JOURNAL OF GEOPHYSICAL RESEARCH VOL. 70, NO. 7 APRIL 1, 1965

Ion Cyclotron Whistlers

D. A. GURNETT[1] AND S. D. SHAWHAN[1]

University of Iowa, Iowa City

N. M. BRICE[2] AND R. L. SMITH

Stanford University, Stanford, California

Abstract. An experimental study of the proton whistler, a new VLF phenomenon observed in satellite data, is presented, and an explanation of this new effect is given. The proton whistler appears on a frequency-time spectrogram as a tone which starts immediately after the reception of a short fractional-hop whistler at the satellite and initially shows a rapid rise in frequency, asymptotically approaching the gyrofrequency for protons in the plasma surrounding the satellite. It is proposed that the proton whistler is simply a dispersed form of the original lightning impulse and that the dispersion can be explained by considering the effect of ions on the propagation of an electromagnetic wave in the ionosphere. The propagation of a wave in a multicomponent plasma for frequencies of the order of the ion gyrofrequencies is discussed. In the ionosphere it is found that, in addition to the right-hand polarized whistler mode, the left-hand polarized mode (ion cyclotron wave) is also a possible mode of propagation for certain ranges of frequencies and altitudes. Between each two adjacent ion gyrofrequencies there is a frequency for which both modes of propagation are linearly polarized. These frequencies are called the crossover frequencies. A wave propagating in the ionosphere changes polarization at the altitude where the wave frequency is equal to a crossover frequency. This polarization reversal provides the mechanism by which an upgoing whistler can become an ion cyclotron wave. We show that the proton whistler is an ion cyclotron wave which occurs via this polarization reversal process. The crossover frequency can be measured from spectrograms of proton whistlers and is used to determine the fractional concentration of H^+ in the plasma surrounding the satellite. Near the altitude and frequencies for which polarization reversal occurs, it is shown that the right-hand polarized wave and the ion cyclotron wave may be strongly coupled. For frequencies of the order of the ion gyrofrequencies, this coupling process plays an important part in determining what regions of the ionosphere are accessible to waves from a given source location.

I. INTRODUCTION

A new very low frequency (VLF) phenomenon associated with the propagation of whistlers in the ionosphere has been found in the VLF recordings from the Injun 3 and Alouette satellites [*Smith et al.*, 1964; *Brice*, 1964]. The new effect appears on a frequency-time spectrogram as a tone which starts immediately after the reception of a short fractional-hop whistler at the satellite and initially shows a rapid rise in frequency, asymptotically approaching the gyrofrequency for protons in the plasma surrounding the satellite. This new phenomenon and the nomenclature that will be used here are shown in Figure 1.

It is proposed that this new phenomenon is simply a dispersed form of the original lightning impulse and that the dispersion can be explained by considering the effects of ions on the propagation of an electromagnetic wave in the ionosphere. In this paper the characteristic tone occurring immediately after the reception of a whistler is called a 'proton whistler' (see Figure 1). The designation proton whistler has been adopted because the phenomenon is postulated to be a whistler[3] and the frequency of the tone is found to asymptotically approach the gyrofrequency for protons at the point of observation. Here the short fractional-hop whistler which immediately precedes the proton whistler is called an 'electron whistler' (see Figure 1). The designation electron

[1] National Aeronautics and Space Administration Graduate Research Fellow.

[2] Now at Carleton University, Ottawa, Ontario, Canada.

[3] VLF phenomena whose source of energy is a lightning impulse are classified as whistlers [*Gallet*, 1959].

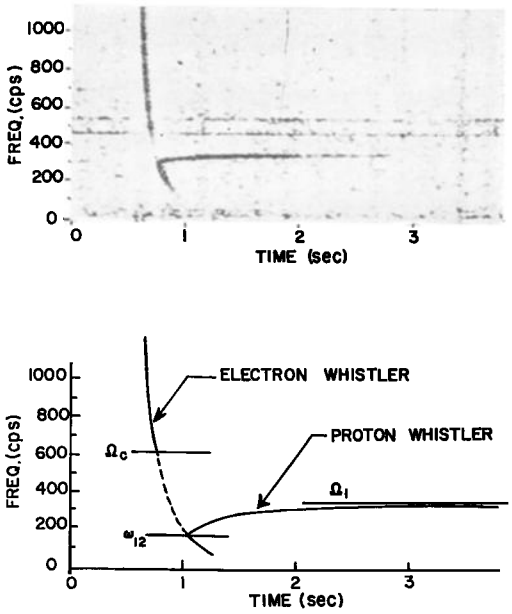

Fig. 1. Spectrogram showing a proton whistler and nomenclature.

whistler has been adopted because the electron gyrofrequency is the upper frequency limit for which the right-hand polarized whistler mode can propagate. The general consistency of this terminology will become evident.

The explanation of the proton whistler is based on calculations of the propagation of a VLF electromagnetic wave in a model ionosphere having several types of ions. Propagation in a plasma containing several ions has been considered by *Hines* [1957], *Buchsbaum* [1960], *Stix* [1962], *Yakimenko* [1962], *Gintsburg* [1963], and *Smith and Brice* [1964]. When the effects of ions are included in the propagation equations, it is found that in addition to the right-hand polarized whistler mode the left-hand polarized mode is also a possible mode of propagation for certain ranges of frequencies and altitudes in the ionosphere. *Stix* [1962] calls the left-hand polarized wave propagating at frequencies less than the maximum ion gyrofrequency an 'ion cyclotron wave.' *Gintsburg* [1963] showed that the ion cyclotron mode has resonances at each ion gyrofrequency. Propagation is possible in the ion cyclotron mode for a band of frequencies below each ion gyrofrequency, and the lower cutoff frequency for each band is greater than the next lower ion gyrofrequency. The bandwidth available depends on the relative concentrations of the various ions. We interpret the proton whistler as an ion cyclotron wave propagating in the frequency band available just below the proton gyrofrequency. Here this wave is called a 'proton cyclotron wave.'

Stix [1962] has indicated that when a plasma has positive ions having different charge to mass ratios the polarization of a wave may change sense of rotation while propagating through a slowly varying medium. *Smith and Brice* [1964] find that between each two adjacent ion gyrofrequencies there is a frequency, called the crossover frequency, for which both modes of propagation are linearly polarized. A wave propagating in a slowly varying medium changes sense of polarization when the wave frequency becomes equal to a crossover frequency. In the ionosphere this polarization reversal provides the mechanism by which an upgoing right-hand polarized electron whistler can become a left-hand polarized proton whistler. Polarization reversal is a necessary point in the explanation of the proton whistler, because, as will be shown in section 4, the predominance of O^+ and He^+ in the lower ionosphere prevents the direct propagation of a lightning impulse to the satellite entirely via the proton cyclotron mode.

In section 4 it is shown that near the altitude and frequencies for which polarization reversal occurs the conditions may exist for strong coupling between the two modes of propagation. When strong coupling exists we can expect an upgoing electron whistler to be split into two upgoing waves. Continuing to higher altitudes, these two waves propagate independently, one in the right-hand polarized mode and the other in the ion cyclotron mode. This coupling process plays an important part in determining what regions of the ionosphere are accessible to waves from a given source location.

In section 5, using a model ionosphere consisting of three ions (H^+, He^+, O^+), we present evidence that proton whistlers are ion cyclotron waves. Group propagation times based on this model ionosphere show that it is consistent to interpret the proton whistler as propagating in the proton cyclotron mode and as being simply a dispersed form of the original lightning impulse. The dispersion of the proton whistler arises from propagation part of the distance to the satellite in the right-hand polarized mode, and part in

the proton cyclotron mode. The crossover frequency can be easily determined from spectrograms of proton whistlers and is used to make estimates of the fractional concentration of H^+ relative to the electron density in the plasma surrounding the satellite.

2. Preliminary Study of Proton Whistlers

From the Injun 3 VLF experiment about 1000 hours of analog VLF data were obtained [*Gurnett*, 1963]. Approximately 10% of these data have been studied to determine the nature of the proton whistlers observed both in the Alouette 1 and the Injun 3 data.

To obtain a random but reasonably sized sample of the analog VLF data, all the satellite revolutions from each ground station, recorded on the eleventh day of each of the ten months that Injun 3 was operable, were analyzed. Parts of sixty-seven revolutions were analyzed aurally. Twenty-three of these revolutions had proton whistlers occurring.

The results of this preliminary study are:

1. The proton whistler is observed only after reception of a short fractional-hop whistler. Figure 2 shows five spectrograms of the electron-whistler trace followed by the proton whistler trace. The electron whistler is always a short fractional-hop whistler; that is, it is a whistler which has propagated only from the base of the ionosphere to the satellite. In all the spectrograms studied there has never been a proton whistler observed without an electron whistler present. When proton whistlers are present they follow every observable short fractional-hop electron whistler. However, on only 36% of the revolutions for which electron whistlers were present, were proton whistlers also present.

2. There is a frequency at which the electron whistler and the proton whistler are coincident in time. Figure 1 shows a trace for which the electron whistler and the proton whistler are obviously coincident in time at a frequency denoted by ω_{12}. As will be shown in section 5, this frequency corresponds to the crossover frequency between the proton gyrofrequency and the He^+ gyrofrequency. Sometimes the electron whistler is cut off at a frequency Ω_e above the starting frequency ω_{12} for the proton-whistler trace (see Figure 1). In these cases it appears that the electron whistler propagation to the satellite is strongly attenuated between frequencies ω_{12} and Ω_e.

3. The proton whistler initially shows a rapid rise in frequency which starts at a frequency ω_{12} and asymptotically approaches the proton gyrofrequency. To determine the value of the asymptotic frequency Ω_1, spectrograms were made on a Sonagraph (Kay Electric Co., Pine Brook, N. J.) and the frequency was measured. The estimated measuring accuracy was ± 15 cps, and measurements were referenced to a standard frequency generated in the satellite. Figure 3 is a plot of the magnitude of the magnetic field B at the satellite against the asymptotic frequency. The solid curve is the calculated proton gyrofrequency. All the experimental frequency values lie within 3½% of the calculated proton gyrofrequency. Most of the frequency values were lower than the proton gyrofrequency, probably because the trace disappeared (about 2 seconds after appearance) before the asymptotic value was reached. The obvious conclusion is that the asymptotic frequency is the gyrofrequency for protons in the plasma surrounding the satellite.

4. The ratio $\omega_{12}/\Omega_1 = \Lambda_{12}$ increases with decreasing altitude, approaching unity at an altitude of approximately 440 km at local midnight and at approximately 640 km at local noon. Figure 2*a–d* is a sequence of spectrograms for proton whistlers around local midnight, 20.87–24.28 local time. With decreasing altitude the value of Λ_{12} increases from approximately 0.52 for (*a*) at an altitude of 874 km to approximately 0.94 for (*d*) (proton whistler trace marked by an arrow) at an altitude of 442 km. For altitudes below 442 km the proton-whistler trace is absent. Note that the proton gyrofrequency is higher for (*c*) than for (*d*), even though (*c*) is higher in altitude; the value of Λ_{12} is, however, higher for (*d*) than for (*c*). This higher Ω_1 value for (*c*) is due to the larger value of B at the satellite; the Λ_{12} value seems to depend, therefore, on altitude and not on B.

Figure 2*e* is a spectrogram of a local noon proton whistler. An altitude sequence of proton whistlers for local noon exhibits the same increasing Λ_{12} with decreasing altitude found in the local midnight data. However, the Λ_{12} values are different for a given altitude, local noon values always being higher. For local midnight Λ_{12} approaches unity at approximately 440 km,

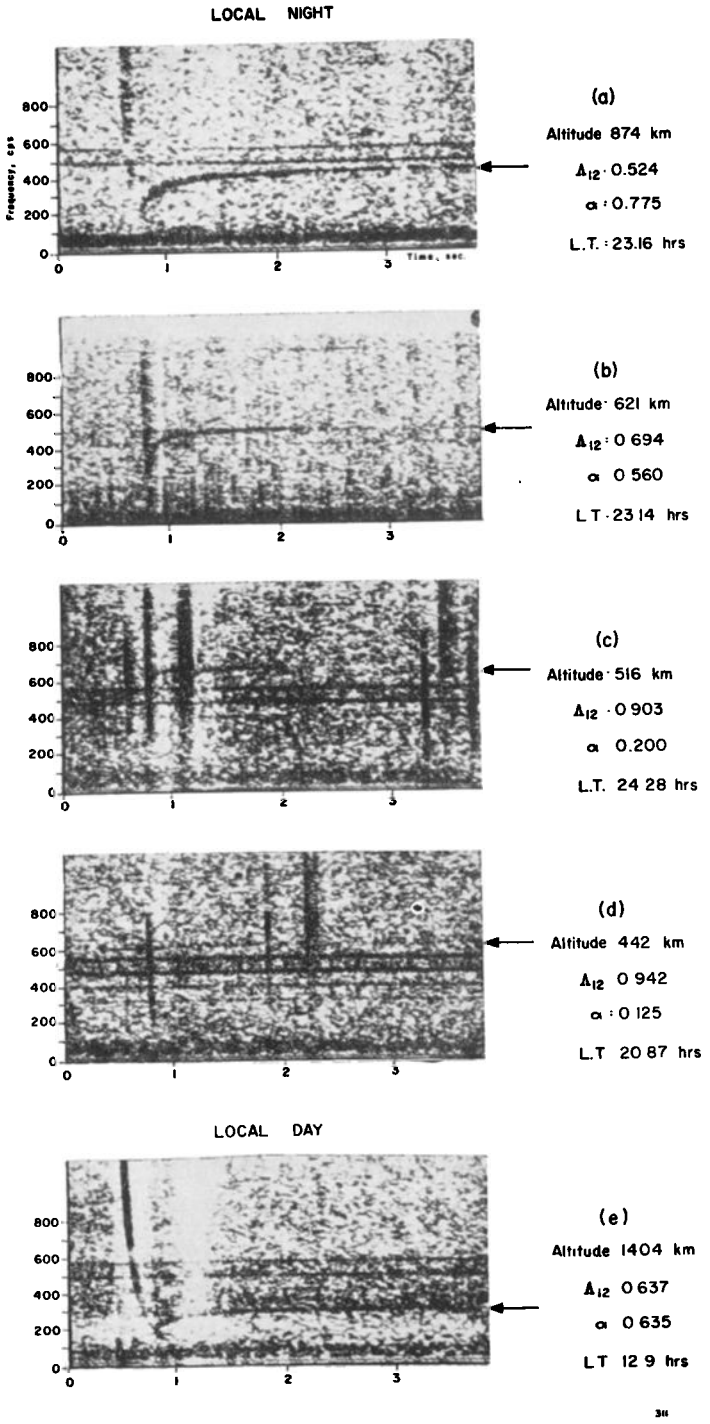

Fig. 2. Altitude and local time dependence of ω_{12}/Ω_1 for proton whistlers.

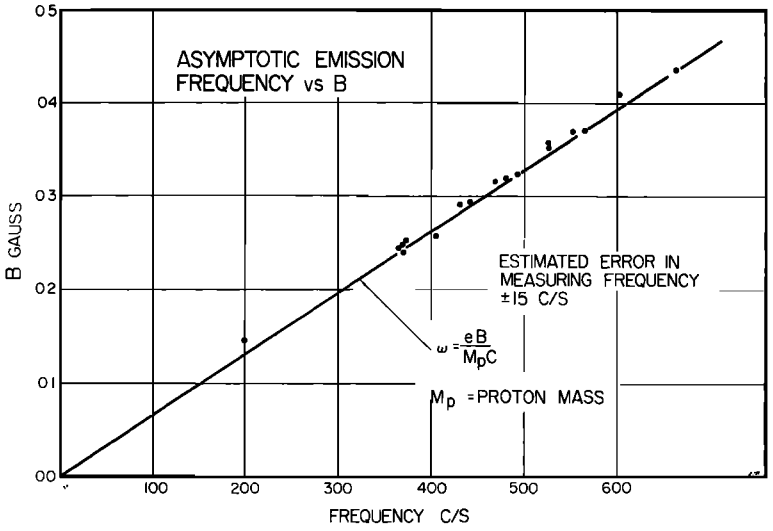

Fig. 3. Asymptotic frequency Ω_1 versus B.

whereas, for local noon, Λ_{12} approaches unity at approximately 640 km. Note that in Figure 2*e* $\Lambda_{12} = 0.64$, which is about the same as $\Lambda_{12} = 0.69$ for (*b*), but (*e*) is at 1404 km, more than twice the altitude of (*b*).

5. The proton whistlers occur more frequently during local night than during local day and have never been observed to occur below 442 km during local night or below 640 km during local day.

In Figure 4 each open circle means that whistlers were occurring on that particular

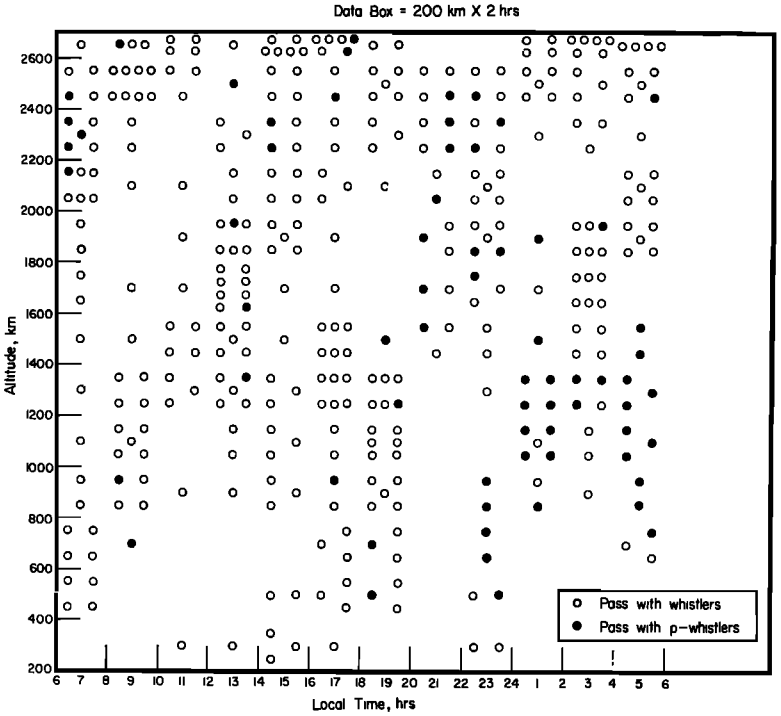

Fig. 4. The occurrence of proton whistlers in altitude–local time space. These data include all latitudes sampled by the satellite.

revolution within a 'box' in the ionosphere 200 km by 2 hours local time. When on a particular revolution proton whistlers were occurring, they are represented by the closed circles. The open circles, therefore, depict the whistler sample density in altitude–local time space with the occurrence of proton whistlers superimposed as the closed circles. These data include all latitudes sampled by the satellite (see Figure 5). From Figure 4 two trends of proton-whistler occurrence can be noted. During local night, 1800–0600, there seems to be a higher frequency of occurrence of proton whistlers than during local day, 0600–1800, whereas the whistler sample density is reasonably constant for local day and night. The proton whistlers occur at lower altitudes during the local night than they do during local day. Although the sample density of Figure 4 was about zero at altitudes below 1000 km around local noon, no proton whistlers were found to occur below 640 km in an independent local noon (1100–1400) sample of the entire altitude range.

6. There is an apparent northern hemisphere–southern hemisphere asymmetry in the occurrence of proton whistlers with proton whistlers occurring more frequently in the northern hemisphere than in the southern. Also, there is an apparent high latitude cutoff in the occurrence of proton whistlers that seems to correspond to the auroral zone.

In Figure 5 the occurrence of electron whistlers (open circles) and proton whistlers (closed circles) is shown in altitude–latitude space. These data include all local times sampled by the satellite. From these data it is evident that proton whistlers occur more frequently (relative to electron whistlers) in the northern hemisphere than in the southern hemisphere. This asymmetry is probably due to the more frequent occurrence of lightning, the source of whistlers, in the northern geomagnetic hemisphere. Thus, in the northern hemisphere we would expect a greater fraction of the total number of whistlers to be short fractional-hop whistlers and consequently a greater fractional occurrence of proton whistlers, relative to the southern hemisphere. Also, from Figure 5 a boundary in

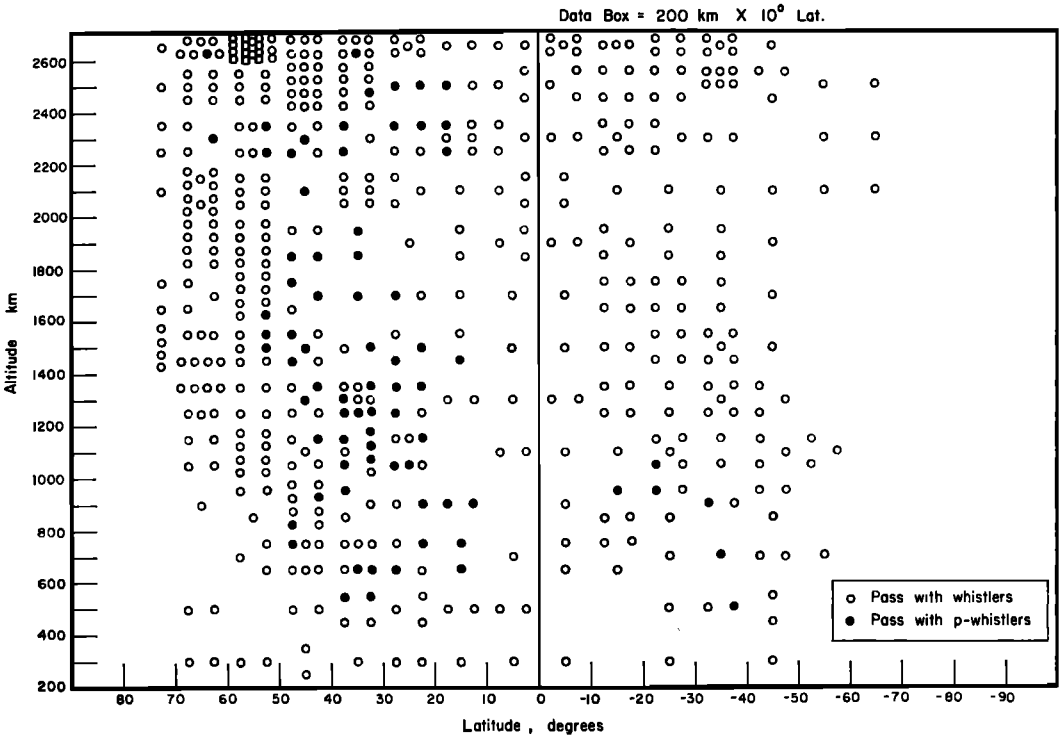

Fig. 5. The occurrence of proton whistlers in latitude–local time space. These data include all local times sampled by the satellite.

proton whistler occurrence can be drawn from approximately 40°N latitude at 600 km to approximately 66°N latitude at 2700 km. This boundary is the apparent high latitude cutoff in the occurrence of proton-whistlers.

3. Propagation Equations for a Multicomponent Plasma

In an effort to explain the proton-whistler phenomenon on the basis of propagation theory, we now consider the effects of ions on the propagation of an electromagnetic wave in the ionosphere. As will be seen, the presence of ions having different charge to mass ratios has important effects on the propagation of an electromagnetic wave in the ionosphere for frequencies of the order of the ion gyrofrequencies.

The refractive index for a multicomponent plasma. The refractive index for a plasma containing several types of ions has been given by *Hines* [1957] and others. Here we use the symbolism introduced by *Stix* [1962]. Following Stix we assume a cold homogeneous plasma immersed in a uniform static magnetic field $\mathbf{B}_0$ and consider all first-order field quantities to vary as exp $[i(\mathbf{K}\cdot\mathbf{r} - \omega t)]$. The refractive index is obtained from the condition for the existence of a nontrivial solution of the following set of homogeneous field equations (see Stix for the detailed derivation):

$$\begin{bmatrix} S - n^2\cos^2\theta & -iD & n^2\cos\theta\sin\theta \\ iD & S - n^2 & 0 \\ n^2\cos\theta\sin\theta & 0 & P - n^2\sin^2\theta \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = 0 \quad (1)$$

The magnetic field $\mathbf{B}_0$ is in the Z direction, and $\mathbf{K}$ is in the X, Y plane at an angle θ with respect to $\mathbf{B}_0$.

The condition that (1) has a nontrivial solution is that the determinant of the matrix be zero. This condition gives the refractive index in the following form:

$$An^4 - Bn^2 + RLP = 0 \quad (2)$$

where

$$A = S\sin^2\theta + P\cos^2\theta \quad (3)$$

$$B = RL\sin^2\theta + PS(1 + \cos^2\theta) \quad (4)$$

$$R = 1 - \sum_K \frac{\pi_K^2}{\omega(\omega + \epsilon_K\Omega_K)} \quad (5)$$

$$L = 1 - \sum_K \frac{\pi_K^2}{\omega(\omega - \epsilon_K\Omega_K)} \quad (6)$$

$$P = 1 - \sum_K \frac{\pi_K^2}{\omega^2} \quad (7)$$

$$S = \tfrac{1}{2}(R + L) \qquad D = \tfrac{1}{2}(R - L) \quad (8)$$

$$\pi_K^2 = 4\pi n_K q_K^2/m_K \quad (9)$$

$$\Omega_K = |q_K B_0/m_K c| \quad (10)$$

$$\epsilon_K = q_K/|q_K| \quad (11)$$

The subscripts K refer to the Kth type of particle having mass m_K, charge q_K (including sign), and number density n_K; π_K and Ω_K are the plasma frequency and gyrofrequency of the Kth constituent. The effects of collisions can be included by replacing m_K with $m_K(1 + iZ_K)$, where $Z_K = \nu_K/\omega$, and ν_K is the collision frequency for a representative particle of type K.

The solutions of (2) can be written in the following form:

$$n^2 = (B \pm F)/2A \quad (12)$$

where

$$F^2 = (RL - PS)^2\sin^4\theta + 4P^2D^2\cos^2\theta \quad (13)$$

The waves propagating along or transverse to $\mathbf{B}_0$, $\theta = 0$ or $\theta = \pi/2$, are of special interest in studying general properties of (12) and are called 'longitudinal waves' and 'transverse waves,' respectively.

$$\begin{aligned} &\text{For} \quad \theta = 0: \quad n_R^2 = R \quad n_L^2 = L \\ &\text{For} \quad \theta = \pi/2: \quad n_x^2 = \frac{2RL}{R + L} \quad n_o^2 = P \end{aligned} \quad (14)$$

The symbols R, L, x, and o stand for right, left, extraordinary, and ordinary.

The wave polarization. The wave polarization is defined in terms of the components of the electric field which lie in the wave front. We choose the Z' axis in the $\mathbf{K}$ direction, the X' axis in the ($\mathbf{K}$, $\mathbf{B}$) plane, and the Y' axis normal to the ($\mathbf{K}$, $\mathbf{B}$) plane. The wave polarization is now defined as

$$\rho = E_{Y'}/E_{X'} \quad (15)$$

From (1) and (15) and using the appropriate coordinate transformation from unprimed to

primed coordinates, we readily find the wave polarization to be

$$\rho = \frac{iDn^2 \cos\theta}{Sn^2 - RL} \tag{16}$$

If ρ_o and ρ_x denote the wave polarization for the two branches of n^2, it can be shown from (12) and (16) that $\rho_o\rho_x = 1$.

The wave polarization gives the time dependence of the component of E perpendicular to the propagation vector. If collisions are neglected ($Z_K = 0$), the wave polarization is always imaginary. When $\rho = i$ the polarization is circular and field quantities varying as exp $[-i\omega t]$ rotate in the right-hand sense relative to the propagation vector. If $\rho = -i$ the polarization is circular and left-handed. For waves along $\mathbf{B}_0(\theta = 0)$, we can verify that the polarization is right circular if $n^2 = R$ and left circular if $n^2 = L$.

Parameter space and wave-normal surfaces. The refractive index (12) for a multicomponent plasma involves the parameters ω, B, π_1, π_2, . . . , π_m, and θ. The dependence of n on these parameters is conveniently studied by investigating the form of $n(\theta)$ at points in an appropriate space having scale lengths proportional to the parameters ω, B, π_1, etc. Such a plot, having an axis for each parameter being considered, is called a 'parameter space' or 'C. M. A. diagram' after the work of *Clemmow and Mullaly* [1955] and *Allis* [1959]. The form of $n(\theta)$ is conveniently represented by a plot of the phase velocity $u = [c/n(\theta)](\mathbf{K}/K)$ in spherical coordinates and is called the 'phase velocity surface' or 'wave-normal surface.' Wave-normal surfaces have rotational symmetry about the static magnetic field. The shape of a wave-normal surface for positive n^2 can be shown to be topologically equivalent to either a sphere, a dumbbell lemniscoid, or a wheel lemniscoid [*Stix*, 1962]. The topological genera to which a given branch of n^2 belongs depends on the coordinates in parameter space.

The terms 'cutoff' and 'resonance' refer to the conditions for which $n^2 = 0$ and $n^2 = \infty$, respectively. The conditions for a cutoff are, from (2): $R = 0$, or $L = 0$, or $P = 0$. The resonances for the longitudinal waves are, from (14), $R = \infty$ and $L = \infty$, and for the transverse extraordinary wave, $S = 0$. The surfaces in parameter space for the cutoffs $R = 0$, $L = 0$, $P = 0$ and the surfaces for the resonances $R = \infty$, $L = \infty$, $S = 0$ have a special significance in determining the topological genera of the wave-normal surfaces. These surfaces are called 'bounding surfaces,' and they divide parameter space into a number of volumes called 'bounded' volumes. *Stix* [1962] shows the important result that for a given branch of n^2 the topological genera of the wave-normal surface is the same throughout a bounded volume. Thus if the wave-normal surfaces are determined at one point in a bounded volume they are topologically correct throughout that volume.

The wave-normal surfaces for the two branches of n^2 are conveniently labeled right (R) and left (L) according to the sense of polarization at $\theta = 0$, or ordinary (o) and extraordinary (x) according to the refractive index at $\theta = \pi/2$. The R, L labeling scheme remains consistent throughout a bounded volume only if the surface $D = 0$ is included as a bounding surface in parameter space. The surface $D = 0$ must be included because the polarization ρ changes sign at $D = 0$ (see equation 16). Wave-normal surfaces can also be consistently labeled fast (f) or slow (s) according to the phase velocities at angles between 0 and $\pi/2$.

4. Propagation of Ion Cyclotron Waves in a Model Ionosphere Having Three Ions

The parameter space for a plasma having three ions. We now specialize this investigation to a plasma consisting of a neutral mixture of three types of singly charged positive ions (H^+, He^+, O^+) and electrons. These four constituents are the predominant charged particles found in the ionosphere above 200 km [*Bourdeau*, 1963]. Since we are interested in effects at frequencies of the order of the proton gyrofrequency, we assume that the wave frequency is less than the geometric mean of the electron gyrofrequency Ω_e and the proton gyrofrequency Ω_1, i.e., $\omega^2 < \Omega_e\Omega_1$. We also make the useful observation that in the ionosphere, above approximately 200 km, we can assume $\Omega_e\Omega_1 \ll \pi_e^2$. For the preliminary analysis the plasma is assumed to be collisionless ($Z_K = 0$).

Before discussing the wave-normal surfaces in parameter space, it is illuminating to consider the refractive index squared for the waves at $\theta = 0$ and $\theta = \pi/2$. Introducing the normalized frequency parameter $\Lambda = \omega/\Omega_1$, Ω_1 = proton gyrofrequency, and the concentration param-

eters $\alpha = n(H^+)/n_e$, $\beta = n(He^+)/n_e$, and $\gamma = n(O^+)/n_e$, we can write R, L, S, D, and P as follows ($\omega \ll \omega_e$):

$$R = 1 + \left(\frac{\pi_e}{\Omega_e}\right)^2 + \frac{\pi_e^2}{\Omega_1\Omega_e}\left[\frac{1}{\Lambda} - \frac{\alpha}{\Lambda(1+\Lambda)} - \frac{\beta}{\Lambda(1+4\Lambda)} - \frac{\gamma}{\Lambda(1+16\Lambda)}\right] \quad (17)$$

$$L = 1 + \left(\frac{\pi_e}{\Omega_e}\right)^2 - \frac{\pi_e^2}{\Omega_1\Omega_e}\left[\frac{1}{\Lambda} - \frac{\alpha}{\Lambda(1-\Lambda)} - \frac{\beta}{\Lambda(1-4\Lambda)} - \frac{\gamma}{\Lambda(1-16\Lambda)}\right] \quad (18)$$

$$S = 1 + \left(\frac{\pi_e}{\Omega_e}\right)^2 + \frac{\pi_e^2}{\Omega_1\Omega_e}\left[\frac{\alpha}{1-\Lambda^2} + \frac{4\beta}{1-16\Lambda^2} + \frac{16\gamma}{1-256\Lambda^2}\right] \quad (19)$$

$$D = \frac{1}{\Lambda}\frac{\pi_e^2}{\Omega_1\Omega_e}\left[1 - \frac{\alpha}{1-\Lambda^2} - \frac{\beta}{1-16\Lambda^2} - \frac{\gamma}{1-256\Lambda^2}\right] \quad (20)$$

$$P = 1 - \frac{1}{\Lambda^2}\left(\frac{\pi_e}{\Omega_1}\right)^2 \quad (21)$$

$$\alpha + \beta + \gamma = 1 \quad (22)$$

Using these equations the refractive index was squared for the longitudinal R and L waves and the transverse x wave has been calculated as a function of the frequency parameter Λ and is shown in Figure 6. The concentration parameters have been fixed for this illustrative case, with $\alpha = 0.8$, $\beta = 0.15$, $\gamma = 0.05$, and $\pi_e/\Omega_e = 1$. The refractive index squared for the transverse ordinary wave ($n_o^2 = P$) is a large negative quantity for the conditions being considered ($\omega^2 < \Omega_1\Omega_e$, $\pi_e^2 \gg \Omega_1\Omega_e$) and is not shown in Figure 6.

The bounding surfaces shown in Figure 6 are Λ values (frequencies) corresponding to the cutoff $L = 0$, the resonances $L = \infty$ and $S = 0$, and the condition $D = 0$. Since only the parameter Λ is being considered, the bounding surfaces in parameter space reduce to points on the Λ axis and bounded volumes become intervals on the Λ axis. The resonances for the longitudinal L wave are at the gyrofrequencies for H^+, He^+, and O^+, corresponding to $\Lambda = 1$, $\Lambda = 1/4$, and $\Lambda = 1/16$. These resonances are called the 'ion cyclotron resonances.' The resonances for the transverse x wave are at Λ values which are roots of $S = 0$. These transverse x-wave resonances are called 'hybrid resonances' and have been discussed by *Auer* [1958], *Buchsbaum* [1960], *Allis* [1963], and *Smith and Brice* [1964]. Cutoffs for the longitudinal L wave and transverse x wave are at Λ values which are roots of $L = 0$. The refractive index squared for the longitudinal

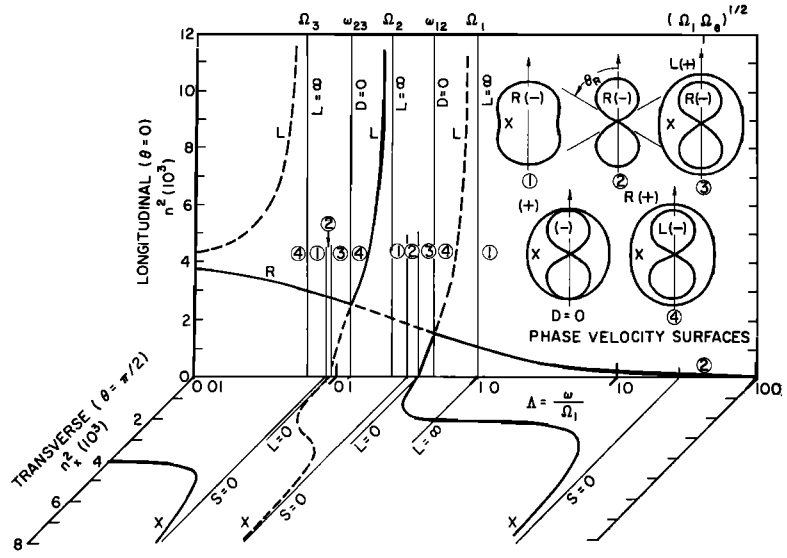

Fig. 6. Refractive index squared for the R, L, and X waves, and wave-normal surfaces in each bounded volume ($\alpha = 0.8$, $\beta = 0.15$, $\gamma = 0.05$).

R wave is positive for $\omega < \Omega_e$. From Figure 6 and equation 12, it is seen that the refractive indices of the L, R, and x waves are equal at the bounding surface $D = 0$.

The frequencies corresponding to the roots of $D = 0$ have been discussed by *Smith and Brice* [1964] and are called the crossover frequencies. *Stix* [1962] finds that the condition $D = 0$ has a solution only when the plasma has positive ions having different charge to mass ratios, and only for frequencies greater than the minimum ion gyrofrequency and less than the maximum ion gyrofrequency. The crossover frequency between the H^+ and He^+ ion cyclotron resonances is denoted ω_{12}, and between He^+ and O^+ is denoted ω_{23}, etc.

Having illustrated the cutoff and resonance frequencies for a particular set of parameters, we now determine how these frequencies depend on the parameters α, β, and π_e/Ω_e. The parameter γ is not an independent parameter, since it is determined by the condition that the plasma be electrically neutral (equation 22). The bounding surfaces $L = 0$, $S = 0$, and $D = 0$ have solutions for Λ which depend on α, β, and π_e/Ω_e. First, we consider the bounding surface $D(\Lambda,\alpha, \beta) = 0$. We note that π_e/Ω_e is not a parameter for this bounding surface, since it is a multiplicative factor in (20). If we consider Λ a constant, (20) becomes a linear equation in α and β. The linear relation in α and β, defined by $D(\Lambda, \alpha, \beta) = 0$, has been plotted for various Λ values and is shown in Figure 7. If values for α and β are given, the Λ values satisfying $D(\Lambda, \alpha, \beta) = 0$ can be determined by interpolating between the Λ contours.

It would be convenient if the technique used to obtain a contour plot of $D = 0$ could be applied to the equations $L(\Lambda, \alpha, \beta, \pi_e/\Omega_e) = 0$ and $S(\Lambda, \alpha, \beta, \pi_e/\Omega_e) = 0$. We note that between the ion gyrofrequencies the functions $L(\Lambda)$ and $S(\Lambda)$ cross the zero axis with a steep slope, and the small term $1 + (\pi_e/\Omega_e)^2$, in (18) and (19), can be omitted when solving for the roots of $L(\Lambda) = 0$ and $S(\Lambda) = 0$, provided that the roots are between the ion gyrofrequencies and that $\pi_e^2 \gg \Omega_e/\Omega_1$. With this simplification, the surfaces $L = 0$ and $S = 0$ are independent of the parameter π_e/Ω_e and can be plotted as contours in the α, β plane, as shown in Figures 7 and 8.

The bounding surfaces $L(\Lambda, \alpha, \beta) = 0$, $S(\Lambda, \alpha, \beta) = 0$, and $D(\Lambda, \alpha, \beta) = 0$ intersect the $\gamma = 0$ plane in parameter space as shown in Figure 9. The $\gamma = 0$ plane gives the parameter space for a two component plasma having a fractional concentration α of H^+ ions and a fractional concentration $1 - \alpha$ of He^+ ions. Using the contour plots for $L = 0$, $S = 0$, and $D = 0$, the dependence of the critical frequencies (cutoff, resonance, etc.) on the concentrations can be visualized. Later in this Section, when a model ionosphere is considered, the bounding surface contours in Figures 7 and 8 are especially useful for determining the cutoff, resonance, and crossover frequencies as a function of altitude.

The topological shapes of the wave-normal surfaces have been determined for each bounded volume in (Λ, α, β) space. The wave-normal surfaces for the two branches of n^2, labeled according to the L, R scheme, have the four topological configurations sketched in Figure 6. The sketches have been numbered 1 through 4, and the wave-normal surfaces appropriate for each bounded volume are indicated by the corresponding sketch number. Figure 9 also shows wave-normal surfaces, and these are numbered to agree with Figure 6.

It is evident from Figure 6 that at the bounding surfaces only four types of transitions occur, these transitions being repeated for each ion. The transitions at these bounding surfaces have been discussed by *Stix* [1962] and by *Smith and Brice* [1964]. These transitions are discussed only briefly here. First, at $L = \infty$ the outer wave-normal surface labeled R suffers no discontinuities from crossing the bounding surface, and the crossing is an intact transition. The inner wave-normal surface labeled L in sketch 4 suffers a destructive transition when the vertex angle θ_R for the dumbbell lemniscoid goes to zero as $L = \infty$ is approached from the left. Since the wave-normal surface labeled L is profoundly influenced by the ion cyclotron resonance at $L = \infty$, *Stix* [1962] chose to call this wave an 'ion cyclotron wave.' Here we call the ion cyclotron wave just below the proton gyrofrequency a 'proton cyclotron wave,' the ion cyclotron wave just below the He^+ gyrofrequency a 'He^+ cyclotron wave,' etc. Stix shows that if $K_\perp$ is held constant $K_\parallel$ approaches infinity as the ion cyclotron resonance is approached; thus, the ion cyclotron wave is refracted toward the surface where $\Omega_K = \omega$ and

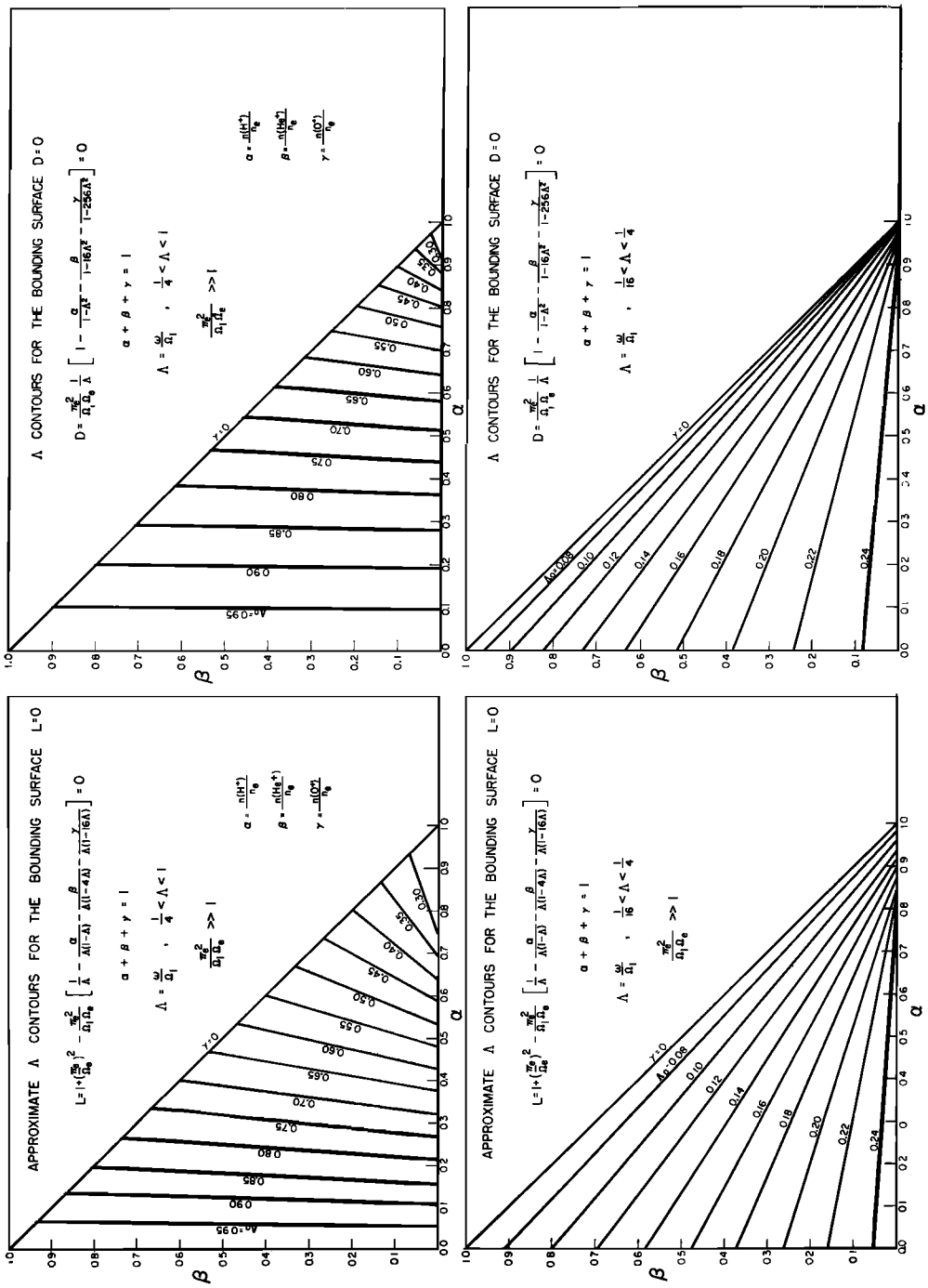

Fig. 7. The bounding surfaces $D = 0$ and $L = 0$ in parameter space.

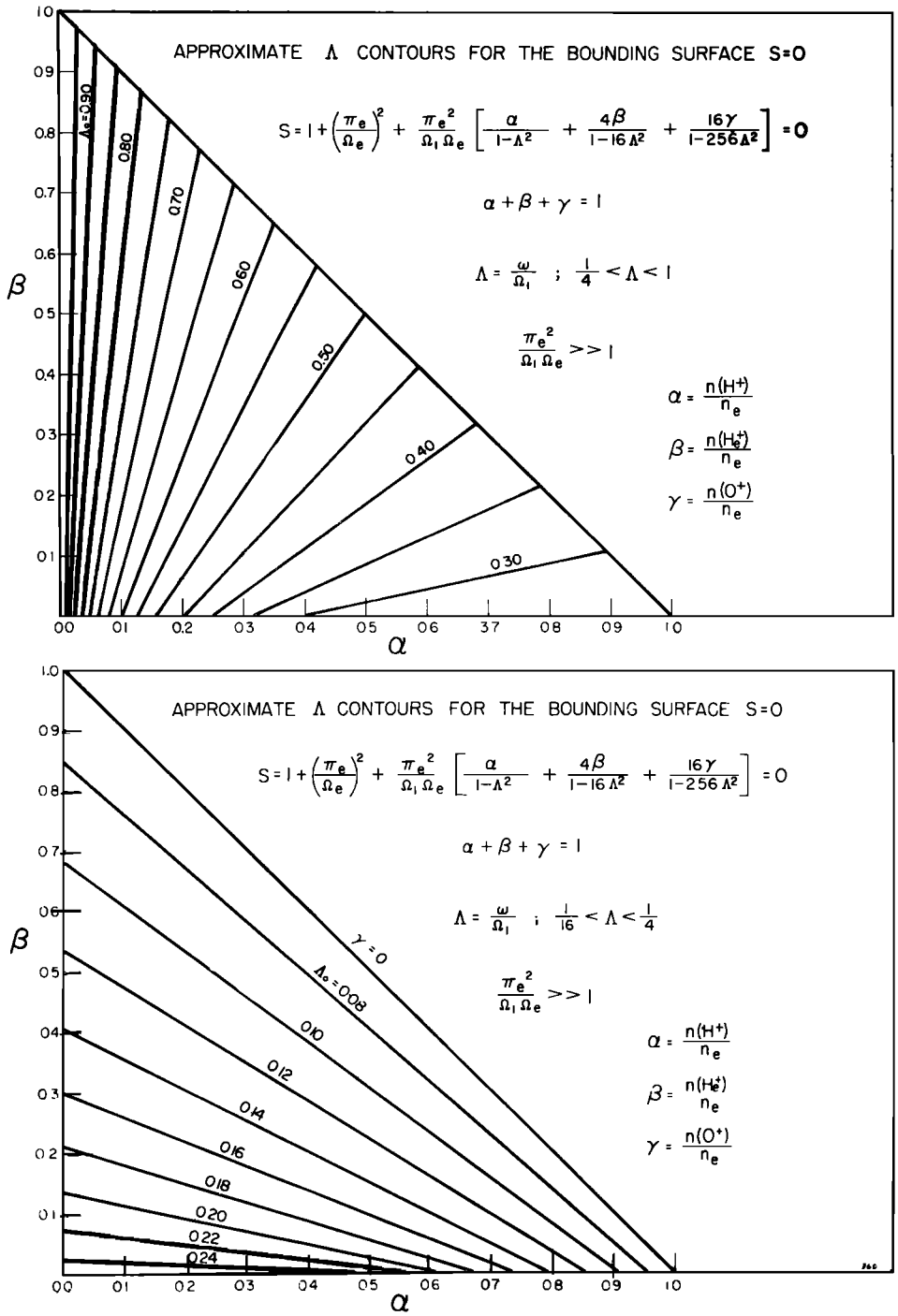

Fig. 8. The bounding surface $S = 0$ in parameter space.

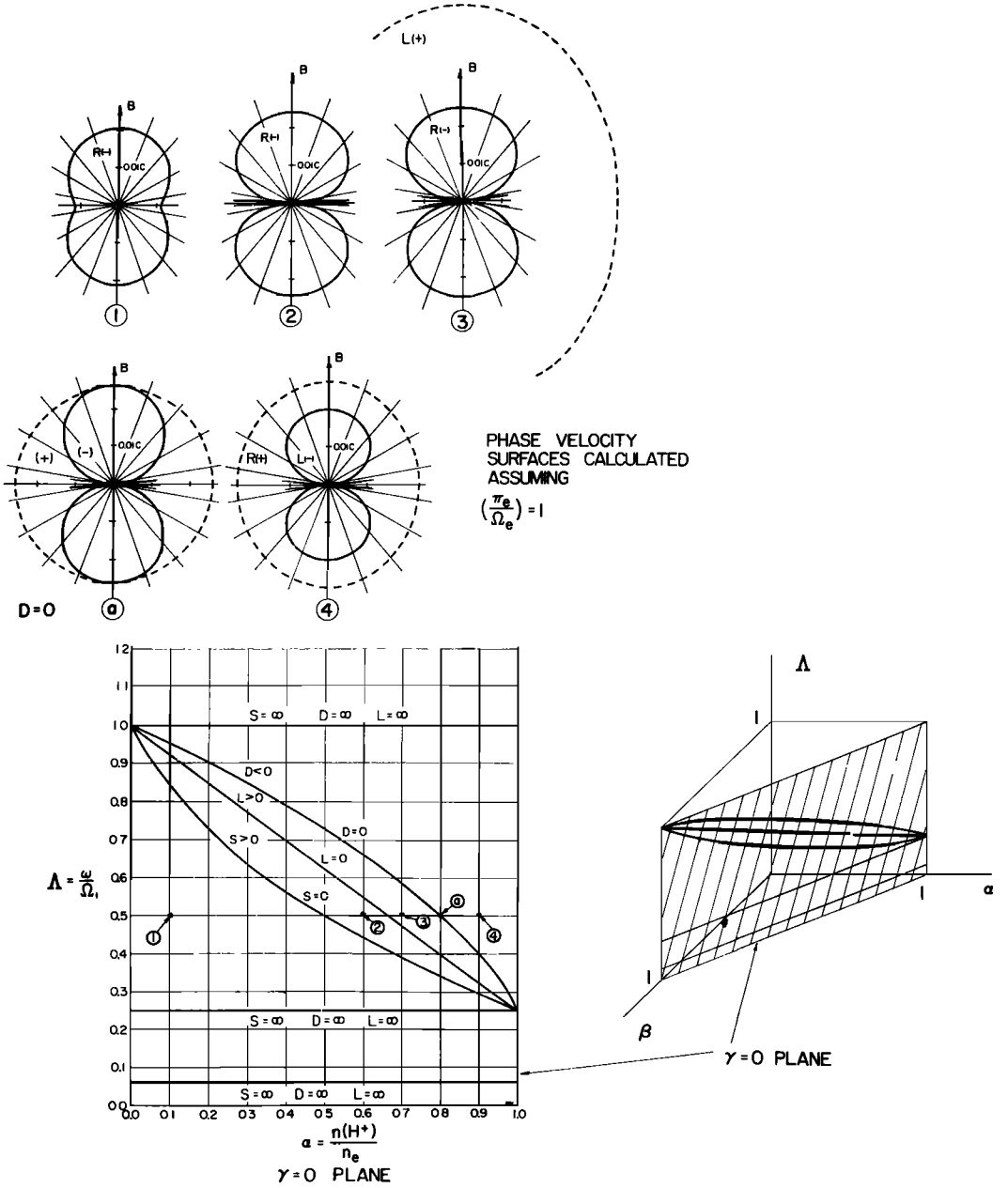

Fig. 9. The bounding surfaces for a two-ion plasma (H^+, He^+) and wave-normal surfaces in each bounded volume.

the ion cyclotron resonance takes place at $\theta = 0$.

The second type of transition evident in Figure 6 occurs at each hybrid resonance $S = 0$, separating regions having wave-normal surfaces labeled R and shown in sketches 1 and 2. It is evident from the refractive index for the transverse x wave that the phase velocity at $\pi/2$ goes to zero as the bounding surface $S = 0$ is approached. As the surface $S = 0$ is crossed from left to right the topological sphere in sketch 1

deforms continuously into a dumbbell lemniscoid. The transition at $S = 0$ is a reshaping transition.

The third type of transition occurs at each cutoff surface $L = 0$. Cutoffs separate regions having wave-normal surfaces shown in sketches 2 and 3. In sketch 3 the inner surface labeled R is a dumbbell lemniscoid and suffers no discontinuities due to crossing a bounding surface $L = 0$. The outer surface labeled L is a topological sphere. As the bounding surface $L = 0$ is approached from the right, a refractive index goes to zero for all angles, and the real wave-normal surface labeled L transforms into an imaginary surface at $L = 0$. Thus the wave-normal surface labeled L suffers a destructive transition at the bounding surface $L = 0$.

The fourth type of transition occurs at the bounding surface $D = 0$ and separates regions having wave-normal surfaces shown in sketches 3 and 4. In the neighborhood of $D = 0$ both the plus and minus branches of n^2 have real wave-normal surfaces. The outer (fast) wave-normal surface, a topological sphere, corresponds to the plus branch of n^2. The inner (slow) wave-normal surface, a dumbbell lemniscoid, corresponds to the minus branch of n^2. As the surface $D = 0$ is crossed, the wave-normal surfaces for the two branches both change in a continuous fashion and suffer no discontinuities due to the crossing. Therefore, the transition at the surface $D = 0$ is an intact transition for both branches of n^2. At the surface $D = 0$ it can be verified from (12) that the fast mode is isotropic and that the wave-normal surfaces are tangent at $\theta = 0$. From the polarization equation 16 it is evident that the right (R) and left (L) labels for the fast and slow modes are interchanged at $D = 0$. For the case $\theta = 0$ the wave polarization ρ is discontinuous and undefined at $D = 0$. For all angles $\theta \neq 0$ the wave polarization for the slow mode approaches zero in a continuous fashion as $D = 0$ is approached. At $D = 0$ the wave polarization for the slow mode is zero ($\rho_- = 0$), and so we have

$$E_{y'}' = \rho\, E_{X'}' = 0 \qquad \text{for} \quad D = 0, \quad \theta \neq 0$$

Thus the electric vector for the slow mode is linearly polarized in the (**K**, **B**) plane for all angles $\theta \neq 0$. The wave polarization for the fast (isotropic) mode at $D = 0$ is undefined using (16), since $Sn^2 - RL = 0$ and $D = 0$. Hence, the homogeneous field equation 1 must be investigated to determine the fields. We deduce that the electric field for the fast wave at $D = 0 (\theta \neq 0)$ is linearly polarized in the direction normal to the (**K**, **B**) plane and perpendicular to the electric field of the slow wave at $D = 0$. The results just presented for the wave polarization at $D = 0$ hold only for a collisionless plasma. In the next subsection we find that collisions strongly influence the wave polarization in the vicinity of $D = 0$.

In Figure 6 the refractive index squared for the longitudinal and transverse waves has been plotted in dashed and solid lines. As the parameter Λ is changed, the solid lines follow the refractive index at $\theta = 0$ and $\pi/2$ for the wave-normal surface, which deforms in a continuous manner at the bounding surfaces. The dashed lines have the same significance. It is evident that the fast and slow wave-normal surfaces touch only at the points where both $D = 0$ and $\theta = 0$, and that the accessibility of various points in (Λ, α, β) space, via a given continuously deformable wave-normal surface, is limited by cutoffs and resonances. In the next subsections we find that coupling between the two modes of propagation near ($D = 0$, $\theta = 0$) can extend the region of accessibility in the (Λ, α, β) space.

The critical coupling condition. The differential equations which describe the propagation of a plane wave in a horizontally stratified ionosphere can be written as a set of four simultaneous linear differential equations [*Budden*, 1961]. Such a system of equations, in general, has four independent solutions. This system of equations is said to be written in coupled form if the dependent variable having the highest order derivative is different for each equation and if the number of dependent variables is equal to the number of equations.

In a given equation the terms in the dependent variable having the highest order derivative are called the principal terms, and the remaining terms are called coupling terms. It often happens that if the parameters of the medium vary slowly with altitude the coupling terms can be neglected for some ranges of altitudes. Under these conditions the equations become four separate equations which can be solved approximately using the Wentzel-Kramers-Brillouin method. In using the WKB method, we solve for the waves in a homogeneous medium having the same param-

eters as the actual ionosphere at some given altitude. In general, four solutions of the form $\exp [i(\mathbf{K}\cdot\mathbf{r} - \omega t)]$ are obtained, two upgoing waves of opposite polarization and two downgoing waves of opposite polarization. These solutions are called characteristic waves. Approximate WKB solutions resembling the four characteristic waves can be constructed if the dispersion relation $K(\omega)$ is known for all altitudes and directions of propagation.

The method of characteristic waves fails at levels for which the coupling terms are not small. *Forsterling* [1942] derived the following second order coupled equations for vertical incidence. The derivation uses the result that $\rho_o\rho_x = 1$, mentioned in section 3.

$$\begin{aligned} F_0'' + F_0(n_0^2 + \psi^2) &= \psi' F_x + 2\psi F_x' \\ F_x'' + F_x(n_x^2 + \psi^2) &= \psi' F_0 + 2\psi F_0' \end{aligned} \tag{23}$$

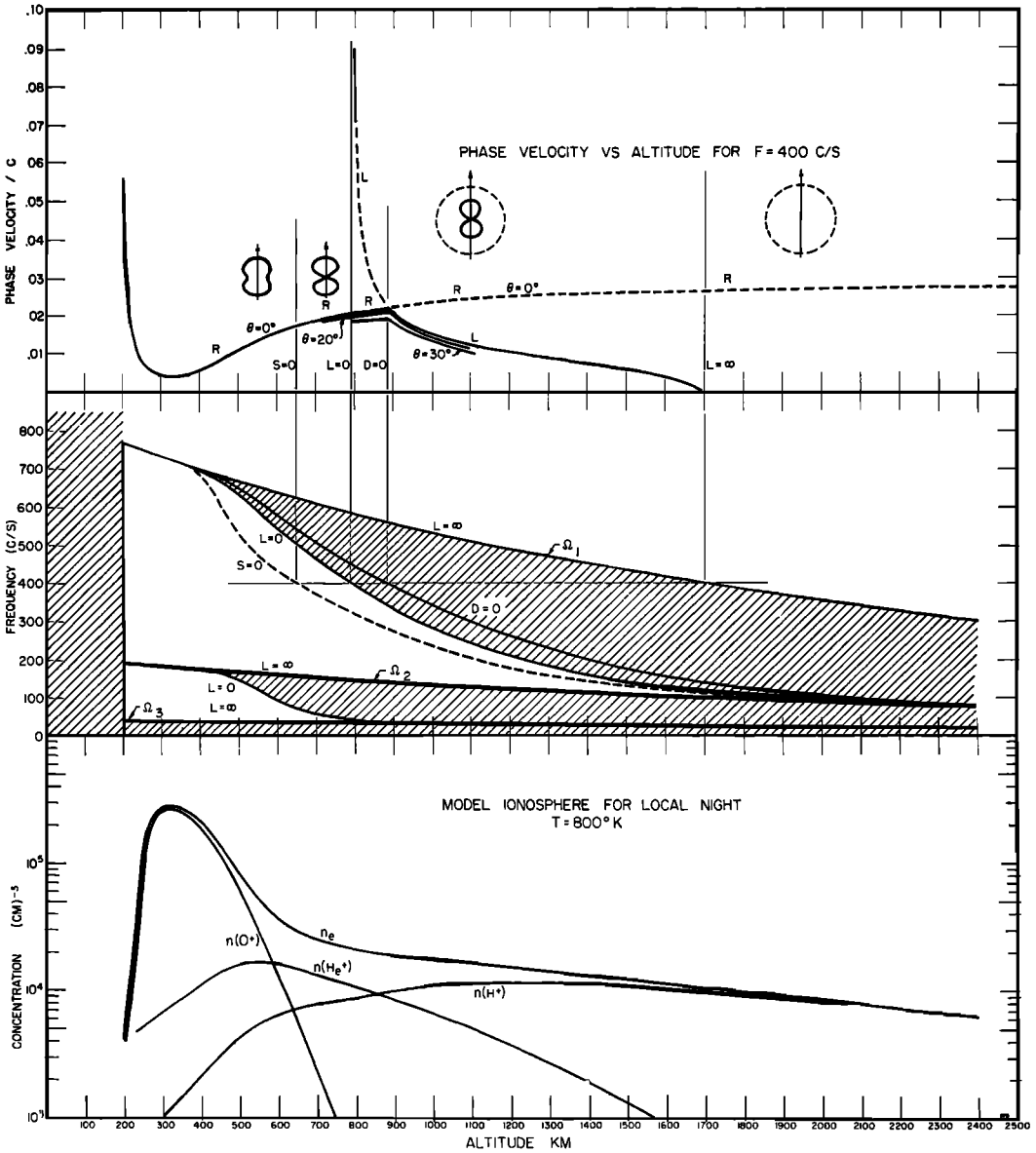

Fig. 10. The bounding surfaces and phase velocities at 400 cps for a model ionosphere.

$$\psi = \frac{1}{2K}\frac{d\rho_0}{dZ}\frac{1}{(\rho_0^2 - 1)}$$

$$= \frac{1}{2K}\frac{d}{dZ}\log\frac{\rho_0 - 1}{\rho_0 + 1} \qquad (24)$$

The dependent variables F_o and F_x are proportional to the electric fields for the two modes of propagation, here called ordinary (o) and extraordinary (x). The variable ψ which serves to couple the two equations is called the 'coupling parameter' and is a function of the wave polarization ρ. If the coupling parameter ψ is negligibly small, as it would be in a homogeneous medium where ρ is constant, (23) shows that the ordinary and extraordinary waves propagate independently. When $\rho = \pm 1$ the coupling parameter is infinite if $d\rho/dZ$ is finite. This condition is called 'critical coupling,' and the ordinary and extraordinary waves are strongly coupled in regions near $\rho = \pm 1$. *Clemmow and Heading* [1954] have formulated the coupled equations for the general case of oblique incidence [see *Budden*, 1961]. At VLF frequencies it is often sufficient to consider only vertical incidence in the lower ionosphere, since a wave incident on the ionosphere from below is refracted to nearly vertical incidence upon entering the ionosphere.

The WKB approximation is valid whenever the relative change in the refractive index per wavelength in the medium is small. This condition can be written [*Allis*, 1963]:

$$\frac{1}{n^2}\frac{dn}{dZ}\frac{c}{\omega} \ll 1 \qquad (25)$$

This condition is clearly violated near altitudes where $n = 0$. This situation also arises when we consider the coupled wave equations. At altitudes where $n \cong 0$ or dn/dZ is large, there is strong coupling between upgoing and downgoing waves [see *Clemmow and Heading*, 1954]. Physically this corresponds to reflection at any altitude where $n = 0$.

Since the wave polarization is always imaginary for a collisionless plasma, we conclude that the critical coupling condition $\rho_o = \rho_x = \pm 1$ cannot occur except for cases where ρ is undefined: ($\theta = 0$, $D = 0$), ($\theta = \pi/2$, $n_x^2 = P$). Therefore, we consider a plasma having collisions. For a plasma having collisions the quantities R, L, P, etc., become complex quantities when m_K is replaced by $m_K(1 + iZ_K)$. These quantities will be written as the sum of their real and imaginary parts, $R = R_0 + iR_1$, etc. If (5), (6), (7), and (8) are expanded to first order in the parameters Z_K, it is evident that the real part of these equations is the same as for the collisionless case. Then the surfaces $L_0 = 0$, $D_0 = 0$, etc., are the same as those discussed earlier in this section.

The critical coupling condition for a plasma having collisions can be stated in terms of the refractive index as follows, from (16):

$$n^2 = \rho RL/(\rho S - iD\cos\theta) \qquad (26)$$

It is evident that the critical coupling condition $\rho_x = \rho_o = \pm 1$ implies that $n_x = n_o$, or, from (12),

$$F^2 = H^2\sin^4\theta + 4P^2D^2\cos^2\theta = 0 \qquad (27)$$

where $H = H_0 + iH_1 = RL - PS$.

We now assume that the collision frequency is small enough that

$$P_1^2 \ll P_0^2 \quad R_1^2 \ll R_0^2 \quad L_1^2 \ll L_0^2 \quad H_1^2 \ll H_0^2$$

Omitting the second-order terms in small quantities, the real and imaginary parts of $F^2 = 0$ are:

$$H_0^2\sin^4\theta + 4P_0^2[D_0^2 - D_1^2]\cos^2\theta = 0 \qquad (28)$$

$$H_1H_0\sin^4\theta + 4P_0^2D_1D_0\cos^2\theta = 0 \qquad (29)$$

Equation 28 has a solution if and only if $-D_1 \leq D_0 \leq D_1$. Since D_1 is a small quantity, $D_1^2 \ll R_0^2 + L_0^2$, we see that critical coupling can occur only in the immediate neighborhood of the bounding surface $D_0 = 0$. Equations 28 and 29 can be solved for the critical coupling angle θ_c and for a critical coupling surface in parameter space.

$$\sin^2\theta_c\tan^2\theta_c = -4P_0^2D_1D_0/H_1H_0 \qquad (30)$$

$$\frac{D_1}{D_0} - \frac{D_0}{D_1} + \frac{H_0}{H_1} = g(\alpha, \beta, Z_K, \Lambda) = 0 \qquad (31)$$

Given α, β, Z_K, and Λ as functions of altitude, (31) can in principle be solved for the critical coupling altitude h_c. This altitude is very close to the altitude where $D = 0$, since $D_0^2 \leq D_1^2 \ll R_0^2 + L_0^2$. The values of α, β, Z_K, and Λ at altitude h_c can then be substituted into (30) to give the critical coupling angle θ_c.

Using the assumption that $H_1^2 \ll H_0^2$, we can write (31) as $D_0/D_1 = H_1/H_0$. Thus D_0 can be eliminated from (30) to give θ_c in terms of quantities which are easily estimated in the vicinity of $D_0 = 0$.

$$\sin^2 \theta_c \tan^2 \theta_c = \frac{-4P_0^2 D_1^2}{(R_0 L_0 - P_0 S_0)^2}\Bigg|_{D_0=0} \quad (32)$$

or

$$\theta_c^2 \simeq 2\,|D_1|/|R_0| \quad \text{if} \quad |R_0| \ll |P_0|$$

To illustrate the magnitude of the critical coupling angle, consider the $D = 0$ condition illustrated in Figure 9 at $\Lambda = 0.5$ and $\alpha = 0.8$. If we let $Z_1 = Z_2 = 10^{-2}$, then $\theta_c = 10.7°$. Thus, if collisions had been included, the wave-normal surfaces shown in Figure 9 at $D = 0$ would intersect for a cone of directions making an angle $\theta_c = 10.7°$ with the magnetic field. This cone will be called the critical coupling cone. Two characteristic waves propagating with **K** near the critical coupling cone are strongly coupled in the vicinity of $D = 0$.

Bounding surfaces, accessibility, and mode coupling for a model ionosphere. We now consider the propagation of a plane wave in a model ionosphere consisting of a horizontally stratified, neutral mixture of H^+, He^+, O^+, and electrons. The electron and ion density profiles that have been assumed for these calculations are shown in Figure 10*a*. This ionospheric model has been chosen as representative of a midlatitude ionosphere during local night. The ion density profiles have been sketched using the method of *N. Smith* [1964]. The equal concentration altitudes and the temperature agree with measurements taken by *Donley* [1963] during local night. The electron density profile is a qualitative model presented by *Bourdeau* [1963] for a local time of 0200. Figure 10*a* shows that each of the three ionic constituents (H^+, He^+, O^+) predominate in a different altitude region. The thickness of each of these regions is strongly dependent on the ionospheric temperature. For daytime conditions, the temperature is higher and the equal concentration points move to higher altitudes. In an experimental situation we can expect departures from the model assumed here, since the composition of the ionosphere depends on local time, latitude, season, etc. *Bourdeau* [1963] has given an excellent brief report on the composition of the ionosphere in which he has included references to many experimental studies.

Using the ion density profiles, the critical frequencies (cutoff, resonance, and crossover) have been plotted as a function of altitude in Figure 10*b*. Crossing a critical frequency in this plot corresponds to crossing a bounding surface in the parameter space discussed earlier in this section. The critical frequencies have been labeled by the corresponding equation for the bounding surface, $L = \infty$, $L = 0$, $S = 0$, and $D = 0$. The resonances labeled $L = \infty$ are the gyrofrequencies for H^+, He^+, and O^+. The critical frequencies labeled $L = 0$, $S = 0$, and $D = 0$ have been obtained from the bounding surface contours in Figures 7 and 8. At each altitude the ion density profiles give values for the relative concentrations α and β. For these values of α and β the roots of $L(\Lambda) = 0$, $S(\Lambda) = 0$, and $D(\Lambda) = 0$ can be found from the bounding surface contours. The critical frequencies are calculated from these roots using $\omega = \Lambda\Omega_1$.

From our earlier discussion it is evident that two real wave-normal surfaces exist in the crosshatched region between the cutoffs $L = 0$ and the resonances $L = \infty$. The crosshatched regions below Ω_1, Ω_2, and Ω_3 are the regions of propagation for the proton cyclotron wave, the He^+ cyclotron wave, and the O^+ cyclotron wave, respectively. In the regions not crosshatched, only one real wave-normal surface exists, and it is labeled R. It is important to note how the region of propagation for the ion cyclotron mode depends on the ion concentrations. If the ionospheric temperature were increased, the equal concentration points would move to higher altitudes and the regions of propagation for the ion cyclotron waves would move up correspondingly. It is evident that the predominant concentration of O^+ and He^+ in the lower ionosphere prevents the direct propagation of a lightning impulse to high altitudes entirely via the proton cyclotron mode.

We now consider what regions of the ionosphere are accessible to waves from a source having a fixed frequency and located, first, at high altitudes (above 2500 km) and, second, below the base of the ionosphere (200 km). To be specific we consider the frequency 400 cps. In Figure 10*c* the phase velocities for 400 cps are shown as a function of altitude. The dashed lines give the phase velocity at $\theta = 0$ for the fast mode (the outer wave-normal surface between $L = 0$ and $L = \infty$). The wave-normal surface (dashed) for this mode is a topological sphere and deforms in a continuous manner for all altitudes above $L = 0$. Thus, for a source of 400-cps downgoing waves, located at 2500 km, all altitudes above $L = 0$ are accessible via the mode having

the dashed wave-normal surface. At the cutoff $L = 0$, a down-going (fast) wave is reflected, since the refractive index goes to zero at this altitude (infinite phase velocity). At the altitude $D = 0$, the wave polarization changes sign and the label for the fast mode changes from R to L for down-going waves. For the moment, the possibility of mode coupling near $D = 0$ is ignored.

The solid lines in Figure 10*c* give the phase velocity at $\theta = 0$, 20°, and 30° for the slow mode (the inner wave-normal surface between $L = 0$ and $L = \infty$). The wave-normal surface for this mode (solid lines) is real for all altitudes below the resonance $L = \infty$ and deforms in a continuous manner over this range of altitudes. Thus, for a source of up-going waves located at the base of the ionosphere, all altitudes up to the resonance $L = \infty$ are accessible via the mode having phase velocities shown with solid lines in Figure 10*c*. For an up-going wave ($\theta \neq 0$) starting at the base of the ionosphere, the wave polarization ρ approaches zero in a continuous manner as the surface $D = 0$ is approached. At $D = 0$ the wave is linearly polarized in the (**K**, **B**) plane. As the surface $D = 0$ is crossed the wave polarization changes sign and the label for the slow wave-normal surface changes from R to L. Proceeding to higher altitudes, the wave propagates via the wave-normal surface labeled L, the proton cyclotron mode. In the next section it is shown that this proton cyclotron wave, between $D = 0$ and $L = \infty$, gives the proton-whistler traces described in section 2. As the proton cyclotron wave approaches the altitude where the proton gyrofrequency is equal to the wave frequency (400 cps), the phase velocity approaches zero and the wave energy is absorbed by the plasma.

From the phase velocity plots in Figure 10*c* and the discussion of wave-normal surfaces, it is evident that the wave-normal surfaces shown by solid lines touch but never intersect those shown by dashed lines. The wave-normal surfaces touch at the one point $D = 0$, $\theta = 0$. If we neglect mode coupling and consider the case $D = 0$, $\theta = 0$ to be of measure zero, we must conclude that regions of the ionosphere above $L = \infty$ (for a given frequency) are not accessible to up-going waves starting at the base of the ionosphere. Experimentally it is well known that regions high in the ionosphere, well above $L = \infty$, are accessible to waves having sources below the ionosphere. The observations of multiple hop whistlers having frequencies continuous down to 200 cps illustrate this point. We conclude that the mode coupling considerations discussed earlier are an important aspect of wave propagation in the ionosphere at these frequencies.

Using the equations in the preceding subsection, we can consider mode coupling for the special case of vertical incidence by including collisions in the model ionosphere presented in Figure 10. This case is of interest because waves incident on the base of a horizontally stratified ionosphere at angles not too oblique are refracted to nearly vertical incidence in the lower regions of the ionosphere where $n \gg 1$. For the frequency 400 cps in Figure 10*c*, the critical coupling angle at $D = 0$ has been calculated to be $\theta_c = 4.1°$. The collision frequencies at $D = 0$ were calculated using equations of *Nicolet* [1953], assuming Coulomb collisions. If we assume that the magnetic field is nearly vertical, then for vertical incidence the **K** vector at $D = 0$ is in a direction near the critical coupling cone and the two modes of propagation are strongly coupled near $D = 0$. Thus, an up-going wave originating at the base of the ionosphere generates a second wave of opposite polarization as it passes through the coupling region near $D = 0$. At altitudes above $D = 0$ these two wave propagate independently, one following the dashed phase velocity curve labeled R in Figure 10*c*, the other following the solid curve labeled L. The wave following the curve labeled L is an ion cyclotron wave and is absorbed at the resonance $L = \infty$. The other wave polarization follows the curve labeled R to altitudes well above $L = \infty$. From this example we see that mode coupling has made regions at high altitudes, above $L = \infty$, accessible to waves originating below the ionosphere. If the direction of the magnetic field had been such that $\theta \gg \theta_c$ at $D = 0$, then the coupling would have been much weaker, and regions above $L = \infty$ would have been essentially inaccessible.

For many cases of interest the assumption of vertical incidence does not apply, and we must consider mode coupling for oblique incidence. The coupled equations for oblique incidence have been investigated by *Clemmow and Heading* [1954] but have not been considered here. It seems likely that the critical coupling condition

$\rho_o = \rho_x = \pm 1$ may continue to apply in the general case and that the features illustrated here for vertical incidence will continue to appear.

5. Evidence that Proton Whistlers Are Ion Cyclotron Waves

In this section we consider the evidence supporting the claim that the proton whistlers described in section 2 are ion cyclotron waves.

Group propagation times for an impulse source below the ionosphere. The spectrograms in Figures 1 and 2 give the group propagation time from the source of the lightning impulse to the satellite as a function of frequency. Analytically this can be written as a line integral:

$$\tau(\omega) = \frac{1}{c} \int n_g \, ds \tag{33}$$

$$n_g = n + \omega(dn/d\omega) \tag{34}$$

We would like to calculate group propagation times for the model shown in Figure 10 and compare the results with spectrograms of proton whistlers. Since there is no way of knowing the exact ray path for the whistlers received by Injun 3, we must make a reasonable assumption about this path. We assume that along the ray path the **K** vector is always directed approximately along the magnetic field line, that is, we assume quasi-longitudinal propagation. For quasi-longitudinal propagation the refractive index is approximately $n^2 = R$ or $n^2 = L$, depending on whether the wave-normal surface is labeled R or L, respectively.

The group velocity c/n_g for longitudinal propagation ($\theta = 0$) is discontinuous at $D = 0$, where the refractive index changes from $n^2 = R$ to $n^2 = L$. For quasi-longitudinal propagation ($\theta \neq 0$) the group velocity changes continuously in the region very near $D = 0$ from that given using $n^2 = R$ to that given using $n^2 = L$. Therefore, for quasi-longitudinal propagation we approximate (33) by using the group refractive index for longitudinal propagation.

The group refractive index for longitudinal propagation is easily calculated from (5), (6), and (34) and is given by

$$n\, n_g = 1 - \tfrac{1}{2} \sum_K \frac{\pm \epsilon_K \pi_K^2 \Omega_K}{\omega(\omega \pm \epsilon_K \Omega_K)^2} \tag{35}$$

The plus signs apply when $n^2 = R$, and the minus signs apply when $n^2 = L$. If we assume that $n \gg 1$, then $n\, n_g \gg 1$ (since $n_g \geq 1$), and we can omit the 1 on the right side of (35). With this assumption, and omitting $1 + (\pi_e/\Omega_e)^2$ in (17) and (18), the group refractive index can be written in the form $n_g = \Gamma f(\Lambda, \alpha, \beta)$, where

$$\Gamma^2 = \pi_e^2/\Omega_1\Omega_e$$

The function $f(\Lambda, \alpha, \beta)$ is shown in Figure 11 for $\gamma = 0$ and various values of α. It is evident that n_g is a well-behaved function for the mode labeled R, whereas for the mode labeled L (the proton cyclotron mode) the group refractive index has an infinity at the proton gyrofrequency ($\Lambda = 1$) and at the cutoff frequency where $L(\Lambda) = 0$.

Group propagation times to various altitudes have been calculated for an impulse source below the ionosphere as a function of frequency. These integrations (equation 33) were performed for two cases, both assuming quasi-longitudinal propagation and using the model ionosphere in Figure 10*a*. For the first case, it is assumed that mode coupling does not occur. The group propagation time is the sum of the group propagation time via the right-hand polarized mode plus the group propagation time via the proton cyclotron mode after the polarization reversal at $D = 0$. For these calculations we assume that the wave crosses the $D = 0$ surface between Ω_2 and Ω_3 with no polarization reversal. Thus, we do not consider the He^+ cyclotron wave. Effects because of the He^+ cyclotron wave can occur only below the maximum frequency for which this mode can propagate, about 170 cps. For the second case, mode coupling was assumed to occur at $D = 0$. Above $D = 0$ two group propagation times must be calculated: a group propagation time for the right-hand polarized wave generated by the coupling near $D = 0$, and a group propagation time for the proton cyclotron wave which results from the polarization reversal at $D = 0$.

The calculated group propagation times are shown in Figure 12 for five altitudes, with and without mode coupling at $D = 0$. The effect of the infinite group refractive index at Ω_1 for the proton cyclotron wave is clearly evident. For the case without mode coupling the crossover frequency ω_{12} separates the proton cyclotron wave from the right-hand polarized wave. We see that there is a range of frequencies between

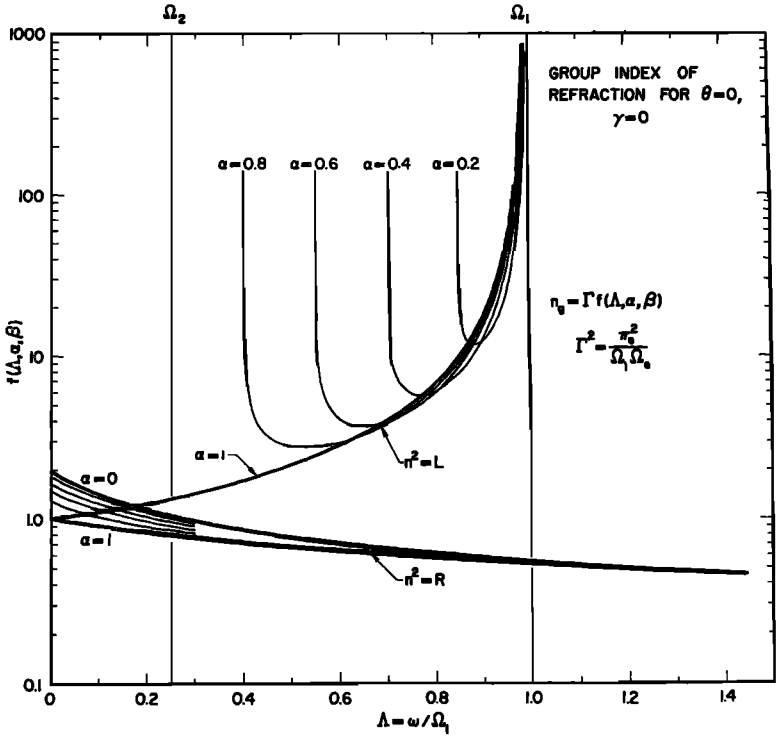

Fig. 11. The group refractive index for the R and L longitudinal waves.

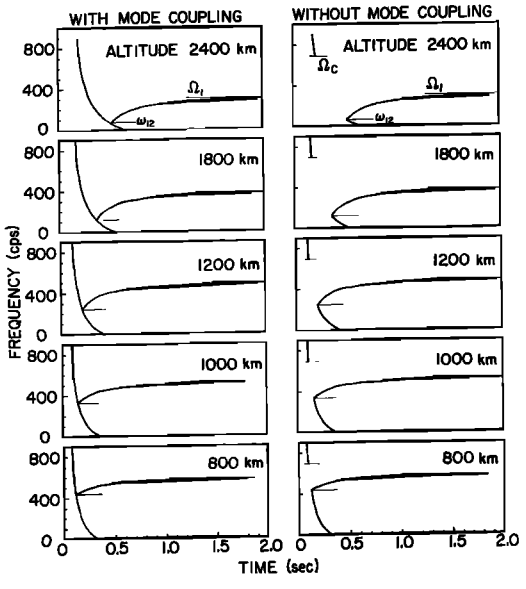

Fig. 12. Calculated propagation delay times for an impulse source below the base of the ionosphere.

Ω_1 and Ω_e which do not reach the altitude being considered. The cutoff frequency Ω_e is the maximum frequency along the path for which polarization reversal occurs. Thus, Ω_e is the proton gyrofrequency at the altitude where the bandwidth of the proton cyclotron wave goes to zero, about 690 cps for the model in Figure 10*b*.

Comparing the delay time curves with the spectrograms of proton whistlers in Figure 2, we make the following observations:

(a) The calculated delay time curves for the proton cyclotron wave are in good quantitative agreement with the proton-whistler traces observed in the Injun 3 data.

(*b*) The frequency ω_{12}, for which the electron whistler and the proton whistler are coincident in time, is the crossover frequency at the altitude of the satellite.

(c) The dependence of the ratio $\omega_{12}/\Omega_1 = \Lambda_{12}$ on altitude is qualitatively the same for the observed proton whistlers and the model calculations.

(d) The cutoff frequency Ω_c for the electron whistler, observed for some cases in the Injun 3 data, is the maximum frequency along the path for which polarization reversal can occur.

(e) The intensity of the electron-whistler trace between Ω_c and ω_{12} is determined by the strength of the coupling near $D = 0$.

These observations suggest that the simple quasi-longitudinal propagation model considered here is substantially correct and that the proton whistler is a proton cyclotron wave occurring via the polarization reversal at $D = 0$. For a propagation path which is not quasi-longitudinal, the essential features of the group delay times for up-going waves are not changed, because the polarization reversal condition does not depend on θ. The propagation path does, however, determine how close $\mathbf{K}$ is to the critical coupling cone near $D = 0$, and hence determines the intensity of the electron-whistler trace between Ω_c and ω_{12}.

If the polarization reversal between Ω_2 and Ω_3 were included, a He^+ cyclotron wave would occur between Ω_2 and ω_{23} and would be qualitatively similar to the proton whistler between Ω_1 and ω_{12}. It has been impossible to positively identify any trace in the Injun 3 data as a He^+ cyclotron wave because of the poor frequency response below 200 cps.

Because these new whistlers are ion cyclotron waves, we have called them ion cyclotron whistlers. Since an ion cyclotron mode exists for a band of frequencies below each ion gyrofrequency, the ion cyclotron whistler traces can be named after the type of ion whose gyrofrequency gives the upper frequency limit of propagation for that wave. Thus, we arrived at the term proton whistler. Since the right-hand polarized mode can propagate for all frequencies below the electron gyrofrequency, we have called the right-hand polarized component of a whistler an electron whistler.

Determining the relative concentration α from the crossover frequency. We now consider how proton whistlers can be used to estimate the concentration parameter $\alpha = n(H^+)/n_e$. A proton whistler provides two parameters which can be measured with good accuracy. These parameters are the proton gyrofrequency and the crossover frequency. The proton gyrofrequency is of little interest, since it can be calculated with excellent accuracy in the lower ionosphere. The crossover frequency is a function of the parameters α and β, and can be used to obtain information about these parameters. The crossover frequency is the solution of $D(\Lambda, \alpha, \beta) = 0$. Substituting a known value of $\Lambda_{12} = \omega_{12}/\Omega_1$ into this equation gives an equation for α as a function of β. This equation is plotted in Figure 7, and we see that, given Λ_{12}, α is nearly independent of β. Thus, by measuring Λ_{12} the relative concentration $\alpha = n(H^+)/n_e$ at the satellite can be estimated using the contour plot in Figure 7.

Using proton whistlers, α values have been determined for a variety of altitudes during local night (2300–0500), and during local day (1100–1400). These α values are presented in Figure 13 and will be compared with existing knowledge of the composition of the ionosphere as further evidence that the explanation of the proton whistler given here is correct.

The experimental values of α shown in Figure 13 for local night decrease with decreasing altitude, as would be expected from an equilibrium model of the ionosphere. A curve giving α as a function of altitude for the nighttime model ionosphere used earlier is shown for comparison with these data. The temperature for the nighttime model ionosphere is 800°K. The agreement between the proton-whistler data and this model is reasonable, although the fractional concentration of H^+ is somewhat larger than that used in the model.

The experimental α values obtained during local daytime are shown in Figure 13. The daytime α values are compared with a model ionosphere after Anderson and Francis [*Walt*, 1964]. Again the α values are somewhat larger than those given by the model. By comparing the daytime and nighttime data we see that α values of about 0.6 occur at higher altitudes during the local day (about 1300 km) than during the local night (about 700 km). This diurnal variation is as would be expected, since the equal concentration points for the local day are at higher altitudes than those for local night. The experimental fact that proton whistlers are observed at lower altitudes and more frequently during local night than during local day (see Figure 4) is also in agreement with the expected diurnal variation of the bounding surfaces in Figure 10*b*.

The occurrence of proton whistlers in latitude is shown in Figure 5. It seems likely that the cutoff at high latitudes (above 55° latitude)

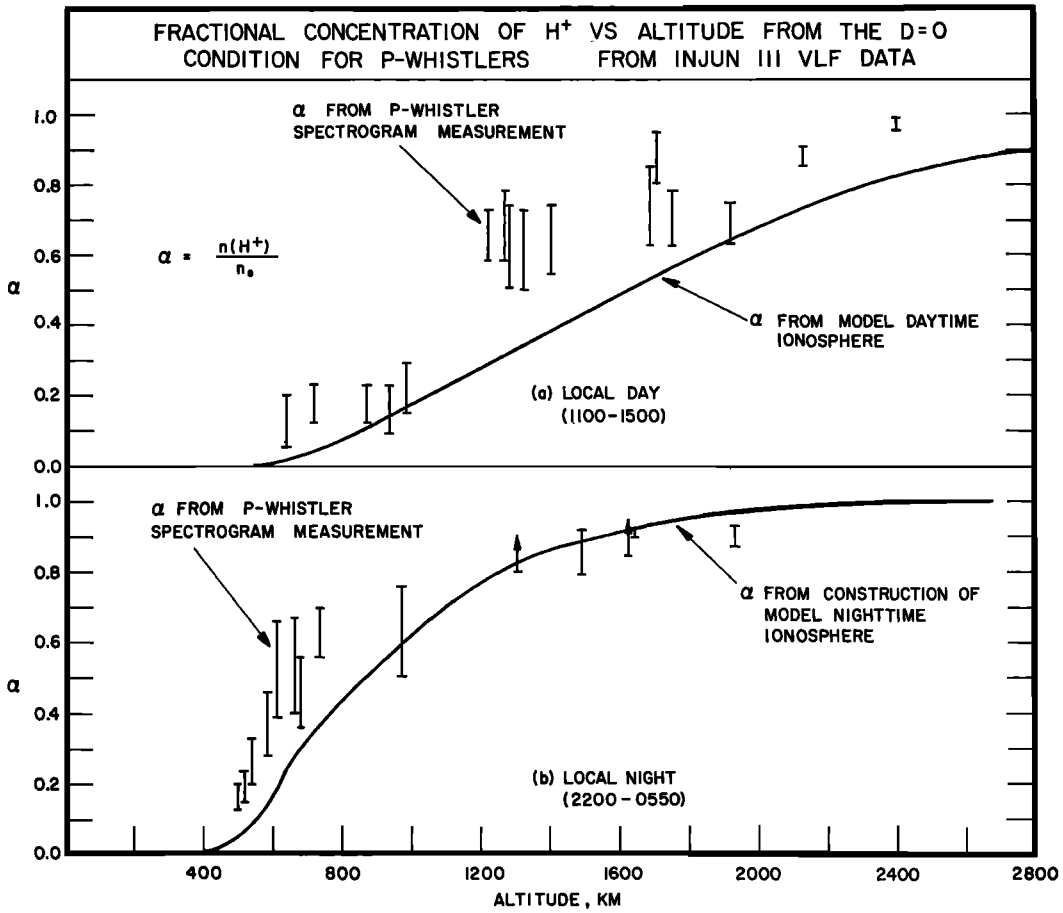

Fig. 13. Values of α determined from the crossover frequency.

may be due in part to the increasing plasma temperature near the auroral zone and the resultant higher altitudes of the helium and oxygen ion layers.

6. Conclusion

In this paper an experimental study of the proton whistler was presented and an explanation of this new effect given. It was proposed that the proton whistler is simply a dispersed form of the original lightning impulse and that the dispersion can be explained by considering the effects of ions on the propagation of an electromagnetic wave in the ionosphere.

A model ionosphere consisting of three ions (H^+, He^+, and O^+) was introduced, and a physically reasonable ion density profile was assumed. For this model the propagation of waves having frequencies of the order of the ion gyrofrequencies was investigated. It was found that for a band of frequencies below each ion gyrofrequency a left-hand polarized wave can propagate. These waves have been called the proton cyclotron wave, the He^+ cyclotron wave, and the O^+ cyclotron wave, corresponding to the type of ion whose gyrofrequency provides the upper frequency limit for each band. This nomenclature follows that of Stix, who called all these waves ion cyclotron waves. The bandwidths of the ion cyclotron waves depend on the relative concentrations of the ions. Because the bandwidths depend on the ion concentrations, the ion cyclotron wave can propagate only for certain ranges of frequencies and altitudes in the ionosphere. The proton whistler was interpreted as a proton cyclotron wave.

It was found that in the ionosphere, at an

altitude determined by the equation $D = 0$, an upgoing right-hand polarized wave changes polarization and becomes an ion cyclotron wave. Polarization reversal appeared as a crucial feature in the explanation of the proton whistler because, as was shown, the predominance of O^+ and He^+ in the lower ionosphere prevents the direct propagation of a lightning impulse to the satellite entirely via the proton cyclotron mode.

An investigation of mode coupling for the special case of vertical incidence showed that if collisions are included the condition $D = 0$ plays an important part in mode coupling theory. It was found that for a small cone of directions around **B** and at an altitude very near the polarization reversal altitude ($D = 0$) the two modes are critically coupled. In the ionosphere, if an upgoing right-hand polarized wave crosses the polarization reversal altitude ($D = 0$) with **K** near the critical coupling cone, a wave of opposite polarization is generated, and above $D = 0$ there are two waves of opposite polarization, each propagating independently. One of these, the ion cyclotron wave, suffers a destructive transition when the wave frequency becomes equal to the ion gyrofrequency. The other wave above $D = 0$, a right-hand polarized wave, can propagate to considerably higher altitudes.

To show how these results apply to the proton whistler, the group propagation times to various altitudes were calculated as a function of frequency, assuming an impulse source below the base of the ionosphere. The propagation was assumed to be quasi-longitudinal. The calculated delay time curves agreed well with spectrograms of proton whistlers and strongly suggested that proton whistlers are ion cyclotron waves. The propagation theory also suggested that a He^+ cyclotron wave could result from a lightning stroke via a process identical to that by which the proton whistler occurs. If and when such an effect is found, it might very well be called a helium whistler. Similar effects might even be expected below the O^+ gyrofrequency. Since these new whistlers are ion cyclotron waves, we have called them ion cyclotron whistlers.

The theory presented here has provided a method of determining the concentration parameter $\alpha = n(H^+)/n_e$ at the satellite from proton-whistler data. Experimental values of α, determined by using proton-whistler data, were shown to be in reasonable agreement with present models of the ion concentration in the ionosphere. This application of propagation theory to a multicomponent plasma shows that radio measurements by satellites at ion gyrofrequencies can provide an excellent diagnostic tool for studying the composition of the ionosphere.

Acknowledgments. We would like to express our appreciation to Professor J. A. Van Allen for his interest in this study and to Professor R. A. Helliwell for his helpful discussions. We also thank Professor J. F. Curtis for the use of the Speech Pathology and Audiology Department's Sonograph.

The research at the State University of Iowa was supported by the Office of Naval Research under contracts N9onr 93803 and Nonr 1509(06) and at Stanford University by the National Aeronautics and Space Administration grant NsG 174-61.

References

Allis, W. P., Waves in a plasma, *Mass. Inst. Tech. Res. Lab. Electronics Quart. Prog. Rept., 54*(5), 1959.

Allis, W. P., S. J. Buchsbaum, and A. Bers, *Waves in Anisotropic Plasmas,* Massachusetts Institute of Technology Press, Cambridge, 1963.

Auer, P. L., H. Hurwitz, Jr., and R. D. Miller, Collective oscillations in a cold plasma, *Phys. Fluids, 1,* 501, 1958.

Bourdeau, R. E., Rocket and satellite investigations of the ionosphere (abstract), *Trans. Am. Geophys. Union, 44,* 443, 1963.

Brice, N., Discrete very low frequency emissions from the upper atmosphere, *SEL 64-088, Radiosci. Lab., Stanford Electronics Lab., Stanford Univ.,* Stanford, Calif., August 1964.

Buchsbaum, S. J., Resonance in a plasma with two ion species, *Phys. Fluids, 3,* 418, 1960.

Budden, K. G., *Radio Waves in the Ionosphere,* Cambridge University Press, Cambridge, England, 1961.

Clemmow, P. C., and J. Heading, Coupled forms of the differential equations governing radio propagation in the ionosphere, *Proc. Cambridge Phil. Soc., 50,* 319, 1954.

Clemmow, P. C., and R. F. Mullaly, The dependence of the refractive index in magnetoionic theory on the direction of the wave normal, in *The Physics of the Ionosphere,* p. 340, The Physical Society, London, 1955.

Donley, John L., Experimental evidence for a low ion-transition altitude in the upper nighttime ionosphere, *J. Geophys. Res., 68,* 2058–2060, 1963.

Försterling, K., Über die Ausbreitung elektromagnetischer wellen in einem magnetisierten me-

dium bei senkrechter incidenz, *Hochfr. Elek., 59,* 110, 1942.

Gallet, R., The very low-frequency emissions generated in the earth's exosphere, *Proc. IRE, 47* (2), 211–231, Feb. 1959.

Gintsburg, M. A., Low-frequency waves in multicomponent plasma, *Geomagnetism Aeronomy, 3,* 610–614, 1963.

Gurnett, D. A., Very low frequency electromagnetic emissions observed with the ONR/SUI satellite Injun 3, *SUI Rept. 63-27, State Univ. Iowa,* Iowa City, August 1963.

Hines, C. O., Heavy-ion effects in audio-frequency radio propagation, *J. Atmospheric Terrest. Phys., 11,* 36–42, 1957.

Nicolet, M. J., The collision frequency of electrons in the ionosphere, *J. Atmospheric Terrest. Phys., 3,* 200–211, 1953.

Smith, Newbern, A simple method of drawing ion-density profiles in an atmosphere in diffusive equilibrium, paper 4-4-7, 1964 Spring URSI Meeting, Washington, D. C., April 15–18, 1964.

Smith, R. L., et al., An ion gyrofrequency phenomenon observed in satellites, *Nature, 204* (4955), 274, Oct. 17, 1964.

Smith, R. L., and Neil Brice, Propagation in multicomponent plasmas, *J. Geophys. Res., 69,* 5029–5040, 1964.

Stix, T. H., *The Theory of Plasma Waves,* McGraw-Hill Book Company, New York, 1962.

Walt, Martin, The effects of atmospheric collisions on geomagnetically trapped electrons, *J. Geophys. Res., 69,* 3947–3958, 1964.

Yakimenko, V. L., Oscillations in a cold plasma containing two ion species, *Sov. Phys., Tech Phys., 7* (2), 117–124, 1962.

(Manuscript received January 8, 1965.)