

A Method for the Determination of Lower Ionosphere Properties by Means of Field Measurements on Spherics¹

F. B. Harris, Jr., and R. L. Tanner

Contribution from Stanford Research Institute, Menlo Park, Calif.

(Received November 24, 1961; revised February 1, 1962)

The propagation of audiofrequency and sub-audiofrequency waves between the earth and an ionosphere whose conductivity varies continuously with altitude is considered in detail. The fields are represented in terms of two scalar potentials satisfying appropriate wave equations in spherical coordinates.

It is shown, on the basis of existing data on the ionosphere, that waves in this frequency range can be considered to be confined to a thin, but not sharply bounded, spherical shell about the earth. Greatly simplified radial wave equations in dimensionless form are derived incorporating this approximation. Solutions of these equations are given for two regions, viz, for the low-altitude region where $\sigma/\omega\epsilon_0 \ll 1$ and, in the case of certain restricted types of conductivity profile, for the high-altitude region where $\sigma/\omega\epsilon_0 \gg 1$. An iterative method is presented, based on an integral equation, which makes possible a computation of the radial wave function in the transition region and a joining of interior and exterior solutions for the propagating TM mode. The result is a direct mathematical relationship between the conductivity profile and the complex propagation constant as a function of frequency.

It is demonstrated that at frequencies above about 50 cycles the propagation constant can be obtained from measurements of the horizontal components of electric and magnetic fields in individual spherics at airplane altitudes, while at lower frequencies the same information can be obtained through ground-based observation of cavity resonance effects in spheric noise. Existing data on these cavity resonances are used to calculate the complex propagation constant as a function of frequency from 6 to 34 cycles. The results, when extrapolated to higher frequencies, predict attenuation rates in excellent agreement with currently available data.

The effect of the diurnal variation in conductivity on observable quantities is briefly examined, and tentative conclusions as to its magnitude are drawn.

1. Introduction

The propagation characteristics of ELF and sub-audiofrequency radio waves in the earth-ionosphere cavity are critically dependent upon the detailed electrical properties of the ionosphere. Most previous theoretical treatments [Budden, 1957; Wait, 1957; Wait, 1958; Wait, 1960 a and b] of the problem have assumed an ionosphere sharply bounded at some definite height, with a conductivity which, from that height outward, either remained constant or varied stepwise or continuously with height. For a number of reasons, such models can never adequately represent the actual physical situation. The same observation has been made recently by Wait [1962] in a paper treating the case of a continuously varying conductivity which becomes constant at high altitude. The work presented below, while confining itself substantially to the ELF region, is an attempt to approach a somewhat more general problem from a different point of view. In particular, this paper is directed toward the possibility of inferring the properties of the lower ionosphere from field measurements taken over a range of frequencies.

For frequencies low compared to the electronic collision frequency—a condition easily satisfied in the VLF range and below—the ionosphere behaves like a medium with unit dielectric constant, unit relative permeability, and a conductivity which is a function of position. The conductivity affects the local behavior of the electromagnetic field only as it enters into the complex relative permittivity ϵ , which can be represented as follows:

¹ This work was sponsored in part by the Geophysical Research Directorate, Air Force Cambridge Research Laboratories, Bedford, Mass.

$$\epsilon = 1 - j \frac{\omega_r}{\omega} \quad (1)$$

where ω_r is a real number, related to the conductivity, which depends upon position but not frequency. At altitudes where (ω_r/ω) is much smaller than unity, the medium behaves essentially like a vacuum. Where this ratio is much larger than unity, it behaves like a good conductor. Although (1) makes no provision for the effect of the terrestrial magnetic field, it has been shown [Wait, 1960 b] that this can be quite simply taken into account at low frequencies through the use of an effective ω_r , dependent on the vertical component of magnetic field.

The transition region, in which (ω_r/ω) is comparable with or at most a few orders of magnitude greater than unity, can play as significant a role in determining the propagation of ELF waves as does the upper region. It is precisely this region which is omitted from sharply bounded models. Moreover, it can be seen from the form of (1) that the height at which the effect of finite conductivity is first felt is progressively lower for waves of lower frequency. On the other hand, because of a phenomenon analogous to the familiar skin effect, the total penetration of the waves into the ionosphere increases with decreasing frequency. When viewed in this light, the concept of a single "ionospheric height," even if this height is permitted to vary with frequency, clearly loses its physical meaning and probably its mathematical utility. Certainly some effort to take into account a continuous variation of ω_r with height is warranted.

Any theory which attempts to relate the conductivity profile of the ionosphere to observed fields must take into account both the real and the imaginary parts of the effective propagation constant as functions of frequency. Unfortunately experimental data [Chapman and Macario, 1956; Jean, 1961] to date, with few exceptions [Balsler and Wagner, 1960a and b], are capable of giving information only about the attenuation characteristics of the wave and yield nothing about its phase velocity. Future experiments on ELF and sub-audiofrequency radiation, if they are to lead to increased knowledge of ionospheric properties, must provide for measurement of both attenuation and phase lag.

2. Scalar Solutions for the Spherically Symmetrical Ionosphere

In the oversimplified, but important, case in which the conductivity depends only on the radius of spherical coordinates, solutions of Maxwell's equations can be found in the form of TM waves (r -component of \vec{H} absent) and TE waves (r -component of \vec{E} absent). Each type of solution can be derived from a single scalar potential. For the TM wave, the fields are given by

$$\vec{H} = \hat{r} \times \nabla \psi \quad (2)$$

$$\vec{E} = \frac{\nabla \times \vec{H}}{j\omega\epsilon_0\epsilon} \quad (3)$$

where ψ is a scalar solution of

$$r^2 \nabla \cdot \left(\frac{1}{r^2 \epsilon} \nabla \psi \right) + k^2 \psi = 0. \quad (4)$$

The symbol k is here used to represent the free-space propagation constant of a wave of the given frequency. For the TE wave, one obtains

$$\vec{E} = \hat{r} \times \nabla \chi \quad (5)$$

$$\vec{H} = -\frac{\nabla \times \vec{E}}{j\omega\mu_0} \quad (6)$$

where χ is a scalar solution of

$$\frac{r^2}{\epsilon} \nabla \cdot \left(\frac{1}{r^2} \nabla \chi \right) + k^2 \chi = 0. \quad (7)$$

Both (4) and (7) are separable in spherical coordinates, giving solutions of the form

$$\psi = \sum_{\nu} \sum_m a_{\nu}^m \Phi_m(\phi) \Theta_{\nu}^m(\theta) R_{\nu}(r) \quad (8)$$

$$\chi = \sum_{\mu} \sum_m b_{\mu}^m \Phi_m(\phi) \Theta_{\mu}^m(\theta) S_{\mu}(r). \quad (9)$$

The functions $\Phi_m(\phi)$ are linear combinations of $\exp(\pm jm\phi)$ where m is zero or an integer, and the $\Theta_{\nu \text{ or } \mu}^m(\theta)$ are appropriate Legendre functions. The radial parts of the solutions satisfy the following equations:

$$\epsilon \frac{d}{dr} \left(\frac{1}{\epsilon} \frac{dR_{\nu}}{dr} \right) + \left[k^2 \epsilon - \frac{\nu(\nu+1)}{r^2} \right] R_{\nu} = 0 \quad (10)$$

$$\frac{d^2 S_{\mu}}{dr^2} + \left[k^2 \epsilon - \frac{\mu(\mu+1)}{r^2} \right] S_{\mu} = 0. \quad (11)$$

The permissible values, or eigenvalues, ν and μ for a given frequency are determined by two boundary conditions. These are, in general, that the fields remain finite at the center of the earth and vanish at infinity. In the case of a perfectly conducting earth, the first condition is replaced by

$$\left(\frac{dR_{\nu}}{dr} \right)_{r=r_0} = 0 \quad (12)$$

$$S_{\mu}(r_0) = 0 \quad (13)$$

where r_0 is the radius of the earth. It should be noted that the eigenvalues for a particular frequency are uniquely determined by the radial equation alone, and are in general complex. The Legendre functions, which can be written $P_{\nu \text{ or } \mu}^m(\pm \cos \theta)$, will therefore in general possess at least one singularity (at $\theta = \pi$ if the + sign is used, or $\theta = 0$ if the - sign is used), corresponding to the fact that a source is needed to drive the system at an arbitrary frequency. There will, however, be certain (complex) frequencies at which one of the eigenvalues becomes equal to an integer n . Since the $P_n^m(\cos \theta)$ have no singularities, these are the natural resonant frequencies of the system, at which oscillation is possible in the absence of a source.

3. Altitude Dependence of the Relative Permittivity

The parameter ω_r depends on the electron density N and the collision frequency ν_c as follows:

$$\omega_r = \frac{e^2}{\epsilon_0 m} \left(\frac{N}{\nu_c} \right) \quad (14)$$

where e and m are the charge and mass of the electron. Existing data [Pierce, 1960; Ratcliffe, 1959] on N and ν_c have been used to construct the curve of figure 1, which shows the probable approximate variation of ω_r with height above the surface of the earth during daylight hours. It is seen that the function, although continuous, is very rapidly varying and reaches extremely high values in the first 100 km. The increase continues thereafter, rather more slowly, until a broad peak at an altitude of 300 or 400 km is reached, where ω_r is several orders of magnitude larger than at 100 km. A relatively slow decrease then sets in, which results in a significant reduction of ω_r only after an altitude of several thousand kilometers is reached.

If the conductivity, given by

$$\sigma = \epsilon_0 \omega_r \quad (15)$$

were even to remain constant from 100 km upward for several hundred kilometers, the skin

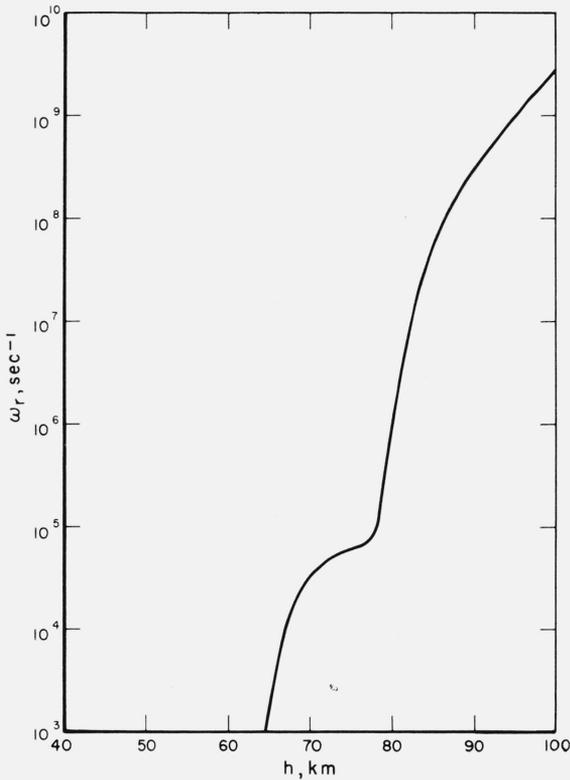


FIGURE 1. *Approximate daytime ionosphere profile.*

depth in that region for ELF radiation would be very small. At the very low frequency of 10 cycles, the skin depth would be only of the order of 1 km. Therefore, it would seem safe to assume that the wave is attenuated essentially to zero within a vertical distance which is small compared with the radius of the earth. Although figure 1 may be in error by an order of magnitude or more at a given altitude, and although the nighttime profile is somewhat different, certainly one can still be confident that all significant phenomena are confined to a shell not more than 100 to 150 km in thickness.

The argument just outlined, together with the fact that the attenuation is also large in the earth itself, leads to an important mathematical simplification in the radial wave equations, which may now be written

$$\epsilon \frac{d}{dr} \left(\frac{1}{\epsilon} \frac{dR_\nu}{dr} \right) + \left[k^2 \epsilon - \frac{\nu(\nu+1)}{\bar{r}^2} \right] R_\nu = 0 \quad (16)$$

$$\frac{d^2 S_\mu}{dr^2} + \left[k^2 \epsilon - \frac{\mu(\mu+1)}{\bar{r}^2} \right] S_\mu = 0. \quad (17)$$

The variable r in the final terms has been replaced by a constant effective radius \bar{r} , which can be chosen to be somewhat larger than r_0 . The form of the relevant solutions will be practically unaffected by this change, even if the independent variable is allowed to go to infinity in satisfying the outer boundary condition.

The simplified forms of the radial equations make it easy to attach a direct physical meaning to the eigenvalues. Equations (16) and (17) are identical to those which are obtained for the potentials in the case of a wave propagating between infinite plane-parallel walls having the same electrical characteristics as those of the earth and the ionosphere, providing only that the following replacements are made:

$$\frac{\nu(\nu+1)}{\bar{r}^2} = k_{\text{TM}}^2 \quad (18)$$

$$\frac{\mu(\mu+1)}{r^2} = k_{\text{TE}}^2 \quad (19)$$

where k_{TM} and k_{TE} are the propagation constants of TM and TE waves, respectively, in a direction parallel to the walls.

4. Dimensionless Form of the Radial Equations

For simplicity of notation, it is convenient to rewrite (18) and (19) in dimensionless form, as follows:

$$\epsilon \frac{d}{dx} \left(\frac{1}{\epsilon} \frac{du_\alpha}{dx} \right) + (\epsilon - \alpha) u_\alpha = 0 \quad (20)$$

$$\frac{d^2 v_\beta}{dx^2} + (\epsilon - \beta) v_\beta = 0 \quad (21)$$

where

$$x = k(r - r_0) \quad (22)$$

$$\alpha = \frac{\nu(\nu+1)}{k^2 r^2} \quad (23)$$

$$\beta = \frac{\mu(\mu+1)}{k^2 r^2} \quad (24)$$

$$u_\alpha(x) = R_\nu(r) \quad (25)$$

$$v_\beta(x) = S_\mu(r). \quad (26)$$

In order to see approximately the nature of the solutions and the order of magnitude of the eigenvalues, it is interesting to consider the case of perfectly conducting earth and ionosphere, separated by a vacuum of thickness h . The solutions would then be of the form

$$u_\alpha = \cos [\sqrt{1 - \alpha} x] \quad (27)$$

$$v_\beta = \sin [\sqrt{1 - \beta} x]. \quad (28)$$

The boundary conditions at $x = kh$ require that

$$\sqrt{1 - \alpha} = \frac{p\pi}{kh} \quad (p = 0, 1, 2, 3 \dots) \quad (29)$$

$$\sqrt{1 - \beta} = \frac{q\pi}{kh} \quad (q = 1, 2, 3 \dots). \quad (30)$$

Since, at frequencies of the order of those under discussion, kh is much smaller than unity, all TE eigenvalues and all but the lowest TM eigenvalue are large and negative. Using (18), (19), (23), and (24), we have for the equivalent parallel-plane propagation constants

$$k_{\text{TM}} \approx -j \frac{p\pi}{h} \quad (31)$$

$$k_{\text{TE}} \approx -j \frac{q\pi}{h}. \quad (32)$$

These modes are all far below cutoff, and are attenuated in horizontal distances small compared with the vertical height of the ionosphere. The lowest TM mode, however, corresponds to

$$\alpha_0 = 1 \quad (33)$$

$$k_{\text{TM}0}=k. \quad (34)$$

The introduction of an earth and ionosphere of reasonably high conductivity will not affect the eigenvalues drastically; α_0 will still be of the order of unity, although it will now be complex, and the remaining eigenvalues will be large and primarily real and negative. The lowest TM mode will still be the only one capable of propagation over significant distances.

The following orthogonality relations are of interest in connection with the application of the radial wave functions to actual situations:

$$\int_{-\infty}^{\infty} \frac{1}{\epsilon} u_{\alpha} u_{\alpha'} dx = 0 \quad (\alpha' \neq \alpha) \quad (35)$$

$$\int_{-\infty}^{\infty} v_{\beta} v_{\beta'} dx = 0 \quad (\beta' \neq \beta). \quad (36)$$

The infinite limits represent integrations into fictitious "flattened" spaces above and below the surface of the earth. This artifice is permissible because of the high conductivity of both bounding media.

An exact solution of either of the radial wave equations, or even an approximate solution by any conventional method, is extremely difficult in any practical case. It is possible, however, to obtain a solution valid below the transition region and, in special cases, a solution valid above that region. In the case of a perfectly conducting earth and finitely conducting ionosphere, the interior solutions are given by (27) and (28), as long as $(\omega_r/\omega) \ll 1$. If the effect of the earth is represented by a wave impedance Z_e —this will be small compared to η_0 —(27) and (28) become

$$u_{\alpha} = \cos \left[\sqrt{1-\alpha} x - j \frac{Z_e}{\eta_0 \sqrt{1-\alpha}} \right] \quad (37)$$

$$v_{\beta} = \sin \left[\sqrt{1-\beta} x - j \frac{Z_e \sqrt{1-\beta}}{\eta_0} \right]. \quad (38)$$

The exterior solutions which can be obtained apply only where (ω_r/ω) is large. One soluble case is that for which ω_r is given by a simple power law. If

$$\epsilon = 1 - j\epsilon_1(x-x_1)^n \approx -j\epsilon_1(x-x_1)^n \quad (39)$$

where n can have any real positive value, the solutions are

$$u_{\alpha} = A(x-x_1)^{1/2} \sqrt{\epsilon} H_{\frac{n+1}{n+2}}^{(2)} \left[\frac{2}{n+2} \sqrt{\epsilon}(x-x_1) \right] \quad (40)$$

$$v_{\beta} = B(x-x_1)^{1/2} H_{\frac{1}{n+2}}^{(2)} \left[\frac{2}{n+2} \sqrt{\epsilon}(x-x_1) \right]. \quad (41)$$

The above expressions are restricted to the region where $(\omega_r/\omega) \gg |\alpha|$. Another soluble case, which would appear to represent more nearly the true state of affairs, is the exponential variation of conductivity:

$$\epsilon = 1 - j\epsilon_1 e^{ax} \approx -j\epsilon_1 e^{ax}. \quad (42)$$

The corresponding exterior solutions are

$$u_{\alpha} = A \sqrt{\epsilon} H_{\sqrt{1+(4/a^2)(\alpha-1)}}^{(2)} \left(\frac{2}{a} \sqrt{\epsilon} \right) \quad (43)$$

$$v_{\beta} = B H_{(2/a)\sqrt{\beta-1}}^{(2)} \left(\frac{2}{a} \sqrt{\epsilon} \right) \quad (44)$$

where the restriction is that $(\omega_r/\omega) \gg 1$. It should be noted that for the propagating mode $(4/a^2)(\alpha-1) \ll 1$ for the real ionosphere, and that therefore the order of the Hankel function becomes essentially unity for that mode.

For many purposes, only the propagating mode is of interest. It happens that, in this exceptional case, a good method exists for joining the interior and exterior solutions, and calculating the eigenvalue. Because the permittivity varies very greatly over distances short compared with a wavelength, no WKB or quasi-WKB method is applicable in the transition region. However, the rapidity of this variation ensures that u_{α_0} will not change greatly from its value at $x=0$ until after the region of validity of the exterior solution is reached. For this reason, a good approximation to the solution in the interior and transition regions can be obtained by transforming (20) into an integral equation. If we let

$$u_{\alpha}(0)=1 \quad (45)$$

then

$$\left(\frac{1}{\epsilon} \frac{du_{\alpha}}{dx}\right)_{x=0} = j \frac{Z_e}{\eta_0} \quad (46)$$

NOTE: Here and in most of what follows, α will be used to denote the lowest eigenvalue only.

The first integration of (20) yields

$$\frac{1}{\epsilon(x)} \frac{du_{\alpha}}{dx} = j \frac{Z_e}{\eta_0} - \int_0^x u_{\alpha}(x') dx' + \alpha \int_0^x \frac{u_{\alpha}(x')}{\epsilon(x')} dx' \quad (47)$$

and a second integration leads to

$$u_{\alpha}(x) = 1 + j \frac{Z_e}{\eta_0} \int_0^x \epsilon(x') dx' - \int_0^x \epsilon(x') \int_0^{x'} u_{\alpha}(x'') dx'' dx' + \alpha \int_0^x \epsilon(x') \int_0^{x'} \frac{u_{\alpha}(x'')}{\epsilon(x'')} dx'' dx'. \quad (48)$$

If x is not too large, an iterative process starting with

$$u_{\alpha}(x'') = 1 \quad (49)$$

will converge to give as accurate a value of $u_{\alpha}(x)$ as may be desired. The eigenvalue may then be calculated by setting the value of $(1/\epsilon u_{\alpha})(du_{\alpha}/dx)$ obtained by means of (48) equal to that obtained from (40) or (43) at some height x_0 , chosen as small as consistent with the validity of the exterior solution.

At sufficiently low frequencies, and of course depending upon the actual shape of the conductivity profile, the first-order approximation in which $u_{\alpha}(x)$ is simply unity can be used, and (47) gives

$$\frac{1}{\epsilon} \frac{du_{\alpha}}{dx} = j \frac{Z_e}{\eta_0} - x + \alpha \int_0^x \frac{dx'}{\epsilon(x')}. \quad (50)$$

Under these circumstances, one obtains

$$\alpha = \frac{x_0 - j \frac{Z_e}{\eta_0} + \left(\frac{1}{\epsilon u_{\alpha}} \frac{du_{\alpha}}{dx}\right)_{x=x_0}}{\int_0^{x_0} \frac{dx}{\epsilon}} \quad (51)$$

where the last term in the numerator is to be calculated from the exterior solution. The term in Z_e can usually be neglected at these frequencies. Moreover, the integral in the denominator is of the order of x_0 , and so a more convenient form of (51) is

$$\alpha = \frac{x_0 + \left(\frac{1}{\epsilon u_{\alpha}} \frac{du_{\alpha}}{dx}\right)_{x=x_0}}{x_0 - \int_0^{x_0} \left(1 - \frac{1}{\epsilon}\right) dx} \quad (52)$$

It should be remembered that (52) represents only a first-order solution, but that the range of applicability of the technique can be greatly extended by repeated iterations of (48).

Equation (52) gives a means for calculating the lowest eigenvalue if the conductivity profile is known. Actually, however, the problem at hand is exactly the reverse—that of calculating the profile if the eigenvalue is known as a function of frequency. The latter information may be obtained experimentally in a number of ways, two of which are described in the following sections. If a reasonable approximation for the upper ionosphere profile is known, (52) permits the calculation of the integral in the denominator as a function of frequency. This information is, in a general sense, a transform of the desired information on the variation of ϵ with height.

The desired information must be obtained by means of an inversion of this transform. Although the transform in question is not among those which have been well investigated in the past, several avenues of approach are open. It is conceivable, of course, that a rigorous analytical method for performing the inversion may be found. It is more probable, however, that recourse will have to be had to some approximate analytical technique. If all else fails, the transform can always be inverted by purely numerical methods.

5. Field of a Vertical Dipole

The atmospheric noise observed at ELF and below is due primarily to the vertical component of current in lightning strokes occurring at relatively large distances from the observer, and can therefore be described in terms of the field of an infinitesimal vertical dipole located just above the surface of the earth. This field is composed of TM waves of all orders with the azimuthal index m equal to zero. If the dipole is located at $\theta=0$, $x=0$ we have

$$\psi(x, \theta) = \sum_{\nu} a_{\nu} P_{\nu}(-\cos \theta) u_{\alpha}(x). \quad (53)$$

It is shown in the appendix that

$$a_{\nu} \sin \nu\pi = \frac{kp}{4 \int_0^{\infty} \frac{1}{\epsilon} u_{\alpha}^2 dx} \quad (54)$$

where p represents the current moment of the source. Except in the immediate vicinity of the source, only the lowest-order wave is significant, and therefore

$$\psi(x, \theta) = \frac{kp}{4 \int_0^{\infty} \frac{u_{\alpha}^2}{\epsilon} dx} \left[\frac{P_{\nu}(-\cos \theta)}{\sin \nu\pi} \right] u_{\alpha}(x) \quad (55)$$

where α denotes the lowest eigenvalue. When $m=0$, as in this case, the only field components present are the following:

$$E_r = j\eta_0 \frac{k\alpha}{\epsilon} \psi \quad (56)$$

$$E_{\theta} = j\eta_0 \frac{1}{r\epsilon} \frac{\partial^2 \psi}{\partial x \partial \theta} \quad (57)$$

$$H_{\phi} = \frac{1}{r} \frac{\partial \psi}{\partial \theta} \quad (58)$$

The eigenvalue can, in principle, be determined by comparative measurements of these three field components in individual sferics. Actually, insofar as the determination of ionospheric properties is concerned, it is the quantity $(\alpha-1)$ which is significant. A comparison

of E_r and H_ϕ is not suitable for several reasons. First, and most important, is the fact that $(\alpha-1)$ enters into the ratio only in second order. Second is the fact that the ratio depends to some extent on θ —that is, on the distance of the observation point from the source. Third is the fact that the horizontal magnetic field measured by a loop antenna will depend upon the direction of incidence of the wave, while the vertical electric field will not. Consequently, two loop antennas will be required to determine the actual magnitude of H_ϕ .

Comparative measurements of E_θ and E_r would yield $(\alpha-1)$ in first order, although the experiment would have to be carried out at an altitude sufficiently great that E_θ were indeed measurable. However, knowledge of the source distance and direction of incidence would still be required.

The best procedure would be to measure and compare E_θ and H_ϕ . The complex ratio of these two field components is entirely independent of the source distance, so long as this distance is great enough for the evanescent modes to be negligible. Moreover, the responses of two suitably oriented antennas will both depend in exactly the same way on the direction of incidence, so that the response ratio will be independent of the location of the source. An additional important advantage is that the ratio is entirely unaffected by the presence of scattering centers such as discontinuities in the ionosphere (again provided the higher-order modes in the reflections can be neglected). The field-component ratio is given by

$$\frac{E_\theta}{\eta_0 H_\phi} = j \frac{1}{\epsilon u_\alpha} \frac{du_\alpha}{dx} \quad (59)$$

If the measurements are made in an airplane at dimensionless height x_a , (37) is applicable, the first-order approximation to $\tan \sqrt{1-\alpha}x$ is valid, and ϵ can be taken as unity, so that one has (with $Z_e \approx 0$)

$$\frac{E_\theta}{\eta_0 H_\phi} = j x_a (\alpha - 1). \quad (60)$$

It should be emphasized that the values of α obtained in this way are those corresponding to the instantaneous conductivity profile above the point of observation. This method is therefore potentially capable of detecting both temporal and geographical variations in the ionosphere. In regions where the magnetic field of the earth is steeply dipping, the results obtained are simply those corresponding to the profile of $(\omega_r)_{\text{eff}}$ [Wait, 1960b], rather than ω_r .

6. Representation of the Field in Terms of Cavity Resonances

Since $(\alpha-1)$ is a quite slowly varying quantity, the field component ratio shown in (60) depends primarily on x_a or, for a given altitude in feet, on the first power of frequency. The same is approximately true of the ratio E_θ/E_r . It is this latter ratio which limits the measurability of E_θ , due to the effect of errors in aircraft attitude. Assuming an altitude of 40,000 feet and an accuracy of 0.1 degree in the determination of aircraft attitude, the probable error in E_θ reaches 10 percent at about 50 cycles. Fortunately, the range below this frequency can be effectively covered through studies of the cavity resonances. A great deal can be inferred from ground-level measurements of the vertical electric field as a function of frequency in this range. However, such data can yield information on ionospheric properties only as averaged over the earth. More complete information could be obtained from simultaneous measurement of the other field components, but the following analysis is based on the vertical field alone, as these are the only data [Balser and Wagner, 1960a and b] presently available.

The theoretical basis for this method is the following expansion:

$$\frac{P_r(-\cos \theta)}{\sin \nu \pi} = \frac{1}{\pi} \sum_{n=0}^{\infty} \frac{(2n+1)P_n(\cos \theta)}{\nu(\nu+1) - n(n+1)} \quad (61)$$

The vertical component of electric field at the surface of the earth, due to a single source, can be found from (45), (55), and (61). It is given by

$$E_r(0, \theta) = \frac{j\nu(\nu+1)}{\omega\epsilon_0\epsilon r_0^2} \psi(0, \theta) = C(\omega)\nu(\nu+1) \sum_{n=0}^{\infty} \frac{(2n+1)P_n(\cos\theta)}{\nu(\nu+1) - n(n+1)} \quad (62)$$

where

$$C(\omega) = \frac{jn_0 p}{4\pi r_0^2 \int_0^{\infty} \frac{u_\alpha^2}{\epsilon} dx} \quad (63)$$

The total vertical field observed at a point is the result of some distribution of such sources over the surface of the earth. It is instructive to examine the case of a uniform distribution of a large number of noncoherent sources. The mean-square value of electric field is given by

$$|E|^2 = \frac{N_p}{4\pi} \int_0^{2\pi} \int_0^\pi E_r E_r^* \sin\theta d\theta d\phi \quad (64)$$

where N_p is the total number of sources. Using (62) and the orthogonality properties of the Legendre polynomials, one obtains

$$|E|^2 = N_p |C(\omega)|^2 \sum_{n=0}^{\infty} \frac{(2n+1)|\nu(\nu+1)|^2}{|\nu(\nu+1) - n(n+1)|^2} \quad (65)$$

For the rapidly varying ionosphere, $C(\omega)$ is roughly inversely proportional to ω .

7. Interpretation of Observed Atmospheric Noise Spectrum

It was pointed out in the preceding section that airborne measurements would lose accuracy at lower frequencies and would cease to be valid below some minimum frequency. In order to determine that frequency, it was necessary to find the approximate relative magnitudes of the horizontal and vertical electric fields to be expected at flight altitudes. The noise spectrum observed by Balsler and Wagner in the range from 5 to 34 cycles, which shows a series of pronounced peaks and valleys, has therefore been interpreted according to the scheme described below. The figure of 50 cycles as the lowest useful frequency for airborne measurements is a result of that analysis.

The development leading to (65) applies to a distribution of noncoherent sources at a single frequency. Actually, the noise energy is distributed throughout the frequency spectrum. The power per unit frequency interval received by an electrically short vertical monopole will be given by

$$\frac{dP}{d\omega} = \frac{F(\omega)}{\omega^2} \sum_{n=0}^{\infty} \frac{(2n+1)(a^2 + b^2)}{[a - n(n+1)]^2 + b^2} \quad (66)$$

where a and b are the real and imaginary parts of $\nu(\nu+1)$, respectively, and $F(\omega)$ is a real function very nearly proportional to the frequency spectrum of the square of the current moment, (p^2) , of the sources.

The problem of using (66) to find α as a function of frequency is not completely determinate, since there is an infinite number of ways to choose the frequency dependence of a , b , and F , all of which will yield the observed spectrum. However, it is inconceivable that a and b could be anything but smooth, monotonic functions of ω , and unlikely that $F(\omega)$, which includes geographical and time averaging of the sources, should be other than smooth and monotonic. The problem thus becomes that of choosing smooth forms for a and b such that the form of F then required to reproduce the spectrum is also smooth.

The first step in the procedure is to plot $\omega^2(dP/d\omega)$ as a function of frequency from the experimental data. The peaks in this function occur because of the resonances of successive terms in the series of (66), and will be located approximately at those frequencies where a is equal to $n(n+1)$. This fact is used to plot several points in the (a, ω) plane. A smooth curve is drawn through these points, and is represented as nearly as possible by a power law, or a combination of power laws. In this way, an initial approximation to $a(\omega)$ is found. Next,

it is assumed that both b and F are constant in the immediate neighborhood of each resonant peak. This permits an approximate determination of b , by means of (66), from the shape of each peak. The points thus found provide a first approximation to $b(\omega)$, which is also represented as a power law. Equation (66) is then used to calculate the initial approximation to $F(\omega)$ over the entire frequency range, which in general shows some irregularities.

From this stage onward, the process is one of "cut and try" to find the smoothest possible $F(\omega)$. It does not appear to be possible, with these data, to obtain a perfectly smooth $F(\omega)$ using this model. There is a tendency for small apparent peaks of $F(\omega)$ to occur near resonances of even n , and small valleys near resonances of odd n . This should not be too surprising, however, as the time-averaged distribution of lightning strokes is not really uniform, but rather is more highly concentrated in a belt centered on the geographical equator. Such a distribution would be expected to couple more strongly to the even-numbered cavity modes. Furthermore, the fixed latitude of the observation point relative to this belt of excitation will tend to favor some modes over others. Still another possible source of discrepancies in the analysis is the fact that no account has been taken of the diurnal variation of ϵ . However, as will be pointed out in the following section, the diurnal variation probably does not seriously affect the validity of the analysis.

The expressions obtained for a and b represent, to first order, an average of day and night conditions. The more laborious part of the analysis has not yet been carried to the point where the results can be presented with complete confidence. However, a reasonably good fit of the data was obtained with the following expressions:

$$\left\{ \begin{array}{l} a=11.4 \left(\frac{f}{20}\right)^{1.9} \quad (f \leq 20 \sim) \\ a=11.4 \left(\frac{f}{20}\right)^{2.0} \quad (f \geq 20 \sim) \end{array} \right\}$$

$$b = -2.22 \left(\frac{f}{20}\right)^{1.7} \quad (67)$$

The expression for a is believed to be accurate within a very few percent, while that for b may be in error by as much as 10 percent, over the range 6 to 34 cycles.

The corresponding source function $F(\omega)$ can be represented very well by

$$F = \text{const.} \times \left\{ 1 + 8.38 \exp \left[-3.15 \left(\frac{f}{20}\right) \right] \right\} \quad (68)$$

This function is shown in figure 2, and the theoretical and experimental noise spectra are displayed in figure 3.

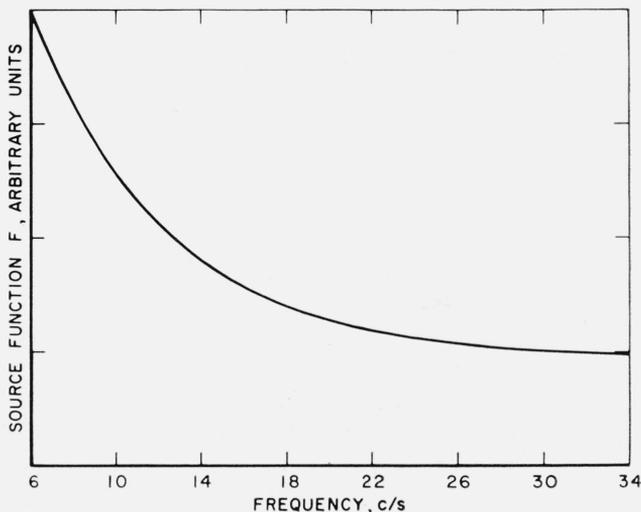


FIGURE 2. Source function F versus frequency.

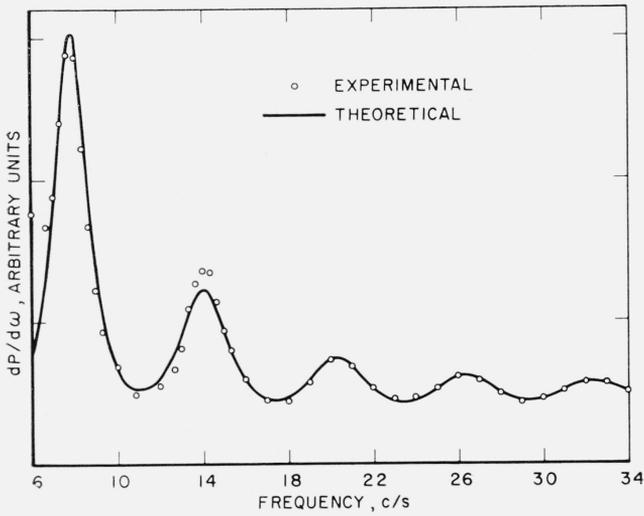
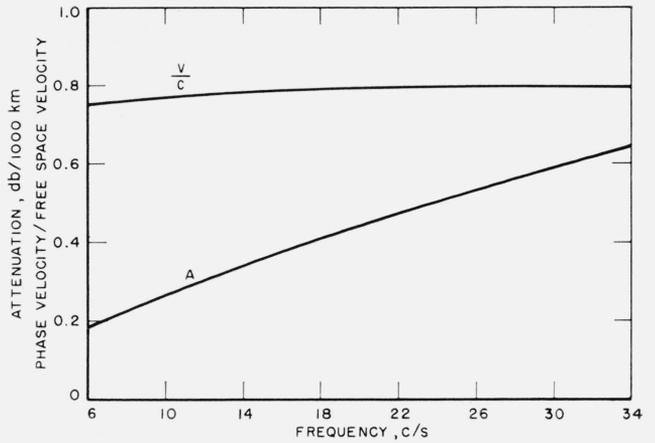


FIGURE 3. *Theoretical and experimental noise spectra.*

FIGURE 4. *Phase velocity and attenuation versus frequency.*



Because $b^2 \ll a^2$, it is possible to write for the equivalent plane-parallel propagation constant

$$k_{TM} = \frac{1}{r} \left(\sqrt{a} + j \frac{b}{2\sqrt{a}} \right). \quad (69)$$

From this, one has

$$A = 0.678 \left(\frac{-b}{\sqrt{a}} \right) \quad \text{db/1000 km} \quad (70)$$

$$\frac{v}{c} = \frac{2.69}{\sqrt{a}} \left(\frac{f}{20} \right) \quad (71)$$

where A is the attenuation constant and v is the phase velocity. The values of these quantities obtained with (67) are shown as functions of frequency in figure 4. Although there is no reason to suppose that the empirical equations (67) should hold outside the range of the original data, it is nevertheless interesting to extrapolate them to higher frequencies. The extrapolated attenuation at $100\sim$ is 1.375 db/1000 km, a value which is intermediate between existing experimental values [Jean, 1961] for day and night at that frequency. The predicted phase velocity remains nearly constant over the whole range, and at $100\sim$ is equal to 0.797 times the velocity of light.

8. Effect of the Diurnal Variation

It is well known that the properties of the ionosphere during daylight hours differ quite markedly from those at night. The question arises as to how the observed cavity resonances are affected by this variation. In the absence of any diurnal variation, each cavity mode is associated with a pair of integer indices m and n , and corresponds to a particular complex resonant frequency s_{mn} (where $s = \xi + j\omega$). The modes are frequency degenerate, as s_{mn} depends only on n . The introduction of an angular dependence of ϵ removes this degeneracy, splitting apart the resonant frequencies of the same n but different m . One would expect this to result in a certain amount of blurring of the observed resonant peaks with a consequent lowering of the apparent Q and raising of the apparent b .

To a first approximation, the diurnal variation can be represented by a discontinuity in ϵ across a plane which divides the spherical coordinate system in half at $\theta = \pi/2$. In each hemisphere, the electromagnetic field can be represented in terms of solutions of the form described in (2) through (7). The boundary conditions which must be satisfied along this plane are that $H_r, H_\theta, H_\phi, E_r, \epsilon E_\theta,$ and E_ϕ be continuous. When the expressions for the field components are combined with the partial differential equations satisfied by the potentials immediately to each side of the discontinuity, and account is taken of the fact that in the neighborhood of $\theta = \pi/2$ the sine of θ is both stationary and equal to unity, we are led to the following system of four functions, the continuity of which is sufficient to ensure that the boundary conditions are satisfied:

$$\eta_0 \frac{\partial \psi}{\partial \theta} - j \frac{\partial^2 \chi}{\partial x \partial \phi} \quad (72)$$

$$\frac{\partial \chi}{\partial \theta} + jr_0 \frac{1}{\epsilon} \frac{\partial^2 \psi}{\partial x \partial \phi} \quad (73)$$

$$\frac{\partial^2 \chi}{\partial x^2} + \epsilon \chi \quad (74)$$

$$\frac{\partial}{\partial x} \left(\frac{1}{\epsilon} \frac{\partial \psi}{\partial x} \right) + \psi. \quad (75)$$

The application of the above conditions to the problem of a vertical dipole radiating at an arbitrary point in one of the hemispheres is extremely complex. The configuration is as shown in figure 5, where the source is located at A , a point not symmetrically placed relative to the discontinuity. The total field consists of (1) a primary TM wave set up by the source in its own hemisphere and describable by means of a potential ψ_0 , (2) reflected waves in the same hemisphere, and (3) transmitted waves in the other hemisphere.

The primary TM potential ψ_0 is independent of azimuth in the coordinate system x, θ', ϕ' , which has OA as its polar axis. Provided A is not too close to the discontinuity, (55) can be used to compute this potential, which can then be expressed in terms of x, θ, ϕ , coordinates

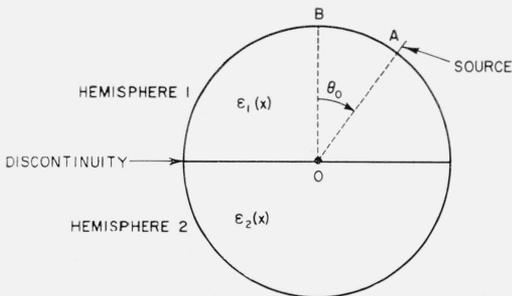


FIGURE 5. Configuration for study of diurnal effect.

centered on OB. The potential does depend upon azimuth in this coordinate system. The reflected and transmitted waves, in general, contain both TM and TE components. The potentials in the hemisphere containing the source are of the form

$$\psi_1 = \frac{kp}{4 \int_0^\infty (u_{1,\nu_0}^2/\epsilon_1) dx} \left[\frac{P_{\nu_0}(-\cos \theta_0 \cos \theta - \sin \theta_0 \sin \theta \cos \phi)}{\sin \nu_0 \pi} \right] u_{1,\nu_0}(x) + \sum_\nu \sum_m a_\nu^m \cos m\phi P_\nu^m(\cos \theta) u_{1,\nu}(x) \quad (76)$$

$$\chi_1 = \sum_\mu \sum_m b_\mu^m \sin m\phi P_\mu^m(\cos \theta) v_{1,\mu}(x) \quad (77)$$

while those in the other hemisphere are given by

$$\psi_2 = \sum_\kappa \sum_m c_\kappa^m \cos m\phi P_\kappa^m(-\cos \theta) u_{2,\kappa}(x) \quad (78)$$

$$\chi_2 = \sum_\lambda \sum_m d_\lambda^m \sin m\phi P_\lambda^m(-\cos \theta) v_{2,\lambda}(x). \quad (79)$$

The reflected wave has $+\cos \theta$ and the transmitted wave $-\cos \theta$, as no singularity must appear at either $\theta=0$ or $\theta=\pi$. The indices 1 and 2 refer to the two hemispheres; ν_0 is the lowest eigenvalue in Hemisphere 1; the ν and μ are eigenvalues in Hemisphere 1, and κ and λ are eigenvalues in Hemisphere 2; $(0, \theta_0, 0)$ are the coordinates of A in the system of OB.

The coefficients in (76) through (79) can, in principle, be evaluated by application of the boundary conditions, which yields a separate set of equations for each value of m :

$$j\eta_0 \sum_\nu a_\nu^m P_\nu^{m'}(0) u_{1,\nu}(x) - m \sum_\mu b_\mu^m P_\mu^m(0) v'_{1,\mu}(x) + j\eta_0 \sum_\kappa c_\kappa^m P_\kappa^{m'}(0) u_{2,\kappa}(x) + m \sum_\lambda d_\lambda^m P_\lambda^m(0) v'_{2,\lambda}(x) = D_m(x) \cos \theta_0 \int_0^{2\pi} P_{\nu_0}(-\sin \theta_0 \cos \phi) \cos m\phi d\phi \quad (80)$$

$$j m \frac{\eta_0}{\epsilon_1} \sum_\nu a_\nu^m P_\nu^m(0) u'_{1,\nu}(x) + \sum_\mu b_\mu^m P_\mu^{m'}(0) v_{1,\mu}(x) - j m \frac{\eta_0}{\epsilon_2} \sum_\kappa c_\kappa^m P_\kappa^m(0) u'_{2,\kappa}(x) + \sum_\lambda d_\lambda^m P_\lambda^{m'}(0) v_{2,\lambda}(x) = D'_m(x) \frac{\sin \theta_0}{\epsilon_1} \int_0^{2\pi} P'_{\nu_0}(-\sin \theta_0 \cos \phi) \sin \phi \sin m\phi d\phi \quad (81)$$

$$\sum_\mu \mu(\mu+1) b_\mu^m P_\mu^m(0) v_{1,\mu}(x) - \sum_\lambda \lambda(\lambda+1) d_\lambda^m P_\lambda^m(0) v_{2,\lambda}(x) = 0 \quad (82)$$

$$-j \frac{1}{\epsilon_1} \sum_\nu \nu(\nu+1) a_\nu^m P_\nu^m(0) u_{1,\nu}(x) + j \frac{1}{\epsilon_2} \sum_\kappa \kappa(\kappa+1) c_\kappa^m P_\kappa^m(0) u_{2,\kappa}(x) = D_m(x) \frac{\nu_0(\nu_0+1)}{\epsilon_1} \int_0^{2\pi} P_{\nu_0}(-\sin \theta_0 \cos \phi) \cos m\phi d\phi \quad (83)$$

where primes denote differentiation with respect to the argument of a function, and where

$$D_m(x) = \left(1 - \frac{1}{2} \delta_{m,0}\right) \left[\frac{\omega \mu_0 P}{4\pi \int_0^\infty (u_{1,\nu_0}^2/\epsilon_1) dx} \right] \left[\frac{u_{1,\nu_0}(x)}{\sin \nu_0 \pi} \right]. \quad (84)$$

Equations (80) through (83) have not yet been studied at great length. It can be seen, however, that severe difficulties exist. A sufficient number of linear equations for the coefficients can be obtained by utilizing the orthogonality properties of the u or the v , but all these equations still have an infinite number of terms.

The case in which no source is present and the two hemispheres are not very different has been briefly considered in an approximate way. The right-hand sides of all the equations are then zero, and oscillations can occur only at certain complex natural resonant frequencies. Each of the corresponding resonant modes has associated with it some particular value of m —that is, there is no mixing in m . For each value of m , modes are possible having an index n which may take on any integer value from m to infinity. These indices correspond to those of the $P_n^m(\cos \theta)$ which one has in the absence of the discontinuity in ϵ . There is, of course, mixing of eigenfunctions in ν , μ , κ , and λ .

The highly tentative conclusions reached can best be explained with reference to the diagram of figure 6, which shows a portion of the complex frequency plane. Point E represents the resonant frequency corresponding to $n=3$ and a perfectly conducting ionosphere located at some finite height. There is a four-fold degeneracy, since the modes for $m=0, 1, 2, 3$ all have the same resonant frequency. Point F represents the complex resonant frequency for an ionosphere having

$$\epsilon = \frac{2\epsilon_1\epsilon_2}{\epsilon_1 + \epsilon_2}. \quad (85)$$

The four-fold degeneracy is still present. Points G, H, J , and K show the way in which the introduction of the discontinuity separates the resonances of the same n and different m . If the vertical and horizontal shifts from E to F are regarded as first order, then, very roughly speaking, the horizontal separation of J and K is of third order, while the vertical separation is of fourth order. Resonances having $(m+n)$ odd tend to be shifted upward and to the right, while the reverse is true when $(m+n)$ is even. The shifts seem too small to impair the validity of the analysis of section 7, although this should not be regarded as definitely established.

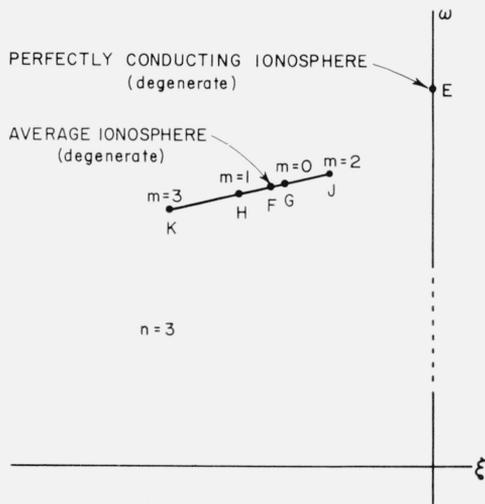


FIGURE 6. Nondegeneracy produced by diurnal variation of ϵ .

9. Appendix

In order to evaluate the coefficients a_ν in (53), the fields on a small vertical cone surrounding the dipole and extending to infinity will be considered. It can be shown that

$$\lim_{\theta \rightarrow 0} P_\nu(-\cos \theta) = \frac{2 \sin \nu\pi}{\pi} \ln \theta, \quad (86)$$

and therefore that

$$\lim_{\theta \rightarrow 0} \psi(x, \theta) = \frac{2}{\pi} \ln \theta \sum_\nu a_\nu \sin \nu\pi u_\alpha(x). \quad (87)$$

In this region the magnetic field is

$$H_\phi = \frac{1}{r} \frac{\partial \psi}{\partial \theta} = \frac{2}{\pi r \theta} \sum_\nu a_\nu \sin \nu \pi u_\alpha(x). \quad (88)$$

On the other hand, the magnetic field must be given by

$$H_\phi = \frac{p \delta(r-r_0)}{2\pi r \theta} = \frac{k p \delta(x)}{2\pi r \theta} \quad (89)$$

where p represents the moment of the source, or the product of current and vertical current path:

$$p = Il. \quad (90)$$

Equating the two expressions for H_ϕ , one obtains

$$\sum_\nu a_\nu \sin \nu \pi u_\alpha(x) = \frac{k p}{4} \delta(x). \quad (91)$$

The earth will be considered to be perfectly conducting, and $u_\alpha(0)$ will be taken equal to unity. Then (91) can be multiplied by $u_{\alpha'}(x)/\epsilon$ and integrated. Making use of the fact that $\epsilon(0) \approx 1$, the orthogonality of the u_α then yields (54).

10. References

- Balsler, M., and C. A. Wagner, Measurements of the spectrum of radio noise from 50 to 100 cycles per second, *J. Research NBS* **64D** (Radio Prop.), 415–423 (July–August 1960).
- Balsler, M., and C. A. Wagner, Observations of earth-ionosphere cavity resonances, Report 34G-0004, M.I.T. Lincoln Lab. (July 12, 1960).
- Budden, K. G., The 'Waveguide mode' theory of the propagation of very-low-frequency radio waves, *Proc. IRE* **45**, 772–774 (June 1957).
- Chapman, F. W., and R. C. V. Macario, Propagation of audio-frequency radio waves to great distances, *Nature* **177**, 930–933 (May 19, 1956).
- Jean, A. G., A. C. Murphy, J. R. Wait, and D. F. Wasmundt, Observed attenuation rate of ELF radio waves, *J. Research NBS* **65D** (Radio Prop.), 475–479 (September–October 1961).
- Pierce, E. T., Propagation of radio waves of frequency less than 1 kc, *Proc. IRE* **48**, 329–331 (March 1960).
- Ratcliffe, J. A., The highest parts of the ionosphere, *Quart. J. Roy. Met. Soc.* **85**, 321–331 (Oct. 1959).
- Wait, J. R., The mode theory of VLF ionospheric propagation for finite ground conductivity, *Proc. IRE* **45**, 760–767 (June 1957).
- Wait, J. R., An extension to the mode theory of VLF ionospheric propagation, *J. Geophys. Research* **63**, 123–135 (Mar. 1958).
- Wait, J. R., On the theory of the slow-tail portion of atmospheric waveforms, *J. Geophys. Research* **65**, 1939–1945 (July 1960).
- Wait, J. R., Mode theory and the propagation of ELF radio waves, *J. Research NBS* **64D** (Radio Prop.), 387–403 (July–August 1960).
- Wait, J. R., On the propagation of VLF and ELF radio waves when the ionosphere is not sharply bounded, *J. Research NBS* **66D** (Radio Prop.), (Jan.–Feb. 1962).

(Paper 66D4–208)