Progress In Electromagnetics Research, *PIER 13, 87–114, 1996*

A UNIFIED THEORY OF IONOSPHERIC PROPAGATION OF SHORT RADIO WAVES WITH SPECIAL EMPHASIS ON LONG-DISTANCE PROPAGATION

B. W. Lu

1. Introduction

The use of short radio waves to achieve long-distance communication and broadcasting has had a long history. There are however still some vague ideas about the mode of long-distance ionospheric propagation of short radio waves. With regard to long-distance propagation between two points on the ground, the distance obtained by the "reflection" of the F2 layer of the ionosphere at grazing incidence, which is about 4000km, is considered the maximum span of a single hop. If the communication distance is longer than the maximum span of a single hop, it is usually considered that the propagation has to be effected by multiple hops with the assistance of the reflection by the earth's surface. But the multiple-hop propagation can not be the only mode of propagation for distances longer than about 4000km. If the multiple-hop propagation were the only mode, then, supposing that we increase the distance from a value less than the maximum span of a single hop to higher values, we would find that the maximum usable frequency (MUF) of the communication circuit first increases with the distance as it should but would suddenly drop to a lower value while the maximum span of a single hop is being crossed, since according to the multiple-hop hypothesis, the mode of propagation would change from a single hop to a double hop. The same situation would occur as the distance increases from a double hop to higher values, and so on. Such phenomena have not been observed in practice. The multiple-hop mode of propagation can not be the predominant mode of long-distance propagation because practice has shown that the actual MUF values of long-distance short radio-wave communication circuits are generally higher than the values predicted by the multiple-hop hypothesis. An empirical method of MUF prediction, the so-called control-point method which has long been in use, has been found to yield results agreeing fairly well with practice, but no explanation has been found in the literature. Apparently, the predominant mode must be one which depends solely on ionospheric refraction without the assistance of ground reflection. In the early fifties, the author of the present paper proposed a theory called the theory of gliding mode propagation (unpublished) which can explain the control-point method of MUF prediction [1,2]. As to the long-distance ionospheric propagation of short radio waves between a point on the ground and an orbiting satellite, it is usually considered infeasible or impractical. In 1961, the first manned satellite was successfully launched by the former Soviet Union. It was noticed that the manned satellite carried a transmitter with a frequency of about 20MHz, apparently for the purpose of utilizing or of testing the possibility of utilizing the ionospheric refraction to achieve long-distance ionospheric propagation between the ground station and the manned satellite. The author realized that such propagation is feasible, and that it can also be explained by the theory of gliding mode propagation. To assess the implications of the Soviet pioneering event, a symposium was organized by the Bureau of New Technology of the Chinese Academy of Sciences. The author attended the symposium and presented a report (now unclassified), a part of which gave an exposition of the theory of gliding mode propagation. In the present paper, the relevant part of that report is reviewed and further research results are introduced.

2. Solution by the Ray Treatment

For the sake of simplicity and for the purpose of quick grasping the main aspects of the problem, we assume for the time being that the effect of the earth's magnetic field can be neglected and that the refractive index of the ionosphere can be regarded as a function of the radial distance r alone. Under such conditions, the refractive index n is given by

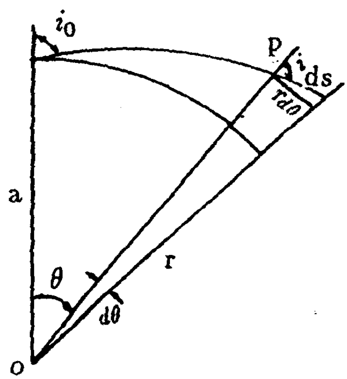

Figure 1. Depiction showing an element ds of the ray trajectory.

$$n^2 = 1 - \frac{N}{N_{max}} \frac{f_o^2}{f^2} \tag{1}$$

where N is the electron density, N_{max} the maximum value of N, f the wave frequency and f_o the plasma frequency of the ionosphere. In Figure 1, an element of the ray trajectory is

$$ds = \sqrt{dr^2 + r^2 d\Theta^2} = \sqrt{r^2 + r'^2} d\Theta \tag{2}$$

where $r' = \frac{dr}{d\Theta}$. According to the Fermat principle,

$$\int nds = \int n(r)\sqrt{r^2 + r'^2} d\Theta = \text{extremum} \tag{3}$$

Let

$$F(r, r') = n(r)\sqrt{r^2 + r'^2} \tag{4}$$

Applying the variational principle, we have

$$\frac{d}{d\Theta}\frac{\partial F}{\partial r'} = \frac{\partial F}{\partial r} \tag{5}$$

From (5), it can easily be shown that

$$r\frac{dn}{dr}+n\left[2-\frac{r(r+r'')}{r^2+r'^2}\right]=\frac{d}{r}(nr)+n\frac{r'^2+rr''}{r^2+r'^2}=0 \tag{6}$$

If $r'\neq 0$, by multiplying (6) by $\frac{rr'}{\sqrt{r^2+r'^2}}$, we get

$$\frac{d}{d\Theta}\left(\frac{nr^2}{\sqrt{r^2+r'^2}}\right)=0 \tag{7}$$

That is, along the ray trajectory,

$$\frac{nr^2}{\sqrt{r^2+r'^2}}=const \tag{8}$$

Since

$$\sin i=r\frac{d\Theta}{ds}=\frac{r}{\sqrt{r^2+r'^2}} \tag{9}$$

Eq. (8) is equivalent to the familiar Snell's law

$$nr\sin i=\text{const.}=a\sin i_0 \tag{10}$$

where i is the angle of incidence formed at an arbitrary point P on the ray trajectory between the radial vector $\vec{OP}$ and the ray element vector $\mathrm{d}\,\vec{s}$, as shown in Figure 1, i_0 the angle of incidence at the earth's surface. For the ionosphere, nr generally has at least one minimum. Suppose that there are no discontinuous surfaces. In the layer considered, there are three possible modes of propagation:

(1) If $r'\neq 0$ and

$$\frac{a\sin i_0}{(nr)_{min}}>1 \tag{11}$$

we have 'reflection', that is, bending back toward the earth by refraction.

(2) If $r'\neq 0$ and

$$\frac{a\sin i_0}{(nr)_{min}}<1 \tag{12}$$

we have penetration.

(3) If $r'=0$, that is $r=const$, from (6) we have,

$$\frac{d}{dr}(nr)=0 \tag{13}$$

Therefore, nr is an extremum, either a minimum or a maximum. According to (10), we should have $i = \frac{\pi}{2}$ at the extremum of nr. Here we have to distinguish between two different cases. If the condition $r' = 0$ were stable, then it would remain to be true and would get perpetual roundearth propagation. If the condition $r' = 0$ were unstable, then, supposing that the round-earth ray trajectory is slightly disturbed, it would deviate further with the result that the outcome would become chaotic. It is therefore interesting to examine whether nr being a maximum and being a minimum are respectively stable or unstable. Suppose that the round-earth ray trajectory at the altitude $r = r_m$ of nr extremum such that

$$\left\{\begin{aligned} r =& r_m + \delta r \\ \frac{dn}{dr} =& (\frac{dn}{dr})_{r_m} + (\frac{d^2 n}{dr^2})_{r_m} \delta r \\ r' =& r'_m + \delta r' = \delta r' \\ r' =& r''_m + \delta r'' \end{aligned}\right\} \tag{14}$$

Neglecting small quantities of the second and higher orders, we have

$$\left[\frac{d^2}{dr^2}(nr)\right]_m \delta r = \frac{n(r_m)}{r_m} \delta r'' \tag{15}$$

At the altitude of $(nr)_{max}$, $[\frac{d^2}{dr^2}(nr)]_{r_m} < 0$, and therefore $\delta r''$ and δr have opposite signs, consequently, once the round-earth ray trajectory is slightly deviated from the altitude of round-earth propagation, it tends to bend toward this altitude, and as a result the propagation is stable. On the other hand, at the altitude of $(nr)_{min}$, $[\frac{d^2}{dr^2}(nr)]_{r_m} > 0$, and therefore $\delta r''$ and δr have the same sign, consequently, once the round-earth ray trajectory is slightly deviated from the altitude of the round-earth propagation, it tends to bend away from this altitude, and as a result, the propagation is unstable. The $(nr)_{max}$ and $(nr)_{min}$ are respectively associated with the trough and the peak of the electron-density profile of the ionosphere. The trough of the ionosphere is usually not pronounced and therefore the attempt to use $(nr)_{max}$ to achieve round-earth propagation is not advisable. We are interested in the $(nr)_{min}$ associated with the peak of the F2 layer. If an incident beam has a very narrow portion the component rays of which just reach the close vicinity of the altitude of $(nr)_{min}$, $r = r_m$, essentially tangential to the circular curve $r = r_m$, these rays tend to propagate along this circular curve for long distances, and at the same time they tend to diverge from this curve both upward and downward as they travel along like in a leaking waveguide. The author, therefore, called such propagation gliding mode propagation. The same problem can also be attacked more in detail in the following way, when ground reflection is not involved, we have for the following three cases the expressions for the angles subtended at the earth's center by the transmitting and receiving points. From Figure 1,

$$d\Theta = \tan i \frac{dr}{r} = a \sin i_0 \frac{1}{\sqrt{n^2 r^2 - a^2 \sin^2 i_0}} \frac{dr}{r} \tag{16}$$

Case 1: For propagation between two points on the ground, the angle subtended by the two end points at the earth's center is given by

$$\Theta_D = 2a \sin i_0 \int_a^{r_1} \frac{1}{\sqrt{n^2 r^2 - a^2 \sin^2 i_0}} \frac{dr}{r} \tag{17}$$

where the upper limit of integration r_1 is the value of r which makes the denominator of the integrand $\sqrt{n^2 r^2 - a^2 \sin^2 i_0}$ equal to 0, i.e.,

$$n(r_1) r_1 - a \sin i_0 = 0 \tag{18}$$

Case 2: For propagation between a point on the ground and an orbiting satellite which is below the altitude of $(nr)_{min}$, the angle subtended is given by

$$\begin{aligned} \Theta_{s1} = 2a \sin i_0 &\int_a^{r_1} \frac{1}{\sqrt{n^2 r^2 - a^2 \sin^2 i_0}} \frac{dr}{r} \\ - a \sin i_0 &\int_a^{r_s} \frac{1}{\sqrt{n^2 r^2 - a^2 sin^2 i_0}} \frac{dr}{r} \end{aligned} \tag{19}$$

where r_s is the radial distance of the satellite.

Case 3: For propagation between a point on the ground and an orbiting satellite which is above the altitude of $(nr)_{min}$, the angle subtended is given by

$$\Theta_{s2} = a \sin i_0 \int_a^{r_s} \frac{1}{\sqrt{n^2 r^2 - a^2 \sin^2 i_0}} \frac{dr}{r} \tag{20}$$

All the integrals (17), (19) and (20) have the common integrand the denominator of which is 0 at $r = r_1$ as shown in (18). The portion of the integration path in the vicinity of $r = r_1$ gives the main contribution to the value of the integrals, the flatter the portion of the ray trajectory in the vicinity of $r = r_1$, the greater the contribution. If the angle of incidence i_0 is such that the ray trajectory reaches the altitude of $(nr)_{min}$, i.e., at $r = r_m$, and becomes tangential to the curve of nr, then, letting i_{0m} denote this particular angle of incidence, we have

$$n(r_m)r_m - a\sin i_{0m} = 0 \tag{21}$$

and in the vicinity of $r = r_m$

$$\begin{aligned} nr - a\sin i_{0m} &= n(r_m)r_m - a\sin i_{0m} + \left[\frac{d}{dr}(nr)\right]_{r_m}(r - r_m) \\ &\quad + \frac{1}{2}\left[\frac{d^2}{dr^2}(nr)\right]_{r_m}(r - r_m)^2 + \ldots \\ &= \frac{1}{2}\left[\frac{d^2}{dr^2}(nr)\right]_{r_m}(r - r_m)^2 + \ldots \end{aligned} \tag{22}$$

If the path of integration contains the point $r = r_m$, the integral $\int \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r}$ is infinite. Infinite subtended angle at the earth's center means round-earth propagation, but since the propagation is unstable, the rays diverge and become chaotic. From (17), (19) and (20), we can show that Θ_D, Θ_{s1}, and Θ_{s2}, can attain any finite positive number greater than 2π if nr has a minimum without requiring to know the specific electron-density profile. For this purpose, we need to consider only the gliding portion of the ray trajectory.

(1) For the case of propagation between two points on the ground and the case of propagation between a point on the ground and an orbiting satellite below $r = r_m$. Remembering that $n^2(r_1)r_1^2 - a^2\sin^2 i_0 = 0$ and $n^2(r_m)r_m^2$ - $a^2\sin^2 i_{0m} = 0$, let i_0 be an arbitrary positive value which is very close to but slightly greater than i_{0m} so that

$$0 < a(\sin i_0 - \sin i_{0m}) \ll 1 \tag{23}$$

In the neighborhood of $r = r_m$ with $r \le r_m$,

$$\sqrt{n^2r^2 - a^2\sin^2 i_0} \simeq \sqrt{2a\sin i_0}\sqrt{\tfrac{1}{2}\left[\tfrac{d^2(nr)}{dr^2}\right]_{r_m}(r_m - r)^2 - a(\sin i_0 - \sin i_{0m})} \tag{24}$$

where $\left[\frac{d^2(nr)}{dr^2}\right]_{r_m}$ is a positive quantity. We know that

$$0 = n(r_1)r_1 - a\sin i_0 \simeq \frac{1}{2}\left[\frac{d^2(nr)}{dr^2}\right]_{r_m}(r_m - r_1)^2 - a(\sin i_0 - \sin i_{0m}) \tag{25}$$

hence,

$$r_1 = r_m(1-\epsilon) \tag{26}$$

where

$$\epsilon = \frac{1}{r_m}\sqrt{\frac{2a(\sin i_0 - \sin i_{0m})}{\left[\frac{d^2(nr)}{dr^2}\right]_{r_m}}} \tag{27}$$

is an arbitrary small positive number. Let p be an arbitrary finite positive number such that $p \ll \frac{1}{\epsilon}$ and at the same time $p < \frac{1}{\epsilon}(1-\frac{a}{r_m})$. In (17) and (19), the integral concerned is $\int_a^{r_1} \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r}$ and let us divide it into two parts as

$$\begin{aligned}\int_a^{r_1} \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r} = &\int_a^{r_m(1-p\epsilon)} \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r} \\ &+ \int_{r_m(1-p\epsilon)}^{r_m(1-\epsilon)} \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r}\end{aligned} \tag{28}$$

The second integral on the right is the one which is relevant to the gliding portion of the ray trajectory. This integral can be written

$$\begin{aligned}\int_{r_m(1-p\epsilon)}^{r_m(1-\epsilon)} \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r} \simeq &\frac{1}{r_m\sqrt{a\sin i_0\left[\frac{d^2(nr)}{dr^2}\right]_{r_m}}} \\ &\times \int_{(1-p\epsilon)}^{(1-\epsilon)} \frac{d(\frac{r}{r_m})}{\sqrt{(1-\frac{r}{r_m})^2 - \epsilon^2}}\end{aligned} \tag{29}$$

Letting $1 - \frac{r}{r_m} = \epsilon\sec\xi$, $d(\frac{r}{r_m}) = -\epsilon\sec\xi\tan\xi d\xi$, we have

$$\int_{(1-p\epsilon)}^{(1-\epsilon)} \frac{d(\frac{r}{r_m})}{\sqrt{(1-\frac{r}{r_m})^2-\epsilon^2}} = \int_0^{\sec^{-1} p} \sec\xi d\xi$$
$$= \ln(\sec\xi + \tan\xi)|_0^{\sec^{-1} p}$$
$$= \ln(p+\sqrt{p^2-1}) \tag{30}$$

This can be any finite positive number according to value of ϵ which in turn depends on the choice of i_0 with respect to i_{0m}.

(2) For the case of propagation between a point on the ground and an orbiting satellite above $r = r_m$. In this case, $i_0 \leq i_{0m}$. In the neighborhood of $r = r_m$ with $r \geq r_m$, it can easily be shown that

$$\sqrt{n^2r^2 - a^2\sin^2 i_0} \simeq \sqrt{2a\sin i_0}\sqrt{\frac{1}{2}\left[\frac{d^2(nr)}{dr^2}\right]_{r_m}(r_m-r)^2 + a(\sin i_{0m} - \sin i_0)}$$
$$= r_m\sqrt{a\sin i_0\left[\frac{d^2(nr)}{dr^2}\right]_{r_m}}\sqrt{(\frac{r}{r_m}-1)^2+\epsilon^2} \tag{31}$$

where

$$\epsilon = \frac{1}{r_m}\sqrt{\frac{2a(\sin i_{0m} - \sin i_0)}{\left[\frac{d^2(nr)}{dr^2}\right]_{r_m}}} \tag{32}$$

The integral of (20) can be divided into three parts as

$$\int_a^{r_s} \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r} = \int_a^{r_m(1+\epsilon)} \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r}$$
$$+ \int_{r_m(1+q\epsilon)}^{r_s} \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r}$$
$$+ \int_{r_m(1+\epsilon)}^{r_m(1+q\epsilon)} \frac{1}{\sqrt{n^2r^2 - a^2\sin^2 i_0}}\frac{dr}{r} \tag{33}$$

where $q \ll \frac{1}{\epsilon}$ and at the same time $q < \frac{1}{\epsilon}(\frac{r_s}{r_m} - 1)$. The last integral on the right is the one which is relevant to the gliding portion of the ray trajectory. To evaluate this integral, letting $\frac{r}{r_m} - 1 = \epsilon \tan\xi$, $d(\frac{r}{r_m}) = \epsilon \sec^2 \xi d\xi$, we have

$$
\begin{aligned}
&\int_{r_m(1+\epsilon)}^{r_m(1+q\epsilon)} \frac{1}{\sqrt{n^2 r^2 - a^2 \sin^2 i_0}} \frac{dr}{r} \\
&\simeq \frac{1}{r_m} \frac{1}{\sqrt{a \sin i_0 \left[\frac{d^2(nr)}{dr^2}\right]_{r_m}}} \int_{r_m(1+\epsilon)}^{r_m(1+q\epsilon)} \frac{d(\frac{r}{r_m})}{\sqrt{(\frac{r}{r_m} - 1)^2 + \epsilon^2}} \\
&= \frac{1}{r_m} \frac{1}{\sqrt{a \sin i_0 \left[\frac{d^2(nr)}{dr^2}\right]_{r_m}}} \int_{\frac{\pi}{4}}^{\tan^{-1} q} \sec\xi d\xi \\
&= \frac{1}{r_m} \frac{1}{\sqrt{a \sin i_0 \left[\frac{d^2(nr)}{dr^2}\right]_{r_m}}} \left[\ln(q + \sqrt{q^2+1}) - ln(1+\sqrt{2})\right]
\end{aligned}
\tag{34}
$$

This can be any finite positive number according to the value of ϵ which in turn depends on the choice of i_0 with respect to i_{0m}.

It is interesting to examine what range of frequency is suitable for gliding mode propagation, we know that at $r = r_m$,

$$
\begin{aligned}
\frac{d}{dr}(nr) &= \frac{d}{dr}\left[r\sqrt{1 - \frac{f_0^2}{f^2}\frac{N(r)}{N_{max}}}\right] \\
&= \frac{1}{\sqrt{1 - \frac{f_0^2}{f^2}\frac{N(r)}{N_{max}}}} \left[1 - \frac{f_0^2}{f^2}\frac{N(r)}{N_{max}} - \frac{f_0^2}{f^2}\frac{r}{2}\frac{d}{dr}\left(\frac{N(r)}{N_{max}}\right)\right] \\
&= 0
\end{aligned}
\tag{35}
$$

i.e.,

$$
\frac{d}{dr}\left(\frac{N(r)}{N_{max}}\right) + \frac{2}{r}\frac{N(r)}{N_{max}} = \frac{2}{r}\frac{f^2}{f_0^2} \tag{36}
$$

This equation was derived specifically for $r = r_m$, but generally it can be approximately applied to the close vicinity of $r = r_m$. To solve the equation, let

$$
\frac{N(r)}{N_{max}} = \frac{1}{r^2} u(r) \tag{37}
$$

then,

$$\frac{du}{dr} = 2\frac{f^2}{f_0^2}r \tag{38}$$

so that

$$u = \frac{f^2}{f_0^2}r^2 + C \tag{39}$$

and

$$\frac{N(r)}{N_{max}} = \frac{f^2}{f_0^2} + \frac{C}{r^2} \tag{40}$$

Let $r = r_p$ be the radial distance of the peak of N, then,

$$C = -r_p^2\left(\frac{f^2}{f_0^2} - 1\right) \tag{41}$$

so that, at $r = r_m$

$$\frac{N(r_m)}{N_{max}} = \frac{f^2}{f_0^2} - \frac{r_p^2}{r_m^2}\left(\frac{f^2}{f_0^2} - 1\right) = \frac{r_p^2}{r_m^2} - \frac{f^2}{f_0^2}\left(\frac{r_p^2}{r_m^2} - 1\right) \tag{42}$$

From (21) and (42), we have

$$n^2(r_m)r_m^2 = \left(1 - \frac{f_0^2}{f^2}\frac{N(r_m)}{N_{max}}\right)r_m^2 = (1 - \frac{f_0^2}{f^2})r_p^2 = a^2 \sin^2 i_{0m} \tag{43}$$

Therefore,

$$f = \frac{f_0}{\sqrt{1 - \frac{a^2}{r_p^2}\sin^2 i_{0m}}} \tag{44}$$

In order to achieve gliding mode propagation, from practical considerations, the angle of inclination $90° - i_0 \simeq 90° - i_{0m}$ should not be less than $5°$, say, i.e., $i = i_{0m}$ not greater than $85°$, and at the same time, $i_0 \simeq i_{0m}$ should not be less than $20°$, say. Suppose that $a = 6370\,\text{km}$ and $r_p = 6370 + 300 = 6670\,\text{km}$, then for $i_0 \simeq i_{0m} = 85°$, we have $\sin i_{0m} = 0.9962$ and $f \simeq 3.25 f_0$, and for $i_0 \simeq i_{0m} = 20°$, we have $\sin i_{0m} = 0.3420$ and $f \simeq 1.06 f_0$. Then, according to the above criteria, the frequency chosen should be in the range

$$3.25 f_0 \geq f \geq 1.06 f_0 \tag{45}$$

approximately, preferably near the higher end.

The discussions given above are chiefly concerned with long-distance gliding mode propagation which is main theme of the present paper. Our theory, however, is a universal one and the gliding mode propagation is an important special case. The basic equations (17), (19) and (20) are applicable to all cases and all distances, under the assumption that n is real and a function of r alone. Let us consider the general problem of oblique incidence, and imagine that we decrease the angle of incidence i_0 from grazing incidence to nearly vertical incidence. The frequency has a fixed value situated in the range as shown in (45). At grazing incidence, by eq. (17), we have the maximum span of a single hop between two points on the ground. As the angle of incidence i_0 is decreased, it can be seen from (17) that, with $D = a\Theta_D$,

$$\begin{aligned}\frac{\partial D}{\partial D(-i_0)} = &-2a^2\cos i_0\int_a^{r_1}\frac{1}{\sqrt{n^2r^2-a^2\sin^2 i_0}}\frac{dr}{r}\\ &-2a^4\sin^2 i_0\cos i_0\int_a^{r_1}\frac{1}{(n^2r^2-a^2\sin^2 i_0)^{\frac{3}{2}}}\frac{dr}{r}\\ &-2a^2\sin i_0\left[\frac{1}{r_1}\frac{1}{\sqrt{n^2(r_1)r_1^2-a^2\sin^2 i_0}}\right]\frac{\partial r_1}{\partial i_0}\end{aligned} \tag{46}$$

Now,

$$\left\{\begin{aligned}&n(r_1)r_1 - a\sin i_0 = 0\\ &\qquad\frac{\partial}{\partial i_0}[n(r_1)r_1] = \frac{\partial}{\partial r_1}[n(r_1)r_1]\frac{\partial r_1}{\partial i_0} = a\cos i_0\\ &\qquad\qquad\frac{\partial r_1}{\partial i_0} = \frac{a\cos i_0}{\frac{\partial}{\partial r_1}[n(r_1)r_1]}\end{aligned}\right\} \tag{47}$$

$$\begin{aligned}\frac{\partial D}{\partial(-i_0)} = 2a^2\cos i_0\Bigg\{&-\int_a^{r_1}\frac{n^2r\,dr}{(n^2r^2-a^2\sin^2 i_0)^{\frac{3}{2}}}\\ &-\frac{a\sin i_0}{r_1\frac{d}{dr_1}[n(r_1)r_1]}\frac{1}{\sqrt{n^2(r_1)r_1^2-a^2\sin^2 i_0}}\Bigg\}\end{aligned} \tag{48}$$

With the decrease of i_0, there are two opposite tendencies co-existing, the first term on the right of Eq. (47) is negative, tending to cause D to decrease, while the second term on the right is positive since

$\frac{d}{dr_1}[n(r_1)r_1]$ is negative, tending to cause D to increase. At first, the former tendency predominates so that D decreases, with the decrease of i_0, until a certain value of i_0 is reached when the two tendencies just balance and D attains its minimum value called the skip distance, and thereafter the latter tendency predominates so that D increases with the further decrease of i_0 as another certain value of i_0 is reached, D equals the initial D_{max} obtained at grazing incidence, so that between D_{min} and D_{max}, for every D there are two different rays corresponding to two different angles of incidence. One of the rays corresponds to the larger angle of incidence i_0 or equivalently to the smaller angle of inclination $90° - i_0$ and is usually called the low-angle (of inclination) ray, and the other corresponds to the smaller angle of incidence i_0 or equivalently to the larger angle of inclination $90° - i_0$ and is usually called the high-angle (of inclination) ray, also often called Pedersen ray. While the low-angle (of inclination) rays have a limit of distance due to the earth's obstruction, namely, the maximum span of a single hop, the high-angle (of inclination) rays can travel to greater distance by further decrease of i_0. If i_0 is further decreased to become very close to i_{0m} as given by (21), we get the gliding mode propagation with the resulting extremely wide coverage of the rays. After having i_0 decreased beyond the gliding mode regime, the rays penetrate deeper into the ionosphere with steeper and steeper slope. From the above discussion, we see that our theory is a universal one and the gliding mode propagation is an important special case with both theoretical and practical interest. It is characterized by the mathematical property of the integrals of the basic equations (17), (19) and (20) which have the same integrand. The main contribution to the values of the integrals is given by the small portion of the integration path with r lying in the close vicinity of r_1 which is the zero of the denominator of the integration, i.e., $\sqrt{n^2r^2 - a^2\sin^2 i_0}$. We have already known that if $i_0 = i_{0m}$ then, $\sqrt{n^2r^2 - a^2\sin^2 i_{0m}}$ is 0 at $r = r_m$ and also $\frac{d}{dr}(nr) = 0$ at $r = r_m$, and then all the three integrals become infinite. Therefore, when i_0 lies in the close vicinity of i_{0m}, we have the gliding mode propagation, and only a very thin bundle of component rays of the incident beam is relevant. If the electron-density profile is given, from (17), (19) and (20), we can calculate the ray trajectories at a specific frequency in the appropriate range for different values of i_0 as shown qualitatively in Figure 2. When Pederson found the high angle (of inclination) ray, he was actually on the verge of being able to discover the gliding mode propagation. Apparently, due to the fact that the gliding mode propagation depends very critically on the parameter, the angle of incidence, it is liable to escape being noticed and also to elude the discovery by computer numerical computation. From the discussions, we see that the gliding mode long-distance ionospheric propagation is actually an example of a type of chaos, now an important wide frontier sphere of scientific investigation.

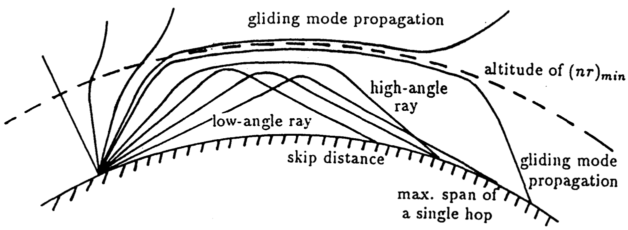

Figure 2. Different ray trajectories for different angles or incidence.

The main difficulty in the application of the gliding mode propagation is that the loss is in general large and the field intensity is in general low, except under a few special situations mentioned below. The loss consists of two categories, one is the loss due to spatial divergence and the other the loss due to absorption, the former being usually more important. Here, we shall discuss only the loss due to spatial divergence. Supposing that the point A on the ground is the transmitting point as above and consider a beam forming an angle Δ_{i_0} in the plane of incidence and an angle $\Delta\phi$ perpendicular to the plane of incidence as shown in Fig. 3. At a unit distance away from the transmitting point along the trajectory, the cross-sectional area is $|\Delta_{i_0}\Delta\phi|$ and at the receiving point $P(r_s, \Theta_s)$,the cross-sectional area is

$$|r_s \frac{d\Theta_s}{di_0} \Delta i_0 \cos i_s r_s \frac{\Delta\phi}{\sin i_s} \sin \Theta_s| = |r_s^2 \frac{\cos i_s \sin \Theta_s}{\sin i_0} \frac{d\Theta_s}{di_0} \Delta i \Delta\phi| \qquad (49)$$

By comparing the two cross-sectional areas, we get the loss due to spatial divergence

$$\Gamma = \ln \left| r_s^2 \frac{\cos i_s \sin i_s}{\sin i_0} \frac{d\Theta_s}{di_0} \right| \quad \text{nepers} \qquad (50)$$

For gliding mode propagation, $|\frac{d\Theta_s}{di_0}|$ is in general large and therefore the field intensity is in general low, but detailed analysis requires the consideration of the various kinds of focusing phenomena, when

$$\left| r_s^2 \frac{\cos i_s \sin \Theta_s}{\sin i_0} \frac{d\Theta_s}{di_0} \right| < 1 \tag{51}$$

Γ is negative and we would get gain due to focusing instead of loss due to spatial divergence. Focusing occurs at the following three points: (1) $i_s = \frac{\pi}{2}$, (2) $\Theta_s = \pi$, and (3) $|\frac{d\Theta_s}{di_0}| = 0$. At a focusing point, Γ theoretically become minus infinity. The first focusing point corresponds to the situation that the ray becomes horizontal at P, the second focusing point corresponds to antipodal propagation, and the third focusing point corresponds to the situation of a single hop at skip distance, not belonging to the gliding mode propagation. For gliding mode propagation, the closer the point P to the altitude of $(nr)_{min}$, the less the value $|\frac{\pi}{2} - i_s|$, so that the loss due to spatial divergence will not be too large to the benefit for the reception of the signal.

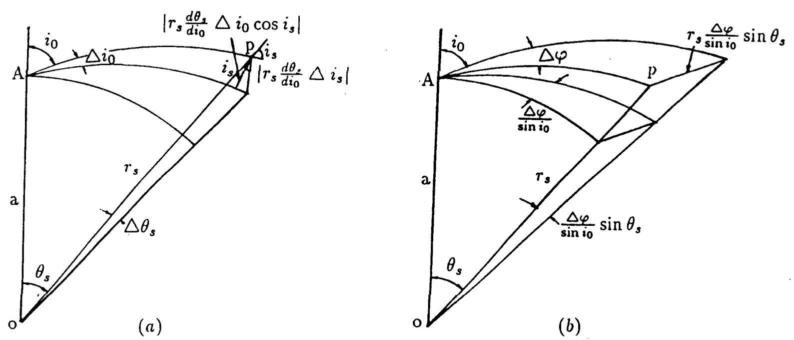

Figure 3. Depiction showing the divergence and focusing effect of a narrow beam.

In the above discussion, it was assumed that n is a function of r alone. In the actual case, n may be a slowly varying function of Θ and the azimuthal angle φ as well, so that the $(nr)_{min}$ surface is somewhat warped. This situation arises if the ionosphere is tilted and if the effect of the earth's magnetic field is taken into consideration, either for the ordinary wave or for the extraordinary wave. Under the condition of a warped $(nr)_{min}$ surface, for a specific long-distance short-radio wave circuit, we can divide the relevant strip of the warped $(nr)_{min}$ surface into a few sections, each of which can be approximately regarded as a portion of a spherical surface with its.center duly shifted. Let us suppose that the transmitter emits a beam propagating to the right and the first section of the relevant strip of the $(nr)_{min}$ surface contains the point which is to the right of the transmitting point by a distance along the earth's surface equal to half of the maximum span of a single hop. This point is termed the control point for the transmitter end. The frequency used must not be greater than the MUF of a single hop of maximum span based on the characteristics of the ionosphere above the control point in the first section to avoid the obstruction of the earth. The transmitted beam consists of a multitude of rays. These rays reach the vicinity of the altitude of $(nr)_{min}$, both slightly above and slightly below this altitude, and thence they continue to propagate in the gliding mode, causing the rays more widely spread. Among these rays, there is always a thin bundle of rays which will reach the vicinity of the altitude of $(nr)_{min}$ of the second section, both slightly above and slightly below this altitude, and so on. Finally, the rays covering almost the whole space both above and below the altitude of $(nr)_{min}$ are obtained. As far as the ray treatment for the propagation between a point on the ground and an orbiting satellite is concerned, no matter whether above or below the altitude of $(nr)_{min}$, the problem can be considered solved by the above discussion. However, for propagation between two points on the ground with the receiving point specified, we have to be sure that the frequency used besides fulfilling the requirement that the incident rays at the transmitting end are not obstructed by the earth, similarly, the returning rays in the last section should also be not obstructed by the earth. The last section should contain the point which is to the left of the receiving point by a distance along the earth's surface equal to half of the maximum span of a single hop, where the characteristics of the ionosphere ensures that the frequency used is also not greater than the MUF of a single hop of maximum span, and the MUF of the whole communication circuit is chosen to be the smaller of the two MUF values calculated for the transmitting end and for the receiving end, as required by the control-point method of MUF prediction.

3. Solution by Full-Wave Treatment

The ray treatment method is an approximate one. The approximation becomes poor as $n^2r^2 - a^2\sin^2 i_0$ approaches zero at the altitude of $(nr)_{min}$. But this may not be important because the field distribution is chaotic anyway. To be complete, in this section, a more exact full-wave treatment is given for the neighboring region of $(nr)_{min}$. To make the problem tractable, we transform the spherical coordinates (r, Θ, φ) as shown in Figure 4 to a new pseudo-rectangular coordinate system (x, y, z) such that [2]

$$\left\{\begin{aligned} x &= r_m\Theta\cos\varphi \\ y &= r_m\Theta\sin\varphi \\ z &= r_m\ln\frac{r}{r_m} \end{aligned}\right\} \tag{52}$$

Hence

$$\left\{\begin{aligned} &r = r_m\exp(\frac{z}{r_m}) \\ &dr = \exp(\frac{z}{r_m})dz \\ &r_m\Theta = \sqrt{x^2+y^2} \\ &r_m d\Theta = \frac{xdx+ydy}{\sqrt{x^2+y^2}} \\ &\tan\varphi = \frac{y}{x} \\ &d\varphi = \frac{xdy-ydx}{x^2+y^2} \end{aligned}\right\} \tag{53}$$

and

$$\begin{aligned} ds^2 =& dr^2 + r^2 d\Theta^2 + r^2\sin^2\Theta d\varphi^2 \\ =& \exp(\frac{2z}{r_m})\left[dz^2 + \frac{(xdx+ydy)^2}{x^2+y^2} + \frac{\sin^2\Theta}{\Theta^2}\frac{(xdy-ydx)^2}{x^2+y^2}\right] \\ =& \exp(\frac{2z}{r_m})\left[dx^2+dy^2+dz^2 - \left(1-\frac{\sin^2\Theta}{\Theta^2}\right)\frac{(xdy-ydx)^2}{x^2+y^2}\right] \\ =& \exp(\frac{2z}{r_m})[dx^2+dy^2+dz^2 - (1-\frac{\sin^2\Theta}{\Theta^2})r_m^2\Theta^2 d\varphi^2] \end{aligned} \tag{54}$$

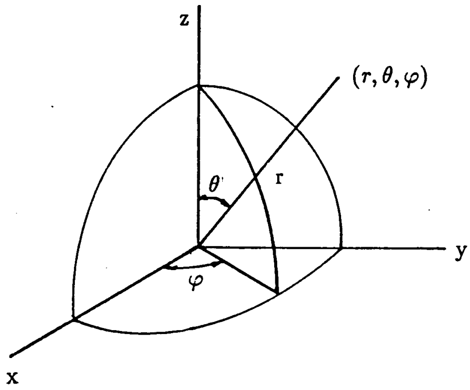

Figure 4. Nomenclature of the spherical coordinate system.

Since by symmetry we have $\frac{\partial}{\partial \phi} = 0$ and since the wavelengths of interest are much smaller than the earth's radius, it is reasonable to assume that the deviation from the plane of incidence is much smaller than the deviation in the plane of incidence, so that we can write approximately even in the whole plane of incidence,

$$\left\{ \begin{array}{l} ds^2 = \eta^2(dx^2 + dz^2) \\ \dfrac{\partial}{\partial y} = 0 \end{array} \right\} \tag{55}$$

where

$$\eta = \exp\left(\frac{z}{r_m}\right) \tag{56}$$

The curl of any vector function $\vec{F}$ is approximately given by

$$\begin{aligned} \nabla \times \vec{F} &= \frac{1}{\eta^3} \begin{vmatrix} \eta\hat{x} & \eta\hat{y} & \eta\hat{z} \\ \frac{\partial}{\partial x} & 0 & \frac{\partial}{\partial z} \\ \eta F_x & \eta F_y & \eta F_z \end{vmatrix} \\ &= -\hat{x}\frac{1}{\eta^2}\frac{\partial}{\partial z}(\eta F_y) + \hat{y}\frac{1}{\eta^2}\left[\frac{\partial}{\partial z}(\eta F_x) - \frac{\partial}{\partial x}(\eta F_z)\right] + \hat{z}\frac{1}{\eta^2}\frac{\partial}{\partial x}(\eta F_y) \end{aligned} \tag{57}$$

Hence, the Maxwell equations in the ionosphere can be written as

$$\left\{\begin{array}{l} \partial z(\eta E_y) - i\omega(\eta\mu_0)(\eta H_x) = 0 \\ \partial z(\eta E_x) - \partial x(\eta E_z) - i\omega(\eta\mu_0)(\eta H_y) = 0 \\ \partial x(\eta E_y) - i\omega(\eta\mu_0)(\eta H_z) = 0 \end{array}\right\} \tag{58}$$

$$\left\{\begin{array}{l} -\partial z(\eta H_y) + i\omega(\eta\epsilon_0)(\eta E_x) = \eta^2 J_x = -e\eta N(\eta v_x) \\ \partial z(\eta H_x) - \partial x(\eta H_z) + i\omega(\eta\epsilon_0)(\eta E_y) = \eta^2 J_y = -e\eta N(\eta v_y) \\ \partial x(\eta H_z) + i\omega(\eta\epsilon_0)(\eta E_z) = \eta^2 J_z = -e\eta N(\eta v_z) \end{array}\right\} \tag{59}$$

where $-e$ is the charge of the electron, N the electron density, $\vec{v}$ the velocity of the electron, and $\exp(-i\omega t)$ is the time dependence convention adopted with the effect of the earth's magnetic field assumed to be negligible. The equations of motion of the electrons are

$$\left\{\begin{array}{l} -i\omega m(\eta v_x) = -e(\eta E_x) \\ -i\omega m(\eta v_y) = -e(\eta E_y) \\ -i\omega m(\eta v_z) = -e(\eta E_z) \end{array}\right\} \tag{60}$$

where m is the mass of the electron. From (59) and (60), we have

$$\left\{\begin{array}{l} -\partial z(\eta H_y) - i\omega(\eta\epsilon_0)n^2(\eta E_x) = 0 \\ \partial z(\eta H_z) - \partial x(\eta H_z) + i\omega(\eta\epsilon_0)n^2(\eta E_y) = 0 \\ \partial x(\eta H_y) + i\omega(\eta\epsilon_0)n^2(\eta E_z) = 0 \end{array}\right\} \tag{61}$$

where $n^2 = 1 - \frac{Ne^2}{m\epsilon_0\omega^2}$. Let us consider the following cases.
Case 1: $E_y \neq 0, H_x \neq 0, H_z \neq 0, E_x = E_z = H_y = 0$, then

$$\left\{\begin{array}{l} -\partial z(\eta E_y) - i\omega(\eta\mu_0)(\eta H_x) = 0 \\ \partial x(\eta E_y) - i\omega(\eta\mu_0)(\eta H_z) = 0 \\ \partial z(\eta H_x) - \partial x(\eta H_z) + i\omega(\eta\epsilon_0)n^2(\eta E_y) = 0 \end{array}\right\} \tag{62}$$

from which it can be shown that

$$\frac{\partial^2 E_y}{\partial z^2} + \frac{\partial^2 E_y}{\partial x^2} + k_0^2\eta^2 n^2 E_y + \frac{1}{r_m}\frac{\partial E_y}{\partial z} = 0 \tag{63}$$

where $k_0^2 = \omega^2 \epsilon_0 \mu_0$. Neglecting the last small term, we have approximately

$$\frac{\partial^2 E_y}{\partial z^2} + \frac{\partial^2 E_y}{\partial x^2} + k_0^2 \eta^2 n^2 E_y = 0 \tag{64}$$

Case 2: $H_y \neq 0, E_x \neq 0, E_z \neq 0, H_x = H_z = E_y = 0$, then

$$\left\{ \begin{array}{l} -\partial z(\eta H_y) + i\omega(\eta\epsilon_0) n^2 (\eta E_x) = 0 \\ \partial x(\eta H_y) + i\omega(\eta\epsilon_0) n^2 (\eta E_z) = 0 \\ \partial z(\eta E_x) - \partial x(\eta E_z) - i\omega(\eta\mu_0)(\eta H_y) = 0 \end{array} \right\} \tag{65}$$

from which it can be shown that

$$\frac{\partial^2 H_y}{\partial z^2} + \frac{\partial^2 H_y}{\partial x^2} + k_0^2 \eta^2 n^2 H_y + (\frac{1}{r_m} - \frac{2}{n}\frac{\partial n}{\partial z})\frac{\partial H_y}{\partial z} - \frac{2}{r_m}\frac{1}{n}\frac{\partial n}{\partial z} H_y = 0 \tag{66}$$

where we know that $\frac{\partial n}{\partial z} = \frac{\partial n}{\partial r}\frac{\partial r}{\partial z} = \eta \frac{\partial n}{\partial r}$ and near $r = r_m$, approximately, $\frac{1}{n}\frac{\partial n}{\partial z} = \eta \frac{1}{n}\frac{\partial n}{\partial r} \simeq -\frac{1}{r_m}$, giving approximately,

$$\frac{\partial^2 H_y}{\partial z^2} + \frac{\partial^2 H_y}{\partial x^2} + k_0^2 \eta^2 n^2 H_y + \frac{3}{r_m}\frac{\partial H_y}{\partial z} + \frac{2}{r_m^2} H_y = 0 \tag{67}$$

Neglecting the last two small terms, we have approximately

$$\frac{\partial^2 H_y}{\partial z^2} + \frac{\partial^2 H_y}{\partial x^2} + k_0^2 \eta^2 n^2 H_y = 0 \tag{68}$$

Let us take Case 2 for example. In order to be able to solve the wave equation analytically, we assume that $\eta^2 n^2 = \eta^2 \left(1 - \frac{e^2 N}{m\epsilon_0\omega^2}\right)$ satisfies the Epstein distribution [3] , namely

$$\eta^2 n^2 = \eta^2 \left(1 - \frac{e^2 N}{m\epsilon_0 \omega^2}\right) = 1 - K_1 \frac{\exp{(\alpha z)}}{1 + \exp{(\alpha z)}} - 4K_2 \frac{\exp{(\alpha z)}}{(1 + \exp{(\alpha z)})^2} \tag{69}$$

where α, K_1 , and K_2 are constants which are free to choose. If $K_1 \neq 0$ and $K_2 = 0$, $1 - \eta^2 n^2$ gives a transition layer and if $K_1 = 0$ and $K_2 \neq 0$, it gives a symmetric layer, as shown in Figure 5. If K_1 and K_2 are both not zero, by adjusting the values of K_1 and K_2 and also of α, we have different shapes of intermediate layers. The wave equation becomes

$$\frac{\partial^2}{\partial z^2}H_y+\frac{\partial^2}{\partial x^2}H_y+k_0^2\left[1-K_1\frac{\exp(\alpha z)}{1+\exp(\alpha z)}-4K_2\frac{\exp(\alpha z)}{(1+\exp(\alpha z))^2}\right]H_y=0 \tag{70}$$

Let

$$H_y = X(x)Z(z) \tag{71}$$

then

$$\frac{1}{Z}\frac{d^2Z}{dz^2}+\frac{1}{X}\frac{d^2X}{dx^2}+k_0^2\left[1-K_1\frac{\exp(\alpha z)}{1+\exp(\alpha z)}-4K_2\frac{\exp(\alpha z)}{(1+\exp(\alpha z))^2}\right]=0 \tag{72}$$

Let

$$\frac{d^2X}{dx^2}+k_0^2\cos^2\beta X=0 \tag{73}$$

where $\cos\beta$ is a constant. Therefore, we can let

$$X=\exp(ik_0x\cos\beta) \tag{74}$$

and

$$\frac{d^2Z}{dz^2}+k_0^2\left[\sin^2\beta-K_1\frac{\exp(\alpha z)}{1+\exp(\alpha z)}-4K_2\frac{\exp(\alpha z)}{(1+\exp(\alpha z))^2}\right]Z=0 \tag{75}$$

Let us introduce

$$u=-\exp(\alpha z) \tag{76}$$

and

$$v=\frac{Z}{f(u)} \tag{77}$$

where $f(u)$ is a function to be determined. We have

$$\left\{\begin{aligned}
&\frac{du}{dz}=\alpha u\\
&\frac{dZ}{dz}=\frac{dZ}{du}\frac{du}{dz}=\alpha u[f(u)\frac{dv}{du}+f'(u)v]\\
&\frac{d^2z}{dz^2}=\alpha^2u^2f(u)\frac{d^2v}{du^2}+[2\alpha^2u^2f'(u)+\alpha^2uf(u)]\frac{dv}{du}\\
&\qquad\quad+[\alpha^2u^2f''(u)+\alpha^2uf'(u)]v
\end{aligned}\right\} \tag{78}$$

Hence,

$$\frac{d^2v}{du^2}+\left[2\frac{f'(u)}{f(u)}+\frac{1}{u}\right]\frac{dv}{du}$$
$$+\left\{\frac{f''(u)}{f(u)}+\frac{1}{u}\frac{f'(u)}{f(u)}+\frac{k_0^2}{\alpha^2u^2}\left[\sin^2\beta+K_1\frac{u}{1-u}+4K_2\frac{u}{(1-u)^2}\right]\right\}v=0 \tag{79}$$

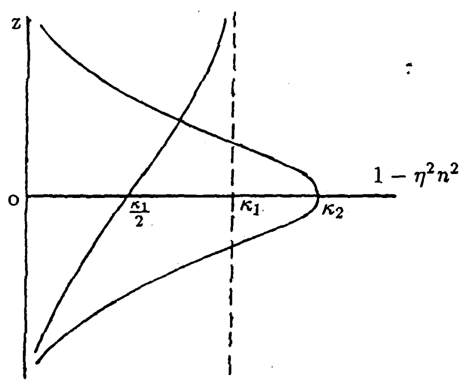

Figure 5. Epstein distribution illustrating forms of transition layer and symmetric layer

Equation (79) can be transformed into the hypergeometric differential equation of the form

$$\frac{d^2v}{du^2}+\frac{C-(A+B+1)u}{u(1-u)}\frac{dv}{du}-\frac{AB}{u(1-u)}v=0 \tag{80}$$

where A,B,C are constants. Comparing equations (79) and (80), we have

$$\frac{f'(u)}{f(u)}=\frac{C-1}{2}\frac{1}{u}-\frac{A+B-C+1}{2}\frac{1}{1-u} \tag{81}$$

so that

$$f(u)=f_0u^{\frac{C-1}{2}}(1-u)^{\frac{A+B-C+1}{2}} \tag{82}$$

where f_0 is a constant. Therefore

$$\frac{f''(u)}{f(u)} + \frac{1}{u}\frac{f'(u)}{f(u)} + \frac{k_0^2}{\alpha^2 u^2}\left[\sin^2\beta + K_1\frac{u}{1-u} + 4K_2\frac{u}{(1-u)^2}\right]$$
$$= \frac{1}{4}\left(\frac{C-1}{u} - \frac{A+B-C+1}{1-u}\right)^2 - \frac{A+B-C+1}{2}\frac{1}{u(1-u)^2}$$
$$+ \frac{k_0^2}{\alpha^2 u^2}\left[\sin^2\beta + K_1\frac{u}{1-u} + 4K_2\frac{u}{(1-u)^2}\right] = -\frac{AB}{u(1-u)} \tag{83}$$

By reducing to common denominator and equating the coefficients of u^0, u^1, and u^2 in the numerator on both sides of the equal sign, we have

$$\left\{\begin{array}{l} (C-1)^2 + 4\dfrac{k_0^2}{\alpha^2}\sin^2\beta = 0 \\ (A-B)^2 + 4\dfrac{k_0^2}{\alpha^2}(\sin^2\beta - K_1) = 0 \\ (A+B)^2 - 2C(A+B) + 2C(C-1) + 4\dfrac{k_0^2}{\alpha^2}(-\sin^2\beta + 4K_2) = 0 \end{array}\right\} \tag{84}$$

Solving, we get

$$\left\{\begin{array}{l} C = 1 + i2\dfrac{k_0}{\alpha}\sin\beta \\ A = \dfrac{1}{2}[1 + i2\dfrac{k_0}{\alpha}\sin\beta + (1 - 16\dfrac{k_0^2}{\alpha^2}K_2)^{\frac{1}{2}} - i2\dfrac{k_0}{\alpha}(\sin^2\beta - K_1)^{\frac{1}{2}}] \\ B = \dfrac{1}{2}[1 + i2\dfrac{k_0}{\alpha}\sin\beta + (1 - 16\dfrac{k_0^2}{\alpha^2}K_2)^{\frac{1}{2}} + i2\dfrac{k_0}{\alpha}(\sin^2\beta - K_1)^{\frac{1}{2}}] \end{array}\right\} \tag{85}$$

Hence

$$Z = f_0 u^{\frac{C-1}{2}}(1-u)^{\frac{A+B-C+1}{2}} v \tag{86}$$

where v is the hypergeometric function. The hypergeometric differential equation has three singular points 0, 1, and ∞. For each of these points, there are two fundamental series expansions convergent in its neighborhood, namely,

(i) about $u = 0$,

$$\begin{aligned} v_1 &= C_1 F(A, B, C, u) \\ v_2 &= C_2 u^{1-C} F(A-C+1, B-C+1, 2-C, u) \end{aligned} \tag{87}$$

(ii) about $u = 1$,

$$\begin{aligned} v_3 &= C_3 F(A, B, A + B - C + 1, 1 - u) \\ v_4 &= C_4 (1 - u)^{C-A-B} F(C - A, C - B, C - A - B + 1, 1 - u) \end{aligned} \tag{88}$$

(iii) about $u = \infty$,

$$\begin{aligned} v_5 &= C_5 u^{-A} F(A, A - C + 1, A - B + 1, u^{-1}) \\ v_6 &= C_6 u^{-B} F(B, B - C + 1, B - A + 1, u^{-1}) \end{aligned} \tag{89}$$

where C_1, C_2, C_3, C_4, C_5, and C_6 are constants, and

$$F(A, B, C, u) = \frac{\Gamma(C)}{\Gamma(A)\Gamma(B)} \sum_{k=0}^{\infty} \frac{\Gamma(A + k)\Gamma(B + k)}{k!\Gamma(C + k)} u^k \tag{90}$$

which is convergent for $|u| < 1$. Although the functions are given in the forms of series expansions which are convergent only within their appropriate definite intervals, each of these series expansions defines an analytic function by analytic continuation which is a solution of the hypergeometric differential equation extending beyond the convergent interval of these series. Since any three solutions of a linear differential equation of the second order are linearly dependent, there is a linear relation which holds for the analytic functions obtained by analytic continuation and therefore it is valid for all values which can be assigned to the three functions. Therefore, it is legitimate and for our purpose more convenient to write the u's in (77), (81), (82) and (83) as -u's, then v_1, v_2, v_3, v_4, v_5, and v_6 are expressed as

(i) about $u = 0$,

$$\begin{aligned} v_1 &= C_1 F(A, B, C, -u) \\ v_2 &= C_2 (-u)^{1-C} F(A - C + 1, B - C + 1, 2 - C, -u) \end{aligned} \tag{91}$$

(ii) about $u = -1$,

$$\begin{aligned} v_3 &= C_3 F(A, B, A + B - C + 1, 1 + u) \\ v_4 &= C_4 (1 + u)^{C-A-B} F(C - A, C - B, C - A - B + 1, 1 + u) \end{aligned} \tag{92}$$

(iii) about $u = -\infty$

$$\begin{aligned} v_5 &= C_5(-u)^{-A}F(A, A-C+1, A-B+1, -u^{-1}) \\ v_6 &= C_6(-u)^{-B}F(B, B-C+1, B-A+1, -u^{-1}) \end{aligned} \tag{93}$$

where

$$F(A, B, C, u) = \frac{\Gamma(C)}{\Gamma(A)\Gamma(B)} \sum_{k=0}^{\infty} \frac{\Gamma(A+k)\Gamma(B+k)}{k!\Gamma(C+k)}(-u)^k \tag{94}$$

We are interested in knowing the relations among the v's. It can be proved that

$$\begin{aligned} &(-1)^A \frac{\Gamma(1-B)\Gamma(A-C+1)}{\Gamma(1-C)\Gamma(A-B+1)} (-u)^{-A} \\ &\qquad F(A, A-C+1, A-B+1, -u^{-1}) \\ &= F(A, B, C, -u) + (-1)^{1-C} \frac{\Gamma(C-1)\Gamma(1-B)\Gamma(A-C+1)}{\Gamma(1-C)\Gamma(C-B)\Gamma(A)} \\ &\qquad (-u)^{1-C} F(A-C+1, B-C+1, 2-C, -u) \end{aligned} \tag{95}$$

As $z \to -\infty, u \to 0^-$ and

$$\left\{ \begin{aligned} &Z \to f_0 e^{ik_0 z \sin\beta} \\ &\qquad + f_0(-1)^{1-C} \frac{\Gamma(C-1)\Gamma(1-B)\Gamma(A-C+1)}{\Gamma(1-C)\Gamma(C-B)\Gamma(A)} e^{-ik_0 z \sin\beta} \\ &H_y \to f_0 e^{ik_0(x\cos\beta + z\sin\beta)} \\ &\qquad + f_0(-1)^{1-C} \frac{\Gamma(C-1)\Gamma(1-B)\Gamma(A-C+1)}{\Gamma(1-C)\Gamma(C-B)\Gamma(A)} e^{ik_0(x\cos\beta - 2\sin\beta)} \end{aligned} \right\} \tag{96}$$

As $z \to \infty, u \to -\infty$ and

$$\left\{ \begin{aligned} &Z = f_0(-1)(-u)^{\frac{B-A}{2}} F(A, A-C+1, A-B+1, -u^{-1}) \\ &\qquad \to f_0(-1)e^{ik_0 z(\sin^2\beta - K_1)^{\frac{1}{2}}} \\ &H_y \to f_0(-1)e^{ik_0[x\cos\beta + z(\sin^2\beta - K_1)^{\frac{1}{2}}]} \\ &\qquad = f_0(-1)e^{ik_0(x\cos\beta' + z\sin\beta')} \end{aligned} \right\} \tag{97}$$

where

$$\left\{\begin{array}{l} k = k_0(1-K_1)^{\frac{1}{2}} \\ \cos\beta' = \dfrac{\cos\beta}{(1-K_1)^{\frac{1}{2}}} \\ \sin\beta' = \dfrac{(\sin^2\beta - K_1)^{\frac{1}{2}}}{(1-K_1)^{\frac{1}{2}}} \end{array}\right\} \tag{98}$$

The first term of the second equation of (96) corresponds to the incident wave and the second term corresponds to the "reflected" wave. The second equation of (97) corresponds to the transmitted wave. From (96), the "reflection" coefficient is

$$R = (-1)^{1-C}\frac{\Gamma(C-1)\Gamma(1-B)\Gamma(A-C+1)}{\Gamma(1-C)\Gamma(C-B)\Gamma(A)} \tag{99}$$

and from (97), the transmission coefficient is

$$T = (-1)^{\frac{3A+B-C+1}{2}}\frac{\Gamma(1-B)\Gamma(A-C+1)}{\Gamma(1-C)\Gamma(A-B+1)} \tag{100}$$

Let us suppose that the incident wave is generated by a source located at $x = 0$ and $z = z_0$ where $|z_0| = -z_0 < r_m$. Neglecting the spread of the beam perpendicular to the plane of incidence, we can approximately write the incident beam before entering the ionosphere as

$$H_y^{inc} = \int_{\beta_1}^{\beta_2} F(\beta)e^{ik_0[x\cos\beta+(z-z_0)\sin\beta]}d\beta \tag{101}$$

where $F(\beta)$ as a function of β is assumed to be known. Comparing the first term of the second equation of (96) with $F(\beta)e^{ik_0[x\cos\beta+(z-z_0)\sin\beta]}$ we have

$$f_0 = F(\beta)e^{-ik_0z_0\sin\beta} \tag{102}$$

Therefore, below the altitude of $(nr)_{min}$ the field of the "reflected" wave is approximately given by

$$\begin{aligned} H_y^{refl}(x,z) = \int_{\beta_1}^{\beta_2} & d\beta F(\beta)e^{ik_0z_0\sin\beta}(-u)^{\frac{C-1}{2}}(1+u)^{\frac{A+B+-C+1}{2}} \\ & \frac{\Gamma(C-1)\Gamma(1-B)\Gamma(A-C+1)}{\Gamma(1-C)\Gamma(C-B)\Gamma(A)}u^{1-C} \\ & F(A-C+1, B-C+1, 2-C, -u) \end{aligned} \tag{103}$$

and above the altitude of $(nr)_{min}$, the field of the transmitted wave is approximately give by

$$\begin{aligned} H_y^{trans}(x,z) = &\int_{\beta_1}^{\beta_2} d\beta F(\beta) e^{-ik_0 z_0 \sin\beta} (-u)^{\frac{C-1}{2}} (1+u)^{\frac{A+B-C+1}{2}} \\ &\frac{\Gamma(1-B)\Gamma(A-C+1)}{\Gamma(1-C)\Gamma(A-B+1)} u^{-A} \\ &F(A, A-C+1, A-B+1, -u) \end{aligned} \tag{104}$$

From (96), (97), and (59), the field in the whole region both below and above the altitude of $(nr)_{min}$ can be computed numerically if $\eta^2(1-\frac{e^2N}{m\epsilon_0\omega^2})$ profile as an Epstein distribution is given.

4. Conclusions and Discussions

Practice has shown that multiple-hop propagation can neither be the only mode nor be the predominant mode of long-distance short radio-wave ionospheric propagation between two points on the ground. Practice has also shown the feasibility of long-distance short radio-wave ionospheric propagation between a point on the ground and an orbiting satellite. Both of these two categories of propagation can be achieved by solely utilizing the refraction of the ionosphere, without considering the presence of the earth. We established a unified theory of short radio-wave ionospheric propagation, with special emphasis on long-distance propagation, which is an important special case both practically and theoretically. Once we assumed that the transmitting point was situated on the ground, we discussed the ray treatment and derived the expressions for the angles subtended by the propagation path at the earth's center for the three cases mentioned before. The integral expressions for the angles subtended at the earth's center has the integrand with $\sqrt{n^2r^2 - a^2\sin^2 i_0}$ as its denominator. The angles of incidence i_0 is a very critical parameter such that if i_0 approaches i_{0m}, where $n(r_m)r_m - a\sin i_{0m} = 0$ with $r = r_m$ equal to the radial distance of $(nr)_{min}$ the values of the integral expressions approach infinity and the distribution of the field becomes chaotic and reaches essentially the whole space of the plane of the incidence both below and above $r = r_m$. This is the so-called gliding mode propagation. It is then proved that to ensure gliding mode propagation, the assumption that n is a function of r alone can be relaxed, and n can be a slowly varying function of Θ and the azimuthal angle φ as well. To be complete, a full-wave treatment is introduced. Experience seems to indicate that echo phenomenon does not cause serious problem. Experiments designed to test the different qualities of communications using the gliding mode propagation are suggested.

Acknowledgment

This project is supported by the National Natural Scientific Foundation of China.

References

1. Lu, B. W. "A unified theory of short radio-wave ionospheric propagation with special emphasis on long-distance propagation," *Presented at PIERS,* California, USA, July 1993.
2. Lu, B. W. "On the problem of radio-wave propagation between a point on the ground and a spacecraft, document of the Sixth Symposium on Astronautics," *Bureau of New Technology,* Academia sinica, December 23, 1961 (in Chinese).
3. Kerr, D. E, *Propagation of Radio Waves,* New York, McGraw-Hill Book Co., 1951.
4. Epstein, P. S., "Reflection of waves in an inhomogeneous absorbing medium," *Proceedings of the National Academy of Sciences,* Vol. 16, No. 10, 1930.