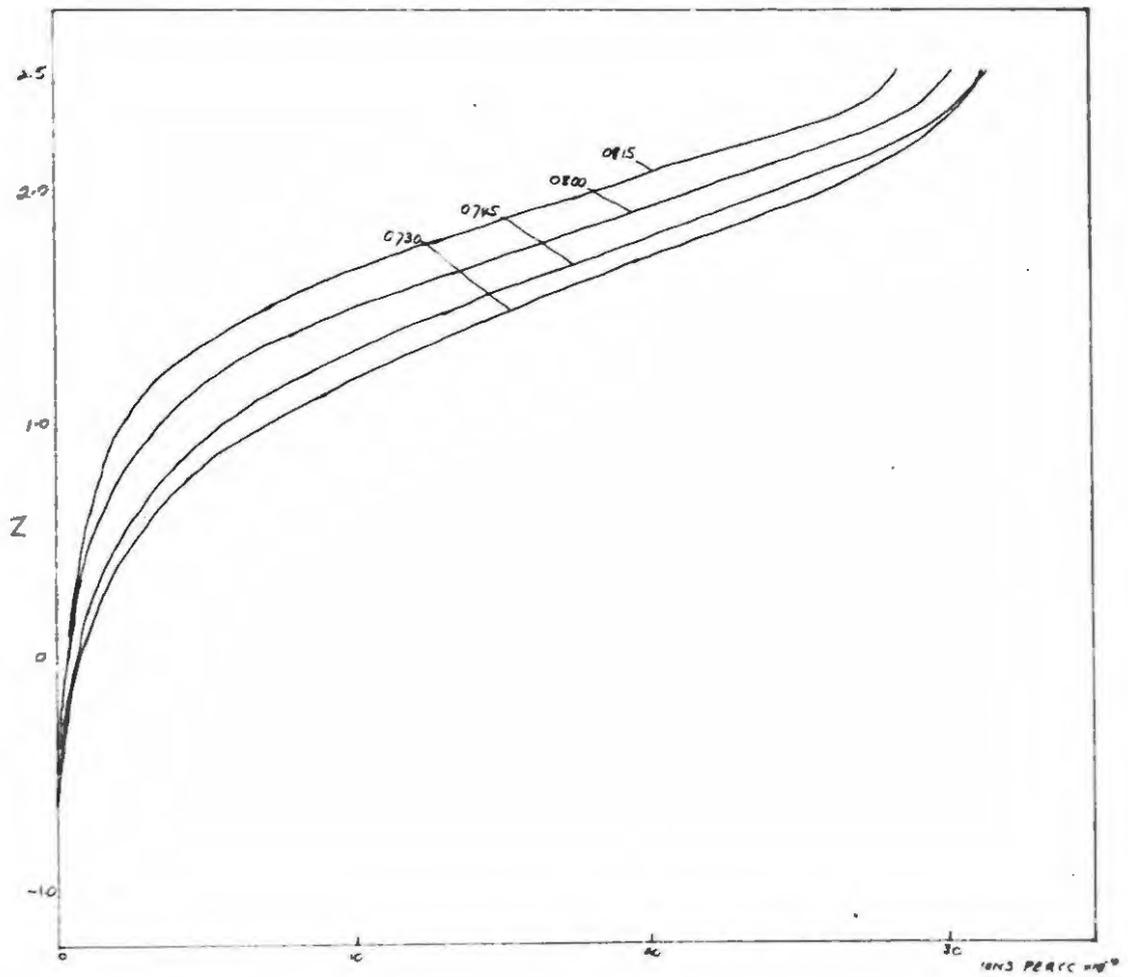


FIG 6.5 O⁺ ION DENSITY AT GIVEN TIMES

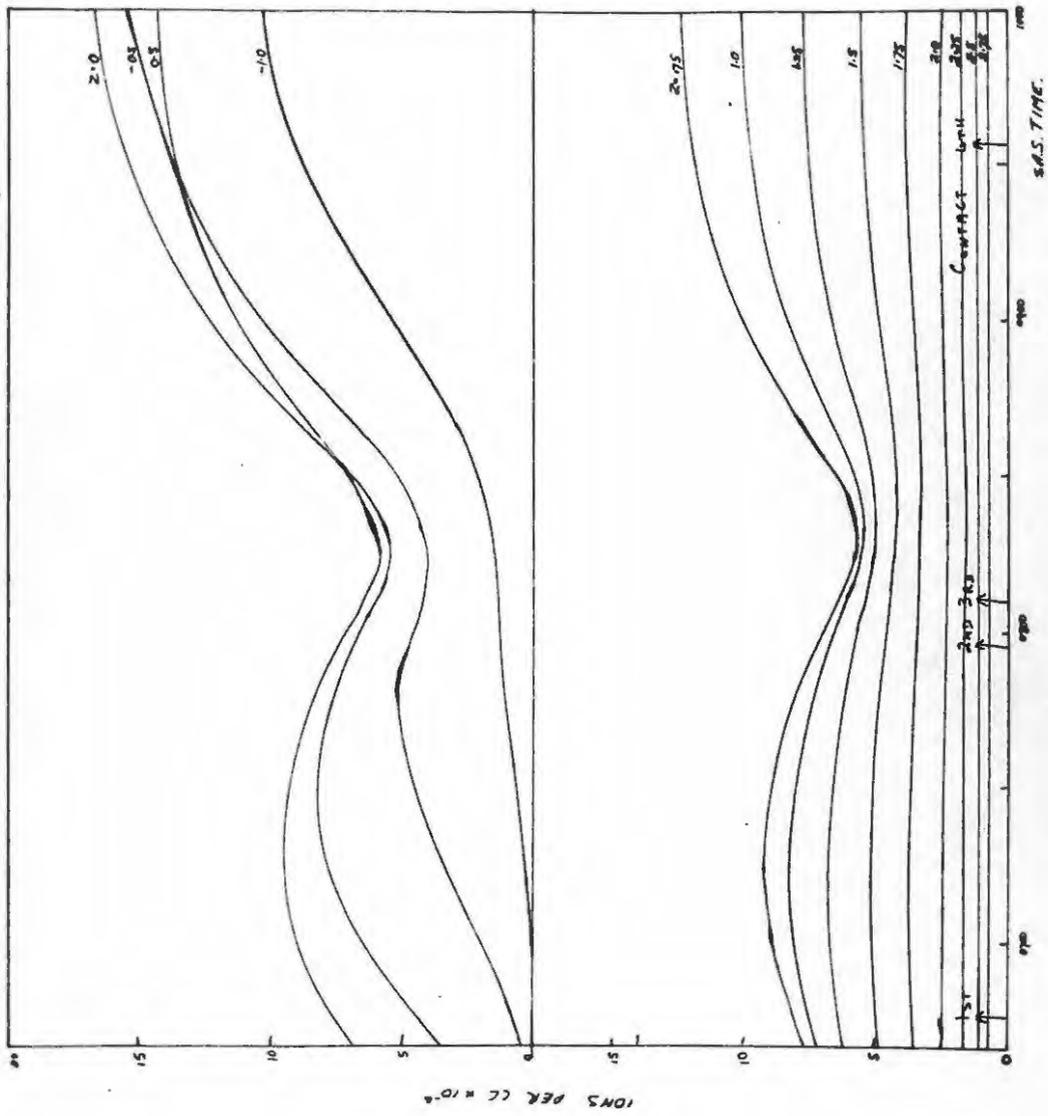


FIG 6.7 CURVES OF THE VARIATION OF OY IONS AT GIVEN HEIGHTS.

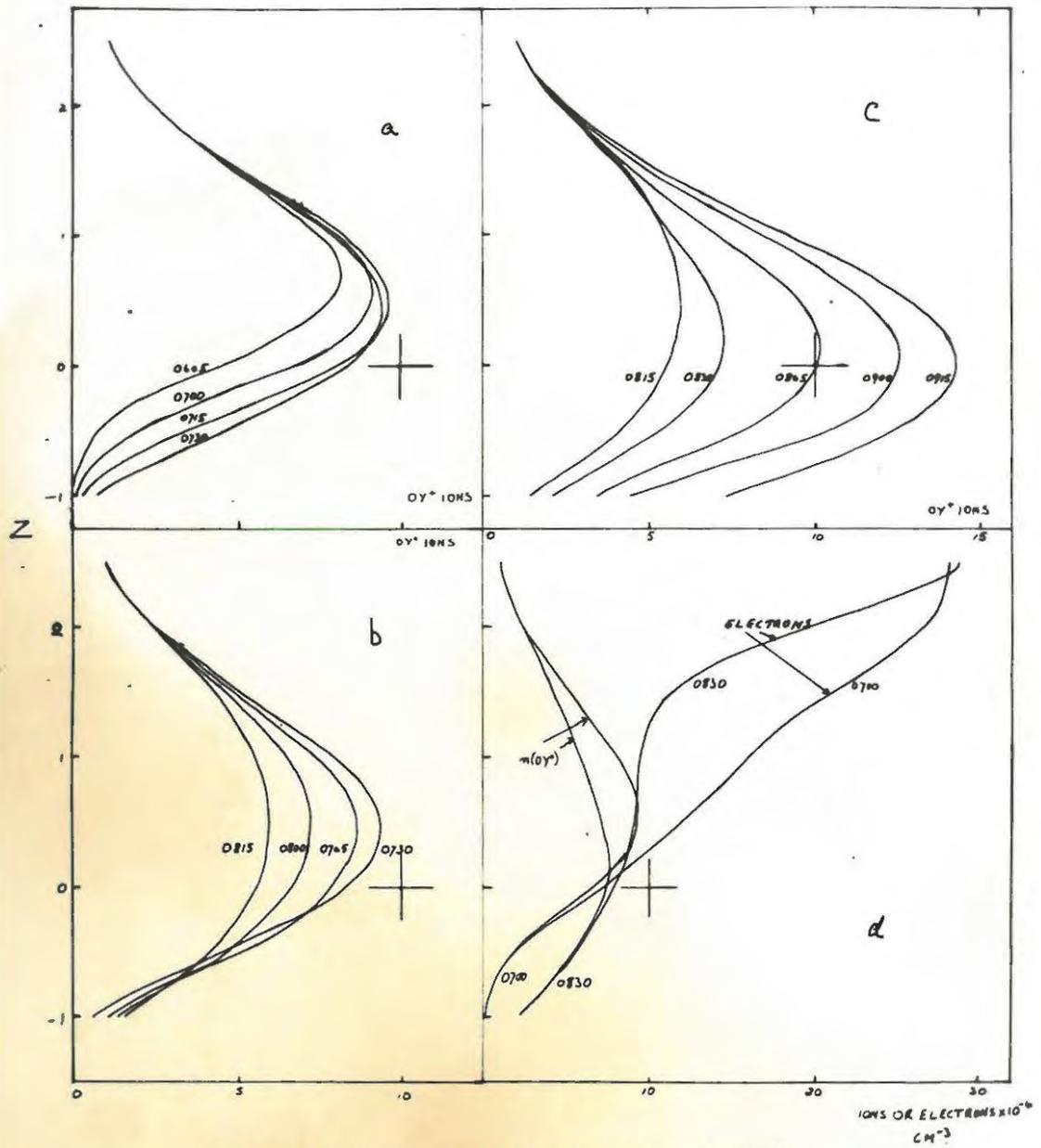


FIG. 8 OY^+ ION AND ELECTRON CONCENTRATION AT GIVEN TIMES

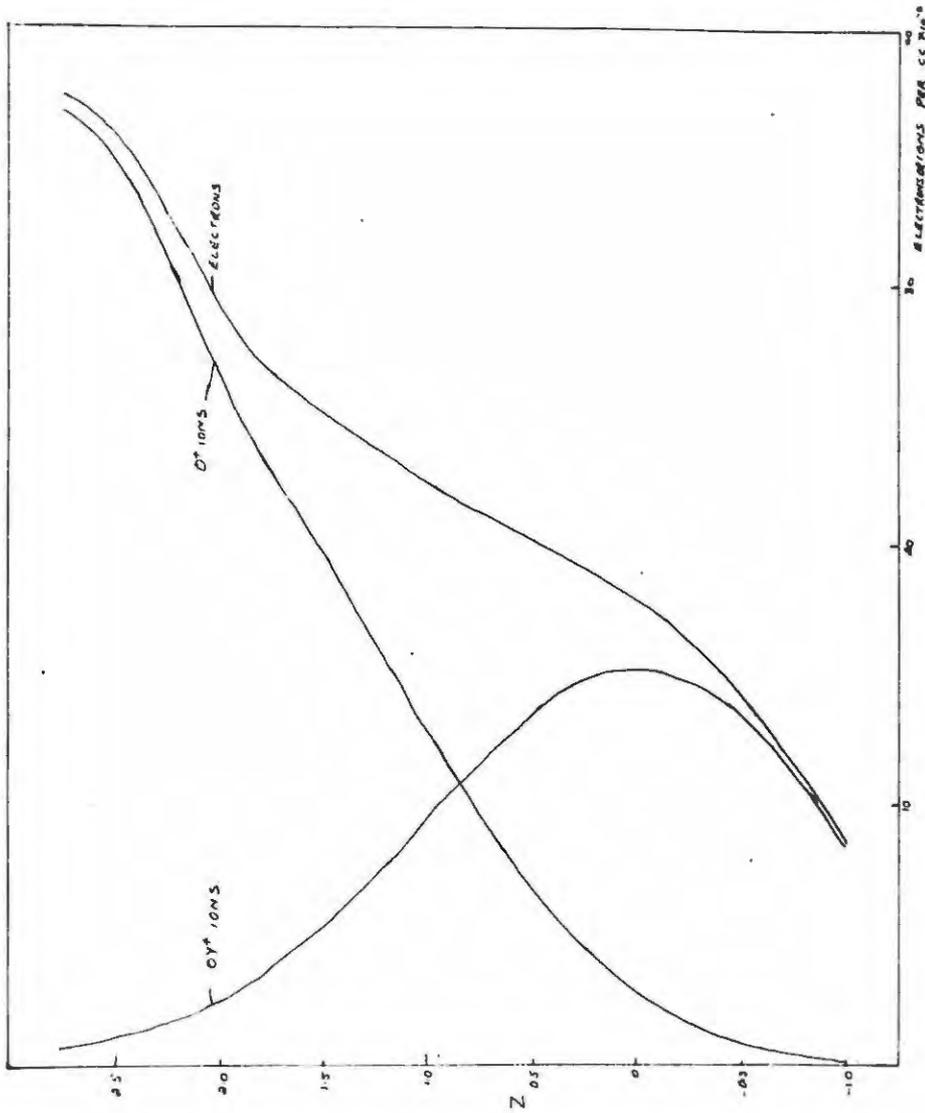


FIG 6.9 HEIGHT DISTRIBUTION OF IONS AND ELECTRONS AT 0925 S.A.S.T. ECLIPSE DAY.

iv. Figure 6.8 (a) and (d) show that a similar downward movement of the peak of OY^+ ions occurs during the early part of the eclipse.

The implications of these observations will emerge during the discussion of the results in the light of the effects observed in the ionograms.

6.4.3 The ionograms.

Ionograms were constructed by the method described in section 1.10 for selected times during the course of the eclipse. They are shown for the times indicated in Figures 6.10, 6.11 and 6.12.

Two much discussed features appear immediately from the study of these ionograms. The first is the occurrence of a more pronounced F1 critical frequency cusp during the main phase of the eclipse. The other is the appearance of an additional critical frequency during the late stages on the ionograms starting at 0925 and continuing until 1000. The behaviour of this cusp is exactly the same as that of the critical frequency $f_{O}F1\frac{1}{2}$ that has been observed during the later parts of eclipses and promises to be one of the most interesting results of the present investigation.

6.5 Comparison with experimental observations.

A comparison with the observational results obtained by McElhinny does not appear to be well advised at this stage for two reasons.

In the first place SZENDREI and McELHINNY (1956) report that

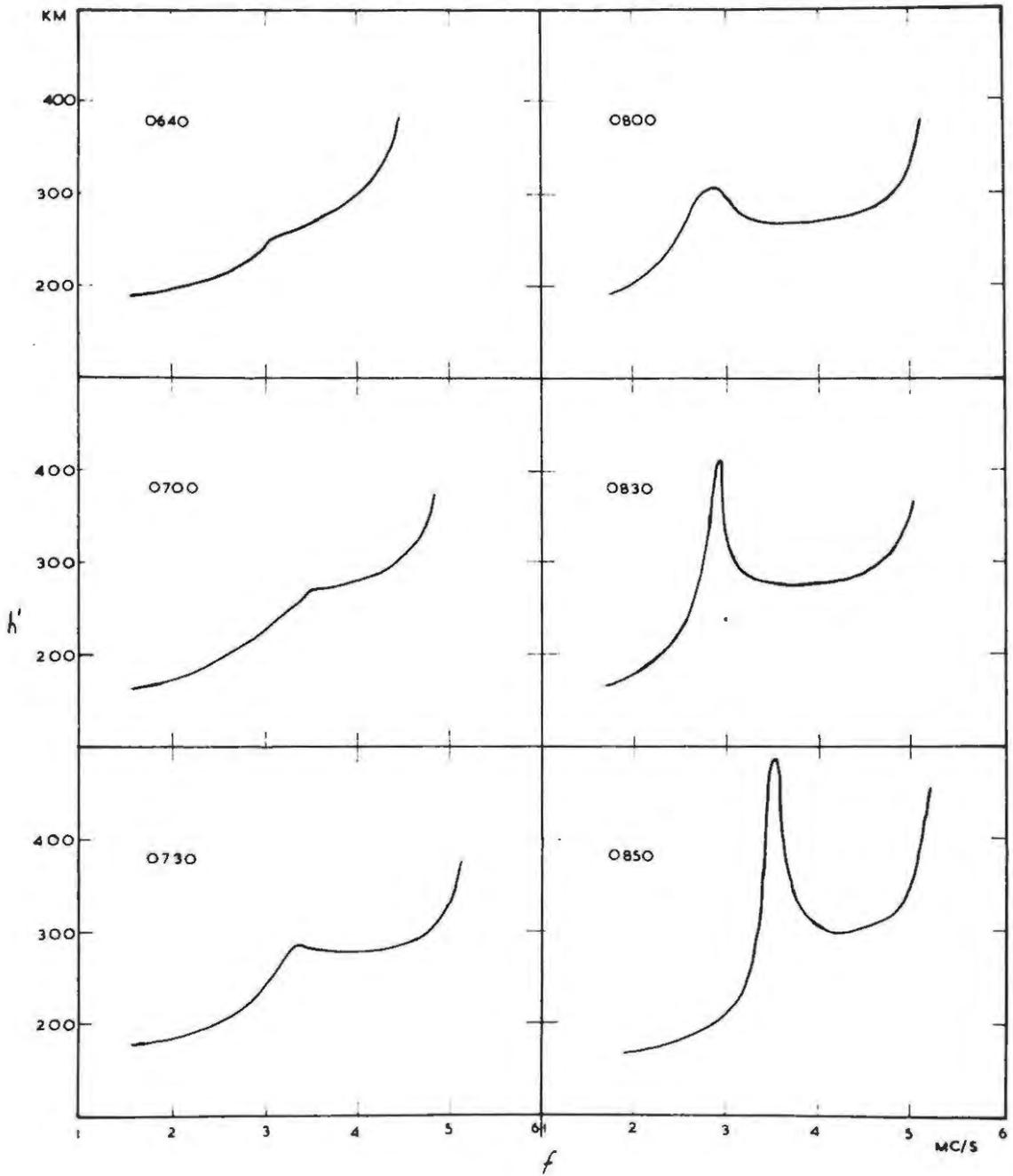


FIG. 6-10 IONOGRAMS CONSTRUCTED FOR TIMES GIVEN

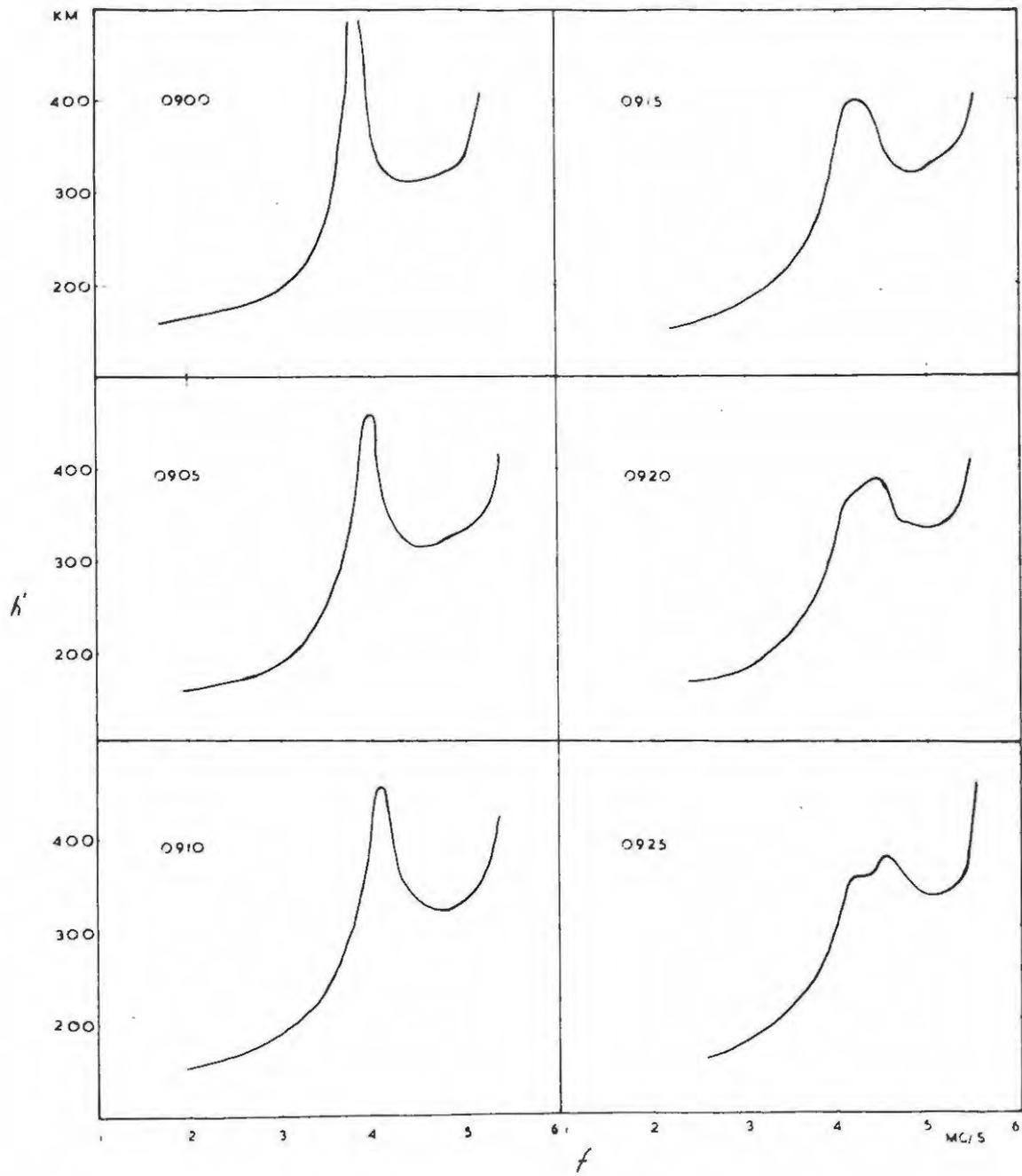


FIG. 6.11 IONOGRAMS CONSTRUCTED FOR TIMES GIVEN

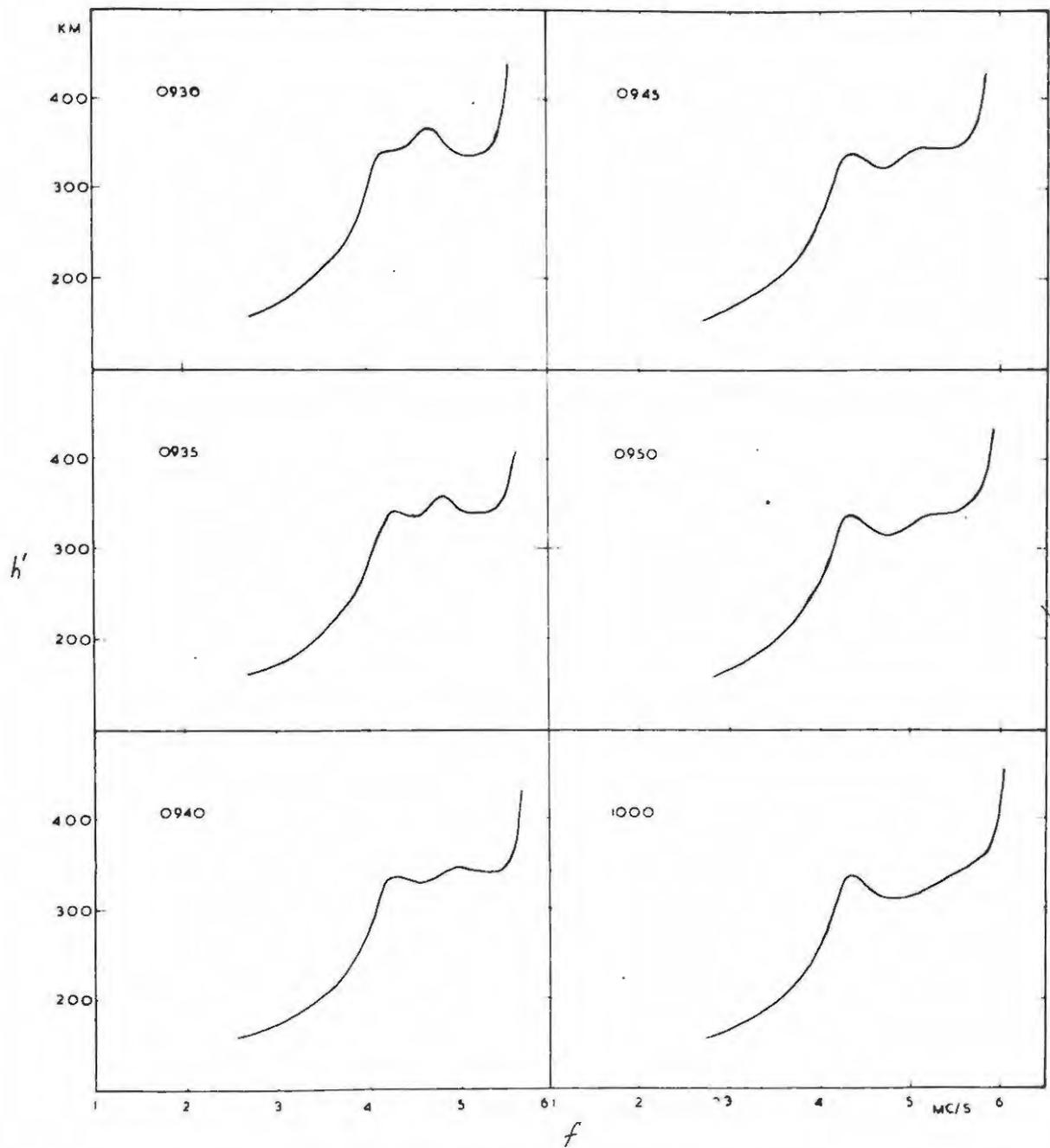


FIG. 6.12 IONOGRAMS CONSTRUCTED FOR TIMES GIVEN

the F1 layer maximum was always well defined on the control days. Furthermore, they state: "During the middle of the eclipse, from 0730 to 0830 approximately, the F1 layer was slightly masked by intermittent Esp, making the determination of the Kelso distributions impossible". No mention is made of a more pronounced f_oF1 cusp during the main part of the eclipse.

The second reason is that no $f_oF1\frac{1}{2}$ critical frequency was observed at Grahamstown during this eclipse. Possible explanations for this discrepancy and a comparison with the eclipse observations for the 25th December, 1954 at Grahamstown are therefore deferred to Chapter 7.

A comparison with eclipse results in which an $f_oF1\frac{1}{2}$ and a more pronounced f_oF1 cusp appear was, however, made to see if the non-equilibrium model shows a close enough similarity to observations. A possible explanation of these phenomena could then be attempted on the basis of the data available in the computation of the model. Ionograms for two such eclipses are shown in Figures 6.13 and 6.14.

Figure 6.13 shows ionogram traces obtained from a paper by MINNIS (1957). These were obtained at Singapore (latitude $15^{\circ}N$) and show a rather close similarity to the theoretically constructed ionograms as far as the more pronounced f_oF1 cusp and the $f_oF1\frac{1}{2}$ critical frequency are concerned.

Figure 6.14 is reproduced from a paper by LEJAY and DURAND (1956) and shows their Figures 7 and 8. These ionogram traces were obtained during the eclipse on 25 February, 1952, at Khartoum

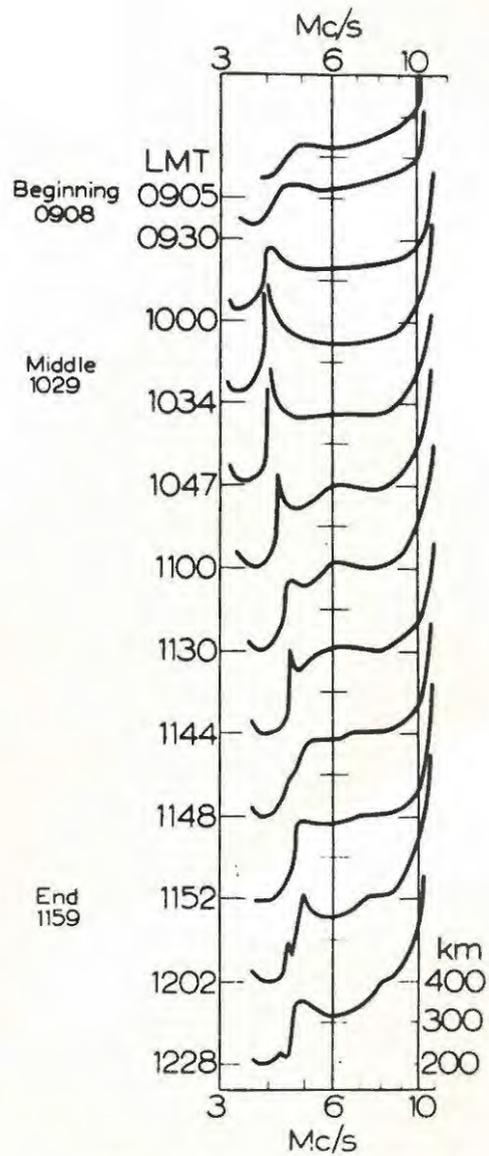


Fig.6.13 *h'f* traces at Singapore during the solar eclipse of 20 June 1955.

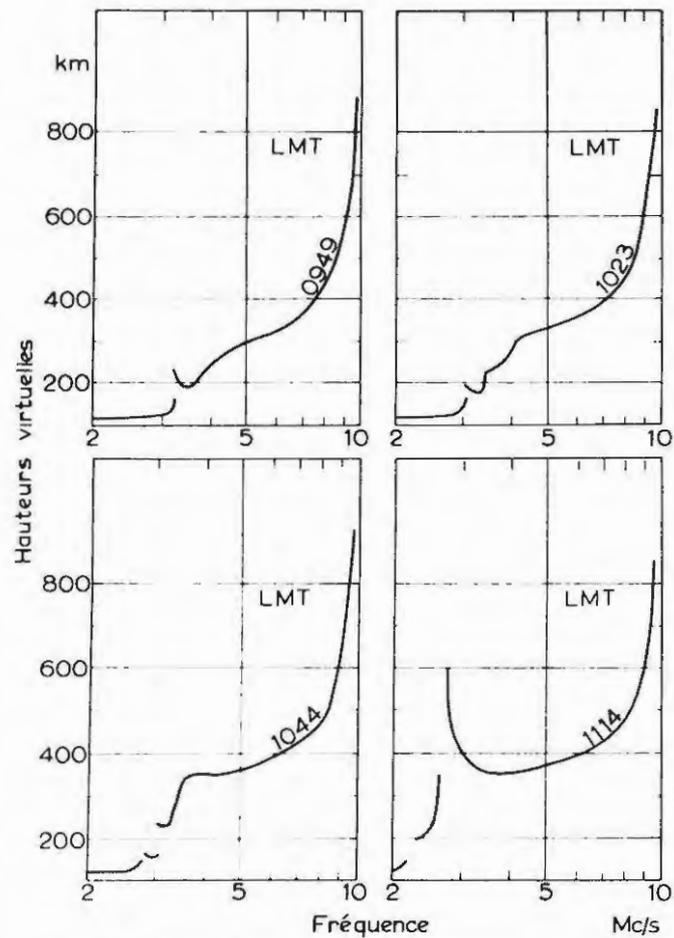


Fig. 7. Observations ionosphériques durant l'éclipse du 25.2.1952 à KHARTOUM avant la totalité.

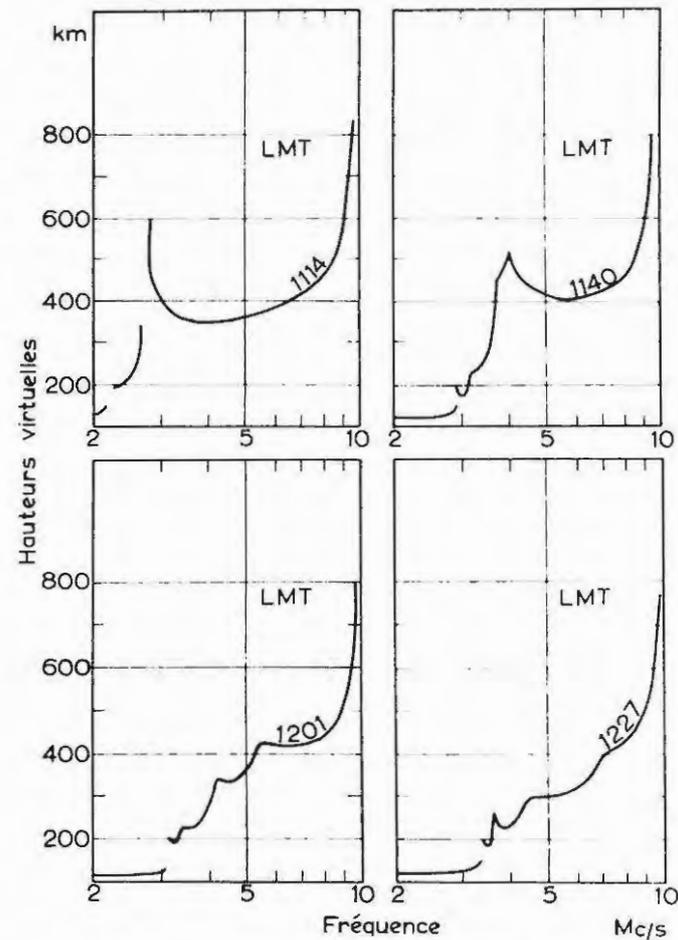


Fig. 8. Observations ionosphériques durant l'éclipse du 25.2.1952 à KHARTOUM après la totalité.

FIG 6.14

(latitude $15^{\circ}35'N$) and also show a pronounced f_oF_1 cusp at 1114 hours. The movement of the $f_oF_1^{1/2}$ critical frequency is illustrated in the ionograms for 1140, 1201 and 1227 hours LMT.

We see that the non-equilibrium model does exhibit some commonly observed eclipse phenomena. It shows the enhancement of the F1 layer during the main phase of the eclipse particularly well. At the time when the F1 layer appears most pronounced, the critical frequency $f_oF_1^{1/2}$ starts making its appearance on the observed ionograms. A broadening of the F1 cusp on the theoretical ionograms corresponds to this, the separation of the two cusps probably only appearing later because of the finite separation of the Titheridge frequencies used in their construction. The present work seems to offer a reasonable way of explaining these phenomena, and we shall now consider them in more detail.

6.6 Enhancement of the F1 layer.

6.6.1 Introduction.

On comparing the ionograms in Figures 6.10 and 6.11 with the partial ionograms in Figure 5.5 it is clear that the F1 critical frequency cusp is more pronounced during the eclipse than on a normal day. This is particularly well shown in the 0900 hour curves. The phenomenon has been observed during many eclipses and RATCLIFFE (1956) has drawn attention to it. It poses a problem because during the course of a normal afternoon the reduction in the production rate of electrons causes the F1 ledge to become less distinct and to disappear completely near sunset.

Why then does it become more distinct when the production rate is decreased during an eclipse, Ratcliffe asks.

6.6.2 The f_oF_1 cusp on an ionogram.

Radio pulses reflected from a region in the ionosphere near an electron density peak usually give rise to an ionogram trace that goes off to very large values of h' on either side of the critical frequency. The values of h' may be greater than that which the ionosonde can record. This causes a clear-cut discontinuity to appear on the ionogram.

When a point of inflection appears in the $N(h)$ distribution, the radio pulses will be retarded to a greater or lesser extent depending on the magnitude of the gradient $\frac{dN}{dh}$ at the point of inflection. If the gradient is fairly large, such as that illustrated by the point marked F1 in Figure 1.1, only a cusp will appear at the critical frequency. If the gradient becomes small at the point of inflection, the cusp on the ionogram will become much more pronounced.

It follows that any mechanism causing the gradient $\frac{dN}{dh}$ at the F1 ledge point of inflection to decrease will therefore cause a more pronounced f_oF_1 critical frequency cusp to appear on an ionogram. This is commonly described as an enhanced F1 layer.

In the subsections that follow we consider three mechanisms that could cause such a decrease in $\frac{dN}{dh}$ at the F1 ledge maximum.

6.6.3 The production function.

Even though a reduction in the rate of electron production causes the F1 ledge to disappear on a normal day, a reduction in q during an eclipse usually results in a more distinct F1 ledge being formed. Part of the explanation is that during an eclipse q reduces in a way quite different from the change on a normal afternoon. Consider equation (2.3) written in the form

$$q = q_0 \exp \left[1 - z - \exp(-z) \cdot \text{Ch}(\chi) \right]. \quad (2.3)$$

The time variation of q (χ is a function of time) at different levels is shown in Figure 6.15. At high levels (say $z = 3$) the production rate reaches a near constant value quite early in the day. At lower levels ($z = 0$) the change in q continues to be appreciable right up to noon. Let us investigate this effect in more detail.

At levels as high as $z = 3$ the factor that determines the number of ion-electron pairs produced per second during the hours near noon is the concentration of the ionizable constituent and the ionization cross-section. In the expression for q the term $\exp(-z) \cdot \text{Ch}(\chi)$ in the exponent of e , the term which takes account of the depletion of the beam of ionizing radiation, is therefore small in comparison with the term $(1 - z)$ which arises from the concentration and ionization cross-section of the ionizable constituent. This means that changes in χ are not very effective in changing q except in the early morning or late afternoon when $\text{Ch}(\chi)$ becomes very large.

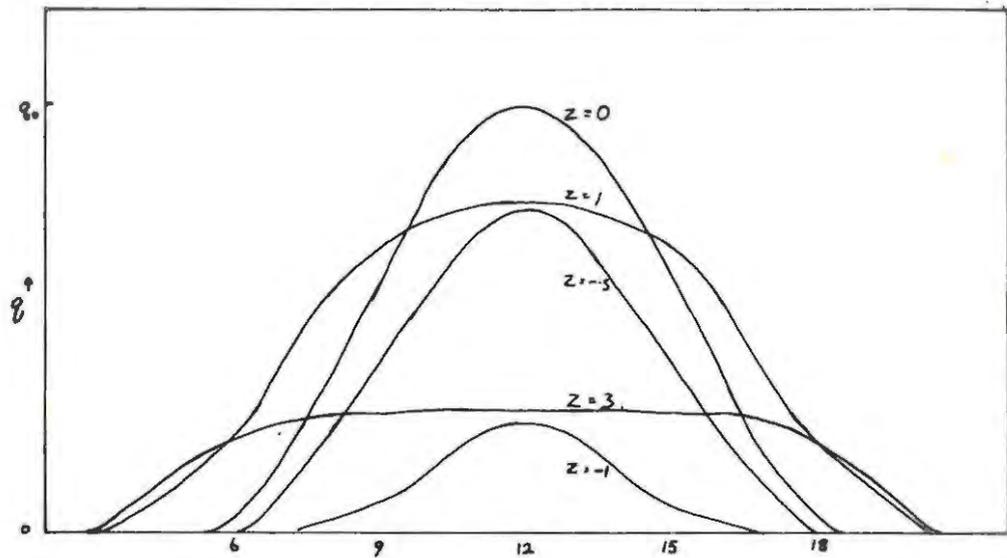


FIG 6.15 VARIATION OF q WITH TIME AT DIFFERENT LEVELS.

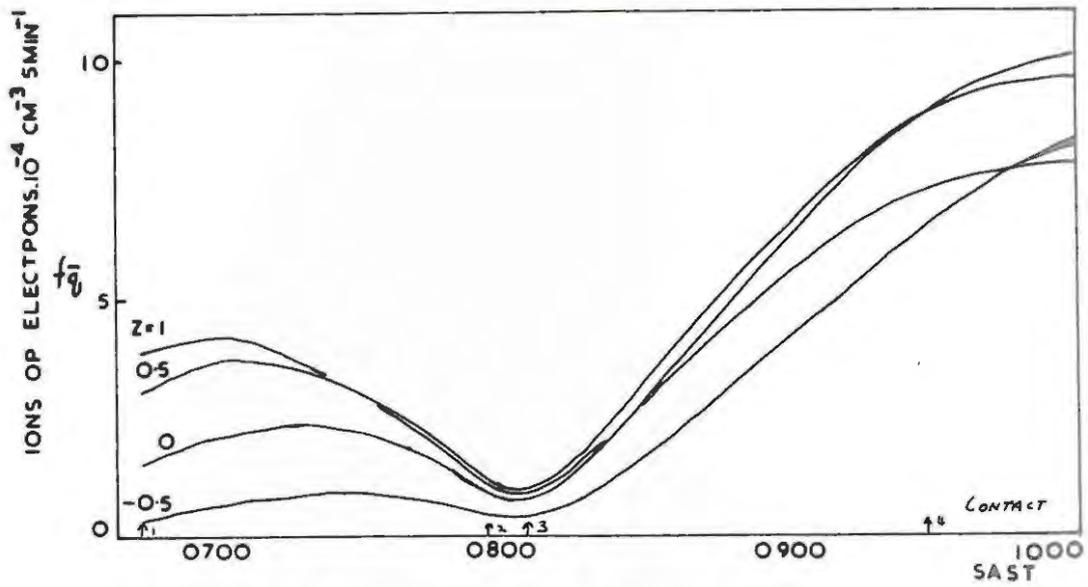


FIG 6.16 TIME VARIATION OF f_q AT GIVEN HEIGHTS

At low levels, below $z = 0$, the factor limiting the production rate is the amount of ionizing radiation available. This in turn depends on its path length in the atmosphere. The term $\exp(-z) \cdot \text{Ch}(\chi)$ therefore becomes predominant and q becomes sensitive to changes in χ even during the hours around noon.

During a normal afternoon the reduction in q is due to $\text{Ch}(\chi)$ becoming large and going to infinity at sunset while I_{∞} , the radiation intensity incident at the top of the atmosphere, and therefore q_0 , remains unaltered. This change in $\text{Ch}(\chi)$ acts to raise the level of peak production, with the production decreasing faster below the peak than above it. Under equilibrium conditions any F1 ledge formed at or near the level $z = 0$ will therefore tend to merge with the ionization above it as the afternoon progresses.

On an eclipse day the obscuration factor f , determined by astronomical considerations, influences I_{∞} so that during the first half of an eclipse, if $\text{Ch}(\chi)$ does not change much, the production rate will be reduced by the same fraction at all levels. If near equilibrium conditions prevail at all levels the ionized layers will not change shape during such a reduction in q . If, however, the decrease in q takes place so rapidly that the condition departs markedly from equilibrium, the relative loss rates will determine any change in the layer shape. We consider this effect in section 6.5.4.

Let us now consider what happens on the model presented here. The eclipse takes place in the morning during a period, 0645 to

0934, when the value of $Ch(\chi)$ is still decreasing fairly rapidly. At levels at or just above the F1 ledge maximum (say $z = 1$ or 1.5) the production rate has reached a magnitude nearer its maximum than at lower levels near the base of the F1 layer (say $z = -0.5$).

When the incident radiation is cut down, electron production at $z = 1.5$ will be cut down very rapidly because any increase due to the change in $Ch(\chi)$ at this level is not large enough to counteract the reduction due to the decrease in $I_{e\infty}$. At $z = -0.5$ on the other hand, the production rate is increasing rather rapidly as the path length of the radiation in the atmosphere diminishes. This means that the initial reduction in the incident radiation due to obscuration is not sufficient to counterbalance the increase. To illustrate this more clearly we have plotted the actual production rates \bar{fq} against time in Figure 6.16 for a few levels round $z = 0$. The crossing over of the lines during the initial phase of the eclipse is due to the downward movement of the production rate peak. This downward movement is continued as can be seen by the crossing over of the lines some time later at 0940.

If the F1 ledge maximum is formed just above the level at which the peak production rate occurs and the loss rate coefficient in this region is approximately independent of height, the electron density distribution will be controlled largely by the distribution of the production function. During the first phase of the eclipse the electron density at the F1 ledge maximum will decrease faster than that at levels below it. This will cause a decrease in the gradient $\frac{dN}{dh}$ and will therefore result in a more pronounced cusp

on an ionogram obtained at this time.

6.6.4 Changes in the loss rate.

The changes in the production rate are not the only factors responsible for the enhancement of the F1 ledge during eclipses. The electron loss rate, depending on the concentrations of electrons and on the loss law operating, also contributes.

Let us first consider the case of a hypothetical F1 layer formed in a region where a quadratic type of loss law is acting below a certain level and a linear type above the level (see section 6.7).

Above this level we shall denote the electron density by N_1 , its rate of change by $\frac{dN_1}{dt}$ and the loss coefficient by γ . Below the level the electron density will be denoted by N_2 , its rate of change by $\frac{dN_2}{dt}$ and the loss coefficient by α .

Under equilibrium conditions, with the same production rate and the same initial electron density above and below the transition level, the resulting loss rates will be equal. This is an oversimplified situation but serves to illustrate the argument.

Consider what happens if the ionizing agent is removed. Measuring the time t from this instant, we find:-

$$\begin{aligned} \text{Above the transition level: } \frac{dN_1}{dt} &= -\gamma N_1 \\ &: \frac{dN_1}{N_1} = -\gamma dt. \end{aligned}$$

78.

$$\text{Integrating} \quad \ln N_1 = -\gamma t + C''$$

$$\text{If at } t = 0, N_1 = N_0 \quad \ln \frac{N_1}{N_0} = -\gamma t$$

$$\text{and} \quad N_1 = N_0 \cdot \exp(-\gamma t) \quad (6.1)$$

$$\text{Below the transition level:} \quad \frac{dN_2}{dt} = -\alpha N_2^2$$

$$\therefore \frac{dN_2}{N_2^2} = -\alpha dt$$

$$\text{and therefore} \quad \frac{1}{N_2} = \alpha t + C'''$$

$$\text{Again, if } N_2 = N_0 \text{ at } t = 0 \quad \frac{1}{N_2} - \frac{1}{N_0} = \alpha t$$

$$\text{so that} \quad N_2 = \frac{1}{\alpha t + \frac{1}{N_0}} \quad (6.2)$$

For the same initial loss rates

$$\left(\frac{dN_1}{dt} \right)_{t=0} = -\gamma N_0 = \left(\frac{dN_2}{dt} \right)_{t=0} = -\alpha N_0^2$$

$$\text{We therefore find } \alpha = \frac{\gamma}{N_0}$$

Substituting for α in (6.2) yields

$$\begin{aligned} N_2 &= \frac{1}{\frac{\gamma}{N_0} t + \frac{1}{N_0}} \\ &= \frac{N_0}{\gamma t + 1} \end{aligned} \quad (6.3)$$

Dividing (6.3) by (6.1) yields

$$\begin{aligned} \frac{N_2}{N_1} &= \frac{\exp(\gamma t)}{1 + \gamma t} \\ &= \frac{1 + \gamma t + \frac{1}{2!}(\gamma t)^2 + \text{positive terms}}{1 + \gamma t} \\ &= 1 + \text{positive terms} \end{aligned}$$

That means $\frac{N_2}{N_1} > 1$ for all values of t .

We see that, under the conditions described, the electron density in the quadratic loss law region will become greater and stay greater than N in the linear loss law region after the ionizing agent is removed.

In the ionosphere the ionizing agent is not removed completely at first contact, but if it decreases rapidly, as it does on this model, the electron density will be determined largely by the loss rate. If the F1 maximum occurs above the transition level, it will decrease faster than the electron density at heights below the transition level and so cause $\frac{dN}{dh}$ to diminish. This results in the enhancement of the f_oF1 cusp on an ionogram obtained at this time. The explanation offered here is basically the same as that advanced by RATCLIFFE (1956,b).

Another additional effect enters in the particular situation that arises on the present model.

At 0700, just after first contact, the F1 layer was barely

visible as a small ledge at the base of the F2 layer. The electron density at and above the F1 maximum had built up to relatively large concentrations in comparison with that in the lower F1 layer. At these higher levels N would therefore be nearer its equilibrium value than below the F1 maximum. A greater loss rate would result at these levels, due to the greater ion and electron densities, than in the lower F1 layer. The resulting distortion in the layer shape would decrease $\frac{dN}{dh}$ at the point of inflection and so cause the F1 cusp to appear more pronounced.

The effect described here would cause the electron density at and above the F1 maximum to decrease to a greater extent than the mechanism mentioned previously could alone account for. If we may refer to the mathematical argument, the situation described here would correspond to N_0 in the linear loss region being greater than the initial electron density in the quadratic region, causing a greater initial loss rate in the linear loss region.

An effect that would counteract the three causes mentioned for the decrease in $\frac{dN}{dh}$ at the F1 maximum is that due to the decreases in the effective loss coefficient with increase in height. In such a region the electron density will decrease faster at the lower levels when the ionizing agent is removed. This is prevented from happening on the present model by the change-over, at a level somewhere in the F1 layer, from a region in which the loss coefficient decreases with height, to one in which it is independent of heights below the level.

Which of these processes are the most important?

6.6.5 The nett rates of change of ion and electron densities.

In an attempt to find out which of these processes are most important during the various stages of the eclipse, the nett rates of change in the density of ions and electrons were plotted against the height z . These plots are shown in Figures 6.17 and 6.18 for the times indicated (These figures can also be found at the back of this thesis). It should be remembered that these nett rates of increase are determined by q and the loss rates.

Consider the distribution of the rates of change at 0700. At $z = -0.5$ there is a relatively large rate of increase in both electron density and OY^+ ions. In this region then, the production must be greater than the loss rate. At the level $z = 0.8$ the nett rate of increase is smaller but the loss rate here is neither predominantly due to the loss rate of O^+ ions nor to that of OY^+ ions. The linear-quadratic loss transition level mechanism can therefore not be largely responsible for the enhancement of the F1 ledge during the early stages of the eclipse. The larger loss rate at the level of the F1 maximum, due to a greater concentration of ions and electrons and the change in the distribution of the production rate, is probably responsible for the enhancement during the period immediately after first contact.

At 0715 these processes are no longer the only significant ones, because in the region between about $z = 0.5$ and $z = 1.5$ there now occurs a nett loss rate of electrons which is largely balanced by a loss rate in O^+ ions. This means that the process for the enhancement of the F1 ledge described first in section

6.5.4 (the one that was argued mathematically) is probably operating.

During the period that starts at approximately 0710 and lasts until just after 0745 the enhancement of the F1 ledge is probably increasingly due to the two processes that arise as a result of the rapid decrease in q so that the enhancement of the F1 ledge is determined largely by the loss rates.

6.6.6 Conclusion.

The phenomenon of the enhancement of the F1 ledge during an eclipse is well represented by the non-equilibrium model. It appears to arise for three reasons, viz. (a) a change in the height distribution of the production rate, (b) a greater loss rate in the region immediately above the F1 maximum due to a linear loss law and (c) a greater loss rate due to greater concentrations of ions and electrons. The relative importance of the linear loss law effect in this respect seems to increase as the first half of the eclipse progresses.

6.7 The upward movement of the base of the F2 layer.

In the period 0715 to 0815 a feature noticeable on the $N(z)$ curves in Figure 6.2 is an upward movement of the base of the F2 layer. The cause of this becomes clear through a consideration of the rates of change of the ion and electron densities during this period, shown in Figure 6.17. On these figures the peak (negative, of course) in the loss rate distribution that occurs at about $z = 1$ at 0715 moves up to $z = 2$ at 0815. Now, a great

enough reduction in q will result in a decrease in the electron and ion densities. Because of the argument presented in section 6.5.4, this means that the loss rate in the linear loss law region gradually becomes greater than that in the quadratic law region. The peak loss rate would therefore be expected to move into a region in which the linear loss law is more predominant, that is upwards, as shown.

A nett increase in O^+ ions and electrons occurs below $z = 1$ at 0815. This is the beginning of the formation of a point of inflection on the $N(z)$ and $n(O^+)$ -h profiles and we consider its development in detail in the next section.

6.8 The $f_oF1\frac{1}{2}$ critical frequency.

6.8.1 Introduction.

We have mentioned that a cusp, corresponding to an additional critical frequency $f_oF1\frac{1}{2}$ between the usual f_oF1 and f_oF2 critical frequencies, has often been observed on ionograms taken during the later part of some eclipses. Under these conditions the $f_oF1\frac{1}{2}$ cusp has been noted to move to higher levels and greater frequencies. This critical frequency has been associated with an independent layer (SHAPLEY, 1956) and has also been interpreted as an effect due to oblique reflections caused by travelling ionospheric disturbances and not as a result of genuine stratifications (HEISLER, 1962). It is observed more often than only during eclipses in certain parts like Japan where it is said to be always observed at sunrise (AONO, 1956).

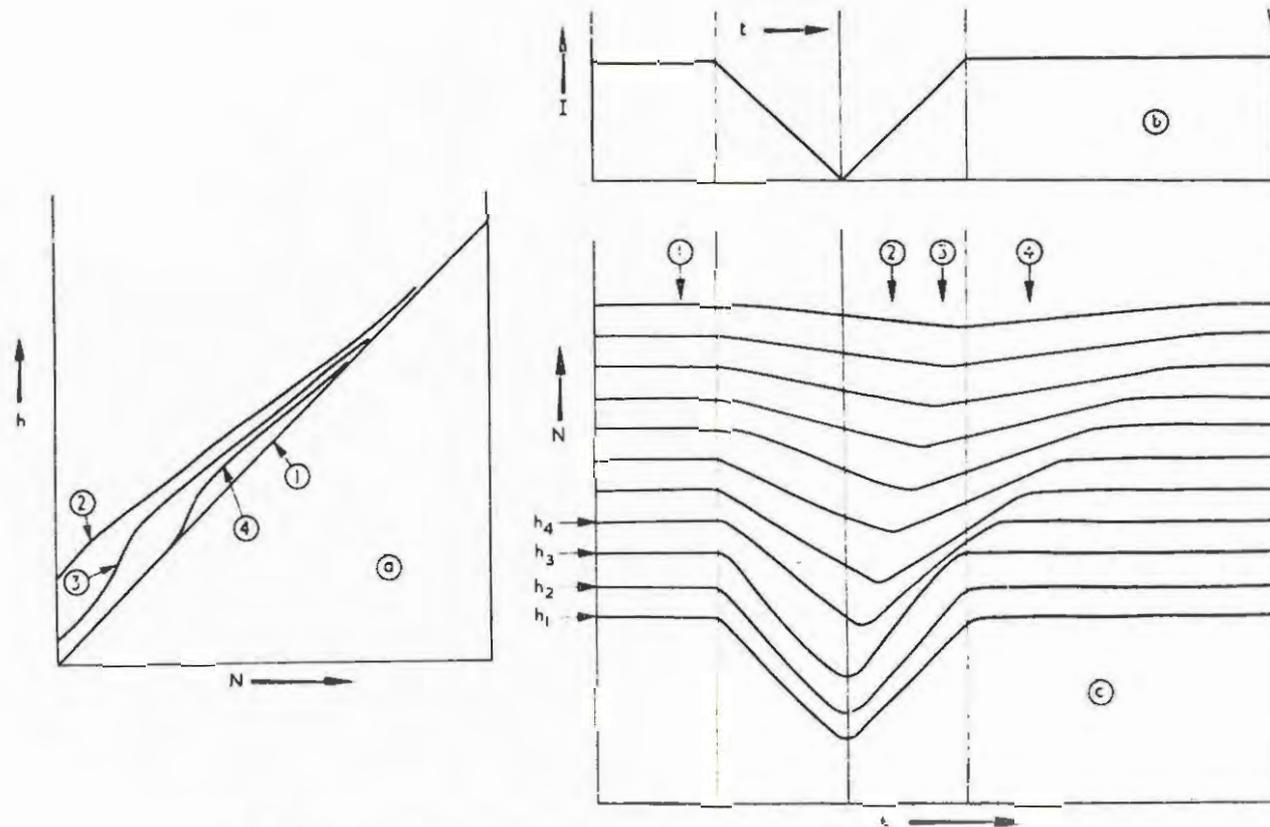


Fig.6.19 To illustrate a possible way in which the extra critical frequency $f_0 F I \frac{1}{2}$ could arise.

In this section we shall consider the $f_o Fl_{\frac{1}{2}}$ cusp as it appears on the model during an eclipse and attempt an explanation for its occurrence under the conditions represented by the non-equilibrium model proposed.

6.8.2 The point of inflection that gives rise to the $f_o Fl_{\frac{1}{2}}$ cusp.

By noting the critical frequency $f_o Fl_{\frac{1}{2}}$ on the ionogram for 0925 in Figure 6.11 we can determine the co-ordinates of the point of inflection on the $N(z)$ curve in Figure 6.3 that gives rise to it. This point of inflection appears to arise out of the $F1$ inflection point at about 0920 and it can be identified at different times during the period 0925 to 0950. It moves to greater electron densities and greater heights after it initially appears just below $z = 1.5$ where it is barely distinguishable from the point of inflection associated with the $F1$ ledge. The $f_o Fl_{\frac{1}{2}}$ cusp on ionograms that results from it, should therefore behave in exactly the same way as the observed $f_o Fl_{\frac{1}{2}}$ cusp does during eclipses.

6.8.3 Identification of the species of positive ions involved.

The $F1_{\frac{1}{2}}$ inflection point in the $N(z)$ distribution should be reflected in the positive ion distribution for the corresponding times. On examining the $n(O^+)-z$ and $n(OY^+)-z$ curves in Figures 6.5, 6.6 and 6.8 it becomes evident that the O^+ ion distribution alone shows a similar point of inflection. This is particularly well shown in the curves for 0925 that appear in Figure 6.9. In this figure a point of inflection occurs at $z = 2$ in both the $N(z)$ curve and the $n(O^+)-z$ curve while none appears in the $n(OY^+)$ curve.

Remembering that the continuity equation for $n(O^+)$ can be solved without referring to the electron density or OY^+ ion concentration, it follows that the O^+ ions are responsible for the $Fl_{\frac{1}{2}}$ point of inflection arising in the $N(z)$ distribution. The occurrence of such an inflection point should therefore be explainable by considering only the changes taking place in the O^+ ion concentration.

We proceed to consider the effects of an eclipse on a layer of electrons in which the effective loss coefficient decreases with height like γ , the effective loss coefficient of the O^+ ions, does in the present non-equilibrium model.

6.8.4 The explanation for an $Fl_{\frac{1}{2}}$ point of inflection proposed by Ratcliffe.

RATCLIFFE (1956) has considered roughly the formation of an inflection point on an $N(h)$ curve in a region in which the effective electron loss coefficient decreases with height. We shall consider his argument here because it offers a basis for explaining the results obtained from the non-equilibrium model. (Figure 6.19 is taken from Ratcliffe's paper).

Consider an equilibrium layer of electrons distributed linearly with height as shown at 1 in Figure 6.19 (a). For the argument's sake the production rate of electrons, q , is assumed to be the same at all heights and to change in the same way at all levels during an eclipse. It may therefore be represented by a $q(t)$ curve such as that in Figure 6.19 (b). If $N(t)$ curves are plotted for

different heights h_1, h_2, h_3, \dots the resulting equal height contours will appear as shown at (c).

At low levels such as h_1, h_2 and h_3 the minimum in electron density occurs at the same time as maximum obscuration because at these levels equilibrium effectively exists between the production and loss rates at all times, due to the large electron loss coefficient. In other words, at these levels the electron density would decay with a very small time constant if the production rate dropped to zero so that the ionospheric layer would exhibit hardly any "sluggishness". The electron density at these heights therefore follows the change in the production rate closely. Due to this no crowding occurs at any stage.

At higher levels the loss coefficient becomes progressively smaller. Consequently the decrease in the electron density is increasingly dependent on the loss rate and does not follow the change in q so closely any longer.

The following rough argument illustrates this effect by analogy: Consider some electrical network in which the output voltage is determined by the magnitude of an applied signal. If the time constant of the circuit is short in comparison with the time constant of any variations in the applied signal, the output voltage will vary in a way similar to that of the applied signal. If, on the other hand, this is not the case, the output voltage will in some measure depend on the time constant of the circuit. This dependence will become more marked as the circuit time con-

stant is increased.

To illustrate the analogy let us take the applied signal to correspond to q and the output voltage to the electron density. At low levels where the characteristic time constant of the variation in the electron density is short, i.e. where short electron relaxation times obtain, the electron density will follow changes in q closely. As the height is increased the change in the electron density will depend more and more on the magnitude of the electron loss coefficient.

Returning to the $N(t)$ curves in Figure 6.19 (c) the minimum in the electron density can be seen to occur at later times at higher levels and is greatest where the loss coefficient is least. This is a consequence of the fact that at the highest levels the electron relaxation time is so large that a reduction in q causes a very gradual decrease in N . Accordingly, by the time that q has reached a minimum the electron density has not departed much from its value before obscuration.

When q starts increasing after maximum obscuration the difference between it and the loss term will at first be negative because N is still relatively near its value before obscuration. Hence the electron density is still decreasing at this time even though q is increasing. As q becomes greater the difference between it and the loss term increases from a negative value to a small positive magnitude and so causes a small rate of increase in N . In other words, the largest departures from equilibrium

occurs at the high levels. At descending heights this departure from equilibrium becomes progressively less and disappears at the lowest levels.

If we now take "cross-sections" of Ratcliffe's $N(t)$ map at the times indicated by 2, 3 and 4 the $N(h)$ curves can be plotted and appear like those marked 2, 3 and 4 in Figure 6.19 (a). As Ratcliffe points out, a point of inflection appears on these curves. This point of inflection moves to higher levels and greater electron densities with time, thus behaving in a way similar to that of the observed $F1\frac{1}{2}$ layer.

6.8.5 Application of this explanation to the present model.

Returning to the results of the present model, the similarity between the situation depicted in Figures 6.19 (c) and 6.4 and correspondingly in Figures 6.19 (a) and 6.6 leads us to believe that the explanation given by Ratcliffe (section 6.7.4) is also applicable to the occurrence of an inflection point in $n(O^+)-z$. Allowances for the differences between the models must, however, be made.

For instance, before the eclipse starts Ratcliffe's model shows an equilibrium $N(h)$ distribution while non-equilibrium obtains on the present model. A result of this is that $n(O^+)$ at about $z = 2.5$ does not stay nearly constant but increases at a nearly constant rate. Secondly, $n(O^+)$ below about $z = 0$ does not fall to zero like N does at the lowest levels in Ratcliffe's model. The reason for this is the fact that the obscuration of the sun on

25 December, 1954 did not reduce q to zero. If allowance is made for these differences the similarity between the two situations is rather close.

6.8.6 An alternative viewpoint.

Another way of looking at the formation of a point of inflection on the $n(O^+)$ curve may be helpful. In Figure 6.17 the $\frac{\Delta n(O^+)}{\Delta t}$ - z curve at 0815 shows a positive peak at about $z = 0.4$. This peak subsequently moves to greater heights. What effect would such an upward movement of the peak have on the $n(O^+)$ - z distributions?

At 0830 the peak in $\frac{\Delta n(O^+)}{\Delta t}$ - z is near $z = 1.0$. During the period 0830 to 0840 the O^+ concentration at levels near $z = 1.0$ should accordingly show a larger increase than at other levels. This results in the formation of a point of inflection on the $n(O^+)$ - z curve at about $z = 1.0$. Such a point of inflection can be seen on the $n(O^+)$ curve for 0840 in Figure 6.6.

This same process continues at successively higher levels at later times as the peak in $\frac{\Delta n(O^+)}{\Delta t}$ - z moves upwards. Consequently the point of inflection also moves up along the $n(O^+)$ - z curve and so shows the same type of behaviour as that of the $Fl_{\frac{1}{2}}$ inflection point on the $N(h)$ curve described in section 6.7.4.

The reason for the existence of a peak in the $\frac{\Delta n(O^+)}{\Delta t}$ - z curve needs clarification. Take as an example the $\frac{\Delta n(O^+)}{\Delta t}$ - z curve at 0930 in Figure 6.18. The peak in this curve occurs at about $z = 2.0$. Below this level the O^+ ion density has built up to

such a magnitude that the loss rate is nearly as large as the production rate, with the result that the nett rate of increase is small. Above this level the production rate falls off with increase in height due to the Chapman distribution of q . The resulting rate of increase is therefore once again not as great as that at about $z = 2.0$.

The argument can be illustrated to some extent by considering the $N(z)$, $n(O^+)-z$ and $n(OY^+)-z$ distributions in Figure 6.20. These distributions are shown for 0930 on an eclipse and a normal day. At levels below $z = 2.0$ the ion and electron densities have built up to near normal day values. The resulting nett rate of increase is more or less the same as on a normal day, a rate of increase which is not very great at 0930. As mentioned, the decreasing value of q with increase in height above $z = 2.0$ now results in a peak in the $\frac{\Delta n(O^+)}{\Delta t} -z$ distribution at about $z = 2.0$.

6.8.7 The electron density.

The assumption of charge neutrality at all levels in the ionosphere at all times, requires the O^+ ion concentration to be balanced by at least an equal electron concentration. Accordingly, if the $F1\frac{1}{2}$ inflection point in the $n(O^+)$ curve occurs at a level some distance above the transition region (see section 6.9) we would expect it to be reflected in the $N(z)$ distribution. This would follow as a result of the fact that above the transition region the loss of electrons is limited by the reaction



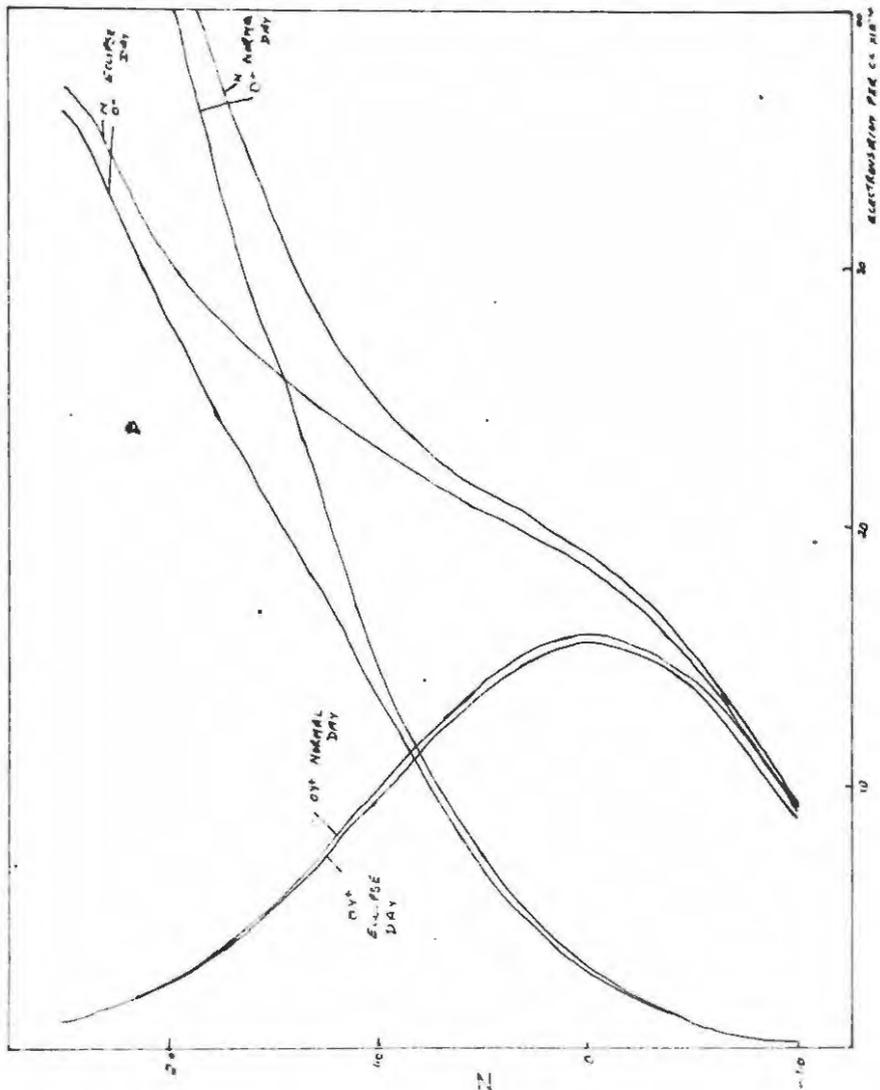


FIG. 6.20 ION AN ELECTRON DISTRIBUTIONS AT 0930 EAST FOR A NORMAL AND AN ECLIPSE DAY.

The electrons are therefore lost with an effective loss coefficient which is equal to γ , the loss coefficient of the O^+ ions. Any phenomena specifically due to the variation in γ should therefore also be evident in the $N(z)$ distribution at these levels.

At heights below the transition region, however, N is mainly balanced by $n(O^+)$. Consequently, when the peak in the $\frac{\Delta n(O^+)}{\Delta t}$ - z curve, giving rise to the formation of the $Fl_{\frac{1}{2}}$ point of inflection, occurs in and below the transition region we would expect the $Fl_{\frac{1}{2}}$ inflection point on the $N(z)$ curve to be increasingly obscured as we proceed to lower levels.

An inspection of the $\frac{\Delta N}{\Delta t}$ - z curves in Figures 6.17 and 6.18 shows that the peak in these curves is much less pronounced than the peak in the $\frac{\Delta n(O^+)}{\Delta t}$ - z curves during the period 0815 to 0845, a period during which the peak occurs in and near the transition region. Accordingly, we would expect the $Fl_{\frac{1}{2}}$ point of inflection on the $N(z)$ curves to be less clearly defined than that on the $n(O^+)$ - z curves during this period. A comparison of the curves in Figures 6.3 and 6.6 shows this to be the case. The point of inflection of the $N(z)$ curve that gives rise to the $f_0 Fl_{\frac{1}{2}}$ cusp appears separated from the Fl inflection point only after 0910, while the $Fl_{\frac{1}{2}}$ inflection point on the $n(O^+)$ curve becomes discernible as early as 0840.

The above argument is clearly illustrated by considering the $\frac{\Delta n(O^+)}{\Delta t}$ - z and $\frac{\Delta N}{\Delta t}$ - z curves at 0845 in Figure 6.18. At this time a point of inflection already occurs in the $n(O^+)$ - z curve in

Figure 6.6 but not yet on the $N(z)$ curve in Figure 6.3.

6.8.8 Discussion.

The non-equilibrium model presented in this thesis shows an $Fl_{\frac{1}{2}}$ point of inflection in the $N(z)$ curves during the later part of the eclipse discussed. The resulting $f_0 Fl_{\frac{1}{2}}$ cusp behaves similarly to $f_0 Fl_{\frac{1}{2}}$ cusps observed during the later parts of eclipses.

By the arguments presented in this section we are led to believe that the explanation for the occurrence of such a point of inflection offered by RATCLIFFE (1956) is probably correct. The contention of HEISLER (1962) that the $f_0 Fl_{\frac{1}{2}}$ cusp originates from oblique reflections from travelling disturbances is not substantiated by the results obtained from the present model.

A discussion of the conditions under which the $Fl_{\frac{1}{2}}$ point of inflection occurs is deferred to Chapter 7.

6.9 Variation in height of the transition level.

Consideration of the $\frac{\Delta N}{\Delta t} -z$, $\frac{\Delta n(O^+)}{\Delta t} -z$ and $\frac{\Delta n(OY^+)}{\Delta t} -z$ curves at 0645 in Figure 6.17 shows a definite point at which the $\frac{\Delta n(O^+)}{\Delta t} -z$ curve crosses over the $\frac{\Delta n(OY^+)}{\Delta t} -z$ curve. Above this level the loss rate of electrons is predominantly determined by the loss rate of O^+ ions. This results from the fact that the electrons are lost through recombination with OY^+ ions which are produced at the same rate as O^+ ions are lost. The $\frac{\Delta N}{\Delta t} -z$ curves therefore tend to merge with the $\frac{\Delta n(O^+)}{\Delta t} -z$ curves with increase in height, as the

loss rate of O^+ ions becomes more and more limiting.

Below the mentioned level the loss rate of O^+ ions is large enough to provide sufficient OY^+ ions to cause the limiting process to be the dissociative recombination of electrons and OY^+ ions,



Since the rate of production of OY^+ ions increases with decrease in height the $\frac{\Delta N}{\Delta t}$ -z curves tend to merge with the $\frac{\Delta n(OY^+)}{\Delta t}$ -z curve at lower levels.

It would seem that a definition of the transition level as that level at which the $\frac{\Delta n(O^+)}{\Delta t}$ -z curve crosses over the $\frac{\Delta n(OY^+)}{\Delta t}$ -z curve might be satisfactory. Such a definition, however, turns out to be ambiguous for the following reason: On inspecting the graphs for O715, a cross-over in these curves occurs at $z = 1.7$ while the loss of O^+ ions can be seen to be the limiting reaction down to a level $z = 0.4$. At this height no cross-over between the lines occurs, but the $\frac{\Delta N}{\Delta t}$ -z curve gets closer to the $\frac{\Delta n(OY^+)}{\Delta t}$ -z curve than the $\frac{\Delta n(O^+)}{\Delta t}$ -z curve does as the height is decreased.

It seems preferable, therefore to define a transition region rather than a transition level. This region is defined as that between the (higher) level where the rate $\frac{\Delta n(OY^+)}{\Delta t}$ is 10% of $\frac{\Delta N}{\Delta t}$, and the (lower) level where $\frac{\Delta n(O^+)}{\Delta t}$ is 10% of $\frac{\Delta N}{\Delta t}$. The change in this region is roughly indicated in Figure 6.21.

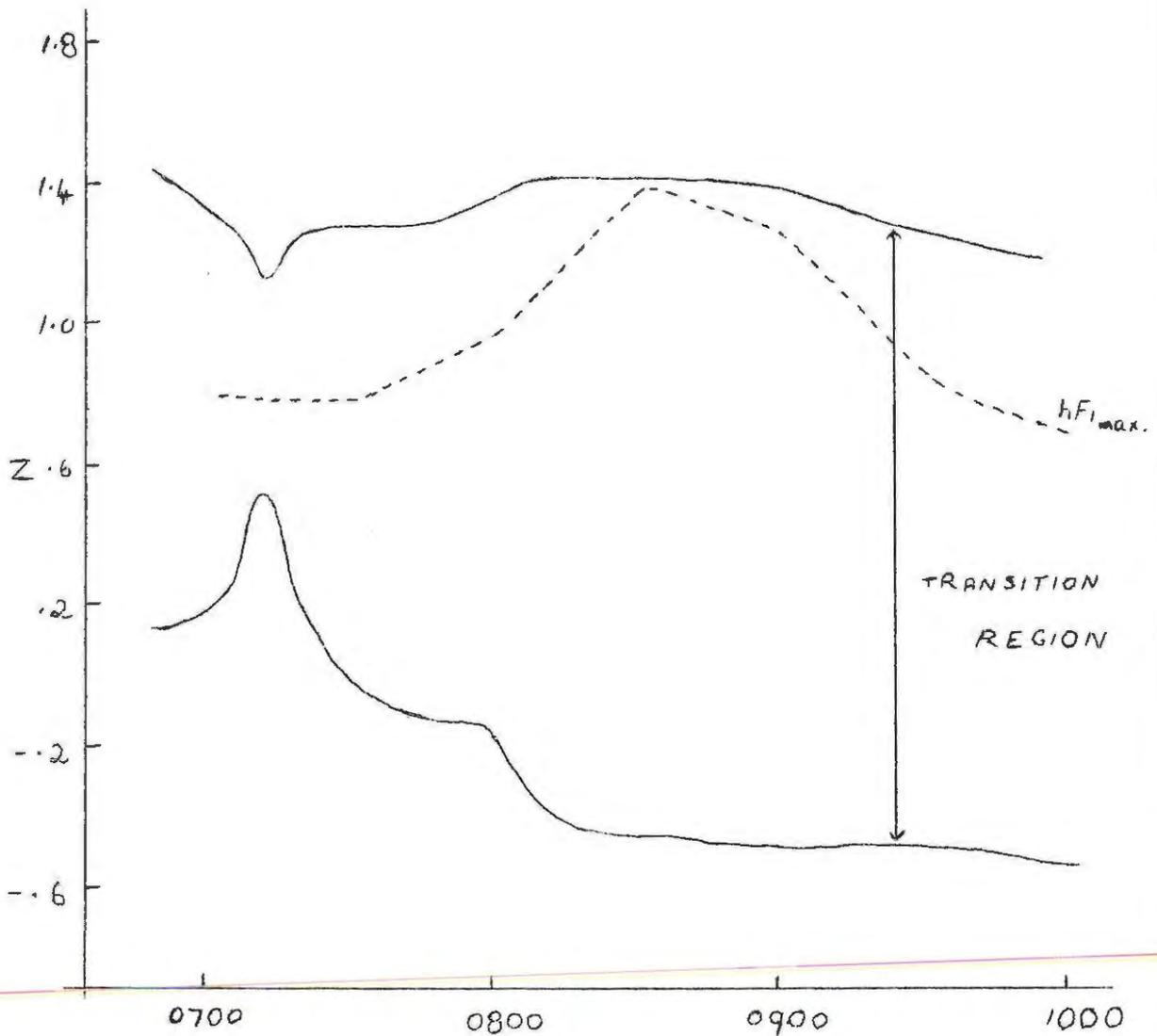


Figure 6.21. Variation in height of the transition region and the height of the F1 maximum.

The position of the F1 ledge maximum during this period is indicated by a dotted line. An increase in the height of this

inflection point occurs just after maximum obscuration, i.e. just in the period during which the $N(t)$ curve shows a minimum. Figure 6.21 makes it clear that the electron density at the F1 maximum should behave in a way much closer to that of a region obeying a linear loss law during the maximum phase of an eclipse than it does during the early and late phases.

6.10 The $N_{\min} - \delta t$ anomaly during an eclipse.

In the analysis of eclipse results it has been usual to assume a purely quadratic loss law for a single species of positive ions in the E and F1 regions. The procedure often adopted would be to assume a value of α and then to compute q from the observed variation of N on control days through the equation

$$q = \frac{dN}{dt} + \alpha N^2.$$

The continuity equation could then be solved step by step with the appropriate production rate f_q to yield the variation of N at the F1 maximum during the eclipse period.

An interesting result in this respect was the following: If the time at which the minimum in N was made to coincide with that observed, the electron density minimum was found to be much too low. It turned out to be impossible to reproduce the observed N_{\min} and the time at which it occurred assuming the same value of α . This has been referred to as the $N_{\min} - \delta t$ anomaly.

By considering two positive ion species BATES and McDOWELL (1957) have shown that different quadratic recombination loss

coefficients for the two species could probably account for the $N_{\min} - \delta t$ anomaly. BOWHILL (1961) has extended this approach to consider more than two species of ions.

It should be pointed out that analyses of eclipse results on the F1 layer should not only take account of more than one species of positive ion, but also of the two different loss laws.

To see what sort of effects are to be expected if this is done, let us consider the variation of the electron density at the F1 maximum as it appears on the model. From Figure 6.21 we obtain the approximate level of the F1 maximum at various times during the eclipse. By referring to Figure 6.1 we can then obtain the variation in N at the F1 maximum that follow.

Between 0700 and 0730 the F1 maximum occurs near $z = 0.75$. Accordingly, the $z = 0.75$ curve in Figure 6.1 should approximately indicate the variation in $N_{F1_{\max}}$. Between 0730 and 0800 an increase in the height of the F1 maximum from about $z = 0.75$ to about $z = 1.0$ takes place on the model. $N_{F1_{\max}}$ should therefore not decrease as rapidly as indicated for the level $z = 0.75$ during this period, but should be close to that indicated by the $z = 1.0$ line at 0800.

During the period 0800 to 0830 the height of the F1 maximum increases from about $z = 1.0$ to about $z = 1.4$. Let us consider two consequences of this.

In the first place the electron density does not fall to such

an extent as it would have, had the F1 maximum stayed at the $z = 1.0$ level. This follows necessarily from the fact that N at the levels above $z = 1.0$ does not attain such low values at minimum as N at $z = 1.0$ does.

The second consequence of the increase in height of $N_{F1_{max}}$ during this period is that the minimum in N at the F1 point of inflection occurs at an earlier time. The argument may be illustrated by making use of diagrams.

Consider the situation in which the electron density at the levels $z = 1.0$ and $z = 1.25$ stays constant during the period in which the height increase of the F1 maximum takes place. The minimum in N then occurs at the beginning of the period considered. This is shown in Figure 6.22 in which $N_{F1_{max}}$ is represented by a broken line and N at $z = 1.0$ and $z = 1.25$ by solid lines. The arrow at 0800 indicates the time at which the minimum in $N_{F1_{max}}$ occurs.

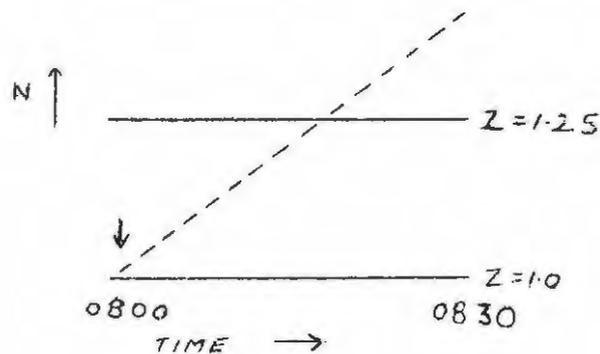


Figure 6.22. Hypothetical $N(t)$ curves showing the variation in $N_{F1_{max}}$.

The situation that arises when a minimum in N occurs at all levels during the period considered is illustrated in Figure 6.23. (The same symbolization as in Figure 6.22 is used).

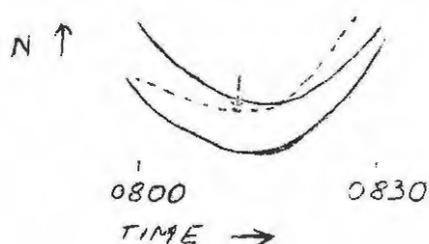


Figure 6.23. Hypothetical $N(t)$ curves for a period during which a minimum in the electron density occurs.

The changes that result when the variation in height on the F1 maximum during a period in which a minimum in N occurs is taken into account, is seen to be such that it could probably to some extent account for the $N_{\min}^{-\delta t}$ anomaly.

Another interesting consequence of the change in height of the F1 maximum might be mentioned in passing. After 0830 a lowering of the height of the F1 maximum occurs. This causes the final recovery of N_{\max}^{F1} to be less rapid than it would have been, had the $F1_{\max}$ stayed at a constant height. It is illustrated in Figure 6.24 in which the relevant portions of the $N(t)$ curves in Figure 6.1 are indicated by solid lines. The broken line shows the probable variation of N_{\max}^{F1} when the changes in height of the F1 maximum is taken into account.

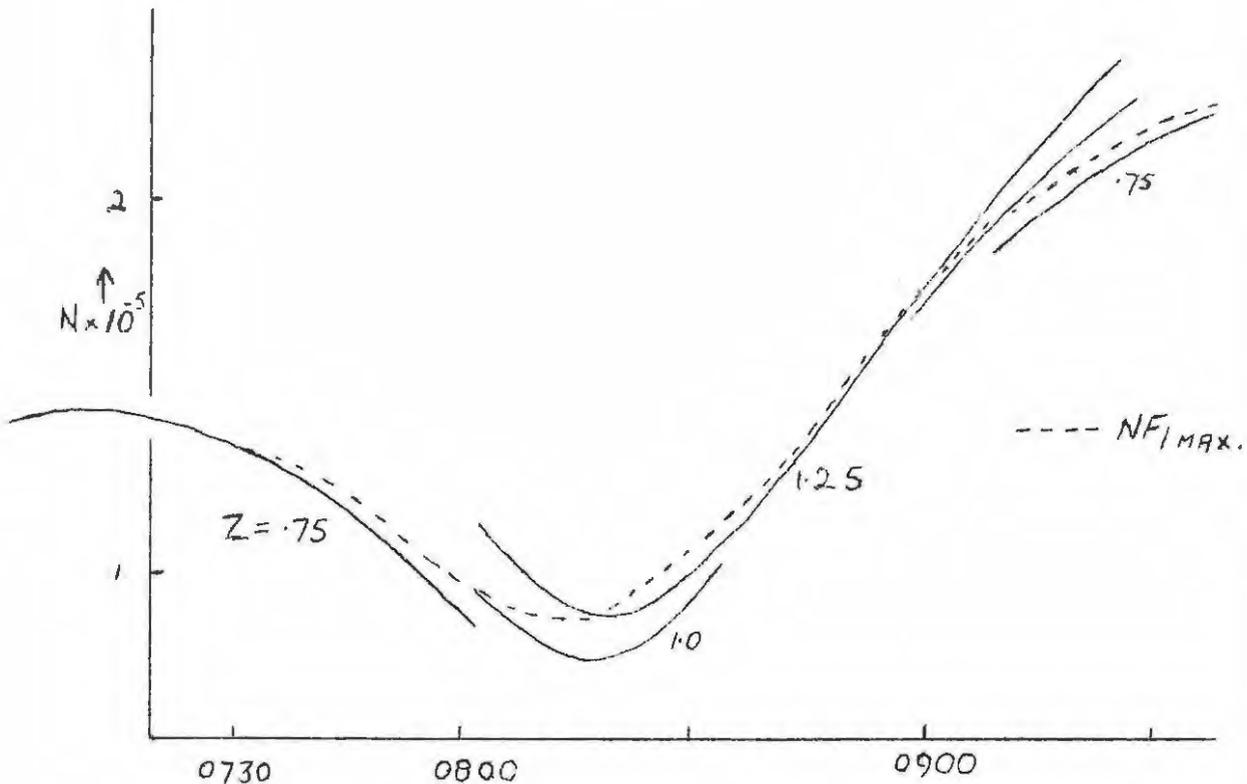


Figure 6.24. Portions of the $N(t)$ curves from Figure 6.1. The broken line indicates the probable variation in $NF_{1 \text{ MAX.}}$

6.11 Discussion.

The eclipse results obtained from the non-equilibrium model presented in this thesis show phenomena such as the enhancement of the F1 layer and the occurrence of an $F1\frac{1}{2}$ layer rather well.

The explanation for the enhancement of the F1 layer in section 6.6 applies mainly to an eclipse that occurs before noon. The distributions of the rates of change in the electron and positive ion densities that were used in this explanation were also used in the explanation for the $F1\frac{1}{2}$ layer, and similarly apply mainly to conditions before noon. RATCLIFFE (1956), in his survey of

eclipse results, records the times at which the eclipses, during which the $F1\frac{1}{2}$ layer was observed, occurred. It might be significant to mention that he failed to draw attention to the fact that all but one of these eclipses occurred before noon. The eclipse that did not take place before noon occurred at noon and was recorded in Japan where, in any case, the $F1\frac{1}{2}$ layer is observed at times other than during eclipses.

For a discussion of the general conditions under which the present model may be applicable we turn to the next chapter.

CHAPTER 7.DISCUSSION.7.1 The problem.

The non-equilibrium model presented in this thesis shows the enhancement of the F1 layer and the occurrence of an $F1\frac{1}{2}$ point of inflection during an eclipse very well. It does not, however, show the same behaviour as the ionosphere showed over Grahamstown on the 25th December, 1954. Under what conditions does the $f_o F1\frac{1}{2}$ cusp then appear on ionograms?

7.2 The conditions under which the $F1\frac{1}{2}$ layer appears.

RATCLIFFE (1956) has investigated this and reports that all the occurrences of the $F1\frac{1}{2}$ layer, except those reported in Japan and one other, were observed in a region in which the dip angle is less than 20° .

It is well known that in this region the F2 peak appears at greater heights than at middle latitudes. The reason is, presumably, that ambi-polar or plasma diffusion of the ionization in the F2 layer is greatly inhibited by the magnetic field which is approximately horizontal there, thus keeping the F2 peak from descending. It should be noted that ordinary diffusion would not be affected in this way. This means that the value of k , which is large for diffusive separation of the ionizable constituent and the molecules which enter into the ion-atom interchange process, will be determined by other factors. If the F2 peak does not descend, the electron density distribution in the lower part

of the F2 layer will be largely controlled by the loss coefficient here. Consequently any effects specifically due to a decrease in the loss coefficient, such as the formation of the $F1\frac{1}{2}$ point of inflection, should be particularly noticeable at low latitudes.

7.3 Middle latitudes.

In middle latitudes the effects of plasma diffusion should be much more noticeable. It, in fact, results in the lowering of the F2 peak to about 2 to 3 scale heights above the F1 maximum. The effects of diffusion should therefore presumably also be detectable in the lower part of the F2 layer. This would probably obscure the formation of an $F1\frac{1}{2}$ point of inflection in the base of the F2 layer in these latitudes. The reason for this is the well known fact that diffusion smooths out irregularities that might occur in the distribution of ionization.

BRIGGS and RISHBETH (1961) describe an analogue computer for the solution of the electron density continuity equation with a (linear) loss coefficient that decreases with height. In one experiment they investigated the changes that result in the $N(h)$ distribution when the production rate varies like it does during an eclipse. They reported no $F1\frac{1}{2}$ point of inflection occurring during the later part of the eclipse. This can probably be explained by the fact that their analogue took diffusion into account.

7.4 The results of the model and McElhinny's observations.

The very pertinent problem of reconciling the results from the model with McElhinny's observations remains.

It has been mentioned that the observed F1 layer showed a more pronounced maximum during the control day period than that which occurs on the model. This may account for part of the discrepancy between the model and the observations during the eclipse for the following reasons: GLEDHILL (1959), in investigating the effects of an eclipse on a model ionosphere, assumed a distribution of ionization very near to that observed by McElhinny, i.e. with a pronounced F1 peak occurring in the N(h) profile. To obtain this distribution under equilibrium conditions a gradient in the effective recombination coefficient was assumed that resulted in a 200-fold increase in the magnitude of α between 300 and 180 km. The solution of the continuity equation for electron density resulted in the formation of a valley between the F1 and F2 peaks. From the work of GLEDHILL and WALKER (1960) it is fairly certain that such a valley did occur above Grahamstown during the eclipse of the 25th December, 1954. The model considered by Gledhill therefore leads to the sort of behaviour of the F1 layer that was observed by McElhinny.

By adapting the present model to yield a distribution closer to that observed, i.e. showing a more pronounced F1 peak, closer agreement with McElhinny's observations could be expected. This was not done in the present case because the model proposed in this thesis leads, among other effects, to the formation of an $F1\frac{1}{2}$ point of inflection. This has its own intrinsic interest and was investigated in the present work.

At least three different ways of bringing about a more pro-

nounced F1 peak on the present model can be suggested.

Firstly, from the work of HIRSH (1959) it is clear that a larger value of the constant $C = \frac{4\gamma_0}{q_0^\alpha}$ will result in a more pronounced F1 ledge. Smaller values of q_0 and α and larger values of γ_0 could bring this about. At present the values of these constants are not accurately enough known to decide which change would be more appropriate. The resulting $N(h)$ distributions would, however, show an increase in N only at great heights above the F1 peak. To illustrate this the $N(z)$ distributions resulting from two choices of C are illustrated in Figure 7.1.

Figure 7.1. $N(z)$ distributions that result from two choices of C .

For such large values of C the F2 peak would occur at great heights. By taking account of diffusion this peak could probably be made to occur at lower levels.

Secondly, a larger gradient in the effective recombination

coefficient would cause a more pronounced F1 ledge. It has been mentioned that a pronounced F1 peak appeared on Gledhill's model in which α increases 200-fold between 300 km and 180 km. The corresponding increase in γ on the model presented here is less than 100-fold. This could be increased by taking $k = 2$, which implies complete diffusive separation between the ionizable constituent (oxygen atoms) and the molecules XY. It also implies the identification of the constituent XY with O_2 . The assumption of complete diffusive separation is hardly justifiable in view of the fairly unassailable reasons put forward by RATCLIFFE (1959) in his discussion of the behaviour of the F2 layer.

In this respect it should be mentioned that a similar more pronounced F1 peak might result if a gradient in the value of the dissociative recombination coefficient α occurs. This could be brought about by a possible temperature or pressure dependence of α . A similar temperature or pressure dependence of γ cannot be precluded. But, to quote BATES and NICOLET (1960): "Quantitative (as distinct from qualitative) prediction of how the effective recombination coefficient varies with altitude is not possible on present knowledge".

A third possible way in which the F1 ledge could be made more pronounced would be to take account of ambi-polar diffusion in the F1 region. Unluckily the final consequence of doing this could not be worked out in the limited time available. A few of the factors that need consideration in this respect might nevertheless be mentioned.

(a) The mean molecular weight of the diffusing plasma comes into the expression for the diffusion coefficient D (see section 2.5.2). This molecular weight is greater for the plasma in the F1 region than for that in the F2 region.

(b) The diffusion coefficient for the O^+ ion-electron plasma diffusing in the parent gas is no longer strictly applicable in the F1 region because the parent gas at this level cannot be considered to consist solely of O atoms.

Geomagnetic control of the peak electron density in the F2 layer is exerted mainly through its control of ambipolar diffusion. It has been mentioned that the height of the F2 peak is lowered by diffusion. A plasma with a greater mean molecular weight than an O^+ ion-electron plasma would come to diffusive equilibrium at a lower level. It might therefore be significant to mention that a geomagnetic control of the F1 peak electron density similar to that of $NF2_{\max}$ has been reported by GHOSH (1955), although the control was not so marked.

7.5 The problem of assuming near complete diffusive separation of the ionizable constituent and the molecular species XY.

If a way could be found to cause a more pronounced peak in the F1 layer without assuming complete or near complete diffusive separation of the ionizable and XY constituents in the ionosphere, the value of k could probably be taken smaller. This would align the present model more closely with that proposed by RATCLIFFE (1959) in which a variable degree of mixing causes k

107.

to assume values between 1.1 and 1.5. It should then be possible to account for many of the observed variations in both the F1 and F2 layers simultaneously.

CHAPTER 8.SUGGESTIONS FOR FURTHER RESEARCH.

The model presented could probably be adapted to yield $N(h)$ distributions very similar to those that were observed by McElhinny over Grahamstown.

If the effects of the eclipse on 25 December, 1954 on such a model were then worked out, it should be possible to compare the theoretical results more readily with observations. Such a comparison could then lead to an estimate of temperature and other changes in the ionosphere which result from an eclipse.

The model presented in this thesis assumes near complete diffusive separation of the neutral gases in the F1 region. This is not well justified. It is suggested that the feasibility of a model in which the above assumption is not made, should be investigated.

SUMMARY.

1. A general survey is made of the theory of the ionosphere.
2. The theory of the formation of the F1 layer is discussed in detail.
3. A model of the F1 layer is proposed in which the equilibrium assumptions $\frac{dN}{dt} = \frac{dn(O^+)}{dt} = 0$ are not made.
4. The continuity equations for ion and electron densities are solved under these conditions.
5. The results of a solution of these equations during a normal day are compared with experimental observations. It is suggested that Kelso scaling of the ionograms, the neglect of diffusion and an insufficiently pronounced F1 ledge on the model could account for the discrepancies that occur.
6. The continuity equations are solved under the conditions that obtain during an eclipse. The results, presented graphically, are compared with observations from low magnetic latitudes.
7. The results show the enhancement of the F1 ledge that occurs during an eclipse. This phenomenon is thought to be due to
 - (a) a change in the distribution of the ionization rate;
 - (b) the fact that at some level in the F1 layer the loss law changes from one that is linear in electron density to one that depends on the square of the electron density;
 - (c) a greater build-up of ionization before the eclipse starts at heights above the F1 ledge than at lower levels.

8. The formation of an $F1\frac{1}{2}$ point of inflection on the electron density profile during the later phase on the eclipse is reported. It is compared with experimental observations. An explanation for its occurrence is attempted on the basis of a gradient in the effective loss coefficient.
9. During an eclipse the model shows changes in the heights of the F1-F2 transition region and the F1 maximum.
10. The $N_{\min} - \delta t$ anomaly that occurs during eclipses is discussed. A possible explanation, on the basis of a change in the height of the F1 maximum, is suggested.
11. The general conditions under which the $F1\frac{1}{2}$ point of inflection has been observed are discussed. The results are compared with observations from temperature latitudes. It is suggested that plasma diffusion obscures it in these latitudes.
12. Possible ways of modifying the model to represent conditions in temperature latitudes are suggested.

REFERENCES.

- | | | |
|---------------------------------------|------|--|
| ALLEN, C.W. | 1948 | Terr. Magn. Atmos. Elec. <u>53</u> . 433. |
| AONO, Y. | 1956 | Solar Eclipses and the Ionosphere.
(Pergamon Press London) p.147. |
| APPLETON, E.V. | 1927 | U.R.S.I. Proc. Washington Assembly. |
| APPLETON, E.V. | 1932 | Jour. Inst. Elec. Engrs. <u>71</u> . 642. |
| APPLETON, E.V. and
BARNETT, M.A.F. | 1925 | Proc. Roy. Soc. A. <u>109</u> . 621. |
| BATES, D.R. | 1955 | Proc. Phys. Soc. Lond. A. <u>68</u> . 344. |
| BATES, D.R. | 1956 | Solar Eclipses and the Ionosphere
(Pergamon Press London) p.184. |
| BATES, D.R. and
MASSEY, H.S.W. | 1946 | Proc. Roy. Soc. A. <u>187</u> . 261. |
| BATES, D.R. and
McDOWELL, M.R.C. | 1957 | J.A.T.P. <u>10</u> . 96. |
| BATES, D.R. and
NICOLET, M. | 1960 | J.A.T.P. <u>18</u> . 65. |
| BOWHILL, S.A. | 1961 | J.A.T.P. <u>20</u> . 19. |
| BRADBURY, N.E. | 1938 | Terr. Mag. <u>43</u> . 55. |
| BREIT, G. and
TUVE, M. | 1926 | Phys. Rev. <u>28</u> . 554. |
| BRIGGS and
RISHBETH, H. | 1961 | Proc. Phys. Soc. <u>78</u> . 409. |
| BUDDEN, K.G. | 1961 | Radio Waves in the Ionosphere (Cambridge)
Chapter 6. p.76. |
| BURKARD, O. | 1956 | J.A.T.P. <u>8</u> . 83. |
| CHALMERS, J.A. | 1962 | J.A.T.P. <u>24</u> . 219. |

- CHAPMAN, S. 1931a. Proc. Phys. Soc. Lond. 43. 26.
- CHAPMAN, S. 1931b. Proc. Phys. Soc. Lond. 43. 483.
- CHAPMAN, S. 1939 Proc. Phys. Soc. 51. 93.
- FERRARO, V.C.A. 1945 Terr. Magn. Atmos. Elect. 50. 215.
- FERRARO, V.C.A. and
OZDOGAN, I. 1958 J.A.T.P. 12. 140.
- GLEDHILL, J.A. 1959 J.A.T.P. 16. 360.
- GLEDHILL, J.A. and
SZENDREI, M.E. 1950 Proc. Phys. Soc. B. 63. 427.
- GLEDHILL, J.A. and
WALKER, A.D.M. 1960 J.A.T.P. 18. 61.
- GLIDDON, J.E.C. 1959 Q.J.M.A.M. 12. 341.
- GLIDDON, J.E.C. and
KENDALL, P.C. 1960 Jour. Geophys. Res. 65. 2279.
- GHOSH, M. 1955 Jour. Geophys. Res. 60. 115.
- HARTREE, D.R. 1931 Proc. Cambr. Phil. Soc. 25. 143.
- HEAVISIDE, O. 1902 Encyclopaedia Brittanica 10th Ed.
33. 213.
- HEISLER, L.H. 1962 J.A.T.P. 24. 483.
- HIRSH, A.J. 1959 J.A.T.P. 17. 86.
- JOHNSON, C.Y.,
MEADOWS, E.B. and
HOLMES, J.C. 1958 Jour. Geophys. Res. 63. 443.
- KAISER, T.R. 1962 J.A.T.P. 24. 865.
- KENNELLY, A.E. 1902 Elec. World. Eng. 39. 473.
- KING, G.A.M. and
LAWDEN, M.D. 1962 J.A.T.P. 24. 565.

- LEJAY, R.P. and DURAND, J. 1956 Solar Eclipses and the Ionosphere (Pergamon Press London) p.85.
- LONG, A.R. 1962 Jour. Geophys. Res. 67. 989.
- MARIANI, F. 1959 J.A.T.P. 16. 160.
- McELHINNY, M.W. 1958 Ph.D. Thesis. Rhodes University.
- McELHINNY, M.W. 1959 J.A.T.P. 14. 273.
- MINNIS, C.M. 1957 J.A.T.P. 10. 229.
- MINNIS, C.M. 1958 J.A.T.P. 12. 272.
- NICOLET, M. 1951 J.A.T.P. 1. 141.
- NICOLET, M. 1960 Physics of the Upper Atmosphere. (Academic Press New York) p.17.
- NICOLET, M. and MANGE, P. 1954 Jour. Geophys. Res. 59. 15.
- RATCLIFFE, J.A. 1956 Solar Eclipses and the Ionosphere (Pergamon Press London) p.1.
- RATCLIFFE, J.A. 1956b J.A.T.P. 8. 260.
- RATCLIFFE, J.A. 1959 Yearbook of the Phys. Soc.
- RATCLIFFE, J.A. and WEEKES, K. 1960 Physics of the Upper Atmosphere. (Academic Press, New York) p.378.
- RATCLIFFE, J.A., SCHMERLING, E.R., SETTY, C.S.G.K. and THOMAS, J.O. 1956 Phil. Trans. Roy. Soc. A. 248. 609.
- REDISH, K.A. 1961 An Introduction to Computational Methods. (The English Universities Press London).
- RISHBETH, H. 1961 J.A.T.P. 20. 277.

- RISHBETH, H. and BARROW, D.W. 1960 J.A.T.P. 18. 234.
- RISHBETH, H. and SETTY, C.S.G.K. 1961 J.A.T.P. 20. 263.
- SEN, H.Y. 1949 Jour. Geophys. Res. 54. 363.
- SETTY, C.S.G.K. 1960 J.A.T.P. 19. 73. and companion papers.
- SHAPLEY, A.H. 1956 Solar Eclipses and the Ionosphere. (Pergamon Press London) p.147.
- SHIMAZAKI, T. 1957 J. Radio Res. Lab. 4. 309.
- SHINN, D.H. and WHALE, H.A. 1952 J.A.T.P. 2. 85.
- SINGH, R.N. and TOLPADI, S.K. 1962 J.A.T.P. 24. 825.
- SZENDREI, M.E. and McELHINNY, M.W. 1956 J.A.T.P. 9. 118.
- TITHERIDGE, J.E. 1959 J.A.T.P. 17. 96.
- VANZANDT, T.E., NORTON, R.B. and STONEHOCKER, G.H. 1960 Some Ionospheric Results obtained during the International Geophysical Year (Elsevier, Amsterdam).
- WALKER, A.D.M. 1962 M.Sc. Thesis. Rhodes University.
- WATANABE, K. and HINTEREGGER, H.E. 1962 Jour. Geophys. Res. 67. 999.
- WILKES, M.V. 1954 Proc. Phys. Soc. Lond. B. 67. 304.
- YONEZAWA, T. 1956 J. Radio Res. Lab. 3. 1.
- YONEZAWA, T. and TAKAHASHI, H. 1960 J. Radio Res. Lab. 7 335.
- YONEZAWA, T., TAKAHASHI, H. and ARIMA, Y. 1959 J. Radio Res. Lab. 6. 21.