NBS

Technical Note

No. 211
Volume 3

CONFERENCE ON NON-LINEAR PROCESSES IN THE IONOSPHERE DECEMBER 16-17, 1963

EDITORS

DONALD H. MENZEL AND ERNEST K. SMITH, JR.

Sponsored By

Voice of America

and

Central Radio Propagation Laboratory
National Bureau of Standards
Boulder, Colorado

U. S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS

NATIONAL BUREAU OF STANDARDS

Technical Note 211, Volume 3

Issued April 19, 1964

CONFERENCE ON NON-LINEAR PROCESSES IN THE IONOSPHERE

DECEMBER 16-17, 1963

Sponsored by
Voice of America
and
Central Radio Propagation Laboratory
National Bureau of Standards
Boulder, Colorado

NBS Technical Notes are designed to supplement the Bureau's regular publications program. They provide a means for making available scientific data that are of transient or limited interest. Technical Notes may be listed or referred to in the open literature.

For sale by the Superintendent of Documents, U. S. Government Printing Office
Washington, D.C. 20402
Price 60¢

PREFACE TO THE PAPER BY V. A. BAILEY

At the Xth General Assembly of URSI in Sydney, in 1952, on the motion of the President, Sir Edward Appleton, a Sub-Commission was set up to prepare for publication a Special Report on the subject of "Ionospheric Wave Interaction." The membership of the Sub-Commission was as follows:

Professor V. A. Bailey, Chairman (Australia);

Mr. J. A. Ratcliffe (Great Britain);

Professor L. G. H. Huxley (Australia);

Dr. M. Cutolo (Italy).

Owing to a serious illness, the Chairman was unable for some time to commence his duties. Then at two informal meetings in 1953 and 1955 it was agreed by the Chairman and Mr. Ratcliffe that the time was not ripe to prepare the Report. However, at a formal meeting of the Sub-Committee in Boulder during the XIIth General Assembly (1957), Mr. Ratcliffe and Professor Huxley indicated that they were unable to continue as members on account of the demands on their time made by other work. The remaining members of the Sub-Committee then agreed to recommend to the Executive that three nominated members be added.

The General Secretary subsequently informed the Chairman that the Sub-Committee was free to co-opt those or any other members of the Union that it desired.

The Chairman then began a first draft of the Special Report for submission to other members of the Sub-Committee, but was unable to complete it in time for the next meeting of URSI. As a result the whole matter was allowed to lapse.

The present communication may perhaps also be regarded as a brief substitute for that Report, prepared by the author alone and brought up to date. It is not intended to be completely exhaustive. He wishes to acknowledge his indebtedness to Dr. R. A. Smith for valuable criticism. of the matter that is here presented in the sections 1 to 5.

Volume III

Table of Contents

---

*Session 1 papers continued from Volume I, Technical Note No. 211.

SOME NON-LINEAR PHENOMENA IN THE IONOSPHERE

V. A. Bailey,

Emeritus Professor, University of Sydney

A paper presented to the Conference on
Non-Linear Processes in the Ionosphere
at Boulder, Colorado, on December 16-17, 1963

ABSTRACT

The historical development is described of the idea that sufficiently strong radio waves can modify the properties of the ionospheric regions which they traverse.

The modified primary properties are the electron temperature and collision-frequency, the rates of electron attachment and recombination and excitation and ionization of molecules.

The resulting changes in the collision frequency and electron density modify the region's absorption coefficient and refractive index for any wave, including the modifying wave itself. These lead to the phenomena of wave-interaction and self-distortion, respectively.

The associated mathematical theory includes the important effects of gyro-waves due to the presence of the Earth's magnetic field. The principal predictions of this theory and their subsequent experimental verification are discussed.

Applications of this theory to the study of the ionosphere and laboratory plasmas are indicated. These include control of the ionosphere and the generation of artificial air-glows (or aurorae) by means of powerful gyro-waves.

The possible modification of the ionosphere by low-frequency extra-terrestrial radio waves is also considered; such waves include geomagnetic disturbances. They may arise from hydromagnetic disturbances in the solar wind and from fluctuating screening by plasma clouds of postulated electric charges on the Sun.

SECTION 1

INTRODUCTION

In a letter to "Nature" B.D.H. Tellegen (1933) reported that the transmissions of the Swiss radio station at Beromunster (650 kc/s) received in Holland appeared to be modulated in the ionosphere by the radiation from the newly erected powerful station at Luxembourg (200 kW, 252.1 kc/s).

This and an earlier report by Butt (1933) are examples of the phenomenon of ionospheric wave interaction* which was subsequently a source of annoyance to European broadcast listeners. The observations reported by B. van der Pol (1935) and others all indicated that a powerful station of sufficiently low frequency can impress its modulation on the wave received from another station (the wanted wave) when the latter traverses ionospheric regions whose projections on the ground lie within about 100-200 kms of the disturbing station. The two waves are now commonly called the <u>disturbing wave</u> and the <u>wanted wave</u>. The observed depth of impressed modulation could amount to about ten percent.

It is worthy of note by applied physicists that the accepted theory of this phenomenon arose exclusively from a "thought experiment" which occurred to the present author while officially examining a doctoral candidate's exposition of the magneto-ionic theory of ionospheric waves. This theory was first mentioned in a letter to Nature (Bailey & Martyn, 1934) and more precisely developed in subsequent

*Also known as the "Luxembourg Effect" & the "Tellegen Effect".

publications (Bailey, 1937a, b, c, d; 1938a, b). A brief account of this development is given in a Supplement to "Nuovo Cimento" (Bailey, 1957).

Since most of the experimental work on wave-interaction has been guided by these mathematical developments it is here desirable to present the theory first.

The two fundamental concepts on which the theory is based are as follows:

I. Any electric field of intensity E (volts/cm) applied to free electrons in thermal equilibrium with a gas at pressure p (m.m.Hg) and temperature T_0 ($^{\circ}$K) will impart energy to the electrons and so raise their temperature to a higher value T. This results in an increase of the rms value u of the random velocities of the electrons and so to a change in ν the mean frequency of collision of an electron with molecules of the gas; in general this change is an <u>increase</u> of the collision frequency.

II. The absorption coefficient k of a radio wave which traverses a particular region of the ionosphere varies with the value of the local collision frequency ν. Except for an extra-ordinary wave of frequency near the local gyro frequency, k increases with ν *.

It follows from I that a powerful radio wave will in general increase the collision frequency in some region of the ionosphere

---

* For the exceptional extraordinary gyro wave, k decreases as ν increases.

and then from II that as a result another wave (not an extraordinary gyro-wave) traversing this region becomes more absorbed by the region.* This increased absorption of the wanted wave will reveal itself (in general) as a decrease in the amplitude of the wave at reception. Thus any modulation of the amplitude of the disturbing wave at a frequency n and to a depth M will impress a modulation of the same fundamental frequency on the wanted wave.

The same concepts I and II lead to the conclusion that a sufficiently powerful wave will modify the absorption of the regions which it traverses and so be received in a distorted form. This phenomenon has been termed "self-distortion".

Besides explaining the observations of Tellegen, Butt, and others, the theory has predicted the following phenomena:

(1) The amplitude of the wanted wave is modulated by a fraction f_i which is given by formula (44) and (45) and consists of a fundamental part of frequency n, phase lag φ, and modulation depth M_1 proportional to M and a harmonic part of frequency $2n$, phase lag φ_2, and depth M_2 proportional to M^2.

(2) The depths M_1 and M_2 of the two impressed modulations decrease in specified manners as the modulation frequency n of the disturbing wave increases.

(3) The depths M_1 and M_2 increase nearly proportionally to the mean power $\bar{P}$ of the disturbing station.

(4) The phase lags φ_1 and φ_2 are equal to the inverse tangents

---

* An extraordinary gyro wave will become less absorbed by the region.

of the ratios $n/G\nu_0$ and $2n/G\nu_0$, respectively, where ν_0 is the value of the electron collision frequency in the unperturbed local ionospheric region and G is a particular constant derived from laboratory data for air; $G \sim 2 \times 10^{-3}$.

(5) The depths M_1 and M_2 decrease as the angular radio frequency ω of the disturbing wave increases within certain limits.

(6) Each of the depths M_1 and M_2 passes through one or two maximum values as the angular radio frequency ω of the disturbing wave is varied about the local angular gyro-frequency Ω, i.e., resonance occurs. Also a radiated power of 1 or 2 kw would then suffice to produce observable interaction with an appropriate wanted wave.

(7) The depths M_1 and M_2 increase as the region of ionospheric reflection of the wanted wave approaches to about 100 km from the ionospheric region vertically above the disturbing station.

(8) A pulsed disturbing wave causes the amplitude of the wanted wave to rise and decay approximately according to a particular law which involves the time-constant $t_0 = 1/G\nu_0$.

(9) The observed modulation of a wave becomes increasingly distorted, at the expense of the lower modulation frequencies, as its power increases, i.e., self-distortion increases with the power.

(10) "Interaction atmospherics" can be impressed on a wanted wave by the low frequency radiation from suitably located electric storms.

(11) Combination tones can be impressed on a wanted wave by two disturbing waves with radio frequencies which differ by an audio-frequency.

The predictions (1) to (9) have been confirmed by subsequent experimental observations. Also (4) has been used by Ratcliffe and Shaw [1948] to determine the height at which a particular wave interaction occurs and (8) has been used to determine the quantity $G\nu$ which relates in part to the mean energy lost by an electron at a collision with an ionospheric molecule. The prediction (11) appears to account for certain observations of Gill [1935].

Also it may be mentioned that the formula for M_1 yielded by the theory (and given below under (45)) correctly accounts for the order of magnitude of the depth of the modulation impressed on Beromunster's wave by the Luxembourg station as observed by van der Pol and van der Mark [1935].

The fundamental concepts I and II have also led to the following predictions:

(12) A specified vertical beam of extraordinary gyro waves with radiated power of 500 kw can generate a visible glow discharge or artificial "aurora" in the E-region at night [Bailey, 1938a, b; 1961].

(13) A vertical beam of extraordinary gyro waves radiated with a power of 500 kw from an array of 40 horizontal dipoles can penetrate into the nocturnal E-region much more easily than a weak beam. Also, this beam suffices to increase the electron density near the 92 km/s level of the nocturnal E-region by an observable amount and thus to determine the frequency of attachment of electrons to molecules near this level [Bailey, 1959].

These results and the corresponding experimental studies by Goldstein et alia (1953a, b; 1955) of the effects of strong micro-waves on gaseous plasma, have led to a proposal for control of the ionosphere by means of radio waves (Bailey and Goldstein, 1958).

We shall now consider in detail according to the following plan the principal theoretical and experimental results of investigations on ionospheric wave-interaction published up to February, 1963.

Section 2 summarises the relevant concepts, methods and experimental observations of the Townsend School which relate to the collisions of electrons with gas molecules or atoms, under the influence of electric fields of effective intensity E volts/cm at gas pressures p m.m.Hg and gas temperatures T_0 (usually about 15° C).

Section 3 summarises the relevant concepts and experimental observations, of the Appleton School and other ionospheric physicists, which relate to the propagation of radio waves in the ionosphere regarded as a magneto-ionic gaseous medium.

Section 4 briefly sets out the mathematical theory of wave-interaction as a logical consequence of the work described in Sections 1 to 3.

Section 5 considers in detail the principal predictions to which this theory leads.

Section 6 describes the experimental tests and verifications of these predictions.

Sections 7 and 8 discuss the use of the theory as an instrument for the study of the ionosphere and electrical gas-discharges, some of its practical applications and some recent developments of the subject.

SECTION 2

COLLISIONS OF ELECTRONS WITH GAS MOLECULES UNDER THE INFLUENCE OF AN ELECTRIC FIELD.

In the absence of an electric field,the mean energy U of an electron in a gas at uniform temperature T_0 is equal to the mean energy U_0 of a molecule* of the gas.

Townsend and his collaborators [Townsend 1947; Healey & Reed, 1940], discovered that in the presence of an electric field E, whether constant or alternating [Townsend 1932], the mean energy of random motion U of an electron is greater than U_0 and that for a given gas, the ratio

$$k = U/U_0$$

is a function of the ratio E/p alone, where p is the gas pressure.

By measuring, in a constant field E, the divergence in the gas of a stream of electrons which passes through a slit, they were able to determine the values of the Townsend factor k corresponding to different values of E/p. Then by deflecting the same stream through a known angle by means of a transverse magnetic field they were able to measure the corresponding values of the drift velocity W of the electrons in the direction of the electric field.

From these known values of k and W and by means of simple+

---

* In this paper the term "molecule" is intended to include "atom"

\+ The more complicated formulae proposed later by various workers have not been found necessary in the present state of experimental knowledge of wave interaction.

approximate theoretical formulae they then calculated the mean free path L of an electron of energy U in the gas at the standard pressure of 1 m.m.Hg and the mean proportion λ of its own energy which such an electron loses at a collision with a molecule.

From the value of L corresponding to a given value of U, the mass m of an electron and the relation $U = \frac{1}{2}m\,u^2$ we can then calculate in succession the corresponding mean random velocity u and collision frequency ν at a pressure p by means of the formulae

$$u = (2U/m)^{\frac{1}{2}} = (2kU_0/m)^{\frac{1}{2}} \qquad (1)$$

$$\nu = u/l = up/L \; , \qquad (2)$$

since the mean free path l at the pressure p is given by

$$l = L/p \; .$$

Also, from the corresponding value of λ we can calculate η the mean energy lost by an electron at a collision by means of the formula

$$\eta = \lambda\, U. \qquad (3)^*$$

The values of k, W, L and λ corresponding to different values of E/p for well dried air were first determined by Townsend & Tizard [1913]. These are also set out in the books by Townsend [1947] and Healey & Reed [1940]. More recent determinations of some of these quantities have been made by Bailey [1925 ; Neilsen & Bradbury, 1937, and Crompton, Huxley & Sutton, 1953]. On account of the limitations set by the experimental methods used the lowest value of the Townsend factor measured in air was $k = 4$; this corresponds to a mean electron energy of 4/27 electron volts. Since the variations of L and η with U for $k < 4$ are not known, the application of the known experimental data to the lower ionosphere makes it

---

* η can also be obtained directly from the drift velocity W by means of the formula $\eta = 1.23\ mW^2$.

necessary to extrapolate these data in some simple way to the range in which $1 \leq k < 4$. For this purpose the simplest two hypotheses to adopt are: (1) L is independent of U and

$$(2) \quad \eta = G\,(U - U_0), \tag{4}$$

where G is a constant determined by the data obtained with the lowest values of the field E. For Air the value $G \sim 2 \times 10^{-3}$ has proved the most successful.

From the first hypothesis and equations (1) and (2) it follows that the collision frequency ν in air always increases with the intensity of the field E in the range of energies considered.

Following Townsend's work on the uniform column in high frequency electrical discharges through gases ("Electrical Discharges", Llewellyn Jones [1953]) the formulae (1) to (4) may be applied to the effects caused by a strong radio wave in the Ionosphere. But the dependence of these effects on the field intensity $\underset{\sim}{E}$, the radio-frequency $\omega / 2\pi$ and the terrestrial magnetic field $\underset{\sim}{H}$ needs further consideration, such as that given below in Section 4 and the Appendix. This involves the determination of $\bar{w}$, the mean work done on an electron over a free path or, alternatively, the mean power $\Xi = N \nu \bar{w}$ expended on the N electrons in unit volume.

SECTION 3

WAVE PROPAGATION IN A GASEOUS MAGNETO-IONIC MEDIUM.

In Appleton's well-known magneto-ionic theory of wave propagation, the medium is a gas which contains N electrons/cc and is pervaded by a uniform, constant magnetic field $\underset{\sim}{H}$; also the mean collision frequency ν of an electron in the gas is taken as independent of the amplitude of the wave.

The complex refractive index $M = \mu - i\chi$ and the polarisation number R which define a mode of propagation of a wave of angular frequency ω in this medium are determined by means of Maxwell's field equations combined with the equation of transport of mean momentum, which governs the momentum $N m \underset{\sim}{v}$ of the electrons in unit volume. When N is uniform and constant, the latter reduces to the equation

$$m \frac{d \underset{\sim}{v}}{d t} + \nu m \underset{\sim}{v} = e \underset{\sim}{E} + \frac{e}{c} \underset{\sim}{v} \times \underset{\sim}{H}, \qquad (5)$$

where e is the charge on an electron in e.s. units and the magnetic vector of the wave field is neglected by comparison with $\underset{\sim}{H}$.

From the dispersion equation yielded by (5) and the field equations, it follows that there are two wave-modes and that for each the corresponding ray becomes attenuated as it progresses, through a factor a_s defined by

$$a_s = \exp \left(-\int_0^s \kappa \, ds \right) \qquad (6)$$

where s is the length of the ray path measured from its beginning and κ is related to χ as follows:

$$\kappa = \omega \chi / c. \qquad (7)$$

The value of the coefficient of absorption κ depends on ν, N, and the angular gyro-frequency Ω.

The theory shows that when

$$\omega^2 >> \nu^2 \tag{8}$$

κ is in general approximately proportional to the product ν N. For the exceptional case, in which the angular wave frequency ω is nearly equal to Ω and the mode is extraordinary, κ is given by Equation (64) and so is nearly proportional to $N\nu / \{ \nu^2 + (\omega - \Omega)^2 \}$. It follows that with increasing ν the attenuation in general increases, but decreases for an extraordinary wave mode with a frequency near the gyro-frequency. The latter may be conveniently termed an extraordinary gyro-wave.

In a slowly varying non-uniform medium a ray theory of propagation may be used. Then along a ray path of length s, the field amplitude is given by

$$E = E_0 \exp(\quad - \alpha)$$

where

$$\alpha = \int_0^s \kappa \, ds.$$

Since the mean power flux density is $\overline{P}$ where

$$\overline{P} = c \mu \overline{E_T^2} / 4 \pi ,$$

where $\bar{E}_T$ is the rms value of the component of $\underset{\sim}{E}$ transverse to the direction of propagation, it follows that

$$\bar{P} / \mu = (\bar{P}_0 / \mu_0) \exp (- 2\alpha) \tag{9}$$

and so

$$\partial (\bar{P} / \mu) / \partial s = - 2 k (\bar{P} / \mu) . \tag{10}$$

SECTION 4

THE MATHEMATICAL THEORY OF IONOSPHERIC WAVE INTERACTION.

From consideration of the results mentioned in section 2, we see that in the equation (5) the quantity ν is not really a constant but must now be regarded as a function of the wave field intensity.

It follows that the classical magneto-ionic formulae for M and R must now be regarded as first approximations which are accurate only when the waves concerned are weak enough to make ν closely equal to the normal collision frequency ν_0 which exists in the unperturbed medium.

The fact that at least h is a function of the field intensity E is sometimes referred to as the "non-linearity" of the medium.*

In order to obtain higher approximations to M and R it is in general necessary to apply lengthy perturbation procedures. So in order to keep the present mathematical theory sufficiently simple for practical purposes we shall relax its rigour to the limits allowed by the uncertainties in the physical data involved.

---

* The experiments of King, described in section 6, show that notable ionospheric self-distortion can occur when the radiated power of the transmitter is at least 200 kw.

4.1 The work w done on an electron by a plane-polarised wave over a mean free path.

It has been shown (Bailey 1937b) that

$$w = \frac{e^2}{2m} E_e^2 \left[A + A \cos 2 \omega t - B \sin 2 \omega t \right] \qquad (11)$$

where E_e is the effective value $E/\sqrt{2}$ of the local electric field intensity $\underset{\sim}{E}$,

$$\left.\begin{aligned} A &= \frac{s^2}{\nu^2 + (\omega - \Omega)^2} + \frac{s^2}{\nu^2 + (\omega + \Omega)^2} + \frac{2c^2}{\nu^2 + \omega^2} \\ B &= -\left(\frac{\omega - \Omega}{\nu^2 + (\omega - \Omega)^2} + \frac{\omega + \Omega}{\nu^2 + (\omega + \Omega)^2} \right) \frac{s^2}{\nu} - \frac{2\omega}{\nu^2 + \omega^2} \frac{c^2}{\nu} \end{aligned}\right\} \qquad (12)$$

ω is the angular frequency of the wave,

$\Omega = H e / m c$ (angular gyro frequency),

$s = \sin \psi$, $c = \cos \psi$

and ψ is the angle between $\underset{\sim}{E}$ and $\underset{\sim}{H}$.

The mean work $\bar{w}$ done during a complete cycle $2\pi / \omega$ and over a free path is therefore given by

$$\bar{w} = \frac{e^2}{2m} E_e^2 A . \qquad (13)$$

In the following sections we shall assume that E_e varies appreciably only during a period of time which is large compared with the period $2\pi/\omega$ of the wave.

4.2 The differential equation for the mean kinetic energy U of an electron.

The equation of transport of energy in the electron gas is

$$\frac{d\,(N\,U)}{d\,t} + N\,\nu\,\eta = N\,\nu\,w\,,$$

so when the wave is not strong enough to change N appreciably we have

$$\frac{d\,U}{d\,t} + \nu\,\eta = \nu\,w\,. \tag{14}$$

When the energy U does not exceed $4U_0$ we have $1 < k \leq 4$ and so we may adopt the formula (4) as a working hypothesis. Then (14) leads to the equation

$$\frac{d\,U}{d\,\tau} + (\nu / \nu_0)\,(U - U_0) = t_0 \nu\,w \tag{15}$$

where

$$\left.\begin{aligned} t_0 &= 1/G\,\nu_0 \\ \tau &= t/t_0\,. \end{aligned}\right\} \tag{16}$$

t_0 is a time-constant and τ is the time measured in units equal to this time-constant.

Also, by (2) and (1)

$$\nu / \nu_0 = (U / U_0)^{\frac{1}{2}}\,(L_0 / L)\,. \tag{17}$$

The dependence of L on U for $U < 4\ U_0$ is not yet known with certainty. But the experiments on air of Townsend and Tizard and of Crompton, Huxley and Sutton indicate that no great error is likely to result from our adopting here the original hypothesis that L is independent of U , i.e.

$$\nu / \nu_0 = (U / U_0)^{\frac{1}{2}} . \tag{17.1}$$

By substitution for ν from (17) or (17.1) in (12) and from (11) we can express νw as a known function of U and E_e^2 . Then (15) becomes a differential equation in the unknown function U with E_e as a given slowly varying function of t , of the form

$$E_e = E_{e1} M\ (t) = E_{e1} M\ (t_0 \tau\ ) \tag{18}$$

where E_{e1} is a constant field. M (t) may be termed the modulation function or modulation envelope.

The resultant differential equation requires some method of approximation for its solution. From such a solution U and (17) or (17.1) we can then obtain an approximate expression for ν as a function of t, E_{e1} , ω , Ω , ψ , G and U_0 or ν_0 .

These expressions for U and ν contain terms which vary slowly with t and terms which are periodic in t with 2ω as their fundamental frequency. Retention of the latter terms in the discussion of wave-interaction and self-distortion would lead to similar periodic terms in these phenomena. But at radio frequencies ω such terms are usually negligible by comparison with the slowly varying terms, so henceforth they will be neglected.

Accordingly we will now consider only the variations of U and ν which take place in a time dt large compared with the wave period $2\pi/\omega$. This requires the symbol w in (14) and (15) to be replaced always by its mean value $\bar{w}$ which is given by (13) or by (48) below.

4.3 The differential equation for the mean collision frequency ν.

It is now convenient to base the discussion on the equation for ν. Thus, on eliminating U from (15) by means of (17.1), replacing w by $\bar{w}$ and using (13) and (18) we obtain

$$\frac{d\sigma}{d\tau} + \frac{1}{2}\left(\sigma^2 - 1\right) = \bar{w} / 2 G U_0 = R M^2 \left(t_0 \tau\right) \tag{19}$$

where

$$\sigma = \nu / \nu_0 , \tag{20}$$

$$R = \frac{e^2}{4 G m U_0} E_{el}^2 A . \tag{21}$$

σ is the collision frequency expressed in units equal to ν_0.

On account of its factor A the quantity R in (19) is a function of σ^2 such that this equation has no obvious solution in finite terms.

However, in order to apply the theory to most of the observations of wave interaction which have been made hitherto it is sufficient to consider only those approximate solutions which correspond to values of $\bar{w}$ / 2G U_o or R M^2 ($t_o\tau$) such that σ varies about some constant value σ_1 by a small fraction of σ_1 , i.e.

$$\left.\begin{aligned} &\sigma = \sigma_1 + \delta , \\ &\text{where} \\ &\sigma_1 = \nu_1 / \nu_o , \qquad |\delta| << \sigma_1 . \end{aligned}\right\} \qquad (22.1)$$

Thus

$$\nu = \nu_1 + \nu_o\delta , \qquad |\nu_o\delta| << \nu_1 \qquad (22.2)$$

These approximate solutions are then found from solutions of the equation

$$\frac{d\,\delta}{d\,\tau} + \sigma_1\,\delta = R_1 M^2\,(t_o\,\tau) + \tfrac{1}{2}(1 - \sigma_1^2) \qquad (23)$$

where

$$R_1 = \frac{e^2}{4\,G\,m\,U_o}\,E_{el}^2\,A_1 \qquad (21.1)$$

and

$$A_1 = \frac{s^2}{\nu_1^2 + (\omega - \Omega)^2} + \frac{s^2}{\nu_1^2 + (\omega + \Omega)^2} + \frac{2c^2}{\nu_1^2 + \omega^2} . \qquad (24)$$

4.4 <u>The collision frequency ν_s in the steady state under the influence of a continuous wave.</u>

Here we have M^2 (t_o τ) = 1 . Then the steady value σ_1 = ν_s/ ν_o

is obtained by setting $d\sigma/dt = 0$ in (19) and solving the resultant quartic in σ^2 or setting $\delta = 0$ in (23) and solving the same quartic in $\sigma_1^{\,2}$.

When E_{el} is not too large* we then deduce the following approximate expression for the steady collision frequency:

$$\nu_s = \nu_0 + R_0 \nu_0 , \tag{25}$$

where

$$R_0 = \frac{e^2}{4G\, m\, U_0} E_{el}^2 A_0 \tag{26}$$

and A_0 is the same as A_1 in (24) with ν_1 replaced by ν_0.

From (25), (26) and (24) it follows that, except when the wave's field vector $\underset{\sim}{E}$ is only slightly inclined to the constant magnetic field vector $\underset{\sim}{H}$, ν_s passes through a maximum value as the wave frequency ω is varied around the gyro frequency Ω. The corresponding mean electron energy U_s behaves similarly.

These resonance effects and their consequences are of great importance, as will be shown below.

4.5 Rise and decay of ν caused by a square pulse of radio waves.

When $\nu = \nu_0$ at $t = 0$ and $M^2 \;(t_0 \tau) = 1$ for $0 \leq t \leq T$ then Equation (23) has $\sigma_1 = 1$ and its right-hand side constant.

The corresponding solution of (23) which describes the rise of ν, is

---

* More definitely when $R_0 \ll \frac{1}{2}$.

$$\nu = \nu_s - (\nu_s - \nu_o)\, e^{-t/t_o} \quad . \tag{27}$$

Similarly we obtain the following solution which describes the decay of ν from the steady value ν_s

$$\nu = \nu_o + (\nu_s - \nu_o)\, e^{-t/t_s} \tag{28}$$

where

$$t_s = 1/G\, \nu_s \; .$$

For the E-region t_o is of the order of 1 millisecond or less. Hence the rise and decay require about 1 millisecond to be practically completed.

Accurate expressions for the rise and decay can also be obtained, since the Equation (19) is now exactly integrable (Bailey, 1937b, Equation (34)).

4.6 <u>Modulation of ν caused by a wave which is sinoidally modulated in amplitude.</u>

For such a wave we have

$$\left.\begin{array}{l} M(t) = 1 + M \sin n t \\ \text{where} \\ 0 < M < 1, \quad n \ll \omega . \end{array}\right\} \tag{29}$$

Then

$$M^2 (t_o\, \tau) = 1 + \tfrac{1}{2}M^2 + 2M \sin (nt_o \tau) - \tfrac{1}{2}M^2 \cos (2nt_o\, \tau). \tag{29.1}$$

If we take

$$\nu_1 = \nu_s$$

for the steady state corresponding to the mean effective force

$$E_{el}(1 + \tfrac{1}{2} M^2)^{\frac{1}{2}}$$

then the constant part of the R.H.S. of (23) vanishes, i.e.

$$R_1 \ (1 + \tfrac{1}{2} M^2) + \tfrac{1}{2} \ (1 - \sigma_1^2 \ ) \ = \ 0, \tag{29.2}$$

and so the equation (23) reduces to

$$\frac{d\delta}{d\tau} + \sigma_1 \delta = 2 \ R_1 M \sin \ (nt_o \tau) - \tfrac{1}{2} R_1 M^2 \ \cos \ (2n \ t_o \tau), \tag{30}$$

where

$$\sigma_1 = \nu_s \ / \ \nu_o \ ,$$

and, by (25),

$$\nu_s = \nu_o + \frac{e^2}{4G \ m \ U_o} E_{el}^2 \ (1 + \tfrac{1}{2} M^2) \ A_o \ \nu_o \ . \tag{31}$$

The general solution of (30) combined with (22.2) leads to the following formula for ν :

$$\nu = \nu_s + \frac{BM \sin \ (nt - \varphi_1)}{(1 + n^2 \ t_s^2)^{\frac{1}{2}}} - \frac{\frac{1}{4} BM^2 \ \cos \ (2nt - \varphi_2)}{(1 + 4 \ n^2 \ t_s^2 \ )^{\frac{1}{2}}} + Ce^{-t/t_s} \tag{32}$$

where

$$\left.\begin{aligned}
B &= \frac{e^2 E_{el}^2 A_s\, t_s}{m^2\, \ell^2} \\[2ex]
t_s &= 1 / G\nu_s \\[2ex]
\tan \varphi_1 &= n\, t_s\,, \qquad \tan \varphi_2 = 2\, n\, t_s\,,
\end{aligned}\right\} \tag{33}$$

A_s is the same as A_1 in (24) with ν_1 replaced by ν_s and C is a constant determined by the condition $\nu = \nu_0$ at $t = 0$.

4.7 Modulation of a wanted wave through the variation of the local collision frequency.

In passing through the ionosphere the field amplitude F of the wanted wave mode is attenuated according to the formula

$$\left.\begin{aligned}
F &= F_0 \exp(-\alpha) \\
\text{where} \qquad & \\
\alpha &= \int_0^s \kappa(\nu)\, ds,
\end{aligned}\right\} \tag{34}$$

s is the length of the wanted wave's path measured form its entry into the ionosphere, F_0 is the field amplitude at entry and $\kappa(\nu)$ is the absorption coefficient corresponding to the value ν of the collision frequency at the point s in the path.

Under those conditions in which the collision frequency varies about some constant value ν_1, i.e.

$$\nu = \nu_1 + \Delta \nu \tag{35}$$

we have

$$\alpha = \alpha_1 + \Delta \alpha \tag{36}$$

where

$$\alpha_1 = \int_0^s \kappa (\nu_1) \, ds$$

and

$$\Delta \alpha = \int_0^s \{ \kappa (\nu) - \kappa (\nu_1) \} \, ds \tag{37}$$

Hence, by (34) and (35)

$$F = F_1 \exp (- \Delta \alpha)$$

where

$$F_1 = F_0 \exp (- \alpha)$$

So at the point s the amplitude of the wanted wave is modulated according to the modulation function $M_i (t)$ where

$$M_i (t) = \exp (- \Delta \alpha)$$

If we set

$$M_i (t) = 1 + f_i (t) \tag{38}$$

the impressed modulation fraction $f_i\ (t)$ is given by

$$f_i\ (t) = \exp\ (-\Delta\alpha\ ) - 1$$

When observation shows that $|f_i\ (t)| << 1$ we will necessarily have the approximation

$$f_i\ (t) = -\Delta\alpha = -\int_0^s \{\ \hbar\ (\ \nu\ ) - \hbar\ (\ \nu_1\ )\ \}\ d\ s \qquad (39)$$

On substituting for ν from (35) and expanding $\hbar\ (\ \nu_1 + \Delta\ \nu\ )$ we obtain

$$f_i\ (t) = -\int_0^s \hbar'\ (\ \nu_1\ )\ \Delta\ \nu\ ds \qquad (40)$$

provided that

$$|\Delta\ \nu\ \hbar''\ (\ \nu_1\ )/\ \hbar'\ (\ \nu_1\ )| \ll 2\ .$$

When $R\ M^2\ (t)$ is sufficiently small for (22.2) to hold true we have, by (35),

$$\Delta\ \nu = \nu_o\ \delta$$

where δ satisfies the equation (23). So, on multiplying (23) throughout by $-\hbar'(\ \nu_1\ )\ \nu_o\ ds$, integrating with respecting to s and using (40) we obtain the following equation for $f \equiv f_i\ (t)$, for a situation in which σ_1 is nearly constant throughout the region concerned.

$$\frac{df}{d\tau} + \sigma_1 f = S_1 M^2 (t_o \tau) + S_2 \tag{41}$$

where

$$\left.\begin{aligned} S_1 &= - \int_0^s R_1 \, \hbar' (\nu_1) \nu_0 \, ds \\ S_2 &= \tfrac{1}{2} \int_0^s (\sigma_1^2 - 1) \, \hbar' (\nu_1) \nu_0 \, ds \end{aligned}\right\} \tag{42}$$

We now apply (41) to determine the modulation fractions f_1 and f_2 of the wanted wave respectively during and after the disturbing square pulse considered in sub-section 4.5 .

Since $\nu = \nu_0$ at $t = 0$ and $M^2 (t_o \tau) = 1$ for $0 \le t \le T$ we then have $\sigma_1 = 1$ and so (41) yields the formula

$$f_1 = S_1 (1 - e^{-t/t_o}) . \tag{43.1}$$

Similarly, if $\nu \doteqdot \nu_s$ at $t = 0$ and $M^2 (t_o \tau) = 0$ for $t > T$ we have $\sigma_1 \doteqdot \nu_s / \nu_0$ and so (41) yields

$$f_2 = S_1 e^{-t/t_s} . \tag{43.2}$$

Next we apply the Equation (41) to determine the modulation fraction f_i impressed on the wanted wave by the sinoidally modulated disturbing wave considered in sub-section 4.6 .

As in 4.6, we now have: $M^2 (t_o \tau)$ given by (29.1), $\nu_1 = \nu_s$ i.e.

$$\sigma_1 = \sigma_s = \nu_s / \nu_0 ,$$

and the equation (29.2) which is satisfied by the root $\sigma_1 = \sigma_s$.

From (42), (29.2) and (29.1) we obtain

$$S_1 M^2 (t_0 \tau) + S_2 = S_1 \{ 2M \sin (nt_0 \tau) - \tfrac{1}{2} M^2 \cos (2nt_0 \tau) \}$$

and so (41) now becomes

$$\frac{df}{d\tau} + \sigma_1 f = 2 S_1 M \sin (nt_0 \tau) - \tfrac{1}{2} S_1 M^2 \cos (2nt_0 \tau) . \qquad (41.1)$$

The impressed modulation fraction is given by the general solution of this equation, namely

$$f_i = M_1 \sin (nt - \varphi_1) - M_2 \cos (2nt - \varphi_2) - Ce^{-t/t_s} \qquad (44)$$

where φ_1, φ_2, t_s are defined under (33),

$$\left.\begin{aligned} M_1 &= B_s M/(1 + n^2 t_s^2)^{\frac{1}{2}} , \\ M_2 &= \tfrac{1}{4} B_s M^2/(1 + 4 n^2 t_s^2)^{\frac{1}{2}} , \end{aligned}\right\} \qquad (45)$$

$$B_s = - 2 \sigma_1^{-1} S_1 \qquad (46)$$

and C is a constant such that $f_i = 0$ when $t = 0$

Since $\nu_1 = \nu_s$ we find from (21.1) and (46) that in a region in which ν_0 and E_1 are constant

$$B_s = - \frac{e^2 E_{el}^2 A_s}{Gm^2 l^2 \nu_s} \int_0^{s_w} \hbar' (\nu_s) ds_w \qquad (46.1)$$

where A_s is the same as A_1 in (24) with ν_s replacing ν_1 and s_w is the length of the path of the wanted wave.

It will be noted that all the quantities in (44) which relate to the wanted wave are contained in the factor B_s which occurs in the quantities M_1 and M_2 .

A simple calculation shows that when, for the wanted wave, $\hbar(\nu)$ is approximately proportional to ν the formula (44) confirms a previously given formula for M_1 (Bailey, 1937, Part 1, Equation (42)).

4.8 Modulation impressed by a Wave Mode.

When it is necessary to take into account the attenuation of each of the wave modes into which a plane wave divides, the theory given above must be modified by using in the place of the formula (13) for $\overline{w}$ a formula for either mode which is obtained as follows (V.A.Bailey, 1937a, Part 1):

The mean power Ξ supplied by the wave to the N electrons in unit volume is given by

$$\Xi = - \frac{\partial \overline{P}}{\partial s_d} \tag{47}$$

where $\overline{P}$ is the mean power flux density and s_d is the length of the path of the disturbing wave from its place of entry.

Since this must be equal to N times the mean power $\nu \overline{w}$ supplied to one electron it follows that

$$N \nu \overline{w} = \Xi \tag{48}$$

Also, when μ varies slowly with s_d we have from (10) and (9)

$$\Xi = 2\, k_d\, \overline{P} = 2\, k_d\, \overline{P}_o \exp\left(- 2\alpha_d \right) \tag{49}$$

where the subscript d to a symbol relates it to the disturbing wave.

Since the amplitude of the wave mode is modulated by the factor $M(t)$ as in (18) therefore

$$\overline{P}_o = \overline{P}_{o1}\, M^2\,(t)$$

where $\overline{P}_{o1}$ is the mean flux density of this mode, over one period, at entry into the ionosphere.

Hence, by (49)

$$\Xi = \Xi_1\, M^2\,(t) \tag{50}$$

where

$$\Xi_1 = 2\, k_d\, \overline{P}_{o1} \exp\left(- 2 \int_o^{sd} k_d\, ds_d\right) \tag{51}$$

and $k_d \equiv k_d\,(\nu)$ when expressed as a function of ν .

On using the value of $\overline{w}$ given by (48) and also (50) the equation (19) assumes the form

$$\frac{d\sigma}{d\tau} + \tfrac{1}{2}\left(\sigma^2 - 1 \right) = R\, M^2\,(t_o\tau) \tag{52}$$

where

$$\sigma = \nu / \nu_o, \qquad \tau = t / t_o = G\nu_o t, \tag{53}$$

$$R = \Xi_1 / 2G\, U_o N \nu_o \sigma \tag{54}$$

As in sub-section 4.3 we shall be content here to consider only those situations in which $R\, M^2\,(t_o\tau)$ is small enough for the condition

(22.1) to hold true. Then (52) may be replaced (approximately) by the equation

$$\frac{d\delta}{d\tau} + \sigma_1 \delta = R_1 M^2(t_0 \tau) + \tfrac{1}{2}(1 - \sigma_1^2) \tag{55}$$

where

$$R_1 = \Xi_1(\nu_1) / 2G\, U_0\, N\nu_1 \tag{56}$$

and $\Xi_1(\nu_1)$ is the value of Ξ_1 in (51) when $\hbar_d = \hbar_d(\nu_1)$ and $\nu_1 \div \nu_0$.

Since (55) has the same form as (23) we can now proceed, as was done with the latter, to derive from it the equation for the modulation fraction $f \equiv f_i(t)$ impressed by the wave mode under consideration when the ionospheric region concerned is nearly uniform in temperature and pressure and σ_1 is nearly constant throughout this region.

This procedure yields the equation (41) for f and the formulae (42) but with R_1 now given by (56), i.e., S_1 is now given by the formula

$$S_1 = -\frac{\nu_0}{2G\, U_0} \int_0^s \Xi_1\, q\, ds \tag{57}$$

where

$$q = \hbar'(\nu_1) / N\nu_1 \tag{58}$$

and Ξ_1 is given approximately by (51) with $\nu = \nu_0$.

We thus see that the formulae (43.1), (43.2) and (44) for the impressed modulations f_1, f_2, f_i respectively are also valid here when S_1 is given by (57).

The only factor in S_1 which relates to the disturbing wave is Ξ_1. For this reason it has been termed the <u>Index of Interaction</u>. Ξ_1 also represents the mean input of power per unit volume of the ionosphere per period of the disturbing wave.

The combined effects of the two modes which constitute a plane wave can be similarly calculated if the distribution of the energy flux between the modes is known.

We now consider the common case (A) in which the wanted wave is an ordinary mode with

$$\hbar (\nu) = \rho N \nu \tag{A}$$

where ρ is nearly constant throughout the region concerned. In this case we generally have

$$\omega^2 \gg \nu^2 \quad \text{and} \quad \omega^2 \gg 4 \pi Ne^2 / m .$$

Therefore, by (58), $q = \rho / \nu_1$ and so q is now nearly constant throughout the region. Hence, by (57) we now have

$$S_1 \doteqdot - \frac{\hbar (\nu_0)}{2GU_0 N\nu_1} \int_0^s \Xi_1 \, ds \tag{57.1}$$

Since, by (50) and (47),

$$M^2 (t) \int_0^s \Xi_1 \, ds = - \int_0^{s_d} \sec \beta \frac{\partial \overline{P}}{\partial s_d} \, ds_d \tag{59}$$

where β is the angle between the elements of path ds_d and ds contained between adjacent wave fronts of the disturbing wave*,

* for which $ds_d = ds \cos \beta$.

we cannot evaluate the integral in (57.1) in finite terms except when β is constant (or nearly so) throughout the region of interaction; this will occur in the following sub-cases A1 and A2 .

A1. Both waves are propagated vertically and then $|\sec \beta| = 1$ throughout.

A2. The disturbing wave is propagated vertically and the region of interaction occurs mainly where the wanted wave is undeviated.

In these cases (59) yields

$$M^2(t) \int_0^s \Xi_1 \, ds = \sec \beta \, (\bar{P}_0 - \bar{P}_{s_d})$$

where $\bar{P}_0$ and $\bar{P}_{s_d}$ are the values of the disturbing wave's flux density at the beginning and end of a length s_d of its path.

Since

$$\bar{P}_0 = \bar{P}_{01} M^2(t) \quad \text{and} \quad \bar{P}_{s_d} = \bar{P}_{s_d 1} M^2(t)$$

where $\bar{P}_{01}$ and $\bar{P}_{s_d 1}$ are the mean flux densities over one period at respectively entry into and exit from the ionosphere, it follows that

---

$$\int_0^s \Xi_1 \, ds = \sec \beta \, (\bar{P}_{01} - \bar{P}_{s_d 1}) ,$$

and so from (57.1) that, in the sub-cases A1 and A2,

$$S_1 \doteqdot - \frac{\hbar(\nu_0) \sec \beta}{2 \, G U_0 \, N \nu_1} (\bar{P}_{01} - \bar{P}_{s_d 1}) , \qquad (60)$$

Under A^2 we have $\sec \beta = \sec i_w$ where i_w is the angle of incidence of the wanted wave. We then obtain from (44) and (60) a formula for the depth M_1 , of the impressed fundamental modulation, which is equivalent to the equation (28) given in Huxley's "Synopsis" (1952) and is originally due to Ratcliffe & Shaw (1948).

4.9 Gyro Interaction: Interaction caused by an approximate gyro-wave.

When the disturbing wave is plane-polarised and has a frequency within 20 percent of the local gyro-frequency, i.e.

$$\left.\begin{aligned} \omega &= \Omega\,(1+\varepsilon) \\ \text{where} \quad |\,\varepsilon\,| &\leq 0.2\,, \end{aligned}\right\} \tag{61}$$

the formula (24) with $\nu_1 = \nu_s$ shows that in a region where

$$\nu_s^2 \ll \Omega^2 \tag{62}$$

the quantity A_s passes through a sharp maximum as the wave's frequency ω is varied about the gyro-frequency Ω .

If follows from (46.1) that B_s behaves similarly and that in consequence each of the modulation depths M_1 and M_2 passed through a sharp maximum. By (21.1), (42) and (43.1) the modulation f_1 impressed by a pulsed gyro-wave will behave similarly.

Since the two component modes of the plane gyro-wave are absorbed at very different rates as they propagate through the ionosphere, a more precise discussion requires the effects caused by each of these modes to be considered separately. We may then use the results obtained in the sub-section 4.8.

For our present purpose it is sufficient to consider only the

effects caused by the much more absorbable extraordinary component*
in the lower E region of the nocturnal ionosphere, at moderate latitudes, for which we have

$$N < 800 , \quad \Omega = 10^7 , \quad \nu \leq 10^6 . \tag{63}$$

It has been shown (Bailey 1938a) that the refractive index μ of the extraordinary mode defined by (61) is closely equal to unity and that its coefficient of absorption k_d is given approximately by

$$k_d = \frac{K\,(1 + \cos^2\theta)\,N\nu}{\nu^2 + \epsilon^2\,\Omega^2} \tag{64}$$

where

$$K = 0.0266,$$

and θ is the angle between the directions of propagation and of the earth's magnetic field.

More recently R.A.Smith (1957) has obtained the following closer approximation:

When

$$N < 500, \quad \Omega = 10^7, \quad 3 \times 10^5 < \nu < 1.2 \times 10^6 , \ |\theta| < 65^\circ,$$

then

$$k_d = \frac{K\,(1 + \cos^2\theta)\,N\nu}{\nu^2 + \epsilon^2\,\Omega^2 - p_o^2\,(A_1\epsilon - A_2\epsilon^2)} \tag{65}$$

---

* When required the weaker effects due to the ordinary component can be calculated in a similar way.

where

$$\left.\begin{aligned} p_0^2 &= 3.19 \times 10^9 \quad N \\ A_1 &= \tfrac{1}{4}\,(2 + 3 \sin^2 \theta\,) \\ A_2 &= \tfrac{1}{2}\,(1 + 3 \sin^2 \theta\,)\,. \end{aligned}\right\} \tag{66}$$

From (64) or (65) it is clear that under the specified conditions $\hbar_d$ passes through a notable maximum value as ω varies about the value Ω. On the other hand (51) shows that Ξ_1 may then pass through either one or two maxima according as the path length s_d is shorter or longer than some particular value. As shown by (57) this behaviour will be reflected in S_1 through the integral

$$\int_0^s \Xi_1 \, q \, ds \quad ,$$

and therefore, by (43.1) and (44), the modulations f impressed by a pulsed or sinoidally modulated gyro-wave will pass through one or two maxima as the disturbing wave's frequency varies about the gyro-frequency. The curves representing S_1, M_1 or M_2 in terms of ϵ, ω, or $\omega/2\pi$, may therefore be called resonance curves.

Whether the extraordinary gyro-wave is transmitted from the ground or from a rocket or a satellite in (or above) the ionosphere, calculation shows that in general the wave is almost completely absorbed within a region of the ionosphere about 6 kms long.

Consequently we may expect the interaction to be greatest when the ionospheric path of the wanted wave lies within this region.

With the gyro-wave transmitted from the ground this region at night lies in a horizontal slab with its lower face at about 90 kms and, for $\epsilon = 0$, its upper face at about 95 kms above the ground. As $|\epsilon|$ increases from 0 to 0.2 the upper face rises from about 95 kms to about 97 kms.

Consequently, when the top of the wanted wave's path lies below 95 kms the resonance curve for the impressed modulation in terms of ϵ possesses only one maximum (dromedarian resonance). But when this top lies above 95 kms the resonance curve possesses two maxima nearly symmetrically situated about the minimum which occurs near $\epsilon = 0$ (bactrian resonance); for increase of $|\epsilon|$ from 0 initially causes the slab to embrace an increasing portion of the wanted wave's path and so increases the interaction.

This deduction is supported by the published curves for the Index of Interaction Ξ_1 at different levels in a specified model of the lower ionosphere (Bailey, 1938a). For these have one maximum at the lower levels and two maxima at the higher levels.

For a particular model of the nocturnal ionosphere and the conditions relating to experiments on gyro-interaction at Armidale, New South Wales, during the period 1950-1954, R.A. Smith (1957) has carried out the computations which are needed to evaluate q ds and the integral in (57).

These computations led to a theoretical resonance curve for the impressed modulation which was in agreement with the observed resonance curves (Bailey, Smith et al, 1952) within the limits of experimental error. Some of these curves are shown in Figure 3 below.

When the disturbing and wanted waves are both propagated vertically the computation of the integral in (57) is notably simplified. The experimental work is also much simplified provided that the power is increased sufficiently to compensate for the reduction in interaction which would otherwise occur.

If also the wanted wave is an ordinary mode under the conditions discussed above, the equation (57) may be replaced by (57.1) and so the theoretical computations then become simpler still.

The impressed modulations pre-determined by means of this theory of gyro-interaction have been fully confirmed within the limits of experimental error by subsequent experiments.

SECTION 5

PREDICTIONS FROM THE MATHEMATICAL THEORY.

The formulae (44), (45) and (46.1) lead to the predictions labelled (1) in the Section 1.

The formulae (45) specify the manners in which M_1 and M_2 decrease with increasing n as stated in the prediction labelled (2). These variations with the modulation frequency are similar to the variation of the current in a choke coil as the frequency of the applied p.d. is varied.

The formula (46.1) shows that when E_{e1}^2 is not too large B_s is nearly proportional to E_{e1}^2 and therefore also to the mean power $\overline{P}$ of the disturbing station. It then follows from (45) that M_1 and M_2 are nearly proportional to $\overline{P}$ and so we obtain the prediction (3),

The formulae (44) and (33) show that

$$\varphi_1 = \tan^{-1} (n/G\nu_s) , \quad \varphi_2 = \tan^{-1} (2n/G\nu_s) .$$

Hence, when the mean power of the disturbing wave is such that ν_s the steady value of the perturbed collision frequency differs only slightly from ν_0 the unperturbed value, we obtain the prediction (4).

The formulae (45) and (46.1) show that M_1 and M_2 are proportional to A_s , where A_s is given by (24) with ν_1 replaced by ν_s . Hence, when ν_s differs only slightly from ν_0 we obtain the predicted behaviour (5) and also the single maxima predicted under the label (6). But the more precise discussion of the sub-section 4.9 shows that when s_d the path of the disturbing wave exceeds a certain value each of the modulations depths passes through two maxima. Also, computations by means of the formulae given in the sub-section 4.9 show that when $\varepsilon = 0$ a mean power of 1 or 2 kW. radiated from a vertical aerial suffices to produce observable interaction. All this is summed up in the prediction (6).

The rays from an average radio station which produce the largest values of E_{e1}^2 in the E-region are inclined at 45° to the vertical. Consequently the ionospheric regions where a wanted wave is most modulated by the disturbing station will be about 100 kms from the ionospheric region vertically above this station. This is the prediction (7).

The formulae (43.1), (43.2), (16) and (33) show that, when the disturbing wave is such that ν does not differ much from ν_0, the modulation impressed by a pulsed disturbing wave rises and falls approximately according to the same exponential law with the time constant $t_0 = 1/G\nu_0$. This yields the prediction (8).

As the mean power $\overline{P}$ of a wave increases the modulations of its amplitude with the same depth M but different frequencies n will, by (32), cause modulations of ν of depths which are greatest for the lowest frequencies. Consequently the modulations of the lowest frequencies will be received with the smallest amplitudes. This self-distortion must increase with $\overline{P}$ and so we obtain the prediction (9). Quantitative discussions of self-distortion have been given by Hibberd (1955, 1957).

The prediction (10) follows immediately from the fact that the low-frequency radiation from an electric storm reaches the nocturnal E-region with great intensity and so considerably increases the local electron collision frequency for periods of the order of one millisecond.

The same general mathematical theory (Bailey, 1937b) which yielded the equation (11), for a single disturbing wave, also yielded a formula (there labelled (35)) for the variation of the collision frequency caused by two or more disturbing waves. The latter formula contained terms representing combination tones and so the prediction (11) immediately follows.

The predicted phenomena (12) and (13) are of great interest and possible future importance for the study of the ionosphere and ionospheric radio propagation. Since no experimental tests of the prediction (12) have as yet been carried out, it is sufficient here to refer to the original publications in which it is discussed in detail (Bailey, 1938, 1959). The predictions (13) are under experimental test by Professor R. A. Smith at the University of New England in N. S. Wales, as indicated in section 8.

SECTION 6

EXPERIMENTAL TESTS OF THE THEORY AND VERIFICATION OF THE PREDICTIONS.

The principal experiments which served to test one or more of these predictions will now be mentioned in their historical order.

Van der Pol and van der Mark (1935) measured the depth of modulation M impressed on Radio Beromunster by Radio Luxembourg at four different frequencies of modulation n. Their observations are represented by the four dots shown in the Figure 1. The theoretical curve representing M_1 as a function of n, with appropriately selected values of the two constants $(B_s M)$ and t_s in (45), is also shown there. The resulting value of $G\nu_0 \div 1/t_s$ then yielded a value of ν_0 which was consistent with the values of ν_0 previously estimated by E. V. Appleton (1932) and S. Chapman (1932) from other experiments. This constituted the earliest verification of the prediction (2) with regard to M_1. Also the measured depths of modulation were later found to agree in order of magnitude with the theoretical value M_1 given in (45), thus verifying a part of the prediction (1).

Under the guidance of E. V. Appleton (1934) and Ralph Stranger (1934) the members of the World Radio Research League in Europe made observations which roughly verified the prediction (7).

In 1935 Baümler and Pfitzer qualitatively verified predictions (2), (5) and (7).

Extensive experiments which were carried out in 1937 with the help of the British Broadcasting Corporation and other European authorities verified at least the last part of the prediction (6) (Bailey, 1937d). During these experiments E. W. B. Gill of Merton College, Oxford, reported observations which appeared to verify prediction (11).

From 1946 onwards M. Cutolo of Naples and his colleagues successfully made extensive observations of gyro-interaction using disturbing waves with powers less than 1 kW. They also succeeded in demonstrating the existence of both double-humped (bactrian) and single-humped (dromedarian) resonance. Thus they verified the whole of the prediction (6).

Commencing in 1946 L. G. H. Huxley and J. A. Ratcliffe and their respective colleagues carried out extensive experiments, both separately and jointly, which verified the predictions (1), (2), (3) and (4) (excepting the proportionality of M_2 to M^2). Moreover, Huxley, Foster & Newton (1947) observed the gyro-interaction caused by the 10 kW station at Stagshaw; this is consistent with the last part of prediction (6). In addition, Ratcliffe & Shaw (1948) directly verified the fundamental concept of the theory that the amplitude of the wanted wave is decreased when the disturbing wave is operating. Also Ratcliffe & Shaw successfully applied the theoretical formula for the phase lag φ_1, given under (33), to determine experimentally the height of the ionospheric region at which the wave-interaction took place. Since the heights so determined were of the same order as the heights postulated in the original theory, we may regard this work as an experimental verification of the prediction (4).

In 1949 experiments on gyro-interaction were commenced in Eastern Australia by V. A. Bailey and colleagues at the University of Sydney in collaboration with R. A. Smith and his colleagues at the University of New England at Armidale, (N.S.W.). These were planned to conform as closely as possible to the published requirements of the theory (Bailey, 1937c and 1938a). Accordingly the arrangement of stations shown in the Figure 2 was adopted. In this the disturbing gyro-wave A was radiated vertically from a horizontal aerial at Armidale, which lies at the mid-point of the 740 km. line that joins Brisbane and Katoomba. The wanted wave B was radiated from Brisbane on 590 kc/s and received at Katoomba. The local gyro-frequency at Armidale was estimated to be 1530 kc/s.

The gyro-wave was radiated 40 times per second as a square pulse of length 1 millisecond and power 36 kW. Its frequency f was kept constant for 2 minutes at each of 8 values lying within ± 23% of the gyro-frequency 1530 kc/s. Each of the cycles 1, 2 and 3 ran through the middle 6 frequencies and lasted 17 minutes. Cycle 4 ran through all the 8 frequencies and lasted 23 minutes.

In order to provide a standard of comparison for the depth of the modulation impressed on the wanted wave B the latter was modulated at its source for 30 seconds in each 2-minute period with a standard tone of 80 c/s to a depth of 5%. This depth was chosen because computations based on our mathematical theory of gyro-interaction had indicated in advance that the experimental arrangement shown in Figure 2 would impress on B modulation depths of the order of 5%. This prediction was nicely verified by the experimental observations.

A detailed account of these Australian experiments on gyro-interaction was published in 1952 by Bailey, Smith, Landecker, Higgs and Hibberd. A briefer account was published by Bailey in 1956. Examples of the experimental resonance curves are shown in the Figure 3. It will thus be seen that these experiments not only verified the predictions (6) and (8) but also confirmed the theory quantitatively in every particular.

The theory of interaction, including gyro-interaction, was also thoroughly verified by the experiments on micro-wave interaction in gas-discharge tubes of Goldstein, Anderson and Clarke (1953) and of Anderson & Goldstein (1955, 1956). In some of these experiments Goldstein was able to demonstrate also the self-distortion of a powerful wave (Private communication, 1955) thus verifying the prediction (9) for a plasma.

The prediction (9), that self-distortion in the Ionosphere increases with the power and does so at the expense of the lower modulation frequencies, was verified by J. W. King (1957)*.

---

* In this publication King inadvertently contradicts the statement on "Ionospheric self-interaction" previously made by Hibberd (1955) namely: "Its existence was first predicted by V.A.Bailey in 1935,..."

In carrying out his experiments he was guided by the quantitative theory of self-distortion developed by F. H. Hibberd (1955, 1956).

Other tests of the theory arose from the experimental study of wave-interaction by J. A. Fejer (1955) in which both the wanted and disturbing waves were pulse-modulated. This study demonstrated the existence of interaction during the day at heights between 70 and 90 km and also that a modification of the theory suggested by Huxley was not valid.

The prediction (5) is now verified by the fact that during the last 30 years the greatest disturbance by interaction backgrounds has been produced by the radio stations with the lowest wave-frequencies.

From the whole discussion given above we see that experimental observations have verified the ten predictions (1) to (9) and (11). The predictions (10), (12) and (13) remain to be tested. We may therefore conclude that in the form presented here the theory of wave-interaction is experimentally established.

SECTION 7

SOME APPLICATIONS TO RESEARCH ON THE IONOSPHERE AND GAS DISCHARGES AND TO RADIO COMMUNICATION.

The experiments mentioned above by van der Pol and van der Mark, Cutolo, Huxley, Ratcliffe & Shaw, Bailey & Smith, Fejer, combined with the theory, have between them served to measure the following important quantities relating to the Ionosphere:

The electron collision frequencies ν , the electron densities N , the local magnetic fields H , the constant G (relating to the mean loss of energy by an electron at a collision with a gas molecule) and the heights at which the wave-interaction occurs.

Similar applications, to the D-region, have been made more recently by Fejer & Vice (1959), Bjelland et al. (1959) and by Barrington & Thrane (1962).

The experiments by Goldstein and his associates mentioned above also demonstrate the value of the phenomenon of wave-interaction and its theory for investigating several fundamental processes in a gas-discharge.

As an illustration of its practical value for radio-communication it may be mentioned that the author has been consulted about it more than once by the leading broadcasting organisations in Europe and the U.S.A. Also, in the U.R.S.I. Information Bulletin No. 73, on p. 59, it was reported that the C.C.I.R. stressed the practical importance of information on non-linear effects in the Ionosphere.

The present trend towards the use of increasingly powerful beams in connexion with the reception of irregularly scattered waves has also compelled attention to the possible occurrence of observable wave-interaction and other non-linear phenomena at very high frequencies.

Lastly, the possibility mentioned by Bailey & Goldstein (1958) of controlling the Ionosphere in various ways by means of powerful radio waves has attracted some attention.

SECTION 8

SOME RECENT DEVELOPMENTS

Filling in the gap between the two situations corresponding to wave-interaction caused by moderate powers and ionospheric airglows caused by extremely high powers and large aerial arrays, an investigation was made (Bailey, 1959) of the possible increases of both the collision frequencies ν and the electron densities N caused by a powerful extraordinary circular gyro-wave in the nocturnal lower E-region and in the daytime D-region. The theory was based on the most reliable available experimental data on the behaviour of free electrons in air with energies up to 1.22 ev . The principal process then freshly introduced was the attachment of electrons to molecules in some collisions.

It was found that some of the resultant effects, of such increases in ν and N , on this wave or on other waves are remarkable in magnitude or in kind. Thus, when the gyro-wave produces near the 92 km. level at night a reduced equivalent electric force (Z/p) equal to 2.5 V/cm per mm Hg (measured at 288° K) it causes notable self-distortion and extremely large wave-interaction; also within 3 milli-seconds its penetration into the E-region increases considerably.

Other possible observable effects which may be caused at night when Z/p has a value between 2.5 and 8 are changes in the visible and infra-red airglow, in parts of some trails of meteors and fast particles of solar or cosmic origin, in the radar echoes from such trails and in the local magnetic elements.

The theory also showed that a pulse of gyro-waves such as that specified above could be used to determine by experiments on the ground the rates of attachment of electrons and other electron-collision data pertaining to the regions near the 92 and 75 km levels.

For these and other reasons it was therefore suggested that experiments be carried out with gyro-wave pulses beamed from an array of at least 80 horizontal dipoles connected to a generator of at least 500 kW power.

As a result of this work and of my recommendation to the Air Force Cambridge Research Center at Bedford (Mass.), the University of New England is under contract with this Research Center to carry out a project to "Modify the Ionosphere by means of Gyro-waves" using an array of 40 dipoles connected to a generator of 500 kW power. The results of this project are being reported by Professor R. A. Smith.

The power density requirements for the excitation of an artificial airglow by means of gyro-waves (Bailey, 1938b) have been reconsidered (Bailey, 1961) in the light of later information. One important conclusion was that an array of 80 dipoles radiating vertically a beam of gyro-waves with a mean power of 500 kW would produce an easily observable enhancement of the nocturnal airglow overhead within a solid angle of 32 square degrees. Another conclusion is that from the same array pulses of power 3500 kW and length about 0.2 ms would each produce a flash of airglow that is 40 times as bright as the night sky.

An extensive report on non-linear phenomena in a plasma was published by Ginzburg & Gurevich in 1960 and is available in an English translation (1961). The theory is based on the Boltzmann equation for the energy distribution function of the electrons, thus following the classical examples of H. A. Lorentz and F. B. Pidduck. The authors point out that "to disclose the physical picture" the simpler "elementary theory" is "convenient and useful". Our present account of the "elementary theory" confirms this opinion.

Similar discussions based on Boltzmann equations for the electronic distribution functions are contained in reports by Megill (1961), Carleton & Megill (1961) and Molmud et al (1962). These are directed towards special problems of considerable interest such as, respectively, the generation of artificial airglows and a method to reduce the electron density in the D-region, by means of powerful radio transmitters on the ground. In all these discussions the use of simplifying approximations was found necessary.

An interesting new way of using wave-interaction for studying the Ionosphere has been proposed independently by Vilenski (1962) and Ferraro, Lee & Weisbrod (1963). This depends on measurement of the phase changes in the wanted wave that are caused by the disturbing wave.*

---

* Ferraro and his co-authors mistakenly attribute the origin of the theory of wave-interaction to its discussion in 1948 instead of to the correct date 1934.

A remarkable new example of the interaction of combination tones of two high-frequency waves with a third (wanted) wave, similar to that mentioned under prediction (11) in Sections 1 and 5, has been reported by Cutolo (1962).

The waves concerned and our theoretical processes were as follows:

(1) Disturbing wave; A carrier wave of frequency $\omega/2\pi$ between 50 and 75 Mc/s was modulated at a frequency of $\Omega/2\pi$ between 1000 and 1300 Kc/s which lies near the local gyro-frequency of 1200 Kc/s; the peak power was 80 kW and the modulation depth m about 42%. Thus the (Yagi) aerial used simultaneously radiated three waves of the angular frequencies $\omega + \Omega$, ω and $\omega - \Omega$. Two pairs of these produced in the Ionosphere electric currents which contained components with frequencies equal to the combination tones $(\omega + \Omega) - \omega$ and $\omega - (\omega - \Omega)$ i.e., equal to the gyro-frequency Ω . These ionospheric current components generated secondary waves of the frequency Ω which in turn caused gyro-resonance variations of the local collision frequency ν .

(2) The wanted waves were selected so as to be reflected at or above the 90 km level in the Ionosphere in accordance with the requirements for gyro-interaction given in our Section 4.9.

Cutolo's experiments demonstrated that in this way modulations of depths from 5% to 20% could be impressed on the wanted waves near the 90 km level of the Ionosphere when the <u>modulation</u> frequency approximated to the gyro-frequency.

Lastly we may consider the waves that could be generated in the interplanetary plasma by the large negative charge on the Sun which has been postulated in recent years to account simultaneously for a number of important phenomena (Bailey, 1960). The evidence in support of this hypothesis has recently been strengthened by the fact that three predictions (based on this hypothesis) that were published in 1960 have now been verified by the data on the interplanetary magnetic field obtained by means of the space probes Pioneer 5, Explorer 10, Mariner 2 and Explorer 12 (Bailey, 1963).

The electrostatic field $\underset{\sim}{E}_s$ due to the Charge $-Q_s$ on the Sun is radial to the Sun and has the magnitude

$$E_s = fQ_s/r^2 \text{ e.s.u.}$$

at the radial distance r, where f is a number between 0 and 1 which represents the partial screening of the solar charge by the interplanetary plasma clouds.

Any fluctuations in the emission by the Sun of these plasma clouds will cause corresponding fluctuations $\underset{\sim}{E}$ in the field $\underset{\sim}{E}_s$. In certain frequency bands the radial components of $\underset{\sim}{E}$ are particularly liable to amplification by the mean static field $\underset{\sim}{\bar{E}}_s$ since they may be regarded as forming longitudinal wave-modes.

These longitudinal waves may be sufficiently strong to produce, or help produce, certain notable non-linear effects in the Earth's Ionosphere which are at present attributed to other causes. For example, some geomagnetic disturbances* may be caused by such longitudinal waves (Heirtzler, 1962).

This possibility deserves further study, both by means of instrumented satellites and theoretically (Bailey, 1948, 1951).

---

* For example the micropulsations of shorter Type A waves (period $\sim$ 1 sec) modulated by Type C waves of periods about 20 secs. (Sci. Am., 206, 128 (March 1962)).

SECTION 9

CONCLUSION

The foregoing account of the theory of wave-interaction and other non-linear processes in the Ionosphere in the form developed at the University of Sydney shows how the ideas and methods of the Townsend School on the motions of electrons in gases and of the Appleton School on magneto-ionic wave propagation have led to a relatively simple mathematical theory which has proved fertile in predictions of new phenomena that have subsequently been verified by experiment.

It may therefore be used, and has been used, with confidence as the basis of new experimental methods for investigating processes in the Ionosphere and in electrical gas-discharges. It has also suggested methods of controlling parts of the Ionosphere for particular purposes by means of beams of radio waves.

The present trend towards increasingly large power-densities in radio beams has made this theory, and its mathematical elaborations, of increasing practical importance.

These elaborations, which are based in part on Boltzmann's equation for the distribution of electron energies, are distinctly onerous, but they serve both to check the simple mathematical theory of the Townsend School* and to study ab initio those phenomena which, like the artificial airglow, depend strongly on the actual distribution of the electron energies. On account of continual and non-uniform fluctuations of the Ionosphere and the lack of precision in the necessary auxiliary laboratory data, it is probable that for some time to come the simple theory* will serve for most purposes quite as well as the elaborations.

The present account also suggests that certain types of low-frequency longitudinal waves of extra-terrestrial origin may produce some modifications of the Ionosphere that are at present attributed to other causes.

---

* Also called, by Ginsburg, the "elementary theory".

REFERENCES

Anderson, J. M. & Goldstein, L., Phys. Rev., 100, 1037 (1955); 102, 388 (1956).

Appleton, E. V., World Radio, November 30 & December 7, 1934.

Appleton, E. V., J.I.E.E., lxxi, 642 (1932).

Bailey, V. A., Phil. Mag., 50, 825 (1925).

Bailey, V. A., Nature, 139, 68 (1937a).

Bailey, V. A., Phil. Mag., 23, 774 (1937b).

Bailey, V. A., Phil. Mag., 23, 929 (1937c).

Bailey, V. A., Nature, 139, 838 (1937d).

Bailey, V. A., Phil. Mag., 26, 425 (1938a).

Bailey, V. A., Nature, 142, 613 (1938b)

Bailey, V. A., Aust. J. Sci. Res., 1, 351, (1948).

Bailey, V. A., Phys. Rev., 83, 439 (1951).

Bailey, V. A., Nuovo Cimento, 4 (Suppl. 4), 1430 (1957).

Bailey, V. A., J. Atmos. Terr. Phys., 14, 299 (1959).

Bailey, V. A., Proc. Roy. Soc. N.S. Wales, 94, 77 (1960).

Bailey, V. A., J. Research N.B.S., 65D, 321 (1961).

Bailey, V. A., Nature, 199, 1029 (1963).

Bailey, V. A. & Martyn, D. F., Nature, 133, 218 (1934).

Bailey, V. A. & Martyn, D. F., Nature, 135, 585 (1935).

Bailey, V. A., Smith, R. A., Landecker, K., Higgs, A. J. and Hibberd, F. H., Nature, 169, 911 (1952).

Bailey, V. A. & Goldstein, L., J. Atmos. Terr. Phys., 12, 216 (1958).

Barrington, R. E. & Thrane, E., J. Atmos. Terr. Phys., 24, 31 (1962).

Baümler, M. & Pfitzer, W., Hochfreq. Elektroak., 46, 81 (1935).

Bjelland, B., Holt, O., Landmark, B. and Lied, F., Nature, 184, 973 (1959).

Butt, A. G., World Radio, April 28, 1933.

Carleton, N. P. & Megill, L. R., "Electron Energy Distributions in Slightly Ionised Gases under the Influence of Electric and Magnetic Fields", N.B.S. Boulder Laboratories, (1961).

Chapman, S., Proc. Roy. Soc., 137, 169 (1932).

Crompton, R. W., Huxley, L.G.H. & Sutton, D. J., Proc. Roy. Soc., A 218, 507 (1953).

Cutolo, M., Nature, 160, 834 (1947).

Cutolo, M., Nature, 166, 98 (1950).

Cutolo, M., Nuovo Cimento, 9, 391 (1952).

Cutolo, M., Istituto di Fisica Tecnica, University of Naples, (June 1962).

Cutolo, M., Carlevaro, M. & Gherghi, M., Alta Freq., 15, 2, p. 111 (1946).

Fejer, J. A., J. Atmos. Terr. Phys., 7, 322 (1955).

Fejer, J. A. & Vice, R. W., J. Atmos. Terr. Phys., 16, 291 (1959).

Ferraro, A. J., Lee, H. S. & Weisbrod, S., J. Geophys. Res., 68, 1169 (1963).

Gill, E.W.B., Private Communication, March 17, 1937.

Ginzburg, V. L. & Gurevich, A. V., Soviet Physics, Uspekhi, 3, 115 and 175 (1961).

Goldstein, L., Anderson, J. M. & Clarke, G. L., Phys. Rev., 90, 151, (1953).

Healey, R. H. & Reed, J. W., "The Behaviour of Slow Electrons in Gases", A. W. A. Ltd., (Sydney, 1940).

Heirtzler, J. R., Scient. American, 206, 128 (March 1962).

Hibberd, F. H., J. Atmos. Terr. Phys., 6, 268 (1955).

Hibberd, F. H., J. Atmos. Terr. Phys., 8, 121 (1956).

Hibberd, F. H., J. Atmos. Terr. Phys., 11, 102, (1957).

Huxley, L.G.H., Proc. Roy. Soc., A200, 486 (1950).

Huxley, L.G.H., Nuovo Cimento, 9 (Suppl. Series 9), 59 (1952).

Huxley, L.G.H., Foster, H. G. & Newton, C. C., Nature, 159 300 (1947).

Huxley, L.G.H., Foster, H. G. & Newton, C. C., Proc. Phys. Soc., 61, 134 (1948).

Huxley, L.G.H. & Ratcliffe, J. A., Proc. I. E. E., 97, 165 (1950).

King, J. W., J. Atmos. Terr. Phys., 10, 166 (1957).

Llewellyn Jones, F., "Electrical Discharges", Reports on Progress in Physics, 16, 216 (1953).

Megill, L. R., "Expected Upper Atmospheric Effects in the Presence of High Frequency Fields with High Power Density", C.R.P.L., N.B.S., Boulder Laboratories (1961).

Molmud, P., Altshuler, S. & Gardner, J. H., "The Study of a Method to Reduce Electron Density in the Ionospheric D-region by means of High-powered Ground-based Transmitters". Space Technology Labs., Inc., Final Report, September 1962.

Neilsen, R. A. & Bradbury, N., Phys. Rev., 51, 69 (1937).

Ratcliffe, J. A., J. I. E. E., 95, 325 (1948).

Ratcliffe, J. A. & Shaw, I. J., Proc. Roy. Soc., A, 193, 311 (1948).

Smith, R. A., Doctoral Thesis, University of Sydney, (1957).

Stranger, Ralph, World Radio, December 14, 1934.

Tellegen, B.D.H., Nature, 131, 840 (1933).

Townsend, J. S., Phil. Mag., 13, 745 (1932).

Townsend, J. S., "Electrons in Gases", (Hutchinson's Publications, 1947).

Townsend, J. S. & Tizard, H. T., Proc. Roy. Soc., A, 88, 336 (1913).

Van der Pol, B., Tijdschrift Nederland Radiogenootschap, 7, 93 (1935).

Van der Pol, B., and van der Mark, J., Tijdschrift Nederland, Radiogenootschap, 7, 12 (1935).

Vilenski, I. M., Radiofizika, 5, 468, (1962).

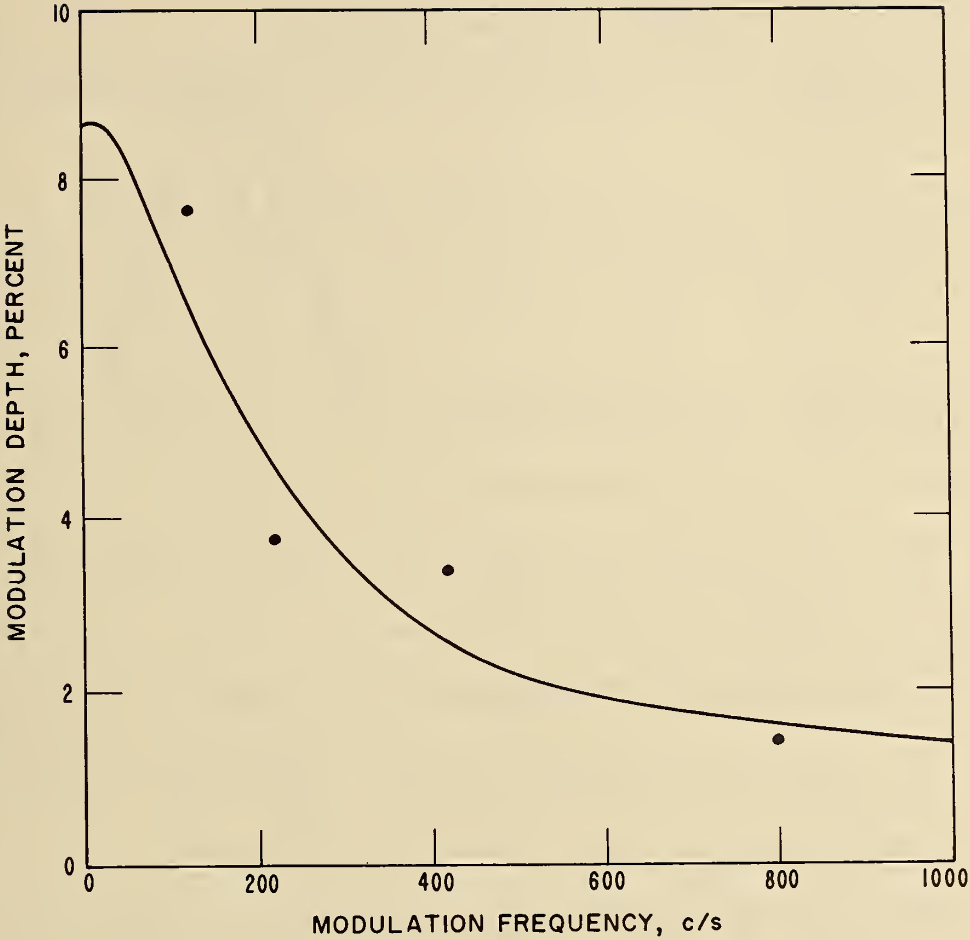


Figure 1

Bailey, V. A., and D. F. Martyn, Nature 135, 585, 1935

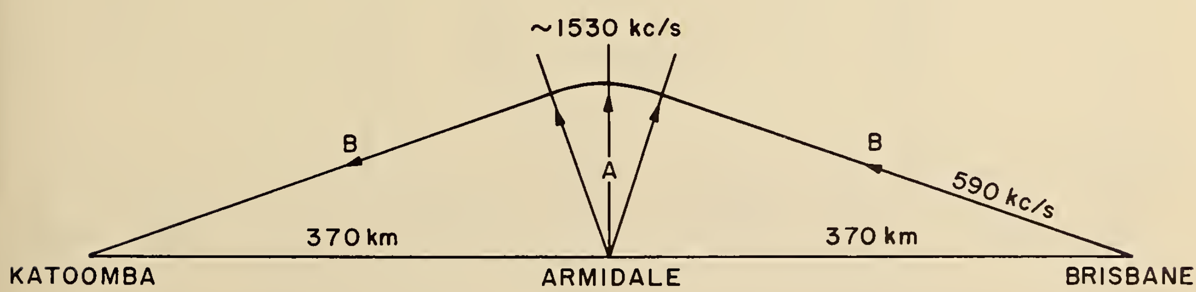


Figure 2

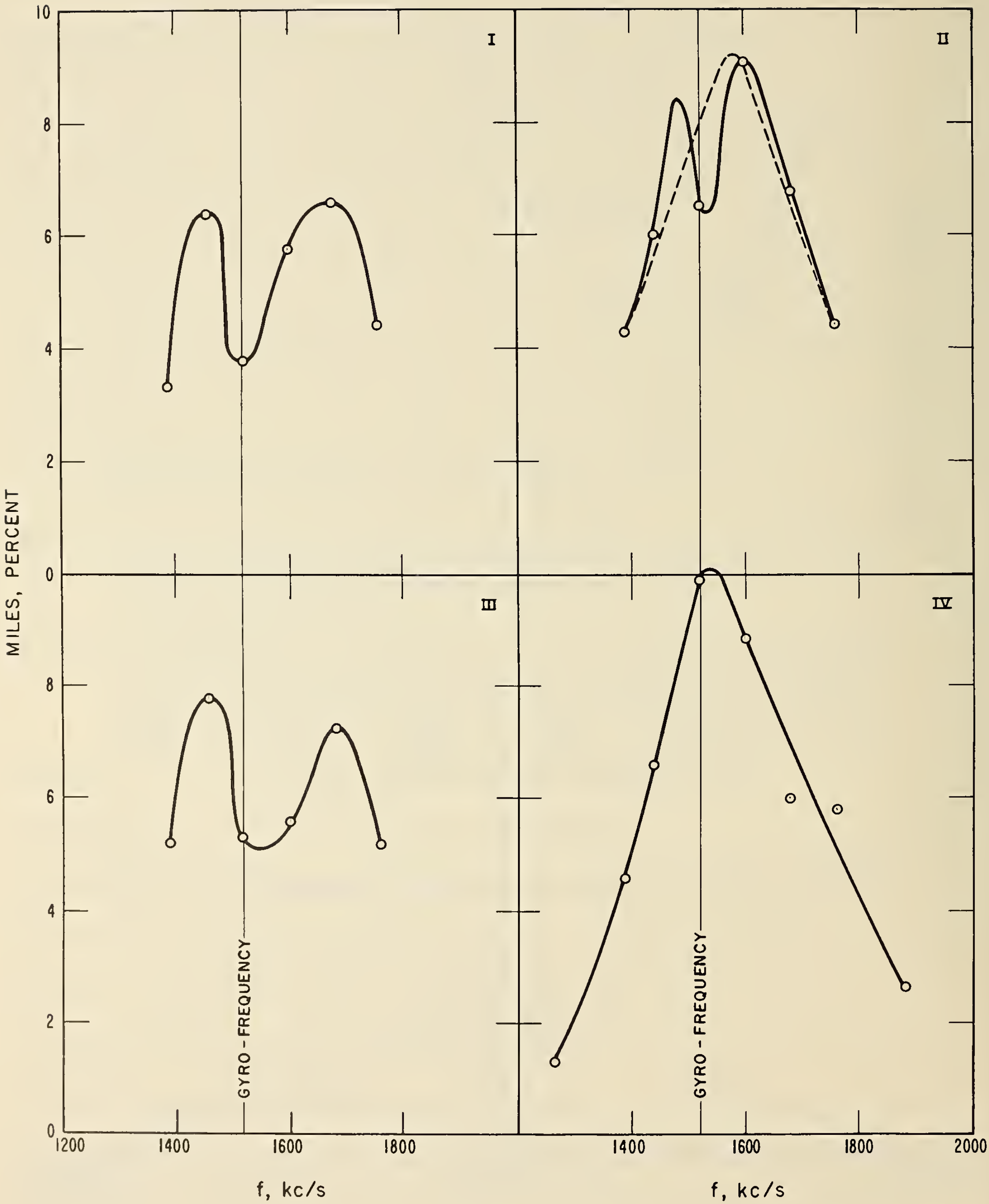


Figure 3

Bailey, V. A., R. A. Smith, K. Landecker, A. J. Higgs, and F. H. Hibberd, Nature 169, 911, 1952.

Remarks by Donald H. Menzel with Reference to Bailey's Comments on Solar Electric Fields

Professor V. Bailey has postulated the existence of a large negative electric charge on the sun and refers to apparent experimental verification of his hypothesis.

I regret to say that I completely disagree with Professor Bailey. Since electrons have velocities higher than those of the ions they would tend to escape more readily and leave the sun with a positive rather than a negative residual charge. But my objection to Bailey's hypothesis rests on the magnitude rather than the sign of the charge. For when the positive (or negative) charge attains a value such that the electrostatic repulsion on the proton (or electron) equals the gravitational attraction, the two oppositely charged particles will thereafter escape in pairs.

This condition requires that

$$\frac{Q\epsilon}{R^2} = \frac{GMm}{R^2} ,$$

wherein Q and M are the respective electric charge and mass of the sun, r the solar radius (which cancels out), and G the constant of gravitation. The quantities ϵ and m refer respectively to the charge and mass of proton (or electron).

Thus the voltage at the surface of the sun is

$$V = \frac{Q}{r} \text{ esu} = \frac{G M m}{\epsilon R} \text{ esu} = 300 \frac{G M m}{\epsilon R} \text{ volts}$$

This equation leads to potentials of

$$V = 6.64 \text{ esu} = 2000 \text{ volts, if positive}$$

or

$$V = 3.6 \times 10^{-3} \text{ esu} = 1.08 \text{ volts, if negative.}$$

These potentials are many orders of magnitude smaller than those postulated by Bailey, 10^{17} volts or higher. No process of producing an electric field could possibly reconcile this disagreement.

Bailey has based his conclusion on the postulate that cosmic-ray energies occasionally attain the figure of 10^{17} electron volts. But if cosmic rays actually derived these energies by falling through a solar electric field, they would be highly directional. One concludes that the most energetic cosmic rays do not derive from solar phenomena. I do not preclude the possibility that cosmic rays of low energy may be produced by an electromotive force resulting from changing solar magnetic fields.

The space probes that Bailey refers to have measured only magnetic, not electric fields. No one denies that the sun possesses high electric currents. But these currents must be galvanic in character and not the result of a rotating charged sun. A sun carrying so high a charge would be violently unstable.

IONOSPHERIC CROSS-MODULATION AT THE GEOMAGNETIC EQUATOR

W. K. Klemperer
Jicamarca Radar Observatory, Apartado 3747, Lima, Peru*

Various experiments to detect radio wave interactions in the lower ionosphere have been carried out using the 22 acre antenna and 4 megawatt transmitter of the Jicamarca Radar Observatory. Although equatorial sporadic-E severely limits observation time, reduction in the amplitude of 3 Mc/s F layer echoes by as much as 25% is readily obtained from the classical Luxembourg effect. Cross-modulation of 50 Mc/s cosmic noise has also been obtained. In this experiment the lower ionosphere is first heated with a 3 millisecond, 4 megawatt rf pulse. A decrease in sky brightness temperature (about $6{,}000^{\circ}$ K, well away from the galactic center) of $30^{\circ} \pm 10^{\circ}$ K is occasionally observed with a recovery time on the order of 2 milliseconds. Both the Luxembourg and cosmic noise observations indicate that the interaction height is above 75 km, and essentially no cross-modulation was observed below 65 km. The lower boundary of the D region thus appears to be substantially higher over the geomagnetic equator than in other parts of the world. This is in agreement with theories of D region formation attributing most of the lower level ionization to cosmic ray and solar particle influx, which is significantly less at the geomagnetic equator.

Three other experiments of an exploratory nature for the detection of "sidebands" or combination frequencies near the electron gyrofrequency were carried out. No such effects were observed.

---

*A cooperative project of the Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colorado, and the Instituto Geofísico del Perú, Lima, Perú.

1. Introduction

The lower ionosphere remains a challenging frontier. Because of atmospheric drag, it is a region inaccessible to satellites, so that one must rely on rocket probes or traditional radiowave methods to obtain information about the environment at altitudes in the 40-90 km range. In a masterful review of all ionospheric knowledge, Ratcliffe and Weekes [1960] published suggested electron density distributions for the D region based on various sources. In their words, the results were somewhat disappointing. The power of radio and rocket techniques has measurably increased over the past years, yet it is still very difficult to obtain D region electron density profiles.

The purpose of this note is to describe some of the radiowave interaction experiments carried out at the Jicamarca Radar Observatory in Peru. This station is the location of an antenna of 84,000 m^2 aperture fed by a 50 Mc/s, 4 megawatt pulse transmitter. Although the frequency is somewhat higher than one would normally employ for cross-modulation work, the field strength attainable at E region heights is sufficiently intense to produce an order-of-magnitude increase in the electron temperature. With this "disturbing transmitter", the detection of cross-modulation using a 3 Mc/s "wanted transmitter" (to use the terminology of the Luxembourg effect) is not at all difficult, and some fairly spectacular effects have been observed. Cross-modulation of 50 Mc/s cosmic noise, on the other hand, is extremely difficult to detect, but some results have been obtained. No effects due to gyrofrequency "sidebands" at 50 Mc/s as reported by Cutolo [1963] have been observed at Jicamarca. The experiments will be described, as well as a few preliminary results which have been obtained. The cosmic noise modulation experiment will form the subject of a later paper.

2. "Luxembourg" Experiments

Fejer [1955] demonstrated that the extremely weak radiowave interactions occurring in the lower ionosphere between pulsed transmissions on about 2 Mc/s could be detected, using a disturbing transmitter of only 5 kw peak power. This technique has been followed by Barrington and Thrane [1962], Bjelland, et al. [1959], Landmark and Lied [1961], Rumi [1962a and 1962b], and Smith [1964].

Cross-modulation effects were obtained using the Jicamarca radar as early as 1961, with only a portion of the antenna completed [Bowles, private communication]. Because of the radar's narrow beam (1° between half-power points when operated on axis and using the full $48\lambda \times 48\lambda$ aperture), ray paths for the wanted wave occasionally fall outside of the heated region due to normal ionospheric layer tilts and irregular reflection effects. To ensure that data would be taken at the proper times, two interferometer systems of 6λ spacing operating at 3 Mc/s were installed on N-S and E-W baselines and operated continuously as direction finders. The wanted wave was generated by a 20 kw pulse transmitter (also at about 3 Mc/s) and beamed vertically using an array of four folded $\frac{\lambda}{2}$ dipoles. This short 50 μsec probing signal was sent out twice as often as the 50 Mc/s disturbing pulse, and F layer reflections were examined on alternate sweeps for evidence of the heating produced. When using short disturbing pulses, a phase sensitive detector was used to search for the weak cross-modulation produced. The sensitivity and stability of the equipment in this mode of operation were sufficiently good to detect as little as 0.3% cross-modulation, using a long integration time of 50 seconds. Usually an integration time of 1 second was used, for which 2% of cross-modulation gave a deflection which was in excess of 10 times the noise level. One of the advantages of working so near the geomagnetic equator is that a single magneto-ionic component can be selected, using simple linear polarization. All the wanted wave antennas were aligned N-S. A serious disadvantage is the presence of equatorial sporadic E (E_{sq}) in the daytime. These

E region irregularities caused by the equatorial electrojet [Bowles et al., 1963; Farley, 1963] forward scatter the wanted wave to such an extent that measurement of cross-modulation becomes difficult or even impossible. This has severely limited observation time. Nevertheless, enough data has been obtained using F layer reflection of the wanted wave to show that usually as much as 12% cross-modulation is obtained using 3 millisecond pulses of 2 megawatt peak power and half of the available antenna aperture for the disturbing transmitter. Spectacular (25%) cross-modulation is observed occasionally. By varying the disturbing pulse length, one can estimate the time constant for heating. It is on the order of 1-2 milliseconds. Two examples of cross-modulation using F layer reflections are shown in figure 1. No effects have been observed at heights of less than 70 km using shorter disturbing pulses, in spite of greatly increased sensitivity (phase-sensitive detection) and diligent searching.

A sizeable increase in sensitivity can be attained by using E layer rather than F layer reflections. Although smooth E layer reflections at our latitude occur but rarely, the wanted wave can then be made to transit the disturbed region twice before much cooling takes place and thus double the amount of cross-modulation produced. Examples of cross-modulation of E layer echoes are shown in figure 2. As with the F layer reflections, the time constants were found to be long; and practically no cross-modulation was detected below 65 km, although the equipment was certainly capable of detecting it. By working close to the E layer critical frequency, cross-modulation in excess of 50% (6 db) could be obtained. Examples of this are shown in figure 3. This indicates that most of the cross-modulation is taking place at the level of reflection, in agreement with the interaction height estimates from the short pulse experiments. When interaction takes place near the level of reflection, the retardation in the medium (difference in group and phase path) enhances cross-modulation according to the relation [Ratcliffe, 1959, p. 116]:

$$- \log \rho = \frac{\nu}{c} (P_g - P_p), \qquad (1)$$

where: ρ = reflection coefficient,

ν = collision frequency for electrons,

P_g = group path,

P_p = phase path,

c = 3×10^8 m/sec.

Under these conditions (copious amounts of cross-modulation and long time constants) near the E layer critical frequency, the experiment can be set up to measure time constants directly by passing three probing pulses through the disturbed region in rapid succession. This is illustrated in figure 4. Note that the echo from the lower height shows a shorter recovery time.

3. Cross-modulation of Cosmic Noise

It occurred to Rumi [Benson, 1963] while working at the Geophysical Institute of the University of Alaska that a simplification of the usual pulse interaction experiment of the type just described should be possible by using cosmic noise as the wanted wave. In addition to removing the requirement for the wanted transmitter, this permits operation during ionospherically disturbed periods (such as strong absorption at high latitudes or E_{sq} at the geomagnetic equator). Observations of the cross-modulation of cosmic noise have been reported by Benson [1962 and 1963].

An experiment to observe this effect at 50 Mc/s was set up in the fall of 1963. The 4 Mw transmitter of the Jicamarca Radar was connected to one linear polarization (NE-SW) of the 84,000 m^2 antenna, and a wide bandwidth receiver was connected to the orthogonal polarization (NW-SE). In this way the entire 2 Mc/s bandwidth capability of the large antenna could be exploited without TR problems, since the attenuation between polarizations is greater than 60 db and the receiver is blanked during the transmitter pulse. Normal incoherent scatter echoes from the F region (Faraday rotated) as well as cross-polarized components of the E_{sq} reflections were removed by a 10 kc/s wide notch filter tuned to the operating frequency.

The sensitivity required to detect modulation of cosmic noise at 50 Mc/s even with such a powerful system approaches the statistical limit of detectability given by

$$\Delta T = \frac{T_s}{\sqrt{B\tau}}, \tag{2}$$

where T_s = sky brightness temperature at 50 Mc/s (minimum ~ 5,000 °K, maximum ~ 45,000 °K),

B = system bandwidth (limited to 2 Mc/s by the antenna),

τ = integration time,

ΔT = rms fluctuations in temperature.

We can neglect receiver noise contributions in comparison with the 5,000 °K minimum sky brightness temperature. τ is limited by the range resolution required and the number of range sweeps which can be conveniently integrated. For example: with a detector time constant of 150 μsec, a sampling interval of 133 μsec, and integration of 10,000 sweeps, τ is about 1.5 seconds. The flucutations in temperature according to (2) would then be at least 3.5 °K for a T_s of 6,000 °K and a two megacycle bandwidth. Under the conditions of the experiment, the fluctuations were about 10 °K, the additional "noise" being due to receiver and integrator gain fluctuations and digitizing errors. Figure 5 shows the change in the cosmic noise level caused by heating with a 3 millisecond pulse. The decrease in sky brightness amounts to only 30 °K in 6,100°.

4. Strong, Non-linear Effects (gyro sidebands)

Three experiments of an exploratory nature for the detection of "sidebands" or combination frequencies of the electron gyrofrequency with the carrier·or carriers were carried out and will be briefly described. The first of these was motivated by a report that Cutolo [private communication] in 1963 had succeeded in detecting the effect on the lower ionosphere of a small 50 Mc/s radar modulated at the gyrofrequency. After the 50 Mc/s and 3 Mc/s pulsed interaction experiments at Jicamarca showed easily detectable results, it was decided to try to duplicate the experiment of Cutolo using higher power. Accordingly, two of the four one-megawatt transmitters were detuned slightly, one to 50.4 Mc/s and the other to 49.6 Mc/s, so that the frequency separation was near the predicted gyrofrequency. The Luxembourg effect experiments were repeated with this arrangement to see if there was any enhancement of the modulation transferred to the 3 Mc/s signal. The frequency of separation was varied from 750 to 900 kc/s in small increments. The pair of one-megawatt transmitters remaining on 50 Mc/s was frequently substituted for the pair separated at the gyrofrequency to show that the normal cross-modulation was produced. No detectable difference (to within 10-20%) could be observed between the two methods of heating the lower ionosphere.

A second experiment was carried out, using the same transmitters (50.4 and 49.6 Mc/s) separated by the gyrofrequency, but with reception at the sum frequency (100 Mc/s). The receiving antenna was a 4 x 4 array of dipoles over a reflecting screen. No effect was detected.

The most sensitive test for a strong, non-linear effect was the search for "sidebands" on the E_{sq} echoes. These echoes are often 100 db above the noise when the entire radar is beamed perpendicular to the magnetic field over Jicamarca. Under conditions of strong E_{sq}, the receiver was detuned and a search was made for echoes at 50 Mc/s $\pm$ 150-900 kc/s. No "sidebands" were detected down to parts in 10^{-7} of the echo power received at 50 Mc/s.

5. Discussion

An interesting feature of the cosmic noise cross-modulation observations (figure 5) is the decrease in sky brightness temperature and the long recovery time. Although the 3 millisecond disturbing pulse increases the electron temperature in the lower ionosphere by an order of magnitude (the rms E field during the pulse at 100 km height near the center of the antenna beam is about 15 volts per meter), the ionospheric contribution to total sky brightness temperature observed at 50 Mc/s remains negligible. A decrease in sky brightness is observed at 50 Mc/s. This is because the ionosphere remains essentially transparent to 50 Mc/s radiation so that any increase in temperature is not directly apparent. An analogous situation is found in the solar corona which, at a temperature of a million degrees Kelvin, has a brightness only a few millionths as great as the $6{,}000^{\circ}$ photosphere. The decrease observed in the 50 Mc/s cosmic noise level is evidently due to an increase in ν, the electron collision frequency, which has the effect of increasing the absorption slightly. The recovery time is related to the time constant for ionospheric heating which in the simple theory is

$$t = 1/G\,\nu \qquad (3)$$

where G is the fractional energy loss per collision and is a small factor x 10^{-3}. Figure 5 shows two time constants, one of approximately 2 milliseconds, corresponding to a collision frequency of 5×10^{5} or a height of about 80 km, and a weaker component with a longer time constant which may be of E region origin.

The approximately 2 millisecond time constant found in the cosmic noise data agrees with the time constant for heating found in the Luxembourg type experiments performed at Jicamarca. By the technique shown in figure 4, the time constant can be determined to considerable precision for several heights. Values are all in the range 1.5 to 2 milliseconds for the echoes from 80 km and above, while an occasional echo from the lower 70 to 75 km height shows a time constant of less than one millisecond. There are some indications that G in (3) may not be a constant but depend on energy [Phelps, private communication] and if so this complicates the analysis considerably.

A comparison can be made between the large cross-modulation produced at 3 Mc/s and the very weak effect detected in the 50 Mc/s cosmic noise. The 3 Mc/s effects were produced using but one-half of the available transmitter power and one-half of the 22 acre antenna. The cosmic noise experiment made use of the whole system. The 50 Mc/s sky brightness changed by 0.02 db. If the 3 Mc/s experiment--which showed, on the average, a 15% change in the F region echo amplitude (1.5 db)--had been performed under the same conditions as the cosmic noise experiment with four times the gain, the cross-modulation would have been 6 db. The ratio of these two numbers is 1:300, which is approximately equal to the reciprocal ratio of the squares of the frequencies (9:2500). This variation as f^{-2} is a characteristic of non-deviative absorption for the case $f > \nu$.

Although only preliminary data is available at the present time, it is of some interest to speculate on the sort of D region electron density profile one might expect to find at the geomagnetic equator that would agree with the observations. It is fairly clear from the preceding experiments that what little ionization does exist below 65 km must be less than that reported at high latitudes [Rumi, 1962b; Benson, 1963] by a large factor or it would have been detected. Figure 6 is a noon-time model, D region profile constructed by Nicolet and Aikin [1960] for a geomagnetic latitude of 50°. Considerably higher electron densities than shown can be produced at low heights from the influx of energetic solar protons in addition to the cosmic ray flux. At the geomagnetic equator we are effectively shielded by the earth's field from all such activity of less than 15 Bev energy (for particles arriving from the zenith). Accordingly there is apt to be less lower

D region ionization over the geomagnetic equator than anywhere else in the world. Whether this is also the reason for the failure to detect gyro "sidebands" or strongly non-linear effects remains to be seen.

Acknowledgments

The author wishes to acknowledge the very substantial support and encouragement given during the course of this work by K. L. Bowles, who showed the feasibility of cross-modulation with the Jicamarca Radar as early as 1961. The author also wishes to thank his colleagues, especially J. L. Green and D. T. Farley, Jr., and personnel of the Instituto Geofísico del Perú for invaluable aid and advice.

References

Bailey, V. A., and D. F. Martyn (1934), The influence of electric waves on the ionosphere, Phil. Mag. 18, 369.

Barrington, R. E., and E. Thrane (Jan. 1962), The determination of D-region electron densities from observations of cross-modulation, J. Atmos. Terrestrial Phys. 24, 31-42.

Benson, R. F. (June 1962), Cross modulation of cosmic noise, J. Geophys. Res. 67, 2569-2572.

Benson, R. F. (May 1963), The cross modulation of cosmic noise: A technique for investigating the disturbed D region, Scientific Rept. No. 1, Contract AF 19(604)-8833, University of Alaska, UAG-R131.

Bjelland, B., O. Holt, B. Landmark, and F. Lied (Sept. 26, 1959), The D region of the ionosphere, Nature 184, 973-974.

Bowles, K. L., B. B. Balsley, and R. Cohen (1963), Field aligned E region irregularities identified with acoustic plasma waves, J. Geophys. Res. 68, 2485-2501.

Cutolo, M. (1963), On some new properties of the lower ionosphere, Report of Il Centro di Studi di Fisica dello Spazio, Univ. of Naples, Naples, Italy.

Farley, D. T., Jr. (Nov. 15, 1963), A plasma instability resulting in field-aligned irregularities in the ionosphere, J. Geophys. Res. 68, No. 22, 6083-6097.

Fejer, J. A. (Dec. 1955), The interaction of pulsed radio waves in the ionosphere, J. Atmos. Terrestrial Phys. 7, 322-332.

Fejer, J. A., and R. W. Vice (Nov. 1959), An investigation of the ionospheric D-region, J. Atmos. Terrestrial Phys. 16, 291-306.

Ferraro, A. J., H. S. Lee, and S. Weisbrod (1963), Phase interaction: A new tool for D-region studies, J. Geophys. Res. 68, 1169-1171.

Landmark, B., and F. Lied (Dec. 1961), Observations of the D region from a study of ionospheric cross modulation, J. Atmos. Terrestrial Phys. 23, 92-100.

Nicolet, M., and A. C. Aikin (1960), The formation of the D region of the ionosphere, J. Geophys. Res. 65, 1469-1483.

Rumi, G. C. (Sept. 1962a), Experiment Luxembourg: Cross modulation at high latitude, low height. Part I: Theoretical aspects, IRE Trans., Antennas Propagation AP-10, 601-607.

Rumi, G. C. (Sept. 1962b), Experiment Luxembourg: Cross modulation at high latitude, low height. Part II: Experimental aspects, IRE Trans., Antennas Propagation AP-10, 601-607.

Smith, R. A. (Dec. 1963), Ionospheric wave interaction using gyro-waves. Part II: The D-region, Paper 4.6, Conference on Non-Linear Properties of the Ionosphere, held at Boulder, Colorado, Dec. 17, 1963.

Figure Captions

Figure 1. Examples of cross modulation F layer reflection.

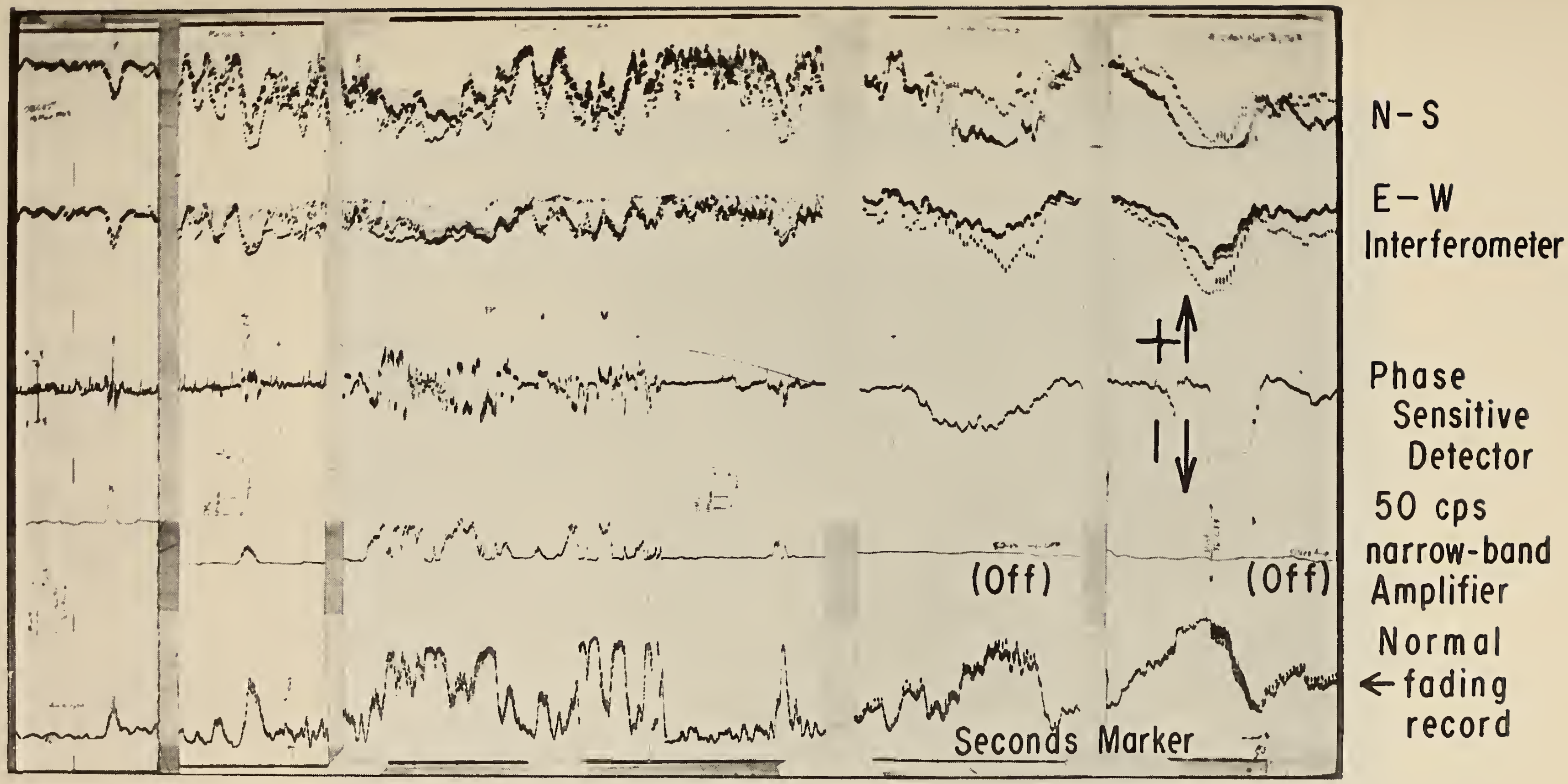
Figure 2. Cross modulation of E layer echoes.

Figure 3. Cross modulation of E layer echoes using two different disturbing pulse lengths.

Figure 4. Triple pulsing (3 E layer echoes in succession).

Figure 5. Recovery of 50 Mc/s sky brightness temperature after 4 megawatt, 3 millisecond disturbing pulse.

Figure 6. Theoretical curves for electron density (after Nicolet and Aikin).



JICAMARCA RADAR OBSERVATORY March 1963 ← Time

Examples of
Cross-Modulation
(LUXEMBOURG EFFECT)

Figure 1

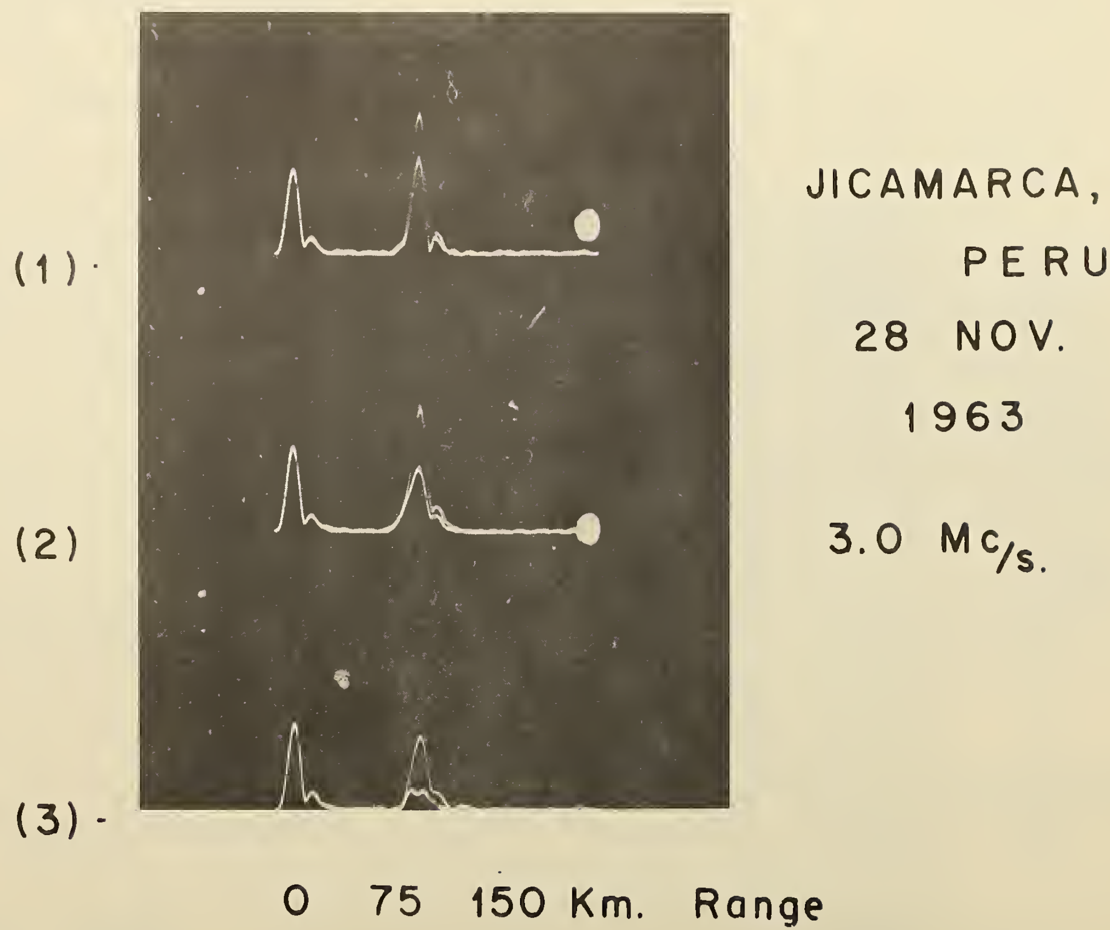


Figure 2. Cross modulation of E layer echoes.

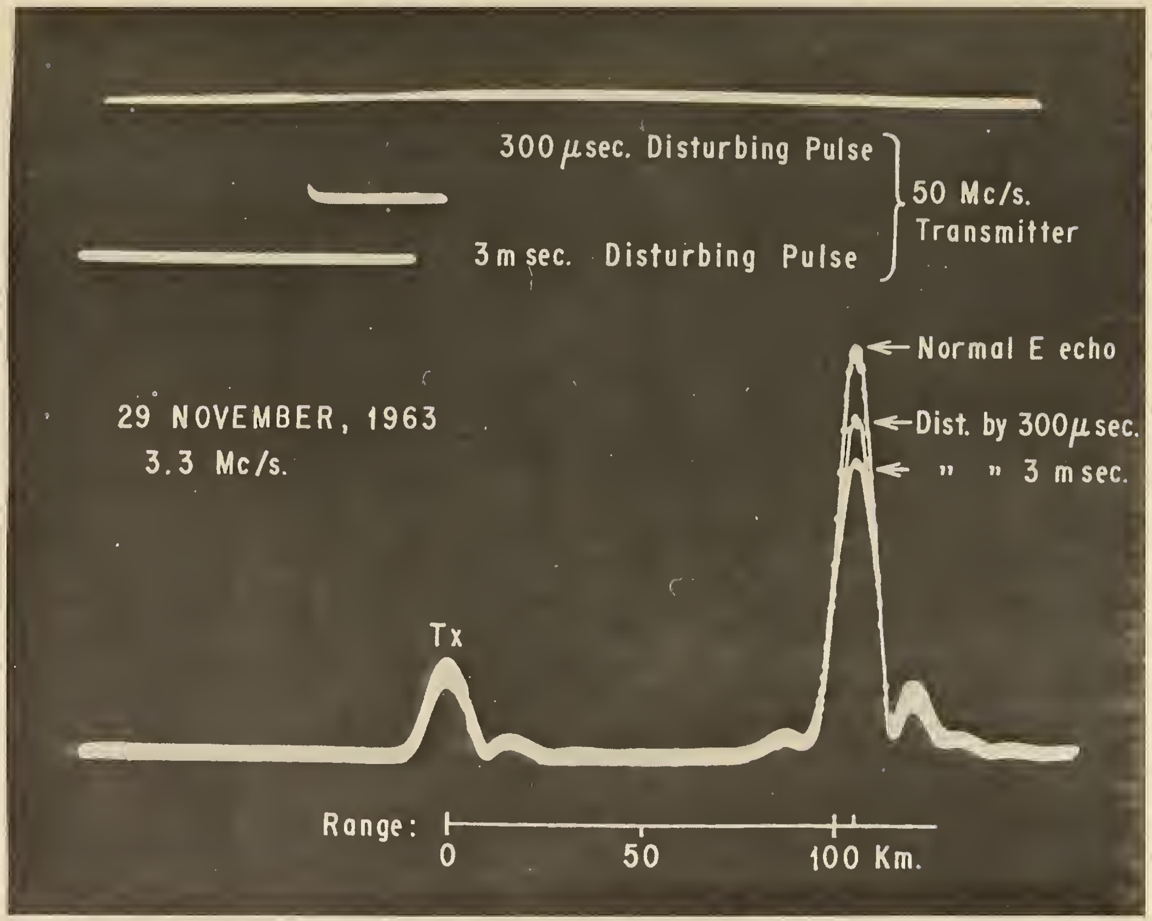


Figure 3. Cross modulation of E layer echoes using two different disturbing pulse lengths.

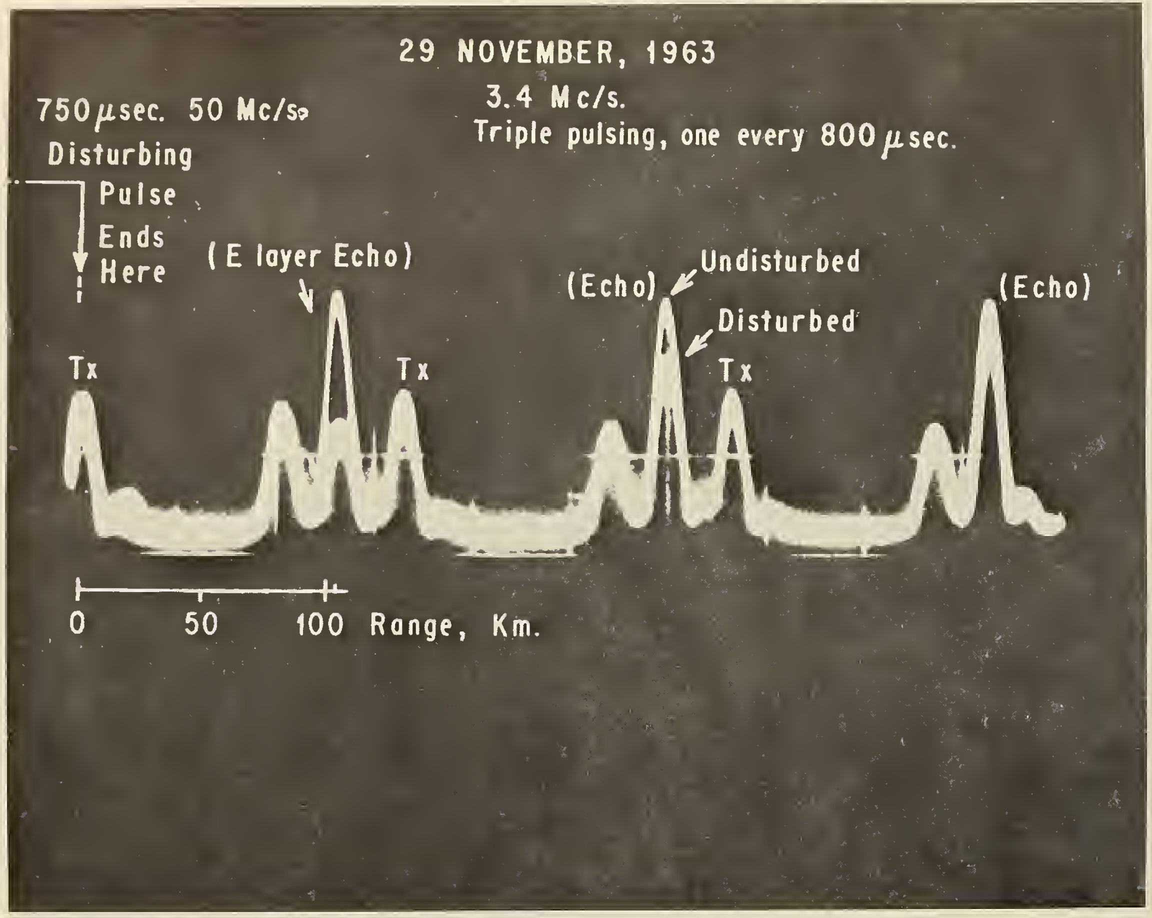


Figure 4. Triple pulsing (3 E layer echoes in succession).

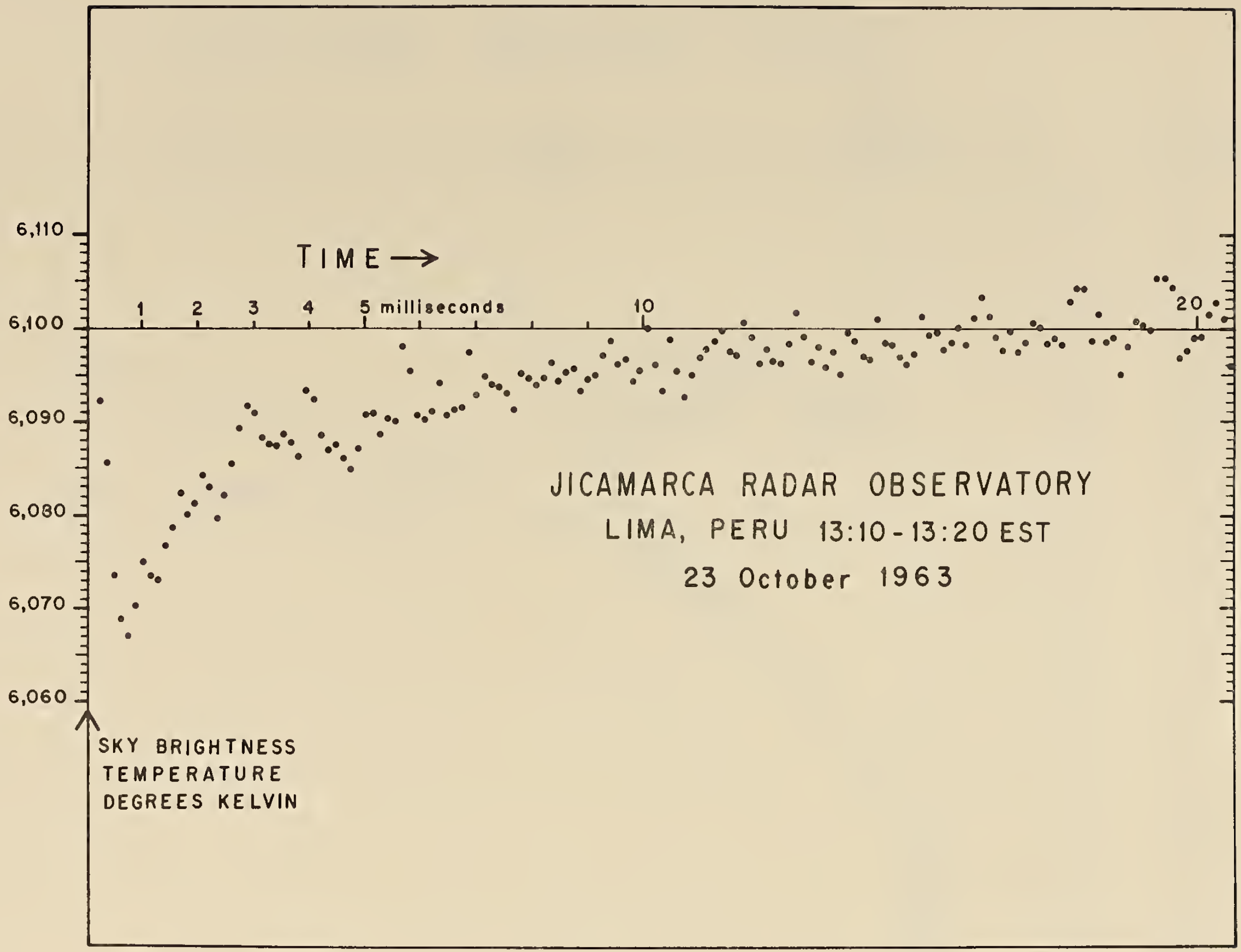


Figure 5. Recovery of 50 Mc/s sky brightness temperautre after 4 megawatt, 3 millisecond disturbing pulse.

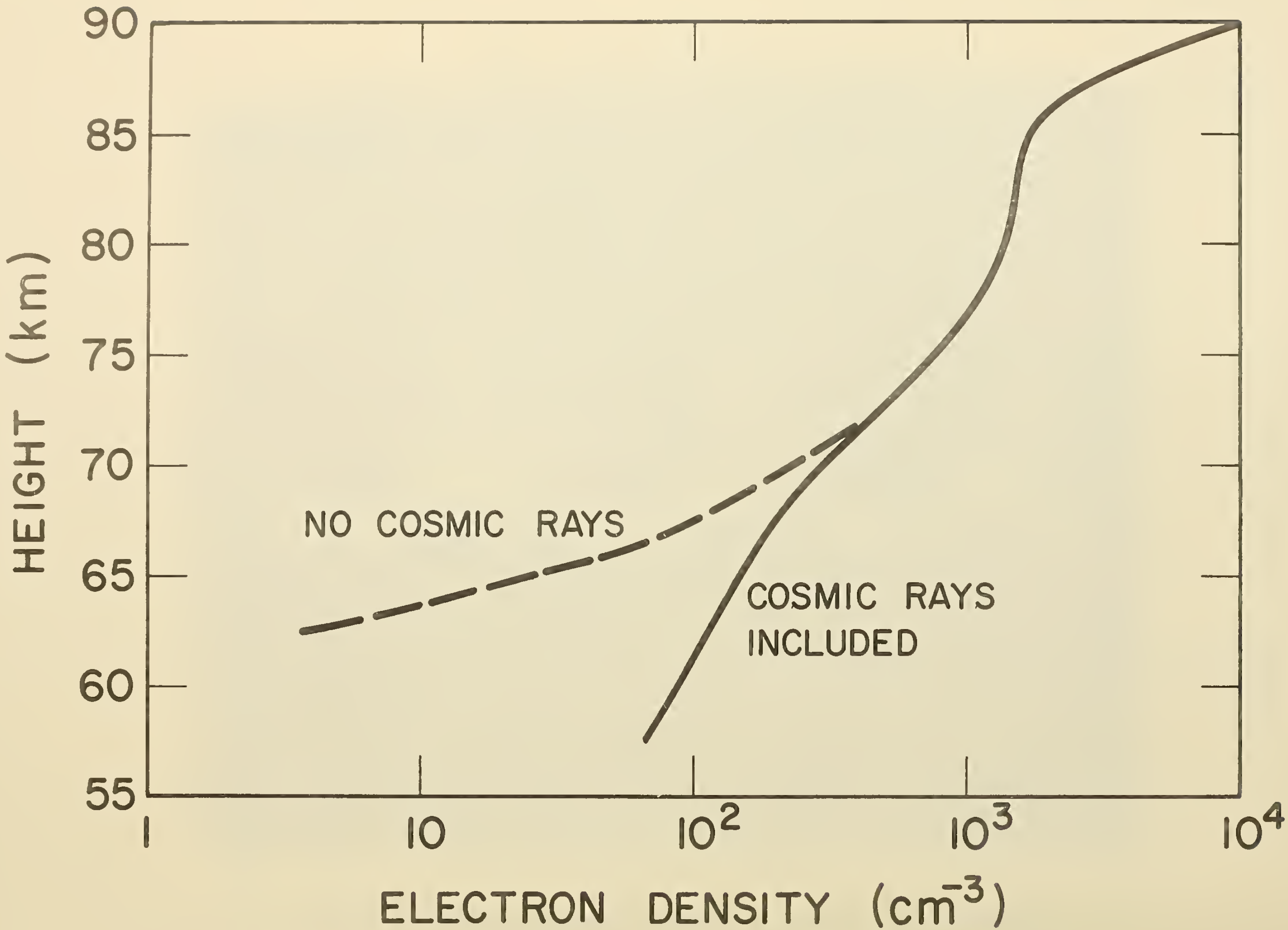


Figure 6. Theoretical curves for electron density (after Nicolet and Aikin).

Editors' Note

Dr. Cutolo remarked, in informal sessions after the meeting, that Dr. Klemperer's experiment was not identical with his, in that it relied on entirely different means for detecting the effect.

VLF Noise Bands Observed by the Alouette I Satellite

J. S. Belrose and R. E. Barrington
Defence Research Telecommunications Establishment
Radio Physics Laboratory
Ottawa, Ontario, Canada

A surprising feature of the signals observed by the VLF receiver aboard the Alouette I Satellite are bands of noise which show systematic variations with the position of the space craft in the geomagnetic field. Sometimes this is a variation of the lowest frequency at which noise is found, sometimes it is a variation of one or more discrete bands, but in all cases it is a variation which shows a frequency decrease with increase in the latitude of the satellite. The most regular noise bands are found within the zone defined by the geomagnetic L values between 2.6 and 3.6, but bands with a stronger dependence on the geomagnetic field are sometimes found at higher latitudes. The polar chorus zone is well defined by the observations between L values of 4 and 9.

These observations are compared with present theories of generation of VLF exospheric noise (hiss) (Dowden, 1962, 1963). The observed spectral distribution of the noise is perhaps more consistent with the doppler shifted cyclotron theory of generation, but the dependence on the geomagnetic field is in better agreement with the travelling wave tube amplification theory. Although both theories can account for narrow, wide and very wide band noise, neither seems to adequately explain all features of the observational data.

1. Introduction

Bursts of VLF ionospheric noise or "hiss", which occur in the audio frequency range 0.1 - 15 kc, and last for periods ranging from several minutes to a few hours, have been studied at DRTE and elsewhere (Ellis, 1959, and Watts, et al., 1963) for several years. Hiss occurs in broad bands or in one or more narrow bands. The narrow bands often have bandwidths of 1 kc or less. The most common type, "isolated bursts", is typical of magnetically quiet conditions. These are also sometimes associated with isolated magnetic bays, and ionospheric radio wave absorption. A second type occurs as an extended series of hiss bursts or noise storms. All major magnetic storms are accompanied by noise storms. Hiss has been correlated with airglow (Duncan and Ellis, 1959), with aurora (Martin et al., 1960), and with abnormal D-region absorption (Watts et al., 1963).

Although the ground observations have revealed much information on the occurrence of the hiss bursts, there are difficulties in interpreting detailed studies of their characteristics. This is partly because ionospheric attenuation affects the noise intensities that can reach a ground based receiver, and partly because propagation under the ionosphere allows observation of hiss generated in a wide range of field line latitudes.

Only by means of a satellite travelling within (or above) the ionosphere, can noise generated on particular field lines be properly studied. In addition there is no ionospheric loss to contend with, and the amplitude-frequency spectra would be as generated. On September 29, 1962, the Alouette satellite was launched into orbit, and, although primarily designed as a top-side sounder, a broad band (0.5 - 10 kc) VLF receiver was included in the pay load. The satellite moves in a nearly circular orbit at an altitude of 1025 $\pm$ 20 km and has an inclination of 80° to the equator. This allows observation of the VLF spectrum over a wide range of latitude, while the nearly circular orbit simplifies the interpretation of the results, e.g., over North America the height of the satellite can for most purposes be considered constant at 1.16 $\pm$ 0.02 earth radius and the position of the spacecraft in the geomagnetic field can therefore very simply be determined. It is the purpose of this paper to discuss, in somewhat more detail, some of the results so far obtained (Barrington, et al., 1963), and to compare these observations with present theories of the generation of VLF noise (Dowden, 1962, 1963).

2. The Observations

In this paper the data are sorted according to the magnetic parameter L (McIlwain, 1961) . The parameter L is constant along a line of force and labels the magnetic shell on which an electron bounces in latitude and drifts in longitude. Numerically L is such that for a perfect dipole field, it would be the equatorial radius to a magnetic shell, expressed in units of earth radii. In comparison of the observations with current theories of the generation of VLF noise (Section 4), it will be necessary to refer to geomagnetic latitude. In the case of a perfect dipole, the magnetic latitude measured to the point where the line of force passing through the satellite cuts the earth's surface (R = 1) is

$$\Lambda = \cos^{-1}\sqrt{1/L}$$

L is however calculated from the real magnetic field as approximated by Finch and Leaton (1957), and therefore Λ is calculated simply from L as indicated above. The discrepancy between Λ and geomagnetic latitude is at most about 2° over North America.

Fig. 1 shows the track of a pole-to-pole pass recorded once per week, which gives full coverage from 80°N to 80°S latitude. Other passes, not so complete in continuous coverage are recorded at other times (about $2\frac{1}{2}$ hours of recordings are made each week); e.g., a pass extending from NE of Bermuda, over Newfoundland, passing midway between Narsarssuak and Godhavn, and ending N and E of Spitsbergen, is recorded in co-operation with ground recordings made by the Danes; and a recording beginning over the Queen Charlotte Islands and ending at about 20°N latitude over the Pacific is recorded at Stanford for comparison with ground recordings made there. The purpose of this figure is to show the asymmetry of the satellite track with respect to the geomagnetic field. The contours are constant magnetic L values at 1025 km (the satellite height). The pass shown covers a range of L values from about 6 in the north to about 7 in the south.

A surprising feature of the signals observed by the satellite VLF receiver are bands of noise which show systematic variations with the position of the space craft in the geomagnetic field. Fig. 2 shows one of the early examples of ionospheric noise which exhibited systematic changes in the spectral distribution of noise with satellite latitude. In this example the noise is sporadic but shows a systematic tendency for the lowest frequency at which noise is found to decrease as the satellite moves to high latitudes, over a range of L between about 2.5 - 3. Above this latitude sporadic noise is found throughout the entire frequency range of the VLF receiver with no well defined lower limit. Fig. 3 is another example, but shows two bands of sporadic noise observed over the same range of L values, both of which decrease as the latitude of the satellite increases. Fig. 4 is an example recorded at higher latitudes.

Fig. 5 is an outstanding example, made on 31 January, 1963, from 1057 to 1148 UT, while the satellite passed through its full range of latitude. A narrow noise band was observed, the low limit of which changed from about $2\frac{1}{2}$ to 9 kc as the satellite moved from L values of about 3.5 to 2.5. To a very high degree the noise observed at the same L value in the northern and southern hemisphere was the same in spite of the fact that the observations in the two hemispheres are separated by at least 30 minutes in time, several degrees in longitude, and the local gyro-frequencies at the points differ by as much as 30%. These facts suggest that such noise bands must be due to conditions in the exosphere which are relatively stable and uniform over a large region.

Other features worthy discussion shown in Fig. 5 are:

(1) The relatively constant noise bands from the lower edge of the receiver pass band (about 400 c/s) to 2.5 kc which were observed when the vehicle was in a region where the magnetic parameter L was greater than 9.

(2) Similar noise bands were found in the region where L was less than 2.5 (even at the equator, not shown, where $L_{min} \simeq 1.1$).

(3) In the region where L lies between 9 and 4, the noise was very erratic both in amplitude and frequency.

We have not recorded enough examples of noise bands and noise storms to come to any firm conclusions about the morphology of such disturbances, but a few comments, particularly in regards to the disturbance pattern of hiss bands shown in Fig. 5 can be made. Ground station recordings of VLF noise made at Ottawa ($L \simeq 3.6$) on a slowly moving paper chart record, employing narrow band receivers centred on 7, 4, and $1\frac{1}{4}$ kc, and an asymmetrical detector circuit to minimize effects on the recordings due to atmospherics are shown in Fig. 6. It is clear that the satellite recording, made at the time shown by the cross-hatched rectangle, was made during a series of noise bursts (in fact noise bursts of much greater amplitude were recorded before and after the making of the satellite record). This series of noise bursts began, after a period of magnetic and VLF noise calm, at about 09 UT on 30 January, and continued in varying degrees of intensity until about 04 UT on 2 February. The noise storm was correlated in time with a geomagnetic disturbance observed in local and planetary k-indices. It seems likely that the electron streams causing the geomagnetic disturbance also produce the hiss, although an experiment to test any possible detailed correlation has not yet been done. A satellite VLF recording made at 0044 - 0049 on 1 February, 1963, while Alouette passed through a range of L values from about 2.5 - 6, is also pertinent to our discussion of the 30 January, 1963, noise storm. Although magnetic activity was less at this time, ground recordings of VLF hiss at Ottawa ($L \simeq 3.6$) showed greater activity (see Fig. 6 which shows records up to just before the making of the satellite recording). The satellite observations did not show on this occasion the two zones clearly evident in the previous recording (Fig. 5); viz. the noise band whose frequency changes with latitude in the range where L is between 2.5 - 3.5, and the auroral disturbance zone where L is between 4 - 9. The satellite observations showed two quasi constant noise bands for L between 3 - 6. The upper one (2.5 - 4 kc) became somewhat erratic above $L \simeq 5$. The lower one (below 1.5 kc) was less regular and weaker.

These results, as well as more detailed comparisons between ground station observations and satellite data reveal that there is no close correlation. This may be due to absorption or total internal reflection of these noisy signals in the ionosphere, the noise arising only in the immediate vicinity of the satellite, or due to the fact that the ground stations can observe VLF signals which have emerged from the ionosphere over a large area.

As mentioned above detailed correlations between particle streams (of the appropriate energy range) and VLF hiss have not yet been made. It is unfortunate that with the Alouette I satellite the VLF receiver and the particle detector recorders are not operated simultaneously. Nevertheless comparisons can be made between the VLF data and average properties of the intensity distributions for particles. The two places where interesting features are sometimes found in VLF data, viz. L between 2.5 - 3.5 and L between 4 - 9, are also places where features are found in the intensity distribution of trapped (and dumped) electrons. McDiarmid et al (1963) have shown that the intensity distribution of trapped electrons (in the lowest energy range observed, viz. $E \geqslant 40$ kev) observed by Alouette exhibits two maxima separated by a minimum at $L = 3.2$ ($\Lambda = 56°$). The minimum is not always present. Large fluctuations occur around $L = 2.2$ ($\Lambda = 48°$), the approximate position of the lower latitude peak (in the so-called slot between the inner and outer radiation belts). The second peak at $L = 5.6$ ($\Lambda = 65°$) corresponds approximately with a peak found in the average intensity of precipitated electrons having similar energies (which occurs at $L = 6.5$ ($\Lambda = 67°$)). Thus the VLF noise bands which exhibit a systematic change with the earth's field are observed at a latitude near the slot between the inner and outer radiation belts, while the higher latitude disturbance zone corresponds approximately with L values at the edge of the magneto-sphere, and the horn of the outer Van Allen belt.

3. Analysis of the Data

In an effort to obtain more quantitative information on these noise bands, several of their features were scaled and plotted as functions of different parameters of the geomagnetic field. It was found that to a good approximation the frequency of the lower edge of the noise band was inversely proportional to the cube of the L parameter at which the noise band was observed for L values between 2.5 and 5. Since the cube of L is inversely proportional to the magnetic field where the L shell crosses the magnetic equator, the frequency of the lower edge of the noise band f_{low} is proportional to the magnetic field B_o or the gyro-frequency f_H at the peak of the line of force on which the noise is observed. This fact is illustrated in Fig. 7 in which results from noise bands observed on several different occasions are combined. At high latitudes, or B_o between 0.003 and 0.01 (L values of 3 to 5), f_{low} varies approximately linearly with B_o but shows no consistent trend from day to day. At lower latitudes B_o between 0.01 and 0.02 (L values between 3.5 to 2.5) f_{low} is not only proportional to B_o but the slope of the points on the (f_{low}, B_o) plots is about the same for all events. Moreover, in this latitude range the points appear to cluster about two frequencies separated by about 3.5 kc.

At very high latitudes, or B_o less than 0.002 (L values between 5.5 to 7), a noise band is sometimes found which shows a systematic increase in frequency with decrease in latitude, but the slope of the points on the (f_{low}, B_o) plot is very much steeper than at lower latitudes.

The times of occurrence of all events so far recorded, although not sufficient in number to be statistically significant, cluster near 10 UT, or 05 LST, which is the time for the maximum of occurrence of the 4 kc noise bursts observed in our longer statistical study.

4. Discussion

It is not possible, at present, to give any final explanation of these noise observations but certain features are suggestive. Recently two theories have been advanced in explanation of the generation of VLF noise or hiss.

The first (Dowden, 1962) explains hiss as amplified Cerenkov emission from electron streams by analogy with the travelling wave tube (TWT process). It was shown that an electron stream in the exosphere acts as a narrow band amplifier if the stream is "weak" but as an over loaded amplifier for "stronger" streams. Fig. 8 shows calculated results based on this theory, in which the interaction parameter s/λ (which is proportional to the amplification of the TWT in nepers) is plotted for stream particles of small pitch guided along the field line terminating at a geomagnetic latitude of 60°. At other latitudes frequencies will scale so the maximum s/λ occurs at the frequency shown by the dotted line. In these respects the theory seems to explain the observational results, i.e.: for geomagnetic latitudes between about 65° and 50°, the theory predicts centre frequencies should range from about 1 to 10 kc. Observed hiss frequencies vary between 3 and 9 kc for L values between $3\frac{1}{4}$ and $2\frac{1}{2}$ ($\Lambda = 56 - 51°$). The maximum amplification which occurs at $0.15\ f_H$ is quite closely the same as observed. This is evident in Fig. 7 where the dashed line is drawn assuming that $f = 0.15\ f_H$ (or $f = 4.2 \times 10^5 B_o$, where B_o is in gauss). The departure between the lower set of data points and the dashed line for L greater than 3.3 is at least partly because a noise band which does not vary with the position of the satellite is usually observed for frequencies below about 2.5 kc, and the noise band whose frequency does vary with latitude disappears into this regular noise band at about L = 3.3.

It is clear, however, from Fig. 8 that frequency response is not symmetrical, in that the amplification falls off much more slowly at the low frequencies than at the high frequencies. Consequently "strong" streams would produce noise having a wider bandwidth but shifted toward low frequencies. The fact that it is the lower edge of the noise bands, rather than their central frequency, which for several events has a constant relation to the geomagnetic field suggests a cut-off mechanism may be involved. Such a cut-off could be due to propagation conditions which determine what frequencies are trapped on a given field line, or may be associated with a cut-off in the spectrum of the energetic particles which radiate in the VLF band.

Doppler shifted cyclotron radiation could produce hiss having such a low frequency limit emission cut-off (Dowden, 1963). In this theory cyclotron radiation from a receding electron in the exosphere can be doppler shifted to a frequency much less than the local gyro frequency. A stream of electrons will produce band limited white noise. Computed amplitude-frequency spectra for hiss generated on field lines of 60° and $63\frac{1}{2}$ geomagnetic by mono-energetic electron streams of 1 kev, 6 kev, and 30 kev are shown in Fig. 9. Thus the soft particle fluxes produce the narrower hiss spectra, and the peak amplitude occurs at $0.55\ f_H$. These spectra exhibit the required feature that the low frequency cut-off does not change much with electron energy (except for high values of electron energy), whereas the upper frequency limit is quite variable. Unfortunately the dependence of the peak frequency on the geomagnetic field is quite different from that found.

Thus neither theory can account for all the detailed features that are observed. The TWT theory is perhaps the most hopeful, and it should receive further study.

5. Acknowledgements

We wish to thank Mr. W. E. Mather for operating the VLF ground station recorders, and for making the spectrogrammes from the satellite data, which are shown in this paper. We are also endebted to Mr. W. E. Thompson for the recordings made with the VLF noise burst receiver.

References

Barrington, R. E., Belrose, J. S. and Keeley, D. A.
VLF Noise Bands Observed by the Alouette I Satellite,
J.G.R., 68, 6339, December, 1963.

Dowden, R. L.
Theory of Generation of Exospheric Very-Low-Frequency Noise (Hiss),
J.G.R., 67, 2223 - 2230, June 1962.

Dowden, R. L.
Doppler Shifted Cyclotron Generation of Exospheric Very-Low-Frequency Noise (Hiss),
Planet. Space Sci., 11, 361 - 369, April 1963.

Duncan, R. A. and Ellis, G. R.
Simultaneous Occurrence of Subvisual Aurorae and Radio Noise on 4.6 kc/s,
Nature, 183, 1618, 1959.

Ellis, G. R. A.
Low Frequency Electromagnetic Radiation associated with Magnetic Disturbances,
Planet. Space Sci., 1, 253 - 258, 1959.

Finch, H. F. and Leaton, B. R.
The Earth's Main Magnetic Field-epoch 1955.
Monthly Notices Roy. Astron. Soc., Geophys. Suppl., 7 (6), 314 - 317, 1957.

Martin, L. H., Helliwell, R. A., and Marks, K. R.
Association between Aurorae and VLF Hiss Observed at Byrd Station, Antarctica,
SEL. Tech. Rept. 1, Stanford University, April 29, 1960.

McDiarmid, I. B., Burrows, J. R., Budzinski, E. E., and Wilson, M. D.
Some Average Properties of the Outer Radiation Zone at 1000 km,
Can. J. Phys., 41, 2064 - 2079, December 1963.

McIlwain, C. E.
Co-ordinates for Mapping the Distribution of Magnetically Trapped Particles,
J.G.R., 66, 3681, 1961.

Watts, J. M., Koch, J. A., and Gallet, R. M.
Observations and Results from the "Hiss Recorder", an Instrument to Continuously Observe VLF Emissions,
J. Res. Nat. Bur. Standards, 67D, 569, 1963.

Figure Captions

Fig. 1 Polar map showing contours of constant L at 1025 km, and the track of particular satellite pass which recorded VLF data discussed in this paper.

Fig. 2 Example of ionospheric noise which exhibits systematic changes in the spectral distribution of the noise with satellite latitude.

Fig. 3 (same caption as for Fig. 2).

Fig. 4 (same caption as for Fig. 2).

Fig. 5 Spectramme of VLF ionospheric noise observed in Alouette 1 on 31 January, 1963, 1057 - 1148 UT. The track of the satellite for this recording is shown in Fig. 1. The captions on the three segments shown refer to the telemetry station where the data were recorded (the equatorial segment is not shown).

Fig. 6 Recording of VLF hiss made on three narrow band noise burst receivers at Ottawa ($L \simeq 3.6$) at the frequencies shown. Time marks are put on the record at hourly intervals. The base line for no (or little) noise can be seen at the beginning of the recordings.

Fig. 7 Plots of f_{low} (the lower edge of the noise band that shows systematic changes in spectral distribution with satellite latitude) versus B_o (the minimum value of the strength of the earth's field on the field line at which the noise was observed) for several different events. The dashed line is drawn for $f = 0.15\ f_H$ (or $f = 4.2 \times 10^5 B_o$) which is discussed in the text.

Fig. 8 Amplitude-frequency spectra in nepers (which is proportional to the interaction parameter s/λ for the TWT process as shown) for stream particles of small pitch guided along the field line terminating at geomagnetic latitude 60°. At other latitudes frequencies scale so that maximum amplification occurs at the frequency given by the dashed curve.

Fig. 9 Amplitude-frequency spectra of hiss generated on field lines of 60° and 63.5° geomagnetic by the doppler shifted cyclotron processes by mono-energetic electron streams having a pitch angle distribution produced by isotropic injection in the equatorial plane. A finite stream width would smooth out the discontinuity which occurs at 0.55 f_H.

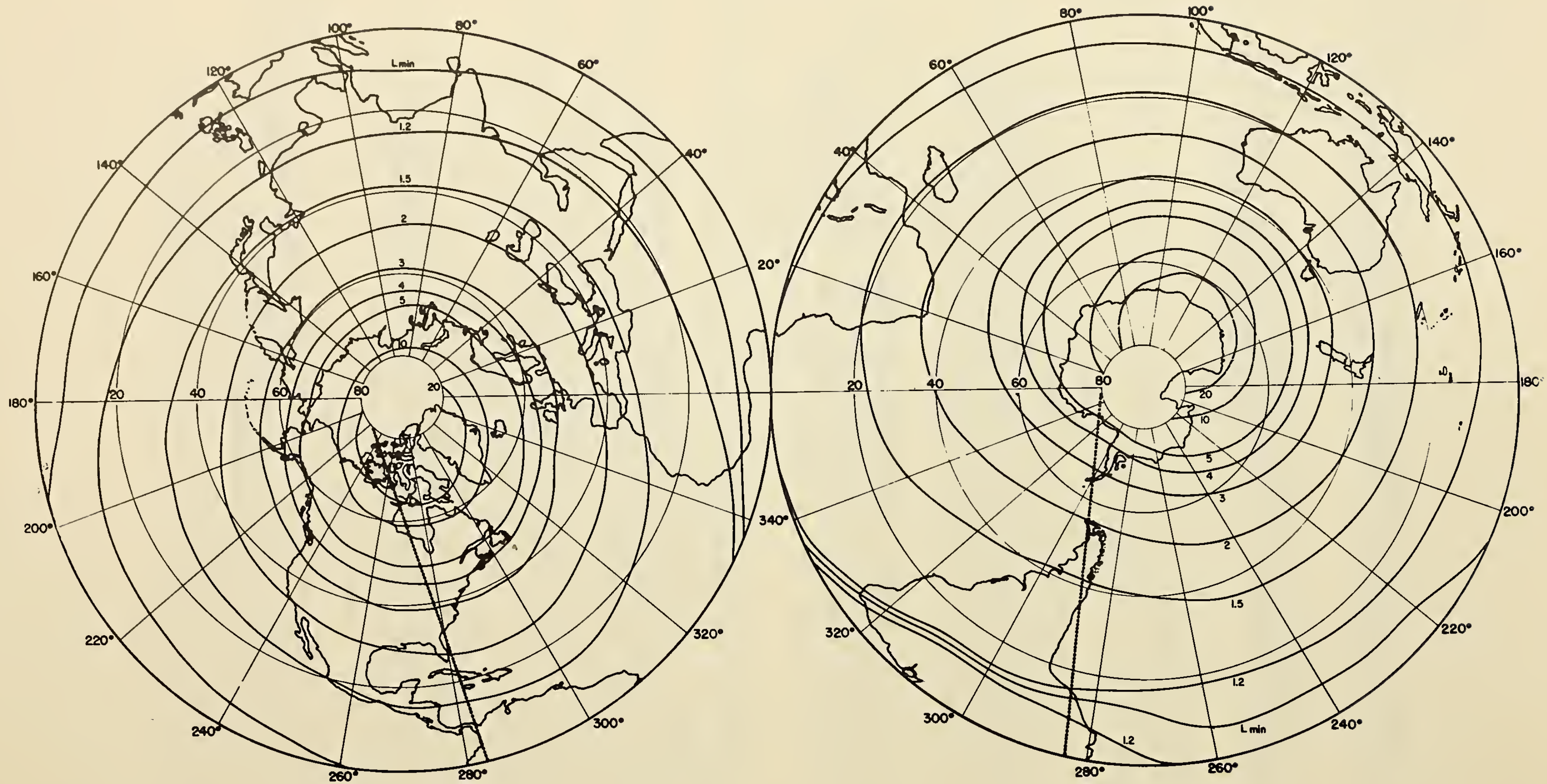


Figure 1. Polar map showing contours of constant L at 1025 km, and the track of particular satellite pass which recorded VLF data discussed in this paper.

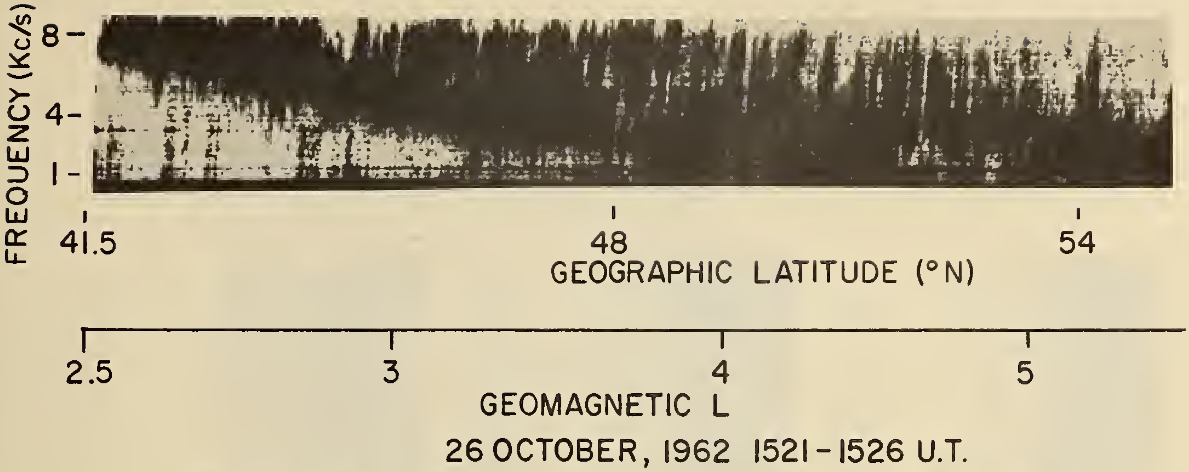


Figure 2. Example of ionospheric noise which exhibits systematic changes in the spectral distribution of the noise with satellite latitude.

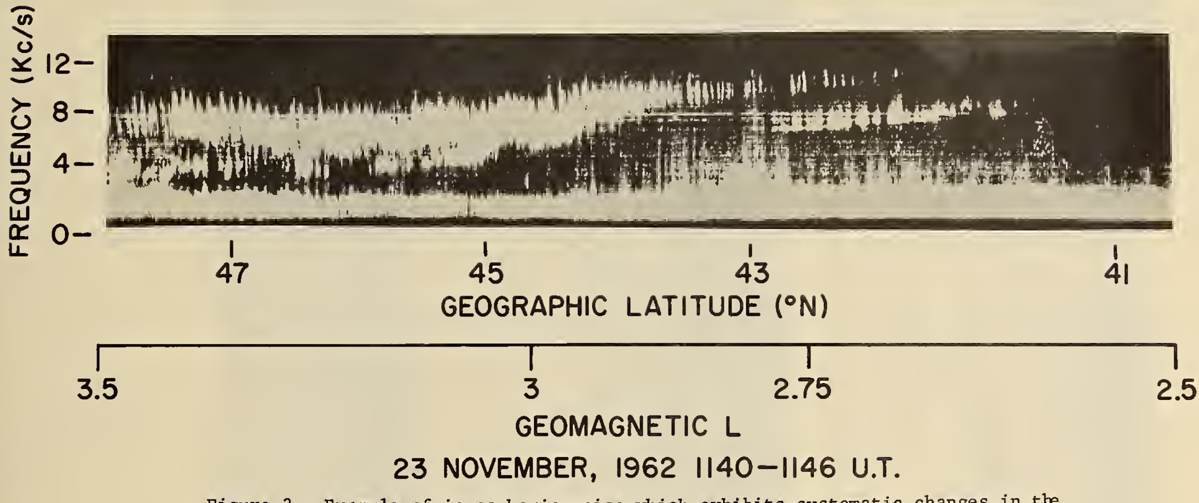


Figure 3. Example of ionospheric noise which exhibits systematic changes in the spectral distribution of the noise with satellite latitude.

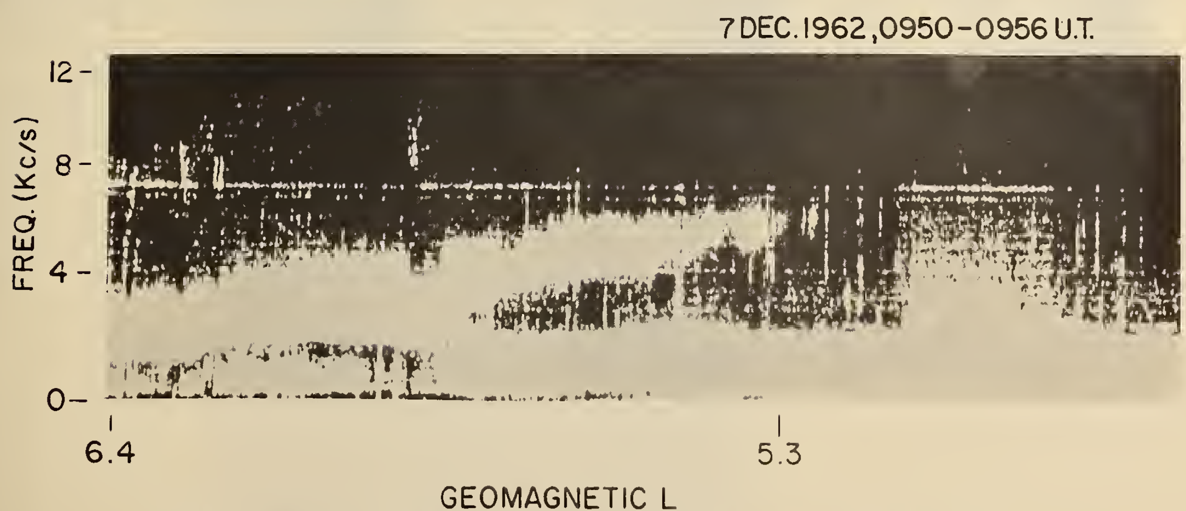


Figure 4. Example of ionospheric noise which exhibits systematic changes in the spectral distribution of the noise with satellite latitude.

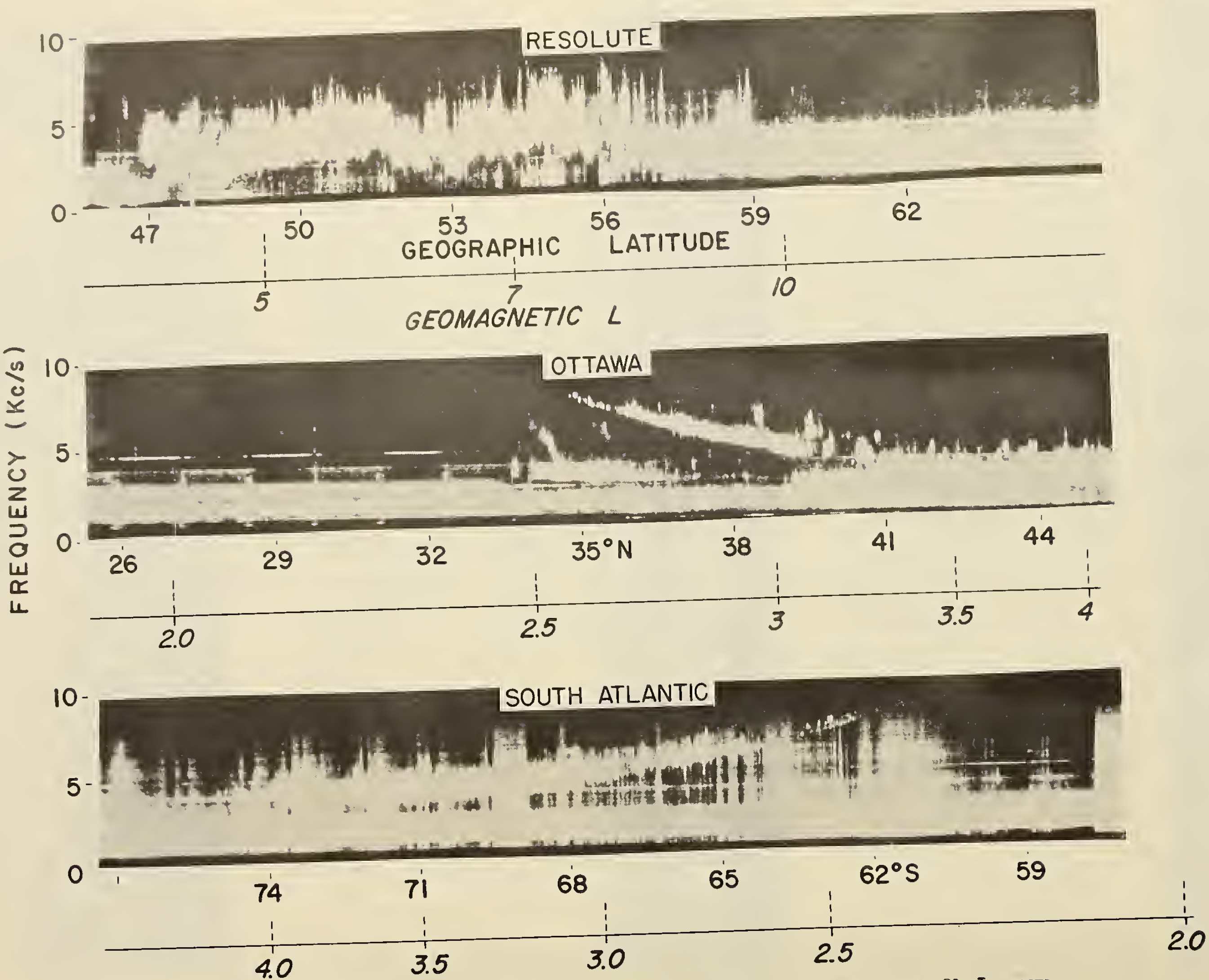


Figure 5. Spectramme of VLF ionospheric noise observed in Alouette 1 on 31 January, 1963, 1057 - 1148 UT. The track of the satellite for this recording is shown in figure 1. The captions on the three segments shown refer to the telemetry station where the data were recorded (the equatorial segment is not shown).

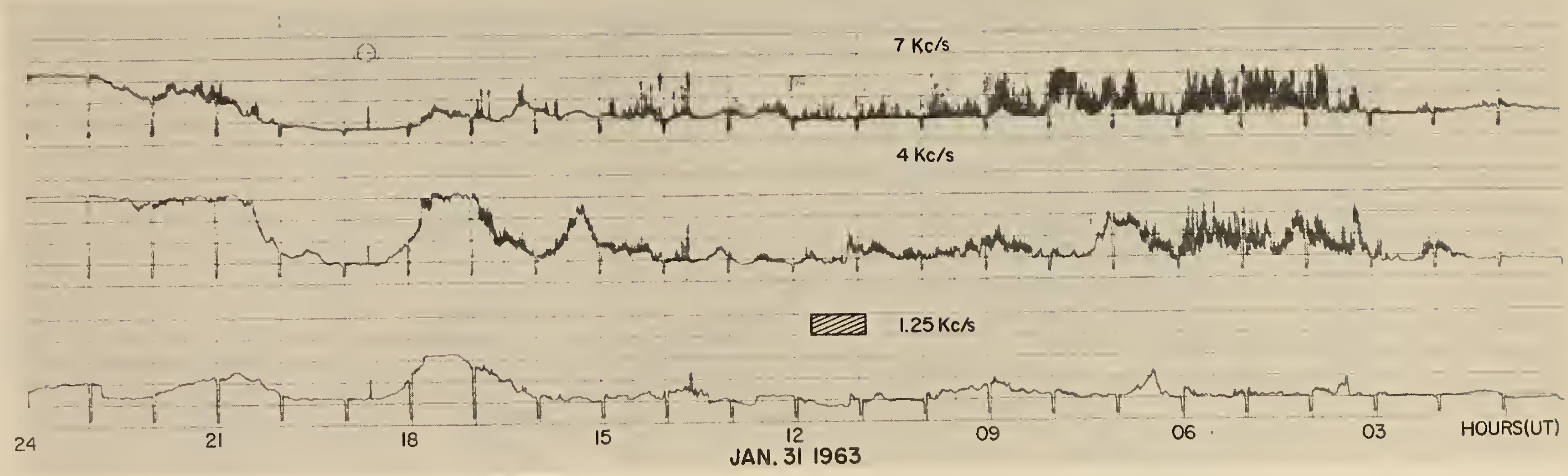


Figure 6. Recording of VLF hiss made on three narrow band noise burst receivers at Ottawa ($L \cong 3.6$) at the frequencies shown. Time marks are put on the record at hourly intervals. The base line for no (or little) noise can be seen at the beginning of the recordings.

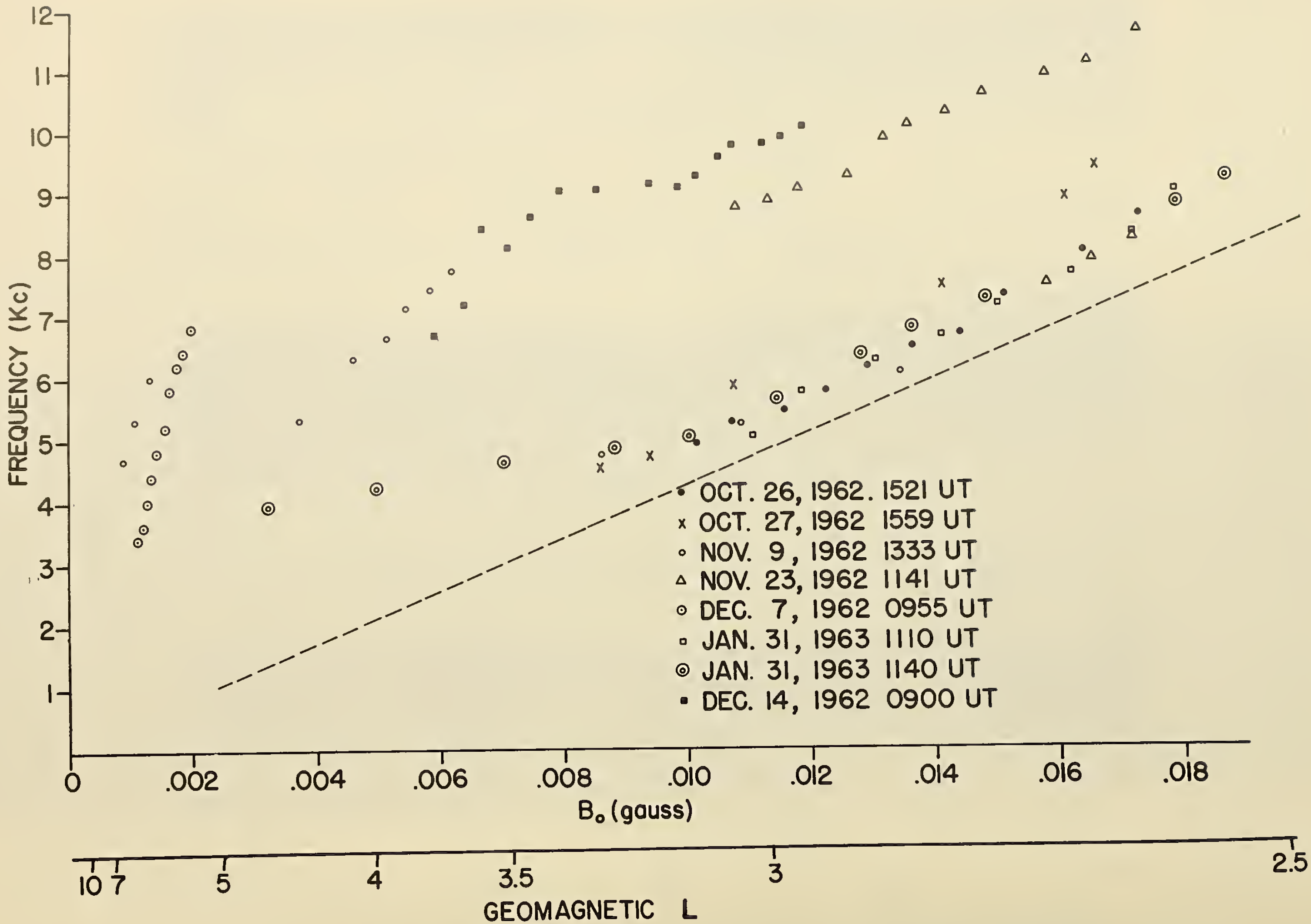
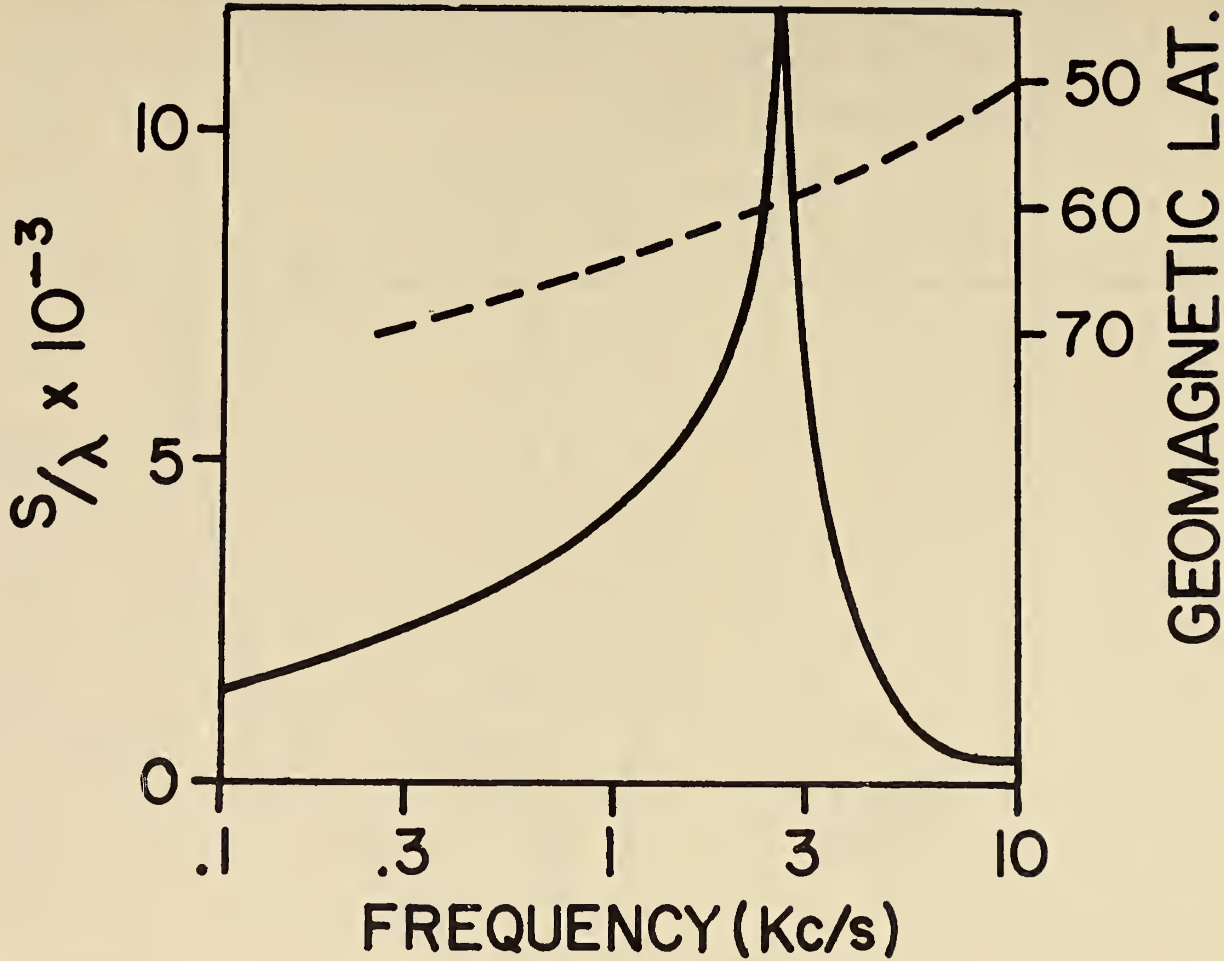


Figure 7. Plots of f_{low} (the lower edge of the noise band that shows systematic changes in spectral distribution with satellite latitude) versus B_0 (the minimum value of the strength of the earth's field on the field line at which the noise was observed) for several different events. The dashed line is drawn for $f = 0.15\ f_H$ (or $f = 4.2 \times 10^5 B_0$) which is discussed in the text.



(after Dowden [1962])

Figure 8. Amplitude-frequency spectra in nepers (which is proportional to the interaction parameter s/λ for the TWT process as shown) for stream particles of small pitch guided along the field line terminating at geomagnetic latitude 60°. At other latitudes frequencies scale so that maximum amplification occurs at the frequency given by the dashed curve.

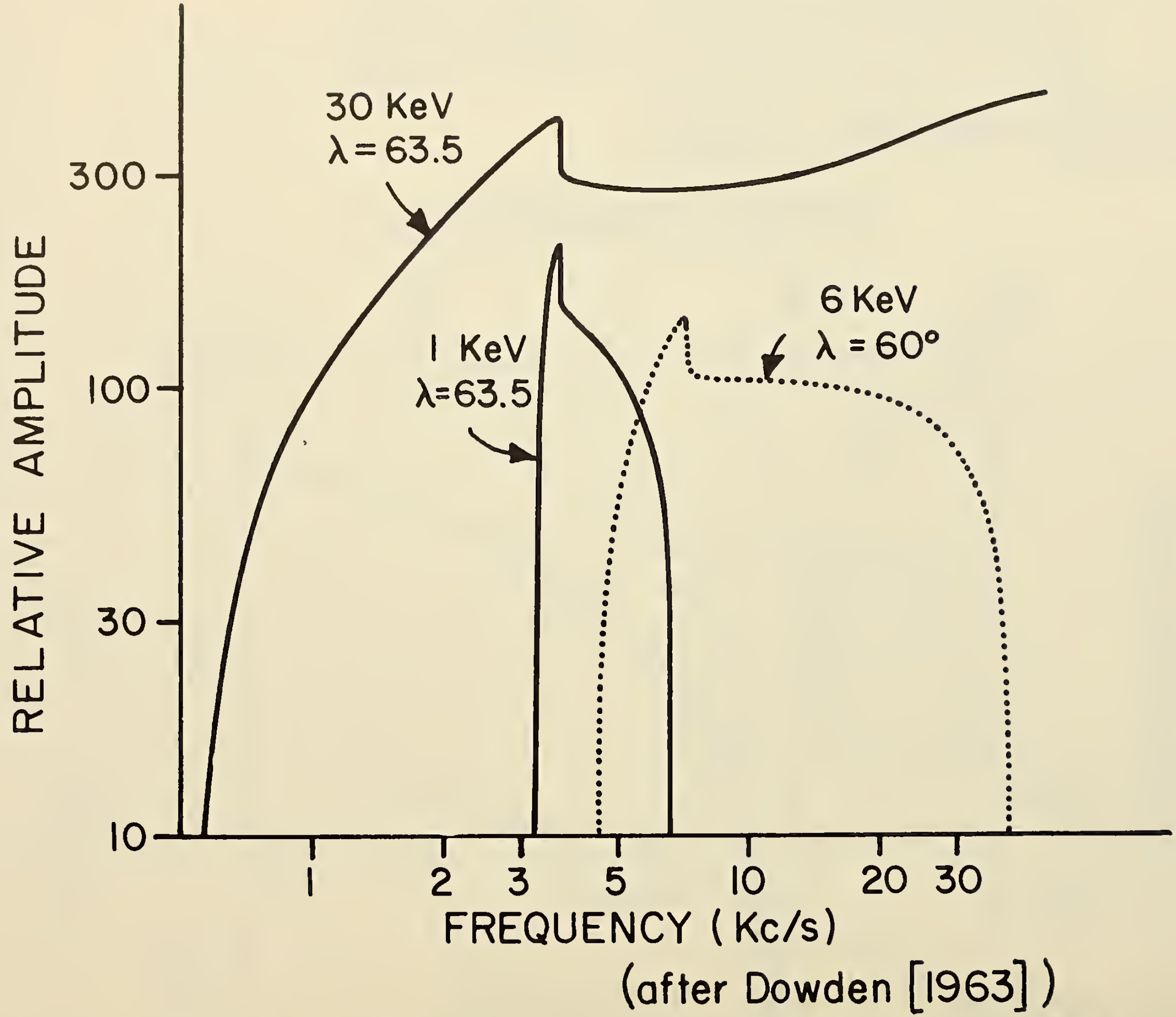


Figure 9. Amplitude-frequency spectra of hiss generated on field lines of 60° and 63.5° geomagnetic by the doppler shifted cyclotron processes by mono-energetic electron streams having a pitch angle distribution produced by isotropic injection in the equatorial plane. A finite stream width would smooth out the discontinuity which occurs at 0.55 f_H.

The Interaction of an Antenna With a Hot Plasma and the Theory of Resonance Probes[1]

J. A. Fejer

Southwest Center for Advanced Studies

P. O. Box 8478, Dallas, Texas

The impedance and the radiation field of a short antenna, which consists of two spherical conductors excited through very thin wires in phase opposition, is calculated. In the calculations the pressure tensor is replaced by a scalar pressure. A discontinuous model of the ion sheath is used.

The losses due to the radiation of electromagnetic and electron-acoustic waves are calculated and are expressed in terms of equivalent series resistances. The operation of resonance probes is discussed. It is shown that their resonant frequency is well below the electron plasma frequency if the probe radius is much larger than the Debye length. The significance of this result to both past and future ionospheric rocket probe experiments is pointed out. The limitations of the present treatment are discussed.

---

[1]Presented at the Conference on Non-Linear Processes in the Ionosphere, Dec. 16, 17, 1963, Boulder, Colo.

1. Introduction

Antennas impedance measurements in the ionosphere during rocket flights show the presence of a much larger resistive component (Whale, 1963) than is predicted by the type of antenna theory, in which the surrounding plasma is treated as cold. In principle the interaction of an antenna with a hot and very tenuous plasma could be treated by solving a combination of the collisionless Boltzmann equation with Maxwell's equations. The mathematical difficulties of solving these equations and staisfying the appropriate boundary conditions appear to be formidable.

In this paper the problem is considerably simplified by the use of certain approximations. The first of these is the use of a scalar pressure (Spitzer 1962; p..24). The problem is further simplified by the rather crude assumption that the plasma surrounding the antenna is sharply bounded and that outside the boundary the plasma is uniform. These approximations are applied here to a simple "spherical" dipole antenna that consists of two spheres excited in phase opposition through wires whose capacity is neglected. For other applications of the same type of approximations the reader is referred to papers by Gould (1959), Fejer (1963), and Nickel, Parker, and Gould (1963). The present analysis and its conclusions are entirely different from those of Whale (1963) who also treated the excitation of electron-acoustic waves by introducing the concept of an isotropic pressure. Whale assumed the existence of interaction between the fluctuating quasi-electrostatic field and the electron-acoustic waves throughout the uniform plasma. In the present treatment the only interaction is taken to occur at the inner boundary (in reality in the sheath). Within the uniform plasma the electron-acoustic wave and the electromagnetic wave (or in the present limit the quasi-electrostatic field) propagate independently (c.f., for example Ginzburg 1961) and therefore cannot interact in terms of the hydrodynamic approximation.

2. The Excitation of Electron-Acoustic Waves

It is only necessary to consider one of the two spheres of the antenna described in the introduction here. It is assumed that the unperturbed electron concentration vanishes for $r < R$, and has the constant value N for $r > R$, where r is the distance from the center of the sphere. This is admittedly a rather artificial model of the unperturbed ion sheath. It effectively assumes an abrupt potential barrier (this could be visualized as a hypothetical double-layer formed from infinitely heavy positive and negative ions) at $r = R$ which prevents the penetration of electrons inside the sphere at $r = R$. It is clear that the radial component of the mean electron velocity must vanish at $r = R$. Immediately inside the sheath there is assumed to be a spherical conductor whose (quasi-electrostatic) perturbation potential is taken to be a harmonic function of the time. (In principle the radius of the conductor could be taken as smaller than R without essentially modifying the analysis; this will not be done here.) Since there is no mean radial motion of the electrons at $r = R$, there is no fluctuating charged surface layer there and therefore both the perturbation potential V and its normal derivative must be continuous at r = R. Inside the plasma the equation

$$mN\partial \underset{\sim}{v}/\partial t = Ne\underset{\sim}{\nabla}V - \gamma KT\underset{\sim}{\nabla}n \tag{1}$$

is taken to be valid where m is the mass, e the absolute value of the charge, N the unperturbed number density and n the perturbation in the number density of electrons (factors $e^{i\omega t}$ are taken for granted in n and in V) K is Boltzmann's constant, T is the temperature, and where the ratio γ of the specific heats is taken as 3 (Spitzer 1962). A combination of the divergence of equation (1) with the equation of continuity (satisfied by the velocity $\underset{\sim}{v}$ of the electron gas)

$$\underset{\sim}{\nabla}.(N\underset{\sim}{v}) = -\partial n/\partial t \tag{2}$$

leads to

$$\frac{\partial^2 n}{\partial t^2} + \frac{Ne}{m}\underset{\sim}{\nabla}^2 V - \frac{\gamma KT}{m}\underset{\sim}{\nabla}^2 n = 0 \tag{3}$$

A combination of equation (3) and Poisson's equation in spherical coordinates and with spherical symmetry:

$$\underset{\sim}{\nabla}^2 V = \frac{\partial^2 V}{\partial r^2} + \frac{2}{r}\frac{\partial V}{\partial r} = \frac{e}{\varepsilon_0}n \tag{4}$$

leads to

$$\underset{\sim}{\nabla}^2 n = \frac{\partial^2 n}{\partial r^2} + \frac{2}{r}\frac{\partial n}{\partial r} = \alpha^2 n \tag{5}$$

where $\alpha^2 = (\omega_N^2 - \omega^2)/u^2$, $\omega_N = (e^2N/\varepsilon_0 m)^{\frac{1}{2}}$ is the electron plasma frequency and $u = (\gamma KT/m)^{\frac{1}{2}}$ is the velocity of electron-acoustic waves in the high frequency limit ($\omega >> \omega_N$).

The general solution, which vanishes at infinity, of the differential equations (4) and (5) has the form

$$n = \frac{C_1}{r} e^{-\alpha r} \tag{6}$$

$$V = \frac{e}{\varepsilon_0 \alpha^2} n + \frac{C_2}{r} \tag{7}$$

where that part of the solution which is associated with the constant C_1, describes an electron-acoustic wave, whereas the part associated with C_2 describes a quasi-electrostatic field (the so-called induction field) which is a good approximation to the electromagnetic wave field within distances very much shorter than a wave length from the antenna.

The boundary condition v = 0 at r = R yield with the aid of equation (1)

$$\left(\omega_N^2 \frac{\partial V}{\partial r}\right)_{r=R} = \left(\frac{e u^2}{\varepsilon_0} \frac{\partial n}{\partial r}\right)_{r=R} \tag{8}$$

Substitution of n and V from equations (6) and (7) into equation (8) leads to the relation

$$C_2 = -\frac{e}{\varepsilon_0} \frac{u^2 \omega^2}{\omega_N^2(\omega_N^2 - \omega^2)} (1 + \alpha R) e^{-\alpha R} C_1 \tag{9}$$

Substitution of C_2 from equation (9) into equations (6) and (7) leads to the following relationship between the perturbation in the number density and the perturbation potential V at r = R:

$$\frac{n(R)}{V(R)} = \frac{\varepsilon_0 \omega_N^2}{e u^2} \left(1 - \frac{R \omega^2}{u(\omega_N^2 - \omega^2)^{\frac{1}{3}}}\right)^{-1} \tag{10}$$

Similarly, the effective, complex capacity of the conductor, defined here as the ratio of the charge on the conductor to the potential V(R) is given by

$$C_{eff} = -\frac{4\pi R^2 \varepsilon_0 (\partial V/\partial r)_{r=R}}{V(R)} = 4\pi\varepsilon_0 R \frac{1 + R u^{-1}(\omega_N^2 - \omega^2)^{\frac{1}{2}}}{1 - R\omega^2 u^{-1}(\omega_N^2 - \omega^2)^{-\frac{1}{2}}} \tag{11}$$

Equations (6), (7), (9), (10), and (11) represent the main results of this paper and a discussion of their significance follows here. Equation (11) may also be used to calculate the impedance $Z_{eff} = (i\omega C_{eff})^{-1}$ of the antenna. The energy loss due to the radiation of electromagnetic waves has not as yet been included in the analysis; this will be done in a subsequent part of the present paper. The resistive part of the impedance Z_{eff} represents only the energy loss due to the "radiation" of electron-acoustic waves.

It is convenient to express the impedance Z_{eff} as a function of the parameters $Y = (4\pi\varepsilon_0 R\omega)^{-1}$ (the free space reactance of a sphere of radius R), $\delta = R u^{-1}\omega_N$ (the ratio of the radius R to $\gamma^{\frac{1}{2}}$ times the Debye length), and $\psi = \omega/\omega_N$ (the ratio of the frequency to the electron plasma frequency). The result is

$$Z_{eff} = -iY \frac{1 - \delta\psi^2(1-\psi^2)^{-\frac{1}{2}}}{1 + \delta(1-\psi^2)^{\frac{1}{2}}} \tag{12}$$

It is clear from equation (12) that Z_{eff} is pure imaginary for $\psi < 1$ or $\omega < \omega_N$. The absence of a loss term is explained in this case by the absence of propagating electron-acoustic waves. It may be seen from equation (12) that the impedance becomes infinitely high at the plasma frequency and vanishes at the frequency where

$1-\delta\psi^2(1-\psi^2)^{-\frac{1}{2}}=0$. More will be said about the significance of this frequency later. At very low frequencies $Z_{eff} = -i\Upsilon(1+\delta)^{-1}$, and thus the impedance is smaller than the free space impedance by a factor $(1+\delta)^{-1}$. It is understandable that the effective capacity of the sphere is larger at low frequencies than the free space capacity because the alternating electric field does not extend to infinity but is confined to within the plasma sheath which in the case of large δ is much smaller than the radius. Figure 1 shows for $\psi<1$ the ratio of the imaginary part of Z_{eff} (the real part is zero for $\psi<1$) to the magnitude Υ of the free space reactance as a function of the ratio $\psi=\omega/\omega_N$ of the frequency to the electron plasma frequency for different values of the ratio δ of the radius to $\gamma^{\frac{1}{2}}$ Debye length. The curve for $\delta=\infty$ represents the well-known approximation in which thermal motions are neglected and the plasma is treated as a dielectric with an effective dielectric constant $\varepsilon_0(1-\omega_N^2/\omega^2)$.

For $\psi>1$ it is convenient to write equation (12) in the form

$$Z_{eff}=\frac{\Upsilon\delta(\psi^2-1)^{-\frac{1}{2}}}{1+\delta^2(\psi^2-1)}-i\,\frac{\Upsilon(1+\delta^2\psi^2)}{1+\delta^2(\psi^2-1)} \tag{13}$$

Figure 1 also shows the real and the imaginary parts of Z_{eff} for (in units of Υ) as a function of ψ, for different values of δ, for $\psi>1$. The case of $\delta=\infty$ again represents the approximation in which thermal motions are neglected and the plasma is regarded as a dielectric. The curves of Fig. 1 show the presence of a resistive component of the impedance for finite values of δ. The resistive, ohmical component represents the energy loss caused by the generation of electron-acoustic waves.

3. The Radiation of Electromagnetic Waves by Two Spherical Conductors Excited in Antiphase

The analysis of this paper can be extended to include radiation losses due to electromagnetic waves. The analysis of electromagnetic radiation by spheres oscillating in antiphase is of course somewhat artificial because in practice antennas resemble cylindrical conductors, rather than two spherical conductors excited in antiphase through thin wires whose capacity is neglected here. It is, however, relatively simple to extend the present analysis to cylindrical conductors, at least in an approximate manner; this extension, whose results may be expressed in terms of Bessel functions, is not carried out in the present paper. The analysis of spheres oscillating in antiphase is considerably simpler and it illustrates the nature of the problem rather well.

If the distance D between the two spheres is much smaller than the electromagnetic wavelength (and at the same time $R \ll D$) then the quasi-electrostatic field at a distance $r \gg D$ on the line connecting the spheres is $2C_2D/r^3$, where C_2/r is that part of the potential given by equation (7) which could be regarded as an approximation to the radiation field of a spherical radiator at distances much smaller than the wave length. At distances much longer than a wavelength and in the plane that perpendicularly bisects the line connecting the two spheres, the magnitude E of the radiation electric field of the dipole antenna is then given by

$$E = C_2 D k^2/r = (C_2 D/r)\left[(\omega^2 - \omega_N^2)/c^2\right] \tag{14}$$

The radiation magnetic field H is given by

$$H = (\varepsilon_o C_2 D/r)\left[(\omega^2 - \omega_N^2)^{3/2}/c\omega\right] \tag{15}$$

where k is the wave number and c is the velocity of electromagnetic waves in vacuum. Equations (14) and (15) were derived by fitting the quasi-electrostatic dipole field to the radiation field of a dipole (Stratton 1941) in a medium with a dielectric constant $\varepsilon_o(1 - \omega_N^2/\omega^2)$. Equations (6), (7), and (9) can then be used to express C_2 in terms of V(R), the alternating potential of one of the spheres or in terms of the current supplied to the spherical conductors $I = i\omega C_{eff}V(R)$ where C_{eff} is given by (11). The Poynting vector $\underset{\sim}{E} \times \underset{\sim}{H}$ may then be expressed in terms of I, and the total radiated power P may be found by integration over a very large sphere. The result is

$$P = (6\pi\varepsilon_o)^{-1} D^2 \omega c^{-3} (\omega^2 - \omega_N^2)^{\frac{1}{2}} I^2 \tag{16}$$

and therefore the resistive component of the antenna impedance (the radiation resistance) due to the radiation of electromagnetic waves is

$$\rho_m = (6\pi\varepsilon_o)^{-1} D^2 \omega c^{-3} (\omega^2 - \omega_N^2)^{\frac{1}{2}} \tag{17}$$

The resistive component due to electron-acoustic waves is given by equation (13) as

$$\rho_a = 2Y \frac{\delta(\psi^2 - 1)^{-\frac{1}{2}}}{1 + \delta^2(\psi^2 - 1)} \tag{18}$$

Equations (17) and (18) show that ρ_m is zero at the electron plasma frequency and increases rapidly with frequency whereas ρ_α is infinitely high at the plasma frequency and decreases rapidly with frequency.

It is interesting to note that the expression (17) for ρ_m is the same as the expression for the radiation resistance of a short antenna of length D (whose capacity is entirely at its extremities) situated in a medium with a dielectric constant $\mathcal{E}_0(1-\omega_N^2/\omega^2)$. This is a significant result in view of the fact that in the calculation of the series reactive part of the antenna impedance the plasma can not be replaced by a medium with a dielectric constant $\mathcal{E}_0(1-\omega_N^2/\omega^2)$.

4. Application to Resonance Probes

The results of the previous section may be used to draw some tentative but very important conclusions about the behavior of resonance probes (Miyazaki et al, 1960). In such probes the change in the collected direct current, caused by the application of a radio frequency voltage is measured as a function of the radio frequency. It has been usually accepted (Miyazaki et al, 1960) as an experimental fact that the change in the collected current shows a resonant increase at the plasma frequency.

In this section the point of view is taken that the change in the collected direct current is due to rectification cause by the non-linear characteristic of a Langmuir probe. The amplitude of the radio frequency variations in the collected current will be proportional to the fluctuations n(R) in the number density (in a more accurate treatment the fluctuation in temperature would also have to be taken into account but these would in any case be proportional to the fluctuations in number density) and therefore equation (10) should really give the characteristics of a resonance probe. Using the parameters δ and ψ equation (10) may be written in the form

$$n(R) = V(R) \frac{\varepsilon_0 \omega_N}{e\, \omega^2} \left(1 - \frac{\delta \psi^2}{(1-\psi^2)^{\frac{1}{2}}}\right)^{-1} \tag{19}$$

Equation (19) shows that n(R) becomes infinitely large at the frequency for which

$$\psi = \left\{ 2^{-1} \delta^{-2} \left[-1 + (1 + 4\delta^2)^{\frac{1}{2}} \right] \right\}^{\frac{1}{2}} \tag{20}$$

This shows that resonance occurs near the electron plasma frequency ($\psi \sim 1$) only when $\delta \ll 1$; for large values of δ the resonance occurs well below the plasma frequency. Fig. 2 shows $(e\,\omega^2/\omega_N^{\,2}\varepsilon_0)\,|n(R)/V(R)|$ as a function of ψ for two values of δ. Fig. 2 illustrates that resonance occurs well below the plasma frequency when the probe radius is much larger than the Debye length. The resonant frequency given by (20) is the same as the previously discussed frequency where the impedance of the probe vanishes.

The present analysis clearly leads to the conclusion that the resonant response of the probe does not occur at the plasma frequency, as laboratory and space experiments are claimed to show, but always below the plasma frequency. Large errors could thus occur in rocket investigations of electron concentration in which the resonant frequency was identified with the plasma frequency, especially if the radius of the probe were much larger than the Debye length. Fortunately a relatively small sphere (R = 1 cm) was used in the ionospheric experiments (Aono et al, 1962) reported so far.

A discussion of the physical nature of this resonance sheds some light on the reasons for the occurrence of the resonance well below the plasma frequency. It is clear from equation (7) that the electric field of the probe consists of two parts. One is simply a quasi-electrostatic field, which at short distances is the approximate form of the radiation field for frequencies above the plasma frequency whereas the other is the field associated with an electron-acoustic wave which is evanescent at frequencies below the plasma frequency.

At very low frequencies the quasi-electrostatic field becomes very small compared to the field of the evanescent electron-acoustic wave. This means that the alternating charge on the conducting sphere is almost perfectly shielded by a suitable (continuous) modification of the sheath. As the frequency is increased, the shielding becomes less perfect and a potential C_2/r appears outside the sheath. The present theory shows that the outside field opposes in phase the field within the sheath. With increasing frequency a situation is reached eventually where the potential drop outside the sheath just balances the potential drop inside so that no exciting voltage is required on the conductor and resonant oscillations occur.

This explanation of the resonance is strongly supported by the full agreement, in one special case, of the present approximate calculation with Landau's (1946) more accurate analysis utilizing the collisionless Boltzmann equation. Landau considered the penetration of an alternating electric field, that is normal to a plane boundary, from empty space on the one side of the boundary into a uniform plasma on the other side. He showed that at frequencies slightly below the plasma frequency the field asymptotically approaches a value which is equal to $(1-\omega_N^2/\omega^2)^{-1}$ times its value at the boundary. The field deep inside the plasma is thus out of phase with the field at the boundary. Equations (6), (7), and (9) of the present paper lead to the same asymptotic value if R and r-R both approach infinity and at the same time satisfy the inequality $R \gg r-R$. Landau also showed that the asymptotic value is approached according to a simple exponential law expressed by his equation (45). Essentially the same exponential law with the same exponent ($-\alpha r$ in equation 6) results from equations (6), (7), and (9) of this paper.

An alternative, simpler but less quantitative, explanation is that the resonance frequency of the probe is lowered by a tight coupling to the medium. For a very small probe which is weakly coupled to the medium, the resonant frequency will be nearly equal to the plasma frequency. A larger probe has a lower resonant frequency because it is more tightly coupled to the plasma.

5. Conclusions

The simple analysis of this paper leads to certain interesting and significant results about the behavior of antennas in plasmas and about the interpretation of observations with the aid of resonance probes in plasmas. It is shown that resonance does not take place at the plasma frequency and that previous measurements made by the resonant probe method may have to be reinterpreted. In principle both the concentration and the electron temperature could be determined by simultaneous measurements of the resonant frequencies of two resonant probes of different size. The same information could be obtained from a single probe if an additional measurement (such as the additional direct current at very low frequencies) was made besides the determination of the resonant frequency. Impedance measurements (not necessarily above the plasma frequency) could also be used to determine the electron concentration and the electron temperature. It is to be remembered, however, that the results of this paper must be regarded as merely semi-quantitative in view of the oversimplified nature of the underlying assumptions.

The present treatment could, in principle, be further refined by using a more sophisticated model of the sheath even if the use of a scalar pressure term were retained. The treatment could also be applied, in principle, to conductors which have other shapes than spherical. The most severe shortcoming of the present approach is undoubtedly the use of a scalar pressure term and it is to be hoped that a calculation with the aid of the collisionless Boltzmann equation will be attempted in the future. It is also to be hoped that the conclusions of the present paper will be submitted soon to an experimental check.

ACKNOWLEDGMENTS - The benefit of discussions with Drs. J. E. Midgley, H. B. Liemohn, and W. B. Hanson is acknowledged. Thanks are due to Miss J. Ligon for assistance with the computations. This research was supported by the National Aeronautics and Space Administration under grant NsG-269-62.

REFERENCES

Aono, Y., K. Hirao and S. Miyazaki (1962), "Profile of charged particle density in the ionosphere observed with rockets," Space Research III, 221 (North-Holland Publishing Co., Amsterdam).

Fejer, J. A. (1963), Scattering of electromagnetic waves by a plasma cylinder, Phys. Fluids. In the press.

Ginzburg, V. L. (1961), Propagation of electromagnetic waves in plasma, (Gordon and Breach, New York).

Gould, R. W. (1959), Experiments on plasma oscillations, Proceedings of the Conference on Plasma Oscillations (Linde Company, Indianapolis, Ind.).

Landau, L. (1946), On the vibrations of the electronic plasma, J. Phys. U.S.S.R., 10, 25.

Miyazaki, S., K. Hirao, Y. Aono, K. Takayama, H. Ikegami and T. Ichimiya (1960), Resonance probe--a new probe method for measuring electron density and electron temperature in the ionosphere--, Rep. Ionos. Space Res. Japan, 7, 148.

Nickel, J. C., J. V. Parker and R. W. Gould (1963), Resonance oscillations in a hot nonuniform plasma column, Phys. Rev. Letters, 11, 183.

Spitzer, L., Jr. (1962), Physics of Fully Ionized Gases (Interscience Publishers, New York and London).

Stratton, J. A. (1941), Electromagnetic Theory (McGraw-Hill, New York and London).

Whale, H. A. (1963), The excitation of electroacoustic waves in the ionosphere, J. Geophys. Res., 68, 415.

CAPTIONS

Fig. 1. The ratios of the real (interrupted line) and imaginary (solid line) parts of the impedance to the magnitude of the free space reactance as functions of the ratio ψ of the radio frequency to the electron plasma frequency.

Fig. 2. The dimensionless quantity $(eu^2/\omega_N^2\varepsilon_0)|n(R)/V(R)|$ which is proportional to the number density fluctuations $n(R)$ produced by a given rf voltage V(R), as a function of the ratio ψ of the radio frequency to the electron plasma frequency.

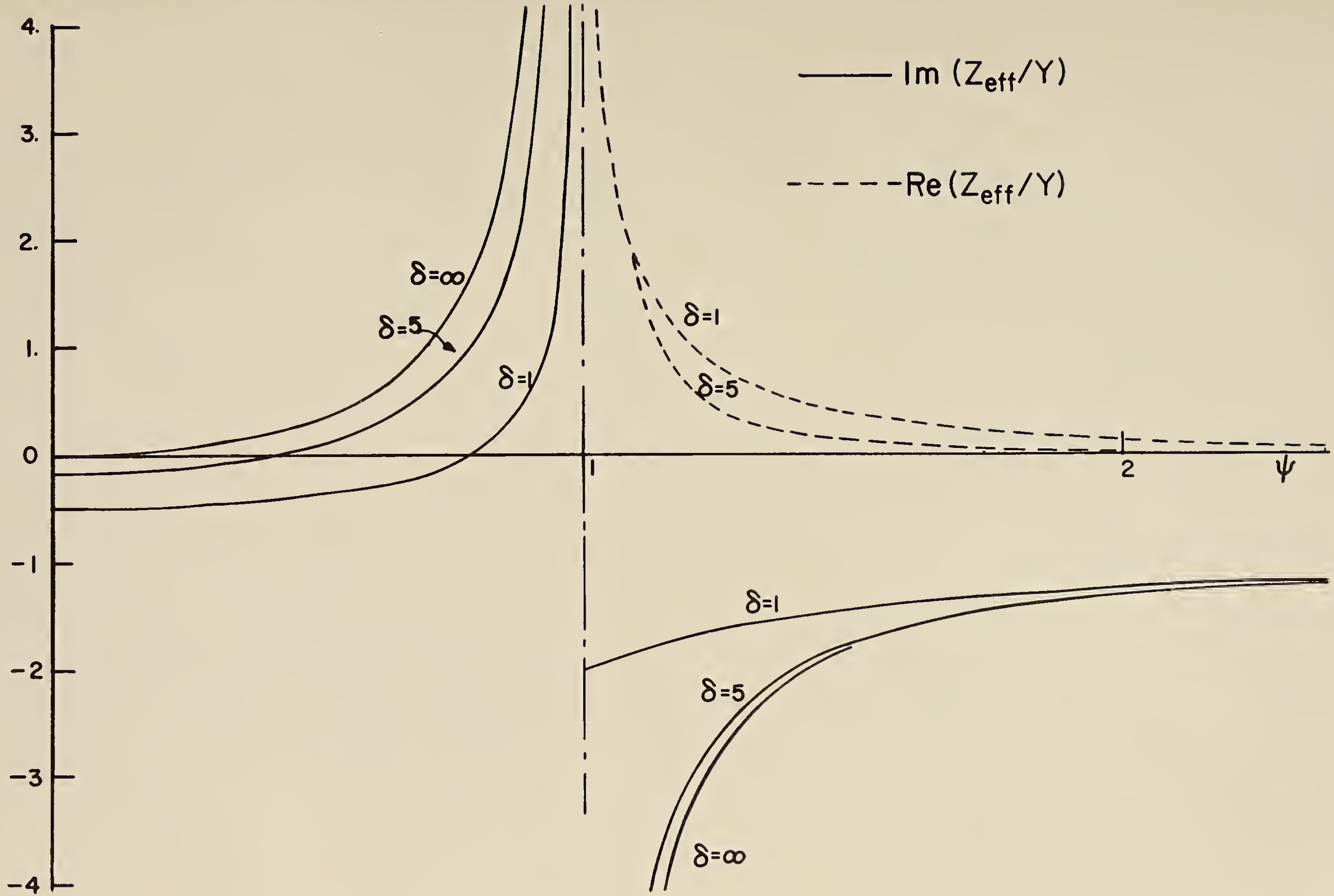


Fig. 1. The ratios of the real (interrupted line) and imaginary (solid line) parts of the impedance to the magnitude of the free space reactance as functions of the ratio ψ of the radio frequency to the electron plasma frequency.

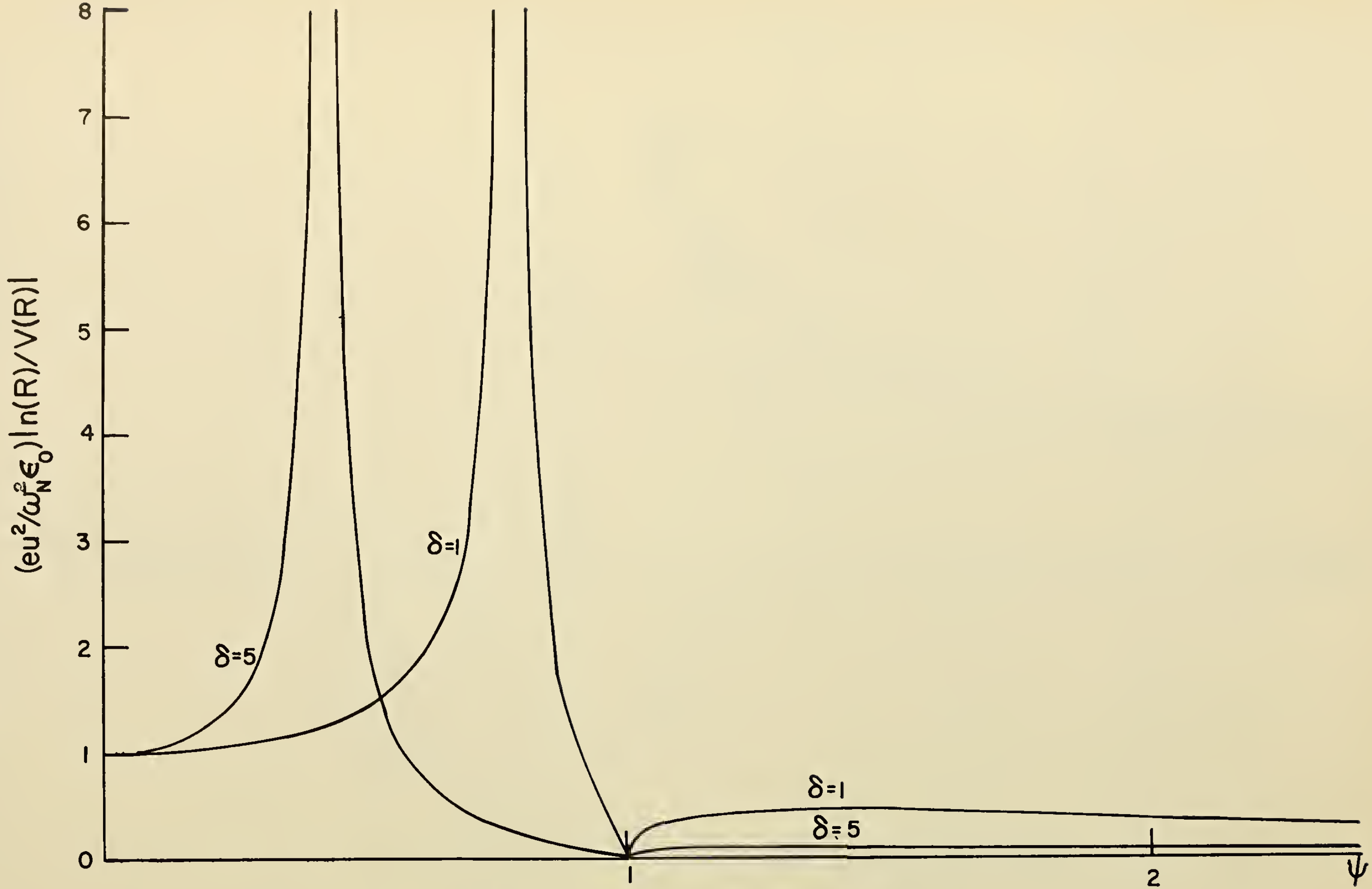


Fig. 2. The dimensionless quantity $(eu^2/\omega_N^2\epsilon_0)|n(R)/V(R)|$ which is proportional to the number density fluctuations $n(R)$ produced by a given rf voltage V(R), as a function of the ratio ψ of the radio frequency to the electron plasma frequency.

DEPARTMENT OF COMMERCE ★ NATIONAL BUREAU OF STANDARDS ★