Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 63D, No. 2, September–October 1959

Effect of Small Irregularities on the Constitutive Relations for the Ionosphere

K. G. Budden [1]

(March 18, 1959)

Irregularities in the ionosphere which are small compared with one wavelength may modify the constitutive relations, and hence, may affect the refractive indices for electromagnetic waves. The modifications are in some ways similar to those which would be introduced into the Appleton-Hartree formula by a Lorentz force. The theory is given first for the case when the irregularities extend only in one dimension, and it is found that even in a loss-free medium the refractive index now has an imaginary part which might be associated with loss of energy from the wave by scattering. The theory for three-dimensional irregularities is then discussed but is more difficult, and a method of successive approximations is used. The results indicate that small irregularities may play an important part in the propagation of very-low-frequency radio waves in the ionosphere. In particular, they may explain why "whistlers" are observed only on comparatively rare occasions.

1. Introduction

In one method of deriving the formulas of the magneto-ionic theory [1],[2] the ionized medium is assumed to have an electric polarization **P** which is related to the electric field **E** through a set of equations called the "constitutive relations." By using Maxwell's equations together with the constitutive relations, the Appleton-Hartree formula for the refractive index may be deduced. The vector **P** is given by:

$$\mathbf{P} = N e \mathbf{r} \tag{1}$$

where N is the number of electrons per unit volume, e is the charge on one electron, and **r** is the average displacement of an electron from its mean position. **P** is thus defined by taking an average over a volume large enough to contain many electrons. When used in Maxwell's equations it is assumed to represent a vector field, continuously distributed in space, and approximately constant over distances which are small compared to a wavelenth. There would be no meaning in speaking of the value of **P** at a specific point in the free space between the electrons. The use of Maxwell's equations implies a "smoothing out" process over a distance which must be large compared with $N^{-1/3}$.

There is now much evidence to show that the ionosphere is an irregular medium, so that the electron-number density N varies from place to place. The effect of these irregularities on the propagation of waves through the medium may be treated by considering the energy scattered from them, or by studying their effect on the refraction and diffraction of the waves. In the present paper, an alternative method is suggested for studying this problem in the special case when the irregularities are very small compared with one wavelength. The electric field **E** and the polarization **P** depend on the electron-number density N and vary from place to place because N varies, but it is possible to find their average values, $\mathbf{E}_0$ and $\mathbf{P}_0$, by a "smoothing out" process similar to that already mentioned. It is the average values which must be used in Maxwell's equations; and, in order to find the refractive index, it is necessary to know how $\mathbf{P}_0$ depends on $\mathbf{E}_0$. A new set of constitutive relations is therefore required giving $\mathbf{P}_0$ in terms of $\mathbf{E}_0$,

[1] Cavendish Laboratory, Cambridge, England. This work was done at the National Bureau of Standards while the author was on sabbatical leave from Cambridge University. The work was supported by the U.S. Air Force and the Bureau.

[2] Figures in brackets indicate the literature references at the end of this paper.

and this is, in general, different from the more familiar relations which give **P** in terms of **E**. Hence, the refractive index must also be modified when small irregularities are present. The purpose of the present paper is to find these modifications in some special cases.

2. Qualitative Description of the Method

The ionosphere is an electrically neutral medium, and it is assumed that the number of heavy positive ions per unit volume is the same as the number of electrons, N. Variations in N from place to place, therefore, occur equally for electrons and positive ions. They are assumed to be statistically stationary, and the average value of N over a very large region is N_0.

Suppose that a uniform electric field, **E**, is applied to the ionosphere. If **E** varies harmonically in time, the electron displacements also vary harmonically with the same frequency, and at any one instant all the electrons are displaced the same distance, **r**, from their mean positions The positive ions are assumed to be so heavy that they do not move. Let x be distance measured in the direction of **E**, and suppose that N varies with x as shown in figure 1. When the electrons are displaced, the effect is to introduce some space charge, which is negative where N is decreasing in the direction of **r**, and positive where it is increasing. The first effect of the space charge is to introduce an additional electric field, $\Delta\mathbf{E}$, whose average value is zero as long as the electrons are all displaced by the same amount. To a first approximation, therefore, the average value, $\mathbf{E}_0$, of the electric field is just the original field **E**, and the average value of the polarization is $\mathbf{P}_0 = N_0 e\mathbf{r}$. In this case, $\mathbf{P}_0$ and $\mathbf{E}_0$ are related by the same constitutive relations as for a homogeneous ionosphere. This results from the assumption that the electron displacements **r** are all the same.

The additional electric field $\Delta\mathbf{E}$ exerts a force on the electrons which changes their displacements. This in turn changes the space charge distribution, which then affects $\Delta\mathbf{E}$. This is a second-order effect since $\Delta\mathbf{E}$ is small, but it will be shown that it causes both **E** and **P** to have new average values, $\mathbf{E}_0$ and $\mathbf{P}_0$, which are related in a different way from the relation between **E** and **P**.

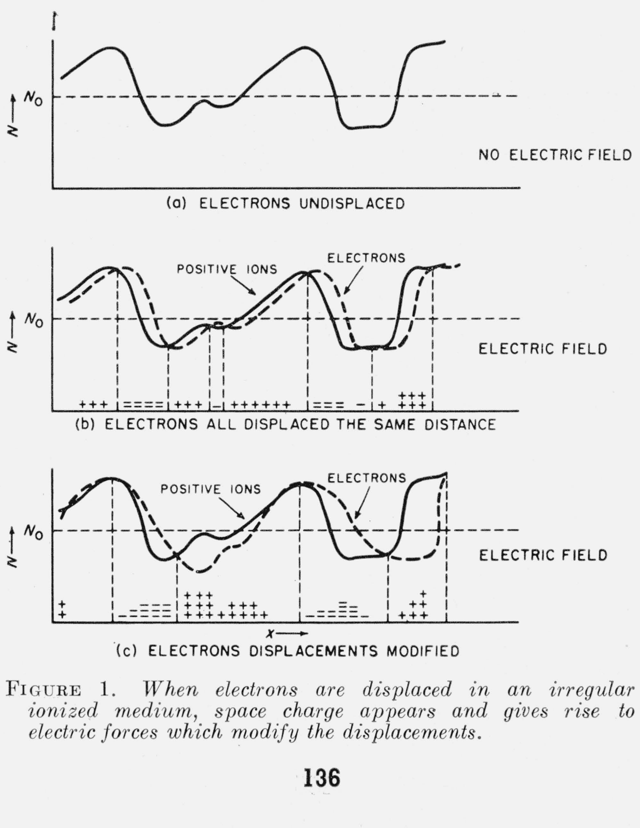

FIGURE 1. *When electrons are displaced in an irregular ionized medium, space charge appears and gives rise to electric forces which modify the displacements.*

3. Comparison With the Lorentz Force

When the magneto-ionic theory was first formulated, it was believed that the electric force on an electron in the ionosphere was made up of two parts. One part, $e\mathbf{E}$, arose from the electric field $\mathbf{E}$ impressed on the medium and the other part, $1/3\ \mathbf{P}e$ (rationalized units), arose from the polarization of the medium itself, and is called the Lorentz force. For many crystalline dielectrics, a Lorentz force must certainly be included, but there is now evidence that in the ionosphere it should be omitted. The strongest evidence is from the theory of "whistlers" which consist of very-low-frequency waves propagated in the extensive ionized region surrounding the earth, in directions close to the lines of force of the earth's magnetic field [5]. The Lorentz force would so modify the constitutive relations for the ionosphere, that these low-frequency waves could not be propagated.

The modification of the constitutive relations that would be introduced by small irregularities is similar in some ways to that which would arise from the Lorentz force. Both effects occur because there is a force acting on the electrons which arises from the polarization $\mathbf{P}$. For the Lorentz force this is simply proportional to $\mathbf{P}$, but for small irregularities it depends on both $\mathbf{P}$ and $\mathbf{E}$. For this reason, the effect of small irregularities is more complicated than that of a Lorentz force, and it will be shown that the additional force on the electron may depend upon the state of polarization of the wave. For circularly polarized waves of very low frequency traveling in the direction of the earth's magnetic field (whistlers), the effect of irregularities may have the same direction as the Lorentz force. This affects the dispersion law for these waves, and if the irregularities are large enough in amplitude, there is a lower limit to the frequencies that can be propagated.

4. Definitions of E and D

For a medium which is not free space, the electric intensity $\mathbf{E}$ is usually defined in terms of a long thin cavity. The component of $\mathbf{E}$ in a given direction is found as follows. A long thin cavity is imagined to be cut in the medium parallel to the given direction. Its length must be so small that the electric state of the medium does not change appreciably within it. For waves in a homogeneous ionosphere this means that the length is small compared to a wavelength. An infinitesimal test charge δq is placed at the center of the cavity and the component $\delta\mathbf{F}$ of the force acting on it in the direction of the cavity is measured. The component of $\mathbf{E}$ in this direction is

$$\lim_{\delta q\to 0} \delta\mathbf{F}/\delta q. \tag{1}$$

For a medium which contains discrete ions and electrons, this definition can still be used provided that the cross section of the cavity is very large compared with $N^{-\frac{1}{3}}$. It is the electric intensity so defined which is to be used in Maxwell's equations. The cavity definition thus effects the "smoothing out" process mentioned in the introduction. A similar argument applies to the electric displacement $\mathbf{D}$ which is defined in terms of the force normal to the plane of a flat, plate-like cavity on a test charge in the cavity. The longest dimension of the cavity must be small compared to a wavelength, λ and the thickness must be large compared with $N^{-\frac{1}{3}}$.

Now consider an ionized medium in which there are irregularities small in size compared to one wavelength. Suppose that the size of the irregularities is given by a length, l, which may be thought of as the average distance between one point of maximum electron density and its nearest neighbor. Then $l<<\lambda$. The electric intensity can now be defined in two ways.

The "local" electric intensity $\mathbf{E}$ is defined in terms of a long thin cavity whose length is very small compared with l, and whose cross section is large compared with $N^{-\frac{1}{3}}$. Its value depends on where the cavity is placed in relation to the maximums of N, and is affected by the local configuration of the space charge. It is this value of $\mathbf{E}$ which determines the force on an electron.

The average electric intensity $\mathbf{E}_0$ is defined in terms of a long thin cavity whose length is very small compared with λ, and whose cross section is large compared with l. In this case the distance of the test charge from the walls of the cavity is very large compared with l, so that the value of $\mathbf{E}_0$ is unaffected by local configuration of the irregularities. It is $\mathbf{E}_0$ which must be used in Maxwell's equations. The cavity definition here effects the "smoothing out" process over a distance small compared to a wavelength but large compared with l.

In a similar way, flat plate-like cavities may be used to define the local value $\mathbf{D}$ and the average value $\mathbf{D}_0$ of the electric displacement. The local electric polarization $\mathbf{P}$ is given by (1). Its average value is denoted by $\mathbf{P}_0$, and is *not*, in general, equal to $N_0\, e\, \mathbf{r}_0$ where $\mathbf{r}_0$ is the average value of $\mathbf{r}$. The following relations may be derived by the standard methods of electromagnetic theory (rationalized units):

$$\mathbf{D}=\epsilon_0\mathbf{E}+\mathbf{P}, \tag{2}$$

$$\mathbf{D}_0=\epsilon_0\mathbf{E}_0+\mathbf{P}_0. \tag{3}$$

$\mathbf{P}$ and $\mathbf{E}$ are related by the same constitutive relations as for a homogeneous medium, but this is not true for $\mathbf{P}_0$ and $\mathbf{E}_0$.

Before attempting the general theory for three-dimensional space, the theory will be illustrated by discussing a simpler problem in one dimension.

5. Irregularities in One Dimension

Suppose that the electron density N is a function only of x, and consider the effect of an electric field parallel to the x direction. The effect of the earth's magnetic field is neglected. Our problem is to calculate the effective dielectric constant of the medium. The reader may find it helpful to imagine that a large sample of the medium is placed between two conducting plates in planes x=constant, forming a parallel plate condenser whose capacity is to be found. $\mathbf{D}$ and $\mathbf{E}$ are both parallel to the x direction and may be written simply as D and E. The electron number density N is given by

$$N=N_0+\Delta N, \tag{4}$$

where ΔN is small compared with N_0. Similarly the displacement r of an electron in the x direction from its mean position is given by

$$r=r_0+\Delta r. \tag{5}$$

Further

$$E=E_0+\Delta E. \tag{6}$$

Each of ΔN, Δr, and ΔE is a function of x, and has an average value of zero.

Now div $\mathbf{D}=0$ so that in this case $dD/dx=0$ and

$$D=D_0=\text{constant} \tag{7}$$

for all values of x. This equation automatically takes care of the effect of space charge. The space charge density is $\rho=-dP/dx$. It gives rise to an additional electric field which must be such that div $\mathbf{D}=0$. It would clearly be incorrect to write div $\mathbf{D}=\rho$, for this would include the space charge twice.

The electron displacement r is related to E thus.

$$r=\gamma E, \tag{8}$$

where γ is a constant which is, in general, complex. If the effect of collisions is neglected, $\gamma=-e/m\omega^2$ where m is the mass of an electron and ω is the angular wave frequency. If the average number of collisions per unit time for an electron is ν, then $\gamma=e/(im\nu\omega-m\omega^2)$. Now $P=Ner=Ne\gamma E$ and hence

$$D=(\epsilon_0+Ne\gamma)E=D_0=\{\epsilon_0+(N_0+\Delta N)e\gamma\}(E_0+\Delta E). \tag{9}$$

This gives

$$\Delta E=\frac{D_0}{\epsilon_0+e\gamma N_0+e\gamma\Delta N}-E_o, \tag{10}$$

whose average value must be zero.

It is now necessary to make some further assumptions about the irregularities ΔN. Let $\Delta N=\xi$. It is assumed that ξ has a Gaussian distribution about zero, which means that the probability, $p(\xi)\,d\xi$, that ξ lies in the range ξ to $\xi+d\xi$ is given by

$$p(\xi)=\frac{1}{\sqrt{2}a\pi^{1/2}}\,e^{-\xi^2/2a^2}, \tag{11}$$

where

$$a^2=\overline{\Delta N^2}. \tag{12}$$

(A quantity with a bar over it denotes the average value of that quantity.) Then the average value of ΔE is

$$\overline{\Delta E}=\frac{1}{a\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{D_0}{\epsilon_0+e\gamma N_0+e\gamma\xi}\,e^{-\xi^2/2a^2}\,d\xi-E_0 \tag{13}$$

which must be zero. The integral is evaluated in appendix A. Let

$$\sqrt{2}ae\gamma=\zeta;\ \frac{\epsilon_0+e\gamma N_0}{\sqrt{2}ae\gamma}=\Omega. \tag{14}$$

Then the result of equating (13) to zero is:

$$\frac{E_0}{D_0}=-\frac{1}{\zeta}\left\{i\pi^{1/2}e^{-\Omega^2}-2F(\Omega)\right\}, \tag{15}$$

where

$$F(\Omega)=e^{-\Omega^2}\int_0^{\Omega}e^{u^2}du. \tag{16}$$

If electron collisions are neglected, γ is real and the integrand in (13) is then real at all points on the real ξ axis, but has a pole on this path. The principal value of the integral is then purely real. If γ has a small imaginary part, however, the integral has an imaginary part which does not tend to zero as $I(\gamma)\rightarrow 0$. Thus the limiting value of the integral is not the same as the principal value. Now it is impossible for the collision frequency to be exactly zero so that clearly the limiting value is the correct one to use when collisions are neglected.

Now

$$1+N_0e\gamma/\epsilon_0=\epsilon', \tag{17}$$

where ϵ' is the dielectric constant of a homogeneous medium with the electron number density N_0. The true dielectric constant is $\epsilon=D_0/\epsilon_0E_0$. Let

$$\overline{\left(\frac{\Delta N}{N_0}\right)^2}=\frac{a^2}{{N_0}^2}=\beta. \tag{18}$$

Then (14) gives

$$\zeta=(2\beta)^{1/2}\epsilon_0(\epsilon'-1);\ \Omega=(2\beta)^{-1/2}\epsilon'/(\epsilon'-1), \tag{19}$$

which may be substituted in (15).

If electron collisions are neglected:

$$\epsilon' = 1 - X, \text{ where } X = \frac{N_0 e^2}{\epsilon_0 m \omega^2}, \tag{20}$$

and (15) gives

$$\mu^2 = \epsilon = \sqrt{2} X \beta^{1/2} \{ i\pi^{1/2} e^{-\Omega^2} - 2F(\Omega) \}^{-1}. \tag{21}$$

This is the formula for the square of the refractive index in the one-dimensional case when small irregularities are present. It replaces the formula $\mu^2 = 1 - X$ which would be true for a homogeneous medium. The quantity X is the same as Appleton's [1] x and is proportional to the electron density. In figure 2 curves are plotted showing how the real and imaginary parts of μ depend upon X for various values of β.

Equation (21) shows that μ^2 has an imaginary part even when electron collisions are neglected. Hence, some energy would be lost from an electromagnetic wave although there is no physical mechanism for converting electric energy into heat. It is known that irregularities cause scattering of energy from the original wave, and this must account for the imaginary part of μ^2.

When $X \gg 1$, Ω is large since β is small. The term, $e^{-\Omega^2}$ may then be neglected, and $F(\Omega)$ may be replaced by the first two terms of its asymptotic expansion (eq A7, app. A). If powers of β higher than the first are neglected, the expression (21) becomes

$$\mu^2 \approx 1 - \frac{X}{1 + \beta X/(X-1)}, \tag{22}$$

which may be compared with the expression for μ^2 in a homogeneous medium when the Lorentz term is included, namely:

$$\mu^2 = 1 - \frac{X}{1 + (1/3)X}. \tag{23}$$

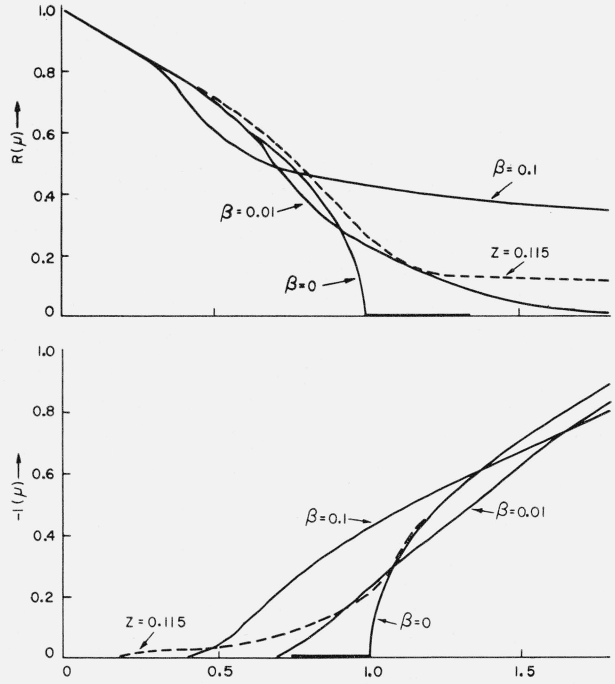

FIGURE 2. *The effect of small irregularities on the variation of $R(\mu)$ and $-I(\mu)$ with X.*

β is a measure of the strength of the irregularities. For the full curves electron collisions are neglected. For the dashed curves, $Z = \nu/\omega = 0.115$, and $\beta = 0$.

The factor $\beta/(X-1)$ in (22) is analogous to the Lorentz factor, 1/3, in (23). The effect of the irregularities has the same sign as a Lorentz force (since $X \gg 1$), but gets progressively smaller as X increases.

6. Irregularities in Three Dimensions, General Theory

There appears to be no direct way of extending the method of the last section to deal with irregularities in three dimensions. In one dimension the equation, div $\mathbf{D}=0$, has the simple integral $\mathbf{D}=$constant, but there is no correspondingly simple result for three dimensions. In this and the following sections, therefore, a solution by successive approximations will be attempted.

Cartesian coordinates x, y, z, are used. Let $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ denote unit vectors parallel to the x, y, and z axes respectively. Then the position of a point (x, y, z) is given by the vector $\mathbf{s}=\mathbf{i}x+\mathbf{j}y+\mathbf{k}z$. The electron number density N is given by (4) where ΔN is now an irregular function of x, y, z. ΔN may be expressed as the sum of many Fourier terms thus:

$$\Delta N=\sum_K (\Delta N)_K, \tag{24}$$

$$(\Delta N)_K=A_K \cos\,(\mathbf{K}\mathbf{s}-\phi_K), \tag{25}$$

where

$$\mathbf{K}=u\mathbf{i}+v\mathbf{j}+w\mathbf{k}. \tag{26}$$

The summation is really a triple sum over all values of u, v, w. A_K and ϕ_K are the amplitude and phase, respectively, of the Fourier term.

7. First Approximation

As a first approximation it is assumed that the electric field $\mathbf{E}$ in the medium is the same at all points and equal to $\mathbf{E}_1$. Then each electron is displaced from its mean position by the same amount $\mathbf{r}_1$ given by

$$\mathbf{r}_1=\mathbf{T}\cdot\mathbf{E}_1, \tag{27}$$

where $\mathbf{T}$ is a tensor which can be found from the equation of motion of an electron. Because of the irregularities ΔN, the displacement $\mathbf{r}_1$ results in the appearance of space charge which in turn gives an additional electric field $\Delta\mathbf{E}_1$. The contribution $\Delta\mathbf{E}_1{}^{(K)}$ to $\Delta\mathbf{E}_1$ from one Fourier component in (24) will now be found.

A single Fourier component $(\Delta N)_K$ is a one-dimensional variation of ΔN and may be treated as in section 5. The additional electric field $\Delta\mathbf{E}_1{}^{(K)}$ must be normal to the planes of constant ΔN for this component and so is parallel to the vector $\mathbf{K}$. At this stage $\mathbf{r}_1$ is taken as a constant, so that the average value of $\mathbf{E}_1$ must remain unaffected by the irregularities. Hence, $\overline{\Delta\mathbf{E}_1{}^{(K)}}=0$. For components parallel to $\mathbf{K}$ we have $\epsilon_0\, E+P=D=$constant, so that $\epsilon_0\, \Delta E+\Delta P=0$. Now $\Delta P=e(\Delta N)_K r_1{}^{(K)}$ where $r_1{}^{(K)}$ is the component of $\mathbf{r}_1$ parallel to $\mathbf{K}$. Hence

$$\Delta E_1{}^{(K)}=-\frac{e}{\epsilon_0}A_K \cos\,(\mathbf{K}\mathbf{s}-\phi_K)\,\frac{(\mathbf{r}_1\mathbf{K})}{\mathbf{K}^2}\cdot\mathbf{K}. \tag{28}$$

Here the factor $(\mathbf{r}_1\mathbf{K})/|K|$ is $r_1{}^{(K)}$, and the factor $\mathbf{K}/|K|$ is a unit vector in the direction of $\mathbf{K}$.

The total additional field $\Delta\mathbf{E}_1$ from the first approximation is given by:

$$\Delta\mathbf{E}_1=\sum_K \Delta\mathbf{E}_1{}^{(K)}. \tag{29}$$

8. Second Approximation

The effect of the appearance of the additional electric field $\Delta\mathbf{E}_1$ is to give the electrons an additional displacement,

$$\Delta \mathbf{r}_1 = \mathbf{T} \cdot \Delta \mathbf{E}_1. \tag{30}$$

Both $\Delta \mathbf{r}_1$ and $\Delta \mathbf{E}_1$ are functions of position. $\Delta \mathbf{r}_1$ gives rise to further space charge which in turn results in a further contribution $\Delta \mathbf{E}_2$ to the electric field. The average value of $\Delta \mathbf{E}_1$ is zero, but the average value $\overline{\Delta \mathbf{E}_2}$ of $\Delta \mathbf{E}_2$ is not in general zero, and its value must now be found.

The resultant electron displacement is $\mathbf{r}_1 + \Delta \mathbf{r}_1$, which must be substituted for $\mathbf{r}_1$ in eq (28). This gives:

$$\Delta \mathbf{E}_2{}^{(K)} = -\frac{e}{\epsilon_0} A_K \cos(\mathbf{K}\,\mathbf{s} - \phi_K) \frac{\{(\mathbf{r}_1 + \Delta \mathbf{r}_1) \cdot \mathbf{K}\}}{\mathbf{K}^2} \cdot \mathbf{K}, \tag{31}$$

and $\Delta \mathbf{E}_2$ is given by

$$\Delta \mathbf{E}_2 = \sum_K \Delta \mathbf{E}_2{}^{(K)}. \tag{32}$$

$\Delta \mathbf{r}_1$, and therefore $\Delta \mathbf{E}_2{}^{(K)}$, contain a summation over all values of $\mathbf{K}$, and so (32) contains a double summation. Now take the average of (32). Clearly $\cos(\mathbf{K}_1 \mathbf{s} - \phi_{K_1}) \cos(\mathbf{K}_2 \mathbf{s} - \phi_{K_2})$ has the average value zero when $\mathbf{K}_1 \neq \mathbf{K}_2$ and the average value 1/2 when $\mathbf{K}_1 = \mathbf{K}_2$. The value of $\mathbf{r}_1$ may be inserted from (27). Hence:

$$\overline{\Delta \mathbf{E}_2} = \frac{e^2}{2\epsilon_0^2} \sum_K \frac{A_K^2 \{\mathbf{T}\,\mathbf{K} \cdot \mathbf{K}\} \{\mathbf{T}\,\mathbf{E}_1 \cdot \mathbf{K}\}}{\mathbf{K}^4} \cdot \mathbf{K}. \tag{33}$$

The average electric field is now given by

$$\mathbf{E}_0 = \mathbf{E}_1 + \overline{\Delta \mathbf{E}_2}. \tag{34}$$

Next it is necessary to find the average value $\mathbf{P}_0$ of the polarization $\mathbf{P}$. From eq (1):

$$\mathbf{P} = e(N_0 + \Delta N)(\mathbf{r}_1 + \Delta \mathbf{r}_1), \tag{35}$$

and hence

$$\mathbf{P}_0 = eN_0 \mathbf{r}_1 + e\,\overline{\Delta N \Delta \mathbf{r}_1}. \tag{36}$$

From (28), (29), and (30):

$$\Delta N \Delta \mathbf{r}_1 = -\frac{e}{\epsilon_0} \sum_K A_K \cos(\mathbf{K}\mathbf{s} - \phi_K) \frac{(\mathbf{r}_1 \cdot \mathbf{K})}{\mathbf{K}^2} \mathbf{T}\,\mathbf{K} \sum_K A_K \cos(\mathbf{K}\mathbf{s} - \phi_K), \tag{37}$$

and on taking the average:

$$\overline{\Delta N \Delta \mathbf{r}_1} = -\frac{e}{2\epsilon_0} \sum_K A_K^2 \frac{(\mathbf{r}_1 \cdot \mathbf{K})}{\mathbf{K}^2} \mathbf{T}\,\mathbf{K}. \tag{38}$$

Hence from (36) and (27)

$$\mathbf{P}_0 = eN_0 \mathbf{T}\,\mathbf{E}_1 - \frac{e^2}{2\epsilon_0} \sum_K A_K^2 \frac{\{\mathbf{T}\,\mathbf{E}_1 \cdot \mathbf{K}\} \mathbf{T}\,\mathbf{K}}{\mathbf{K}^2}. \tag{39}$$

In eq (33), (34), and (39), $\mathbf{P}_0$ and $\mathbf{E}_0$ are expressed in terms of $\mathbf{E}_1$. Elimination of $\mathbf{E}_1$ gives the constitutive relations between $\mathbf{P}_0$ and $\mathbf{E}_0$.

9. Discussion When T Is a Scalar

If the effect of the earth's magnetic field is neglected, the tensor $\mathbf{T}$ reduces to the scalar γ used in sec. 5. Then eq (33) becomes

$$\overline{\Delta \mathbf{E}_2} = \frac{\gamma^2 e^2}{2\epsilon_0^2} \sum_K A_K^2 \left\{ \frac{(\mathbf{E}_1 \cdot \mathbf{K})}{\mathbf{K}^2} \cdot \mathbf{K} \right\}. \tag{40}$$

The factor $[(\mathbf{E}_1 \cdot \mathbf{K})/\mathbf{K}^2] \cdot \mathbf{K}$ is simply the component of $\mathbf{E}_1$ in the direction of $\mathbf{K}$.

It is now assumed that the irregularities are isotropic. This means that for a given value of $|K|$ the directions of the vectors, $\mathbf{K}$, are uniformly distributed in any hemisphere and the probability of a given value of A_K is independent of the direction of K. Select first those values of $\mathbf{K}$ for which $|K|$ is in a specified range, $d|K|$, and A_K is in a specified range, dA_K, and perform the summation for these values first. Then the last factor in (36) is simply 1/2 $\mathbf{E}_1$. Hence (40) becomes:

$$\overline{\Delta \mathbf{E}_2} = \frac{1}{4} \frac{\gamma^2 e^2}{\epsilon_0^2} \mathbf{E}_1 \sum_K A_K^2. \tag{41}$$

But

$$\frac{1}{2} \sum_K A_K^2 = \overline{(\Delta N)^2}. \tag{42}$$

Hence (34), (41), and (42) give:

$$\mathbf{E}_0 = \mathbf{E}_1 \left\{ 1 + \frac{1}{2} \frac{e^2 \gamma^2 N_0^2}{\epsilon_0^2} \overline{\left(\frac{\Delta N}{N_0} \right)^2} \right\}. \tag{43}$$

Equation (39) becomes:

$$\mathbf{P}_0 = \gamma e N_0 \mathbf{E}_1 - \frac{\gamma^2 e^2}{2\epsilon_0} \sum_K A_K^2 \left\{ \frac{(\mathbf{E}_1 \cdot \mathbf{K})}{\mathbf{K}^2} \cdot \mathbf{K} \right\} \tag{44}$$

which reduces in a similar way to

$$\mathbf{P}_0 = \mathbf{E}_1 \left\{ \gamma e N_0 - \frac{1}{2} \frac{\gamma^2 e^2 N_0^2}{\epsilon_0} \overline{\left(\frac{\Delta N}{N_0} \right)^2} \right\}. \tag{45}$$

The dielectric constant ϵ is given by

$$\epsilon_0 \, \epsilon \, \mathbf{E}_0 = \epsilon_0 \, \mathbf{E}_0 + \mathbf{P}_0. \tag{46}$$

It is convenient to use the expression (17) for ϵ', and the abbreviation (18). Then:

$$\epsilon = \epsilon' / \{ 1 + \tfrac{1}{2} \beta \, (\epsilon' - 1)^2 \}. \tag{47}$$

If electron collisions are neglected and (20) is used, eq (47) becomes:

$$\mu^2 = \epsilon = \frac{1 - X}{1 + \frac{1}{2} \beta X^2} \tag{48}$$

which should be compared with the expressions (21) and (22) obtained for the one-dimensional case. The difference is not surprising since the plane sheets of space charge formed in the one-dimensional problem would be expected to have a larger and different effect from an irregular three-dimensional distribution of charge. It is probable that (48) is unreliable when X is near unity, for then the wave frequency is near the plasma frequency, and it is possible for the electrons to make oscillations of large amplitude. The approximations made in the derivation would then be invalid. This may also explain why the expression (48) is purely real. In a more accurate treatment it should have an imaginary part which is associated with loss of energy from the wave by scattering.

10. Discussion When T is a Tensor

The constitutive relations will now be derived for the special case of a plane wave with its wave normal in the direction of the earth's magnetic field, which is taken as the z axis. The tensor $\mathbf{T}$ is then given by:

$$\mathbf{T} = -\frac{e}{m\omega^2} \frac{1}{U(U^2 - Y^2)} \begin{bmatrix} U^2 & iYU & 0 \\ -iYU & U^2 & 0 \\ 0 & 0 & U^2 - Y^2 \end{bmatrix}, \tag{49}$$

(see for example, Banerjea [2]) where $U=1-iZ$, $Z=\nu/\omega$, ν is the average number of collisions made per unit time by an electron, $Y=\omega_H/\omega$, and $\omega_H/2\pi$ is the magnetic gyrofrequency for electrons. This case is of interest in the study of the propagation of whistlers [5].

In a homogeneous medium, a wave whose normal is parallel to the magnetic field is circularly polarized, and the electric field has no component in the direction of the wave normal. Now $\mathbf{E}_1$ is the first approximation to the electric field of the wave and is, therefore, the field of a wave in a homogeneous medium. Hence, it has no component parallel to the z axis, and (27) and (49) show that $\mathbf{r}_1$ similarly has no z component. Hence we may write:

$$\mathbf{r}_1=\mathbf{i}\,x_1+\mathbf{j}\,y_1. \tag{50}$$

There is strong evidence that irregularities in the ionosphere are elongated in the direction of the earth's magnetic field [4]. To allow for this, eq (26) is replaced by

$$\mathbf{K}=u\mathbf{i}+v\mathbf{j}+fw\,\mathbf{k}, \tag{51}$$

where f is a number less than unity. It will be assumed that the irregularities of ΔN are isotropic with respect to the variables u, v, w. The effect of the factor f is as though a region of truly isotropic irregularities had been stretched in the z direction by a factor $1/f$, and f will therefore be called the "elongation factor".

The expression (33) for $\overline{\Delta\mathbf{E}_2}$ must now be written down for this special case. The notation T_{ij} (i, $j=x$, y, z) will be used for elements of $\mathbf{T}$. Equation (49) shows that $T_{xz}=T_{yz}=T_{zx}=T_{zy}=0$, $T_{xy}=-T_{yx}$. Hence the factor $(\mathbf{T}\,\mathbf{K})\cdot\mathbf{K}$ in (33) becomes

$$(\mathbf{T}\,\mathbf{K})\cdot\mathbf{K}=T_{xx}u^2+T_{yy}v^2+T_{zz}f^2w^2. \tag{52}$$

Equations (27) and (50) show that the factor $\{(\mathbf{T}\,\mathbf{E}_1)\cdot\mathbf{K}\}\cdot\mathbf{K}$ in (33) is given by

$$(x_1u+y_1v)\,(u\mathbf{i}+v\mathbf{j}+fw\mathbf{k}). \tag{53}$$

Now when the summation in (33) is completed, terms involving the products uv, vw, wu will give zero, because ΔN is isotropic with respect to u, v, w. Hence (53) becomes:

$$\{\mathbf{T}\,\mathbf{E}_1\cdot\mathbf{K}\}\cdot\mathbf{K}=x_1u^2\mathbf{i}+y_1v^2\mathbf{j}. \tag{54}$$

Equation (49) shows further that $T_{xx}=T_{yy}$. Hence eq (33) can now be written in full using (52) and (54). It gives

$$\overline{\Delta\mathbf{E}_2}=\frac{e^2}{2\epsilon_0^2}\sum_K\frac{A_K^2}{(u^2+v^2+f^2w^2)^2}\,[\{T_{xx}(u^4+u^2v^2)+T_{zz}f^2u^2w^2\}\mathbf{i}\,x_1+\{T_{xx}(v^4+u^2v^2)+T_{zz}f^2v^2w^2\}\mathbf{j}\,y_1]. \tag{55}$$

It is shown in app. B that

$$\sum_K A_K^2\frac{u^4}{(u^2+v^2+f^2w^2)^2}=\sum_K A_K^2\frac{v^4}{(u^2+v^2+f^2w^2)^2}=3\sum_K A_K^2\frac{u^2v^2}{(u^2+v^2+f^2w^2)^2}=\frac{3}{4}L(f)\,\beta N_0^2 \tag{56}$$

$$f^2\sum_K A_K^2\frac{u^2w^2}{(u^2+v^2+f^2w^2)^2}=f^2\sum_K A_K^2\frac{v^2w^2}{(u^2+v^2+f^2w^2)^2}=M(f)\cdot\beta N_0^2 \tag{57}$$

where $L(f)$, $M(f)$ are known functions of f, and β is given by (18). Hence (55) becomes:

$$\overline{\Delta\mathbf{E}_2}=\frac{e^2}{2\epsilon_0^2}\,\beta N_0^2\{T_{xx}L(f)+T_{zz}M(f)\}\mathbf{r}_1, \tag{58}$$

and (34) (27) give:

$$\mathbf{E}_0=\mathbf{E}_1+\frac{e^2}{2\epsilon_0^2}\,\beta N_0^2\{T_{xx}L+T_{zz}M\}\mathbf{T}\,\mathbf{E}_1. \tag{59}$$

Next, eq (39) must be written down for this special case. By methods similar to the above it can be shown that

$$\sum_K A_K^2 \frac{\{\mathbf{T}\,\mathbf{E}_1\cdot\mathbf{K}\}\cdot\mathbf{T}\,\mathbf{K}}{\mathbf{K}^2}=\beta N_0^2 Q(f)\{\mathbf{i}(T_{xx}x_1+T_{xy}y_1)+\mathbf{j}(T_{yx}x_1+T_{yy}y_1)\} \tag{60}$$

where

$$\sum_K A_K^2 \frac{u^2}{u^2+v^2+f^2w^2}=\sum_K A_K^2 \frac{v^2}{u^2+v^2+f^2w^2}=Q(f)\cdot\beta N_0^2 \tag{61}$$

(see app B)
Equation (60) is now substituted in (39) and (27), (50) are used. This gives

$$\mathbf{P}_0=eN_0\mathbf{T}\,\mathbf{E}_1-\frac{e^2N_0^2}{2\epsilon_0}\,Q\beta\mathbf{T}^2\,\mathbf{E}_1 \tag{62}$$

which may now be combined with (59) to give the constitutive relation.

The relations (59) and (62) are equivalent to:

$$\mathbf{E}_0=\left\{\mathbf{T}^{-1}+\frac{e^2}{2\epsilon_0^2}\beta N_0^2(T_{xx}L+T_{zz}M)\right\}\mathbf{T}\,\mathbf{E}_1, \tag{63}$$

$$\mathbf{P}_0=eN_0\left(1-\frac{eN_0}{2\epsilon_0}\,Q\beta\,\mathbf{T}\right)\mathbf{T}\,\mathbf{E}_1, \tag{64}$$

from which the vector $\mathbf{T}\,\mathbf{E}_1$ is to be eliminated. This then gives

$$\mathbf{E}_0=\mathbf{R}\,\mathbf{P}_0 \tag{65}$$

where $\mathbf{R}$ is to be found. Now β is a small quantity and its square and higher powers will be neglected. Then it can be shown that

$$\mathbf{R}\approx\frac{\mathbf{T}^{-1}}{eN_0}-\frac{\beta}{2\epsilon_0}\left\{X\left(\frac{LU}{U^2-Y^2}+\frac{M}{U}\right)-Q\right\}, \tag{66}$$

and it follows from the expression (49) for $\mathbf{T}$ that

$$\left.\begin{array}{l} R_{xz}=R_{zx}=R_{yz}=R_{zy}=0 \\ R_{xx}=R_{yy};\quad R_{xy}=-R_{yx} \end{array}\right\}. \tag{67}$$

The standard methods of magneto-ionic theory may now be used to find the polarization and refractive indices of the waves which can be propagated. The two polarizations are given by $\rho=\pm i$, so that the two waves are circularly polarized with opposite senses, and the two refractive indices are given by

$$\mu^2=1+\frac{1}{\epsilon_0(R_{xx}\pm iR_{xy})}, \tag{68}$$

which leads to

$$\mu^2=1-\frac{X}{U+\frac{1}{2}\beta\left\{X^2\left(\frac{LU}{U^2-Y^2}+\frac{M}{U}\right)-QX\right\}\pm Y}. \tag{69}$$

11. Application to Whistlers

Whistlers are naturally occurring radio signals propagated in the extensive ionized region surrounding the earth, with the wave normals in directions close to the lines of force of the earth's magnetic field [5]. The frequency is very low (often in the audible range) so that X and Y

are very large compared to unity. The waves are not heavily attenuated so that the effect of electron collisions must be small. Hence in (69) we take $U=1$, $X\gg 1$, $Y\gg 1$. Moreover X/Y is inversely proportional to frequency and is very large at low frequencies, so that $X/Y\gg 1$ [5]. One of the values of μ given by (69) is almost purely imaginary; the other value, with $-Y$ in the denominator, refers to the wave which is believed to explain whistlers, and it is given by

$$\mu^2\approx\frac{X}{Y-\frac{\beta}{2}X^2\left(M-\frac{L}{Y^2}\right)+\frac{1}{2}\beta QX}. \tag{70}$$

If the elongation factor f is not near zero, then $M(f)$ and $L(f)$ are both non-zero, but $L/Y^2\ll M$ since Y is large. X may be neglected in comparison with X^2, so that (70) becomes:

$$\mu^2\approx\frac{X}{Y-\frac{1}{2}\beta MX^2}. \tag{71}$$

The additional term $\frac{1}{2}\beta MX^2$ in the denominator has an effect similar to that which the Lorentz force would have. Waves could not be propagated if

$$\frac{\beta}{2}MX^2>Y, \tag{72}$$

for then the refractive index would be purely imaginary. Hence, irregularities could modify the propagation of whistlers. Equation (72) shows that there is a lower limit to the frequencies which could be propagated, given by

$$(\tfrac{1}{2}\beta Mf_N^4/f_H)^{1/3} \tag{73}$$

where f_N is the plasma frequency and f_H is the gyrofrequency.

If, however, the irregularities are very elongated in the direction of the earth's magnetic field, so that the elongation factor f is very small, then $M(f)$ is small (see app. B), and its effect could be opposed by the term L/Y^2 in the denominator of (70). If f is so small that $L\approx MY^2$, then the effect of irregularities on the propagation of whistlers would be small.

It is well known that whistlers are not observed regularly, but only during comparatively short periods of time. They are believed to start from lightning flashes, so that sources of whistlers are always plentiful. The absence of whistlers for long periods must therefore be explained by some feature of the propagation. It is now tentatively suggested that whistlers cannot be observed when the path they would travel contains irregularities which are not greatly elongated. On the occasions when whistlers can be observed, the irregularities are either absent or greatly elongated in the direction of the earth's magnetic field.

12. Conclusion

It has been shown that irregularities in the ionosphere, which are small compared with one wavelength, may modify the constitutive relations and hence may affect the refractive indices for electromagnetic waves. The modifications are in some ways similar to those which would be introduced into the Appleton-Hartree formula by a Lorentz force. When the irregularities extend only in one dimension the theory can be handled by the statistical method of section 5, and it was found that even in a loss-free medium the refractive index now has an imaginary part which might be associated with loss of energy from the wave by scattering. For the general three-dimensional case, the theory is more difficult and has only been handled by a method of successive approximations. A theory similar to that of section 5 is needed for this case but has not so far been found. The method of section 6 onwards is less satisfactory because of the numerous approximations. The results, however, do indicate that small irregularities may play an important part in the propagation of very-low-frequency radio waves in the ionosphere. In particular, they may explain why whistlers are observed only on comparatively rare occasions.

The author is greatly indebted to the University of Colorado, and to the Directors and Staffs of the High Altitude Observatory and NBS for providing him with the opportunity to work in Boulder.

13. Appendix A: Evaluation of the Integral in Equation (13)

With the notation (14), the integral in (13) becomes

$$\frac{D_0}{a\zeta\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{1}{\Omega+\xi/a\sqrt{2}}\,e^{-\xi^2/2a^2}\cdot d\xi. \tag{A1}$$

Let $\xi/\sqrt{2}a=w$. Then (A1) gives

$$\frac{D_0}{\sqrt{\pi}\,\zeta}\int_{-\infty}^{\infty}\frac{1}{w+\Omega}\,e^{-w^2}\cdot dw. \tag{A2}$$

Now

$$\frac{1}{w+\Omega}=i\int_0^{\pm\infty}e^{-i\tau(w+\Omega)}\,d\tau, \tag{A3}$$

where the top limit is $+\infty$ when $I\,(\Omega)$ is negative, and $-\infty$ when $I\,(\Omega)$ is positive. Here $I\,(\Omega)$ is always positive. Hence

$$\int_{-\infty}^{\infty}\frac{1}{w+\Omega}\,e^{-w^2}dw=i\int_0^{-\infty}e^{-i\Omega\tau}\left[\int_{-\infty}^{\infty}e^{-w^2-i\tau w}dw\right]d\tau$$

$$=i\sqrt{\pi}\int_0^{-\infty}e^{(-1/4\,\tau^2-i\Omega\tau)}\cdot d\tau=2i\pi^{1/2}e^{-\Omega^2}\int_{i\Omega}^{-\infty}e^{-v^2}dv$$

$$=-\pi^{1/2}[i\pi^{1/2}e^{-\Omega^2}-2F(\Omega)] \tag{A4}$$

where

$$F(\Omega)=e^{-\Omega^2}\int_0^{\Omega}e^{u^2}\cdot du. \tag{A5}$$

When (A5) is subsitituted in (A2) it leads to the result (15).

The function $F\,(\Omega)$ has been studied by Mitchell and Zemansky [3] who give the following formulas:

when Ω is small;

$$1-2\Omega F(\Omega)=1-\frac{2\Omega^2}{1}+\frac{(2\Omega^2)^2}{1\cdot 3}-\frac{(2\Omega^2)^3}{1\cdot 3\cdot 5}+\cdots \tag{A6}$$

and when Ω is large;

$$F(\Omega)\sim\frac{1}{2\Omega}\left[1+\frac{1}{(2\Omega^2)}+\frac{1\cdot 3}{(2\Omega^2)^2}+\frac{1\cdot 3\cdot 5}{(2\Omega^2)^3}+\cdots\right]. \tag{A7}$$

14. Appendix B: Evaluation of Coefficents in Equations (55) and (60)

According to eq (51) the vector $\mathbf{K}=\mathbf{i}u+\mathbf{i}v+\mathbf{k}fw$. Its direction cosines are therefore

$$l=\frac{u}{(u^2+v^2+f^2w^2)^{1/2}},\quad m=\frac{v}{(u^2+v^2+f^2w^2)^{1/2}},\quad n=\frac{fw}{(u^2+v^2+f^2w^2)^{1/2}} \tag{B1}$$

These may be written in spherical polar coordinates:

$$l=\sin\theta\cos\phi,\quad m=\sin\theta\sin\phi,\quad n=\cos\theta. \tag{B2}$$

In equation (55) the coefficient $u^4/(u^2+v^2+f^2w^2)^2$ is $l^4=\sin^4\theta\cos^4\phi$. It is assumed that the distribution of the directions of K is the same for all values of A_K, which means that

$$\sum_K A_K^2 l^4 = \overline{\sin^4\theta\cos^4\phi}\sum_K A_K^2. \tag{B3}$$

Since all values of ϕ are equally probable this is equal to

$$\sum_K A_K^2 m^4 = \overline{\sin^4\theta\sin^4\phi}\sum_K A_K^2. \tag{B4}$$

It can be shown that $\overline{\sin^4\phi}=\overline{\cos^4\phi}=3/8$; $\overline{\cos^2\phi\ \sin^2\phi}=1/8$; $\overline{\cos^2\phi}=\overline{\sin^2\phi}=1/2$. Hence:

$$\sum A_K^2 l^4 = \sum A_K^2 m^4 = 3/8\ \overline{\sin^4\theta}\sum A_K^2, \tag{B5}$$

$$\sum A_K^2 l^2 m^2 = 1/8\ \overline{\sin^4\theta}\sum A_K^2, \tag{B6}$$

$$\sum A_K^2 l^2 n^2 = \sum A_K^2 m^2 n^2 = 1/2\ \overline{\cos^2\theta\sin^2\theta}\sum A_K^2, \tag{B7}$$

$$\sum A_K^2 l^2 = \sum A_K^2 m^2 = 1/2\ \overline{\sin^2\theta}\sum A_K^2, \tag{B8}$$

$$\sum A_K^2 n^2 = \overline{\cos^2\theta}\sum A_K^2. \tag{B9}$$

Now let

$$fz=z' \tag{B10}$$

so that

$$\mathbf{K}=u\mathbf{i}+v\mathbf{j}+w\mathbf{k}', \tag{B11}$$

where **i, j, k**′ are unit vectors parallel to the axes of the coordinate system (x, y, z'). Let the spherical polar angles in this system be θ', ϕ. Then

$$f\tan\theta=\tan\theta'. \tag{B12}$$

Now the irregularities are isotropic with respect of u, v, w, which means that all directions of the vector **K** in the system x, y, z' are equally probable. Hence, the probability $p(\theta', \phi)d\theta' d\phi$ that θ', ϕ, lie in given ranges $d\theta' d\phi$ is given by

$$p(\theta', \phi)=(1/4\pi)\sin\theta'. \tag{B13}$$

The corresponding probability that θ,ϕ lie in ranges $d\theta,d\phi$ is $p(\theta,\phi)\ d\theta d\phi$ where

$$p(\theta,\phi)=p(\theta', \phi)\frac{d\theta'}{d\theta}=\frac{f^2\sin\theta}{4\pi(\cos^2\theta+f^2\sin^2\theta)^{3/2}}. \tag{B14}$$

The average values in eq (B5) to (B9) can now be found. For example

$$\overline{\sin^4\theta}=\int_0^\pi \frac{f^2\sin^5\theta}{4\pi(\cos^2\theta+f^2\sin^2\theta)^{3/2}}d\theta\int_0^{2\pi}d\phi. \tag{B15}$$

This reduces to

$$\overline{\sin^4\theta}=L(f)=(1+\tfrac{1}{2}f^2)(1-f^2)^{-2}-2f^2(1-\tfrac{1}{4}f^2)(1-f^2)^{-5/2}\log[\{1+(1-f^2)^{1/2}\}/f] \tag{B16}$$

Similarly it can be shown that

$$\overline{\cos^2\theta\sin^2\theta}=M(f)=-\frac{3}{2}f^2(1-f^2)^{-2}+f^2\left(1+\frac{1}{2}f^2\right)(1-f^2)^{-5/2}\log[\{1+(1-f^2)^{1/2}\}/f] \tag{B17}$$

$$\overline{\cos^2\theta}=P(f)=1-L(f)-M(f) \tag{B18}$$

$$\overline{\sin^2\theta}=Q(f)=L(f)+M(f) \tag{B19}$$

In particular: $L(0)=1$; $L(1)=8/15$; $M(0)=0$; $M(1)=2/15$.
Figure 3 gives curves showing how $L(f)$, $M(f)$ depend upon f.

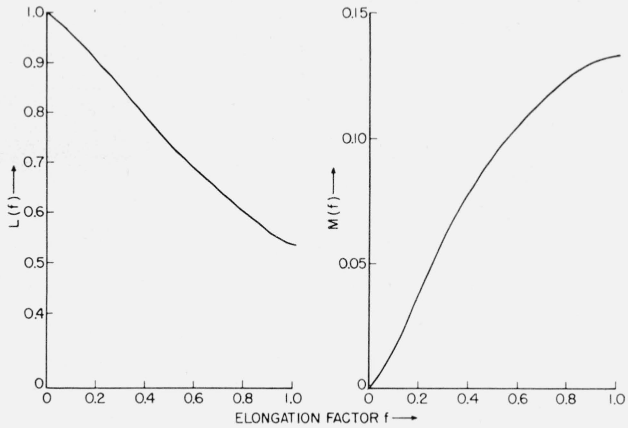

FIGURE 3. *Curves showing how the functions $L(f)$ and $M(f)$ depend on the elongation factor f.*

15. References

[1] E. V. Appleton, Wireless studies of the ionosphere, J. Inst. Elec. Engrs. (London) **71,** 642 (1932).

[2] B. K. Banerjea, On the propagation of electromagnetic waves through the atmosphere, Proc. Roy. Soc. (London) [**A**] **190,** 67 (1947).

[3] A. C. G. Mitchell and M. W. Zemansky, Resonance radiation and excited atoms, Cambridge Univ. Press (1934).

[4] M. Spencer, The shape of irregularities in the upper atmosphere, Proc. Phys. Soc. (London) [**B**] **68,** 493 (1955).

[5] L. R. C. Storey, An investigation of whistling atmospherics, Phil. Trans. Roy. Soc. London Ser. [**A**] **246,** 113 (1953).

BOULDER, COLO. (Paper 63D2–14)