STUDIES

ON

THE PROPAGATION OF ELECTROMAGNETIC WAVES

IN THE IONOSPHERE AND EXOSPHERE

IWANE KIMURA

JULY 1961

STUDIES

ON

THE PROPAGATION OF ELECTROMAGNETIC WAVES

IN THE IONOSPHERE AND EXOSPHERE

IWANE KIMURA

JULY 1961

CONTENTS

Preface

Owing to the outstanding development of the space vehicles, such as rockets, earth satellites and manned rockets, our dream to enter into the cosmos has been realized step by step under our eyes. Then we, investigators are urged to look at the outside space of the ionosphere, although concerning the vicinity of the earth, our survey has only reached the ionosphere or so.

On the other hand in the recent decade, attention has been paid to this same space from a different field. That is the investigation of a very low frequency (VLF) atomospherics, called whistler. It was initiated by L.R.O.Storey who reported the propagation mechanism of the whistlers in 1953. That is to say, the point which attracted our interest is that it was proved that the whistler was originated from lightning flash and propageted in the exosphere following the horseshoe-shaped lines of force of the geomagnetic field. This mechanism resulted in proving the existence of a considerable amount of electrons in the exosphere.

After that we have noticed another peculiar atmospherics, called VLF emissions. And it is known that the VLF emission is originated in the exosphere and is propagated to the ground like the whistlers in the exosphere and ionosphere.

From these trends, the outer space, or outer ionosphere has become an important object of our study. Recently we have called this space "exosphere". The exosphere is filled with plasma of a considerable density, where the earth magnetic field is pervaded, and it is sometimes penetrated by incoming stream of high energetic charged

particles, which often cause the aurora.

Here it must be emphasized that we are able to know the physical property of the space by analysing various phenomena associated with the exosphere.

Rocket and satellite are naturally utilized as the space sounding probe. Intensity of cosmic ray, intensity of the earth magnetic field and solar radiation could be directly measured. Especially the discovery of Van Allen belt was a conspicuous result of the probe. The merit of "direct measurement" by means of the space probe can not be replaced by any means. However in actual case, there is limitation in the number of times of the experiment by the space probe, the space swept by one probe is not so wide, and some physical quantities can not yet be observed by it, from technical reasons.

Then in order to estimate unkown physical quantities in the exosphere the observations of the various phenomena, explained previously, have been needed. Investigation in this field, therefore, have much advanced in the IGY.

In part I of this paper, taking account of the above point of view, a theoretical study of the whistler, especially the problem of its path in the exosphere will be described. Then in part II, the VLF emissions will be theoretically dealt with, since the originating mechanism of the emissions has not yet been definitely known, although many hypotheses have been proposed so far.

As described above, there are many probems concerning the exosphere, whilst as to the underlying ionosphere not a few problems are left unknown, in spite that it has been investigated for a long time. For the propagation of the whistler and VLF emission, the ionosphere

is actually in charge of an important part.

Now, the observation of the ionosphere has been made by means of the impulse method. At night, however, the state of the E layer has yet hardly been clarified. It is because, utilized frequency of the impulse wave is not so low. The author has experimented a new method of ionosphere sounding by using middle frequency broadcast waves without using any impulse wave transmission. In part III this will be explained in detail.

PART I

PROPAGATION OF THE VLF ELECTROMAGNETIC WAVE IN THE IONOSPHERE AND EXOSPHERE

(Propagation of the whistlers)

Chapter 1.1 Introduction

A plasma which is pervaded with magnetic field has an anisotropic property for electromagnetic wave propagation. Such a plasma is usually called the magneto-active plasma. For the propagation of electromagnetic wave, the ionosphere and the exosphere must be regarded as a kind of the magneto-active plasma.

Although the propagation of middle and high frequency electromagnetic waves in the medium have been already much studied, it was quite recent that the phenomena concerning the propagation of VLF (very low frequency) waves in the medium aroused our interest. It was initiated by L.R.O. Storey (1953) who developed the theory of the origin of a kind of atmospherics known as the whistler and clarified it almost perfectly. After that the theoretical and experimental studies of the whistlers have much advanced.

The whistler is an atmospherics characterized by a whistling musical tone, which can be often heard by a long antenna and a subsequent audio frequency amplifier.

According to Storey, the source of the whistlers is the lightning flash occuring near the ground. Electromagnetic waves radiated from the lightning partly reach the ionosphere. The low frequency component of the radiation can penetrate the ionosphere due to the magnetic field and is compelled to propagate along the lines of force

of the geomagnetic field so that it comes back again to the ground. Since the magneto-active plasma has dispersive propaerty, the atmospherics is dispersed during its journey in the exosphere and then heard as the musical tone on the ground. This interpretation is based on the supposition that the exosphere is filled with electrons and protons, the densities of which are fairly large, not with vacuum as believed until quite recently.

As is recognized from the mechanism, the whistler time delay is greatly dependent upon the integrated plasma density and geomagnetic field intensity there. Consequently by analysing the time delays of the whistlers, we can estimate the electron density and the geomagnetic field intensity encountered in the exosphere.

The effect of such parameters on the propagation of the whistler depends upon their ray path, while the ray path itself also related with the parameters. For the sake of simplicity the ray paths have been sometimes supposed to be identical with lines of force of the geomagnetic field.

There are another kinds of atmospherics which are originated in the exosphere and propagate therein down to the ground. These atmospherics are called the VLF emissions and will be described in Part II. As will be seen there, the VLF emissions also have valuable information about the medium parameters, electron density and magnetic field intensity. They naturally propagate in the same mode as the whistler through the ionosphere and exosphere.

Then it is worth investigating the ray tracing of these VLF electromagnetic wave in the magneto-active media. The ray tracing in an anisotropic and inhomogeneous medium is a tedious problem.

When, in general, the direction in which the medium construction varies, is limited in one direction and the direction of magnetic field is constant, the Snell's law is known to be applicable to the wave normal of the electromagnetic wave. And a relation between the wave normal and the ray direction in which wave energy is propagated, is given by a certain equation. Therefore in this case, the ray path can be traced by using the Snell's law.

On the other hand, for the ray tracing of the whistler, the ionized medium may distribute concentrically and the direction of the earth's magnetic field varies with latitudes. Then it is too complicated to be analytically solved.

In what follows an approximate method to treat the problem of ray tracing will be mainly given, after reviewing briefly the properties of the whistler and the mechanism of its origin. Consequently the Part I of this paper is concentrated upon the description of ray tracing of the VLF electromagnetic waves in magnetoactive media.

Some arguments will be given for the absorption of the VLF electromagnetic waves in the ionosphere.

Chapter 1.2 On the whistlers and whistler mode of propagation in a magneto-active plasma

1.2.1 The originating mechanism and the kinds of whistlers

The atomspherics called whistler is composed of a quite peculiar musical tone which continues usually for 1 or 2 seconds, as the name stands for. Generally, these kinds of VLF atmospherics are analysed with a frequency analyser, such as the sonagraph, to see their qualities pictorially. In case of the whistlers, the sonagram, frequency versus time characteristics becomes a smooth curve as shown in Fig. 1.1. That is to say, the time of arrival of the signal uniformly increases with decreasing frequency, and when this curve is rearranged on a f^{-1} vs. t graph, it lies on a straight line, the gradient of which is called dispersion. This value is only one characteristic quantity obtainable from an observed whistler. As will be explained later, the quantity is very important because it is uniquely determined by the electron density and magnetic field intensity over its whole path.

In the exosphere, the medium is transparent due to the magnetic field for the extraordinary mode of VLF electromagnetic waves. And the medium is also anisotropic, so that the waves do not propagate in the wave normal direction, but in the ray direction which can be proved to lie among the wave normal direction and the direction of the earth's magnetic field. Namely the energy of the wave is generally propagated in a direction approximate to the earth's magnetic field. This effect compels the wave to follow the horse-shoe-shaped path defined by the earth's magnetic field.

The magneto-active plasma presents a dispersive property for

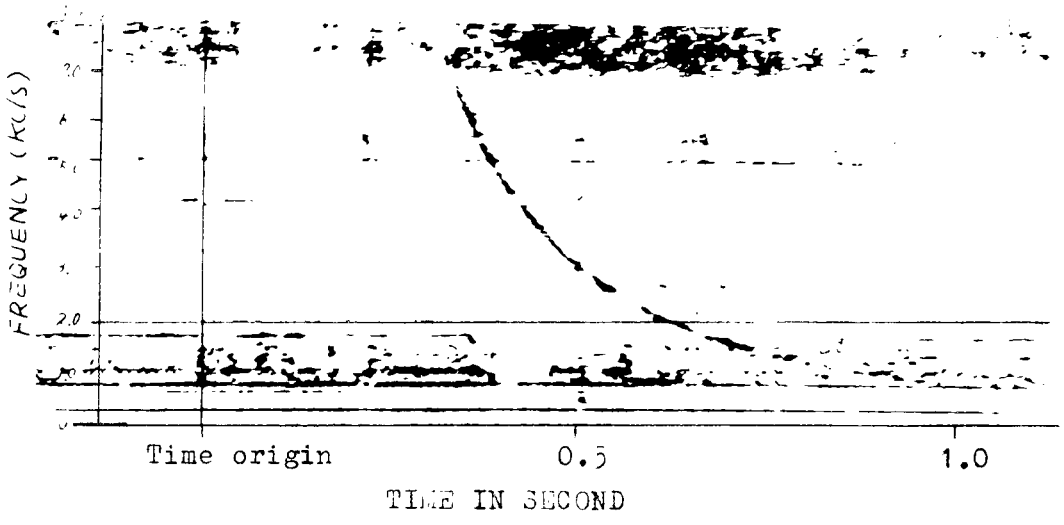


Fig.1.1 An example of the whistler obseved at Toyo..., Japan (After Otsu).

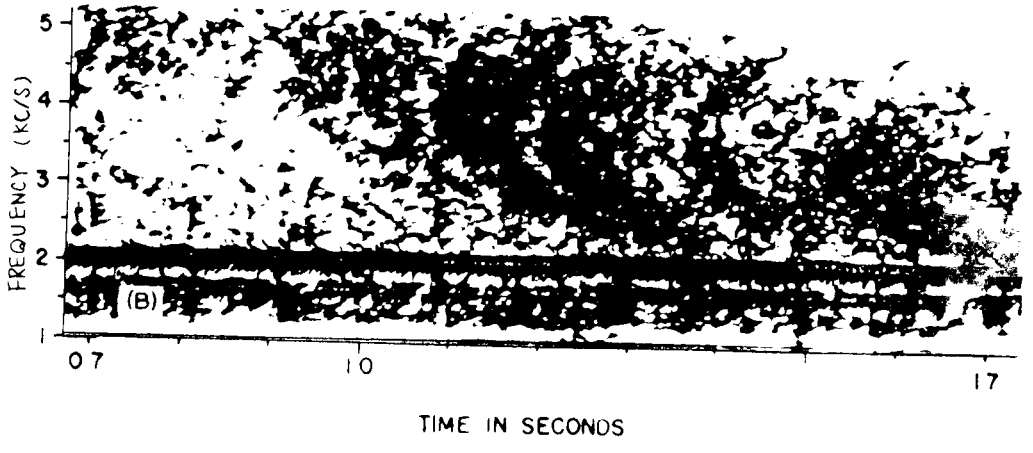


Fig. 1.2 Nose whistler detected at College, Alaska (After Helliwell).

the VLF waves, that is the group velocity of energy propagation varies with frequency; the higher the frequency is, the higher the speed.

On account of the above mentioned properties of the medium in the exosphere, a packet of VLF electromagnetic waves arising from a lightning flash is changed in frequency characteristics and returns back to the ground as a whistling atmospherics, that is a whistler.

Hence the whistler will be heard in the opposite hemisphere to the location of the lightning. In this case, it is called "short whistler". On the other hand, it is possible that the short whistler is reflected at the ground and pursues the same path to reach the hemisphere of the source. This type of whistler is called "long whistler".

The definite difference between the long and the short whistler is that 1) in case of the long whistler the source is located near the receiver, so that the long whistler usually follows a preceding click noise, whereas in case of the short whistler there is no clear click, because the source is located in the opposite hemisphere, and 2) dispersion of the long whistler is approximately twice that of short whistlers observed at the same receiving point.

Like the long whistler, whistler echoes which have traveled many times the unit path by reflections are sometimes observed. These echoes are called whistler train. For the existence of such phenomena the path of the whistler must be exactly symmetrical with respect to the geomagnetic equator, and all echoes travels along the same path. Actually in case of the short type of whistler train, the ratio of the time delays or dispersions of the successive whistlers is approximately 1:3:5: etc., and in case of the long type, the ratio is 2:4:6: etc.

At high latitude, somewhat different type of whistler can be detected(Helliwell et al, 1956). It is of a special feature as shown in Fig. 1.2, which was observed at College, Alaska. This type of whistler is called "nose whistler", standing for the nose shape of the sonagram. This phenomenon is not different from the ordinary whistlers with respect to the origin, as will be mentioned later.

1.2.2 Observation of the whistlers and the properties of the observed whistlers

Since Storey proposed the theory of the origin of the whistler, much observed data have been obtained in many countries of the world. In Table 1.1, the typical stations and their observed whistlers are listed. From these world-wide observations, the following general properties are found.

1) Dispersion (abbreviated as D) of the whistlers greatly depends on the latitude of the observatory; being larger at high latitudes. At low latitude, for example at Toyokawa, Japan D ranges 30 to 50, whereas at high latitude, for example at Washington D.C. it ranges 80 to 100, for the short whistler.

2) For lower latitudes:

a) Only the short whistler has been detected and no repeated whistler, that is the long whistler or the whistler train, has been observed.

b) Every whistler has pure tone.

c) There is a close correlation between D and local values of the the critical frequency of the F_2 layer (Iwai and Otsu, 1958).

Table 1.1 List of observed whistler.

Station	Magnetic dip	Magnetic latitude	Geomagn. latitude	Geomag. dip	Dispersion (D) $sec^{1/2}$ *	Characteristics of whistlers	Observers
College, Alaska	77.5°	66°	64.7°	76.5°		nose whistler, short whistler	Helliwell et al
Hanover N.H.	73°	58.5°	55.2°	71°		short & long whistler	Morgan & Dinger
Washington D.C.	71.5°	56°	50.3°	67.5°	80~100	nose whistler	
Seatle, Washington	70.3°	54.5°	53.6°	69.7°	53~95	short whistler, long whistler	Helliwell
Ireland	68°~69°	51°~52.5°	57~59°	72~73°	65~80	short whistler, whistler pair	Burton & Boadman Potter
Wellington & Dunedin, New Zealand	65° 70°	-47° -54°	-45.4° -50.7°	63.5° 68°	60~100	short & long whist, whist pair & train	Morgan & Mck Allcock
Cambridge, England	67°	50°	54.8°	70.5°	45~80	ditto	Storey
Unalaska, Aleutian Ids	65°	47°	51°	68°	60~100	ditto	Morgan & Mck Allcock
Stanford, California	61.3°	42.5°	43.7°	62.3°	455~925	short whistler	Helliwell
Toyokawa, Japan	48.4°	29.5°	24.5°	42°	30~50	short whistler	Iwai et al
Wakkanai, Japan	59.5°	40.3°	35.5°	55°	42~62	short & long whist, whistler pair	ditto

* The values of D in this table are those for one way whistler, i.e. short whistler

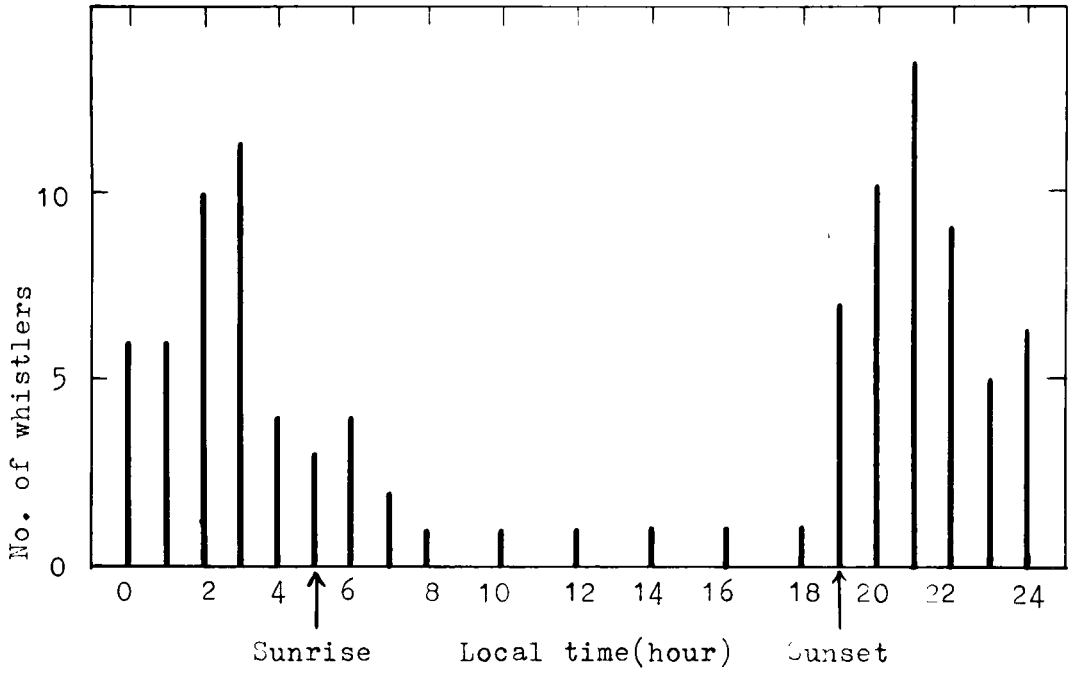


Fig. 1.3 Diurnal variation of the occurrence frequency of the whistlers. (averaged value for July, August, 1956) (after Iwai et al, 1957)

For higher latitudes:

a) All kinds of whistler have been observed.

b) The tone of a whistler is not pure.

c) In the trains of more than ten repetitions, which have been sometimes detected, the loudness reached its maximum at the third or fourth whistler.

3) Daily variation of the occurrence frequency of the whistlers is remarkable. Namely as illustrated in Fig. 1.3, the occurrence frequency rapidly increases just ahead of the ground sunrise and just after the sunset, and usually the whistlers are seldom heard in the day time (Iwai et al, 1957; Kimpara, 1960). This tendency can be explained in terms of the absorption suffered in the ionosphere, which will be stated in Chapter 1.6.

4) Nose whistlers are usually heard at high latitude (Helliwell et al, 1956; Dinger, 1956). The nose frequency. by which we mean the frequency of the crook point on the frequency-time trace as shown in Fig. 1.2, decreases with latitudes.

1.2.3 Whistler mode of propagation of the VLF electromagnetic waves in magneto-active plasma

In case a magnetic field is taken into account, the refractive index n for an electromagnetic wave in a plasma which contains free electrons and protons, is represented by the well-known Appleton and Hartree's formula (Appleton, 1932):

$$n^2 = 1 - \frac{X}{1 - jZ - \frac{\frac{1}{2}Y_T^2}{1 - X - jZ} \pm \sqrt{Y_L^2 + \left(\frac{\frac{1}{2}Y_T^2}{1 - X - jZ}\right)^2}} \qquad (1.1)$$

where the following notations are used (in M.K.S. unit)

e; magnitude of charge of an electron,

m; mass of an electron,

N; electron density (number of free electrons per unit volume),

H; magnetic field intensity,

ν; mean frequency of collisions between a free electron and neutral air molecules,

ε_o; dielectric constant in free space,

μ_o; magnetic permeability in free space,

f; wave frequency,

f_H; gyro-frequency of electrons $= |e| H\mu/2\pi m$,

f_o; electron plasma frequency $= (e^2 N/4\pi\varepsilon_o m)^{\frac{1}{2}}$,

θ ; the angle between the wave normal and the magnetic field,

$X = f_o^2/f^2$,

$Y = f_H/f$,

$Y_L = Y\cos\theta$, $Y_T = Y\sin\theta$,

$Z = \nu/2\pi f$.

In (1.1) the plus and minus signs indicate the ordinary and extra-ordinary modes respectively.

Now we are going to apply it to the whistler. In terms of the quasi-longitudinal approximation (Booker, 1935; Lepechinsky, 1956), the refractive index n is simplified as

$$n^2 = 1 - \frac{f_o^2}{f(f \pm f_H \cos\theta)} . \qquad (1.2)$$

If $f_o^2 \gg f f_H$ holds, the above is more simplified to

$$n^2 = - \frac{f_o^2}{f(f \pm f_H \cos\theta)} . \qquad (1.3)$$

Actually, for the ordinary mode n^2 becomes negative, so that there is no actual propagation of the energy. Hereafter only the extraordinary mode will be the subject of the argument.

In general the phase velocity v_{ph} of a wave in this medium is given by $v_{ph} = c/n$, and the group velocity v_g is given by

$$\left.\begin{aligned} v_g &= c/n' \\ n' &= n + f\frac{dn}{df} \end{aligned}\right\} \qquad (1.4)$$

Then the group velocity for n given by (1.2) becomes

$$v_g = 2c\,\frac{(f_H\cos\theta - f)^{\frac{3}{2}}\cdot\left[f^2(f_H\cos\theta - f) + f f_0^2\right]^{\frac{1}{2}}}{2f^3 - 4f^2 f_H\cos\theta + 2f f_H^2\cos^2\theta + f_H f_0^2\cos\theta} \,. \qquad (1.5)$$

This expression can be simplified if f_0^2 is much larger than $f f_H$, as

$$v_g = 2c\,\frac{f^{\frac{1}{2}}(f_H\cos\theta - f)^{\frac{3}{2}}}{f_H f_0\cos\theta} \,. \qquad (1.6)$$

From the above, it can be understood that v_g becomes maximum at $f = \frac{1}{4}f_H\cos\theta$ and becomes zero at $f = 0$, or $f_H\cos\theta$, and also there is no propagation for $f > f_H$.

The gyro-frequency f_H is clearly determined by the magnetic field intensity H. In case of the whistler propagation, the magnetic field is assumed to be dipole field so that it will decrease like r^{-3} where r is the distance from the center of the earth. Consequently a wave of frequency f larger than the f_H at the apex of its ray path, can not be propagated along such a path.

Especially f being much smaller than f_H, (1.3) becomes

$$n = f_0/\sqrt{f f_H\cos\theta} \,. \qquad (1.7)$$

This is the expression which Eckersley and Storey (1953) used to explain the whistlers.

In an anisotropic medium, the wave normal of a wave is not always identical with the direction of energy propagation. The latter is usually called ray direction, and in the magneto-active medium such as in the ionosphere and exosphere, these directions are functionally related with the direction of the magnetic field as follows. Namely as shown in Fig. 1.4, when the wave normal(N) makes an angle θ from the direction of the magnetic field(H), and the ray direction (R) is measured from the wave normal(N) by the angle χ, reckoned positive if the normal lies between the field and the ray, the relation will be (Bremmer, 1949)

$$\tan\chi = -\frac{1}{n}\frac{\partial n}{\partial\theta} = -\frac{1}{2n^2}\frac{\partial n^2}{\partial\theta} \quad . \tag{1.8}$$

In case $f_0^2/f f_H \gg 1$ and $f \ll f_H$, eqs. (1.7) and (1.8) result in

$$\tan\chi \simeq -\tfrac{1}{2}\tan\theta \tag{1.9}$$

On account of this relation, the angle between the ray direction and the direction of the magnetic field, $(\theta + \chi)$ is confined within $19^\circ 29'$, irrespective of θ as shown in Fig. 1.5. Thus it can be imagined that the wave energy travels in the direction close to the line of force of the magnetic field.

Now let us calculate the time of propagation of the VLF waves in the magneto-active plasma. The group velocity v_g given by (1.6) is the velocity in the wave normal direction. Then the velocity v_{gR} in the ray direction will be expressed by

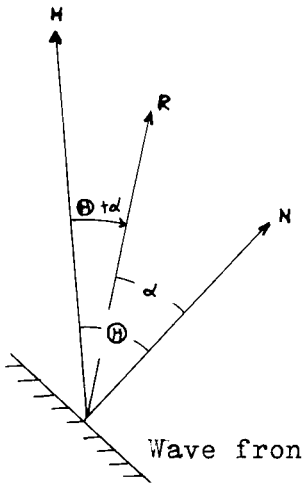


Fig. 1.4 Configulation of magnetic field(H), wave normal(N) and ray direction(R).

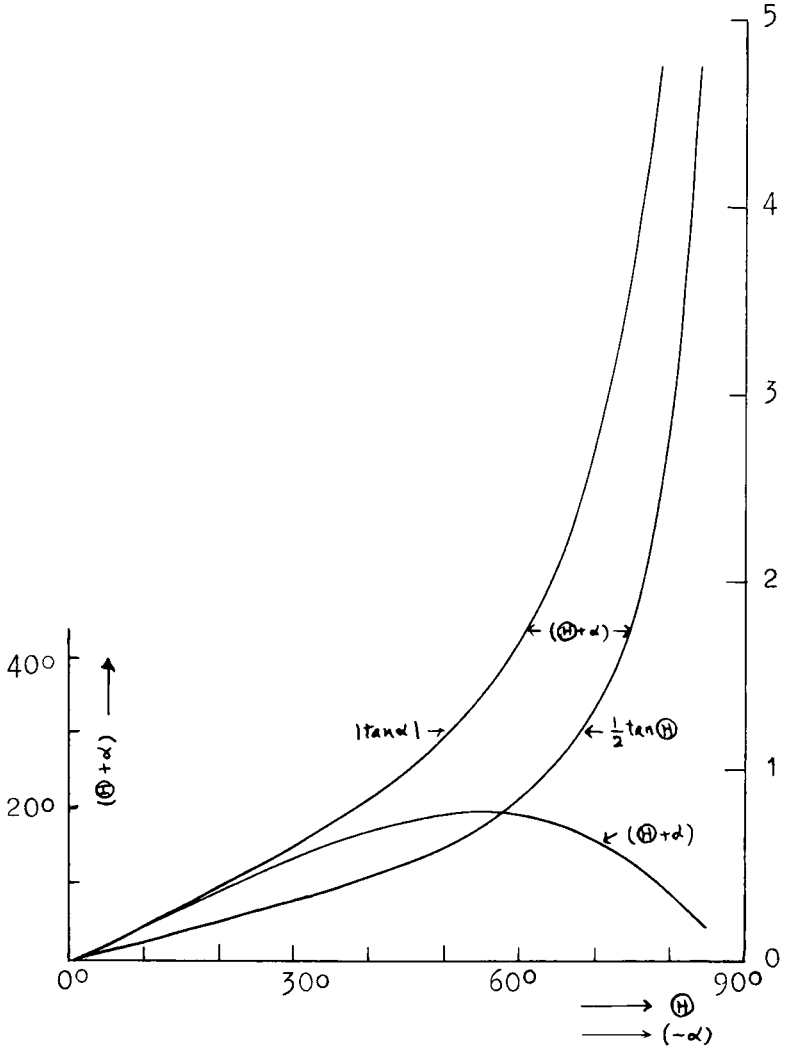


Fig. 1.5 Relation among Θ , α , and (Θ + α).

$$v_{gR} = v_g \sec\alpha = 2c \frac{\sqrt{f f_H}}{f_0} \frac{(\cos\theta - f/f_H)^{\frac{3}{2}}}{\cos\theta \cdot \cos\alpha} . \qquad (1.10)$$

The time of propagation is, then, given by

$$t = \int \frac{ds}{v_{gR}} = \frac{1}{2c} \int \frac{f_0}{\sqrt{f f_H}} \frac{\cos\alpha}{\sqrt{\cos\theta - f/f_H}} \frac{\cos\theta}{\cos\theta - f/f_H} \, ds , \qquad (1.11)$$

where the integration is performed over the whole ray path of its passage. In the above integral the quantity

$$\frac{\cos\alpha}{\sqrt{\cos\theta - f/f_H}} \cdot \frac{\cos\theta}{\cos\theta - f/f_H} , \qquad (1.12)$$

is a function of θ with f/f_H as parameter. This function is shown in Fig. 1.6. Then if f/f_H is sufficiently small, (1.12) is approximately unity for $\theta < 70^\circ$, so that (1.11) becomes

$$t = D f^{-\frac{1}{2}}, \qquad (1.13)$$

where

$$D = \frac{1}{2c} \int \frac{f_0}{\sqrt{f_H}} \, ds , \qquad (1.14)$$

D in (1.14) is a constant value independent of the frequency f, and uniquely determined by the distribution of the electron density and the magnetic field intensity along the ray path. From this relation, the frequency-time characteristics of the whistlers can be clearly understood.

If the frequency f is not sufficiently small as compared with f_H, the propagation time t must be calculated by (1.11).

As is clear from (1.6), the group velocity v_g becomes maximum at $f = \frac{1}{2} f_H \cos\theta$, so that t given by (1.11) will become minimum at a certain frequency f near $\frac{1}{4} f_H \cos\theta$. In this case a nose whistler is produced (Helliwell et al, 1956).

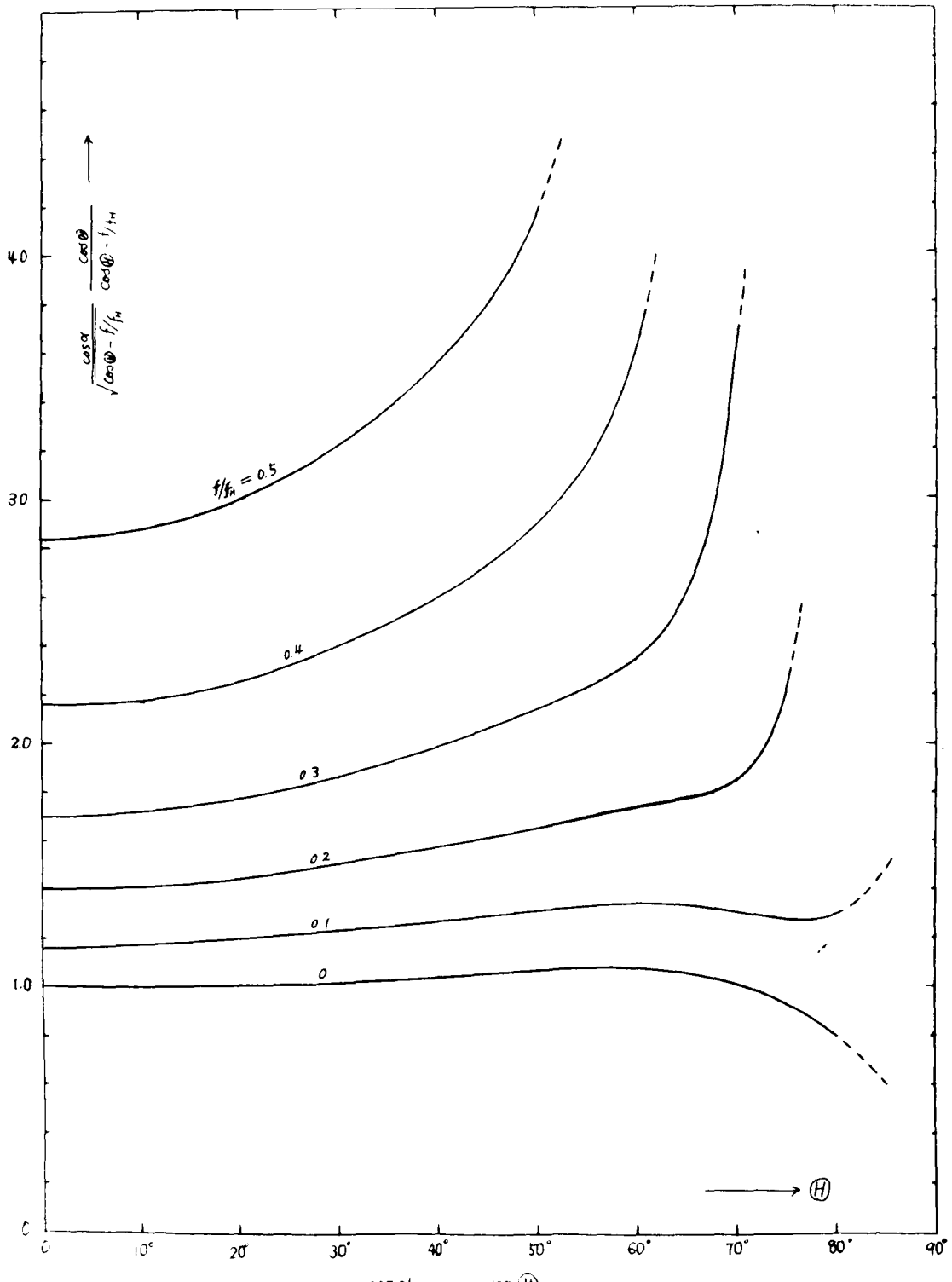


Fig. 1.6 Graph of $\frac{\cos\alpha}{\sqrt{\cos\textcircled{H}-f/f_H}} \cdot \frac{\cos\textcircled{H}}{\cos\textcircled{H}-f/f_H}$ versus Ⓗ, with f/f_H as parameter.

Chapter 1.3 A method of ray tracing of the VLF electromagnetic wave in the ionosphere and exosphere

1.3.1 Fermat's principle and Euler's differential equations

In general, in order to get the ray path of a wave, Fermat's principle is usable within the range in which geometrical optics is approximately applied. This principle says that a ray passing through any medium takes such a path that propagation time in the medium becomes minimum, and it is expressed in the formula,

$$\delta \int n_\alpha \, ds = 0, \tag{1.15}$$

where

$$n_\alpha = n \cos\alpha , \tag{1.16}$$

and α is given by (1.8). ds in (1.15) is the ray path element.

Let us take the ray path as given in the xyz-coordinate system by the two equations,

$$x = x(z), \quad y = y(z), \tag{1.17}$$

then ds is expressed as follows.

$$\left.\begin{aligned} ds &= \sqrt{1 + x'^2 + y'^2}\,dz , \\ \text{where} \qquad x' &= dx/dz, \quad y' = dy/dz, \end{aligned}\right\} \tag{1.18}$$

n_α is a function of x, y, x', y', and z. When we write

$$n_\alpha \sqrt{1+x'^2+y'^2} \equiv F(x,y,z,x',y') , \tag{1.19}$$

(1.15) becomes

$$\delta \int F(x,y,z,x',y')dz = 0 . \tag{1.20}$$

This is transformed into the following two equations by the theory on the calculus of variations.

$$\left.\begin{aligned} \frac{d}{dz}\left(\frac{\partial F}{\partial x'}\right) - \frac{\partial F}{\partial x} &= 0, \\ \frac{d}{dz}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} &= 0, \end{aligned}\right\} \tag{1.21}$$

These equations are called the Euler's differential equations.

If the path exists in a plane and is given in polar coordinate by the following equation

$$r = r(\theta), \tag{1.22}$$

ds in (1.15) is given by

$$ds = \sqrt{r^2 + r'^2}\, d\theta, \tag{1.23}$$

where r represents radius, θ is colatitude, and $r' = dr/d\theta$. Since n_α is a function of r, r', θ, if we write

$$n_\alpha \sqrt{r^2 + r'^2} \equiv F(r, r', \theta), \tag{1.24}$$

the equation (1.15) becomes

$$\delta \int F(r, r', \theta) d\theta = 0. \tag{1.25}$$

This is transformed into the following Euler's differential equation.

$$\frac{d}{d\theta}\left(\frac{\partial F}{\partial r'}\right) - \frac{\partial F}{\partial r} = 0. \tag{1.26}$$

These equations (1.21) and (1.26) are clearly equivalent to the second order non-linear ordinary differential equations, but it is too difficult to be analytically solved in general.

1.3.2 Solution of the Euler's equations

1) Plane case

The Euler's equations (1.21) are to give the path of a ray traversing through arbitrary medium. But it cannot be solved except

in a special case as mentioned before. This special case is the plane case that the constants of the medium (including the direction of magnetic field) vary only in one direction, e.g. the z axis. This case will be applied to getting ray paths in the ionosphere regarded as a planely stratified layer in which the direction of magnetic field is invariable.

We know that this case will be reduced to the Snell's law for wave normal. But the author could not find any papers verifying it in the anisotropic case, except the Bremmer's(1949) in which he has treated it wave-theoretically and has deduced the Snell's law from the boundary conditions between two adjacent different media, planely stratified. So it seems to be worth while to derive the Snell's law through the Euler's differential equations from the Fermat's principle. Therefore a few spaces of this paper will be given for the derivation of the law.

We take the xyz-Cartesian coordinate in which, without missing generality, the z-axis is made as the direction perpendicular to the planely stratified layer and the constant magnetic field $\overrightarrow{OH}$ lies in the xz-plane, at an angle ψ to z-axis as shown in Fig. 1.7. Then, we denote the angle between the ray direction $\overrightarrow{OR}$ and the z-axis by i and the angle between the projection of the ray direction upon the xy-plane and the x-axis by γ. In the same way, the wave normal $\overrightarrow{ON}$ is determined by the angle β and δ.

According to Bremmer, the relation between the wave normal, the ray direction and the magnetic field is given by (1.8) and they lie in one plane. Their configulation is illustrated in Fig. 1.4. The angle between the ray and the magnetic field is then represented by

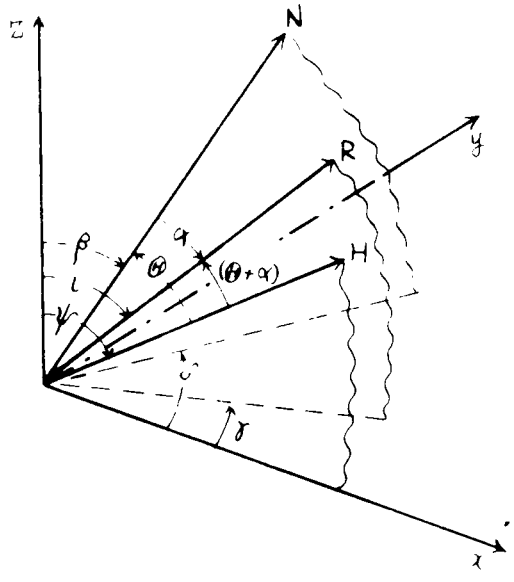


Fig. 1.7 Relation of the directions of magnetic field(H), wave normal(N) and ray direction in three dimensional case.

$(\Theta + \alpha)$ and the following relation holds between $\Theta + \alpha$, i, γ, and ψ.

$$\cos(\Theta + \alpha) = \sin\psi \sin i \cos\gamma + \cos\psi \cos i . \tag{1.27}$$

By the conditions already postulated, n is independent of the coordinates x and y, and also $F(\equiv n\alpha\sqrt{1+x'^2+y'^2})$ does. Then $\frac{\partial F}{\partial x} = \frac{\partial F}{\partial y} = 0$ and (1.21) becomes simply the following set of two equations.

$$\frac{d}{dz}\left(\frac{\partial F}{\partial x'}\right) = 0 \longrightarrow \frac{\partial F}{\partial x'} = \text{const.}, \tag{1.28}$$

$$\frac{d}{dz}\left(\frac{\partial F}{\partial y'}\right) = 0 \longrightarrow \frac{\partial F}{\partial y'} = \text{const.}. \tag{1.29}$$

By (1.19), the equations (1.28) and (1.29) are calculated as

$$\sqrt{1+x'^2+y'^2}\,\frac{\partial n\alpha}{\partial x'} + n\alpha\,\frac{x'}{\sqrt{1+x'^2+y'^2}} = C_1, \tag{1.30}$$

$$\sqrt{1+x'^2+y'^2}\,\frac{\partial n\alpha}{\partial y'} + n\alpha\,\frac{y'}{\sqrt{1+x'^2+y'^2}} = C_2. \tag{1.31}$$

$\frac{\partial n\alpha}{\partial x'}$ and $\frac{\partial n\alpha}{\partial y'}$ are given by

$$\frac{\partial n\alpha}{\partial x'} = \frac{\partial n\alpha}{\partial \Theta}\left\{\frac{\partial \Theta}{\partial i}\frac{\partial i}{\partial x'} + \frac{\partial \Theta}{\partial \gamma}\frac{\partial \gamma}{\partial x'}\right\}, \tag{1.32}$$

$$\frac{\partial n\alpha}{\partial y'} = \frac{\partial n\alpha}{\partial \Theta}\left\{\frac{\partial \Theta}{\partial i}\frac{\partial i}{\partial y'} + \frac{\partial \Theta}{\partial \gamma}\frac{\partial \gamma}{\partial y'}\right\}, \tag{1.33}$$

where

$$\frac{\partial n\alpha}{\partial \Theta} = \frac{\partial(n\cos\alpha)}{\partial \Theta} = n\cos\alpha\left\{\frac{1}{n}\frac{\partial n}{\partial \Theta} - \tan\alpha\frac{\partial \alpha}{\partial \Theta}\right\}$$

$$= -n\alpha\tan\alpha\left(1 + \frac{\partial \alpha}{\partial \Theta}\right), \quad (\text{by (1.16)}) \tag{1.34}$$

and

$$\left.\begin{aligned}\frac{\partial \Theta}{\partial i} &= \frac{\sin i\cdot\cos\psi - \sin\psi\cdot\cos i\cdot\cos\gamma}{\sin(\Theta+\alpha)\left(1+\frac{\partial \alpha}{\partial \Theta}\right)}\\ \frac{\partial \Theta}{\partial \gamma} &= \frac{\sin i\cdot\sin\gamma\cdot\sin\psi}{\sin(\Theta+\alpha)\left(1+\frac{\partial \alpha}{\partial \Theta}\right)}\end{aligned}\right\}, \tag{1.35}$$

which can be obtained from (1.27) at once. And by means of

$$x' = \tan i \cdot \cos\gamma, \quad y' = \tan i \cdot \sin\gamma,$$

$$\text{or} \quad i = \tan^{-1}\sqrt{x'^2 + y'^2}, \quad \gamma = \tan^{-1}(y'/x'),$$

$\frac{\partial i}{\partial x'}, \frac{\partial i}{\partial y'}, \frac{\partial \gamma}{\partial x'}$, and $\frac{\partial \gamma}{\partial y'}$ are calculated as

$$\left.\begin{aligned} \frac{\partial i}{\partial x'} &= \frac{1}{1 + x'^2 + y'^2}\frac{x'}{\sqrt{x'^2 + y'^2}}, & \frac{\partial \gamma}{\partial x'} &= -\frac{y'}{x'^2 + y'^2}, \\ \frac{\partial i}{\partial y'} &= \frac{1}{1 + x'^2 + y'^2}\frac{y'}{\sqrt{x'^2 + y'^2}}, & \frac{\partial \gamma}{\partial y'} &= \frac{x'}{x'^2 + y'^2}, \end{aligned}\right\} \quad (1.36)$$

By using the equations (1.32) to (1.36), the equations (1.30) and (1.31) are transformed into

$$n\cos\alpha\cdot\cos\gamma\left[\sin i - \frac{\tan\alpha}{\sin(\textcircled{H}+\alpha)}\left\{(\sin i\cdot\cos\psi - \sin\psi\cdot\cos i\cdot\cos\gamma)\cos i - \sin\gamma\tan\gamma\sin\psi\right\}\right] = C_1, \quad (1.37)$$

$$n\,\frac{\sin\gamma\cdot\sin i\cdot\sin\textcircled{H}}{\sin(\textcircled{H}+\alpha)} = C_2, \quad (1.38)$$

respectively through a little cumbersome deduction.

These equations are for the ray direction, that is, the ray direction at an arbitrary point on a ray path is determined by these complicated equations. But, when we transform these equations by using the angle β and δ relating to wave normal (see Fig. 1.7), they become respectively

$$n \sin\beta\cos\delta = c_1, \quad (1.39)$$

$$n \sin\beta\sin\delta = c_2. \quad (1.40)$$

From these, we find easily

$$\delta = \text{const.}, \quad (1.41)$$

$$n \sin\beta = \text{const.}. \quad (1.42)$$

This means that the wave normal always lies in a plane parallel to that formed by the incident wave normal and the z-direction, and the

angle of the wave normal with the axis z is given by (1.42). Here we have obtained the Snell's law for the wave normal. The deduction of the equations (1.39) and (1.40) from (1.37) and (1.38) will be listed in the Appendix 1.1.

2) Circularly curved case

When medium is spherically stratified, it is convenient to treat such a case in spherical coordinate system. Though this case is more general than the plane case obtained in the last paragraph more difficult calculations may be required.

Then we will deal with a simpler case; the ray path lies in one plane containing magnetic field and radius vector. This case was treated in 1.3.1 in the polar coordinates (r,θ) and reduced to the Euler's differential equation (1.26).

Now that the ray path is taken as $r = r(\theta)$ in the polar coordinates as shown in Fig. 1.8, then

$$\tan i = r/r', \quad (r' = dr/d\theta) \tag{1.43}$$

$$\text{Ⓗ} + \alpha = i - \psi, \tag{1.44}$$

where i is the angle between the ray direction and the radius vector and ψ is the angle between the magnetic field and the radius. In this case ψ is a function of θ, and Ⓗ or α is a function of r, r', and θ. Then we must solve the equation (1.26) from F as a function of r, r' and θ. As it is difficult to get the solution analytically, the following approximation is devised. As shown in Fig. 1.9, the space is divided into many small sectors, whose boundaries are defined by $\theta_0, \theta_1, \theta_2, \cdots\cdots$ etc. In each sector ψ is taken as constant, i.e. as having the average value in each sector, say $\psi_{01} = \psi(\frac{\theta_0 + \theta_1}{2})$, $\psi_{12} = \psi(\frac{\theta_1 + \theta_2}{2}) \cdots\cdots$etc. And at the boundary of two sectors ψ

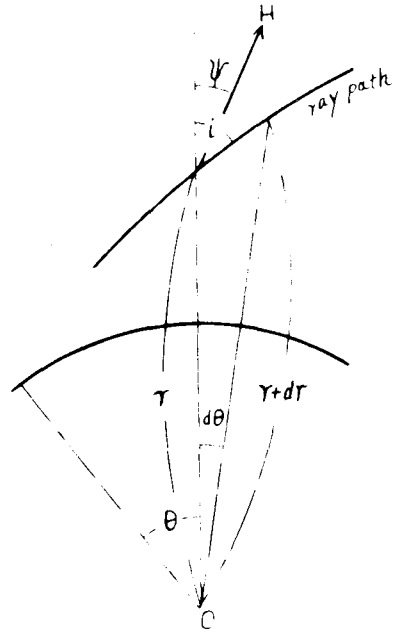


Fig. 1.8 Ray path in polar coordinates.

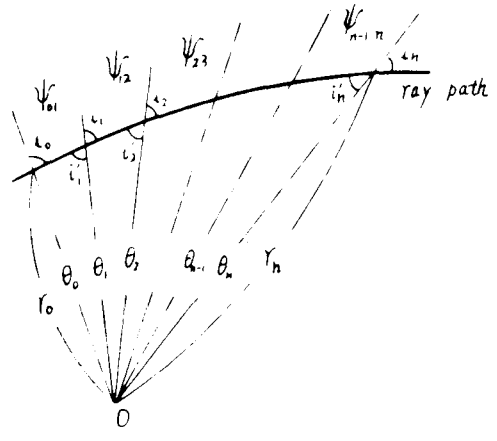


Fig. 1.9 Division of the space into sectors.

changes abruptly and here $\frac{\partial F}{\partial r}$ is taken to be zero.

From now on, we call the propagation within each sector as case "A" and the refraction at the sector boundary as case "B". Application of (1.26) to case "A" determines the ray path in each sector, and to case "B" gives the discontinuous change of the ray direction (i). These two operations are repeated in turn to determine the ray path in the outer ionosphere. And in this case, of course, the more finely the space is divided, the better approximation will result in.

Case "A" (ψ = constant)

As F is a function of r and r', (1.26) is transformed into

$$\frac{d}{d\theta}\left\{r' \frac{\partial F}{\partial r'} - F\right\} = 0 \quad \text{or} \quad r'\frac{\partial F}{\partial r'} - F = \text{const.}. \tag{1.45}$$

And it is calculated by (1.24), then

$$n_\alpha \left\{ - \frac{r^2}{\sqrt{r^2+r'^2}} + \sqrt{r^2+r'^2}\, \frac{r'}{n_\alpha}\frac{\partial n_\alpha}{\partial r'}\right\} = C.. \tag{1.46}$$

Here we know clearly

$$\frac{1}{n_\alpha}\frac{\partial n_\alpha}{\partial r'} = \frac{1}{n_\alpha}\cdot\frac{\partial n_\alpha}{\partial \textcircled{H}}\cdot\frac{\partial \textcircled{H}}{\partial i}\cdot\frac{\partial i}{\partial r'} \quad . \tag{1.47}$$

$\frac{1}{n_\alpha}\frac{\partial n_\alpha}{\partial \textcircled{H}}$ is obtained by (1.34). $\frac{\partial \textcircled{H}}{\partial i}$ and $\frac{\partial i}{\partial r'}$, are calculated from (1.44) and (1.43) as

$$\left.\begin{aligned} \frac{\partial \textcircled{H}}{\partial i} &= \frac{1}{1+\frac{\partial \alpha}{\partial \textcircled{H}}} \quad , \\ \frac{\partial i}{\partial r'} &= -\frac{r}{r^2+r'^2} \quad . \end{aligned}\right\} \tag{1.48}$$

From these equations, we obtain

$$\frac{1}{n_\alpha}\frac{\partial n_\alpha}{\partial r'} = \frac{r}{r^2+r'^2}\tan\alpha \, , \tag{1.49}$$

by using (1.49) and considering the equations

$$\frac{r}{\sqrt{r^2+r'^2}} = \sin i, \quad \frac{r'}{\sqrt{r^2+r'^2}} = \cos i, \quad n\alpha = n\cos\alpha,$$

then (1.46) ultimately becomes

$$nr \sin(i-\alpha) = \text{const.}, \tag{1.50)'}$$

or by (1.44) $$nr \sin(\Theta + \psi) = \text{const.} \tag{1.50}$$

From Fig. 1.10, we can understand that $(\Theta + \psi)$ or $(i - \alpha)$ is equal to β, the angle between the wave normal and the radius vector. Then we find that (1.50) or (1.50)' is rewritten as

$$nr \sin\beta = \text{const.} \tag{1.51}$$

This is the Snell's law corresponding to the wave normal for the spherical case that the constants of the medium change radially.

Case "B" $\left(\frac{\partial F}{\partial r} = 0\right)$

From (1.26), we get

$$\frac{d}{d\theta}\left(\frac{\partial F}{\partial r'}\right) = 0 \quad \text{or} \quad \frac{\partial F}{\partial r'} = \text{const.} \tag{1.52}$$

After the calculation as in case "A" we finally obtain

$$n\cos(i - \alpha) = \text{const.}, \tag{1.53)'}$$

or $$n\cos(\Theta + \psi) = \text{const.}, \tag{1.53}$$

or $$n\cos\beta = \text{const.} \tag{1.54}$$

In spite that (1.54) consists of n and $\cos\beta$ (not $\sin\beta$), we can realize that this equation also is Snell's law as being explained here. In Fig. 1.11, the arrow r indicates a certain radius vector which corresponds to the boundary in this case, and in both side of the boundary the media are completely homogeneous besides the difference

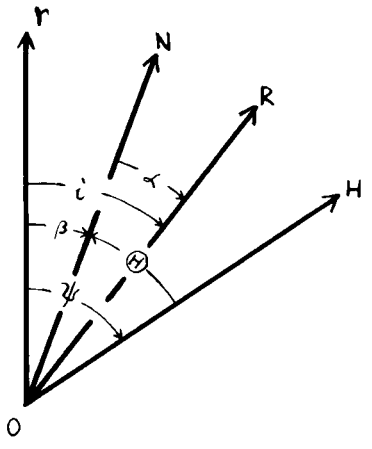


Fig. 1.10 Configulation of radius vector(r), wave normal(N), ray direction(R) and magnetic field(H).

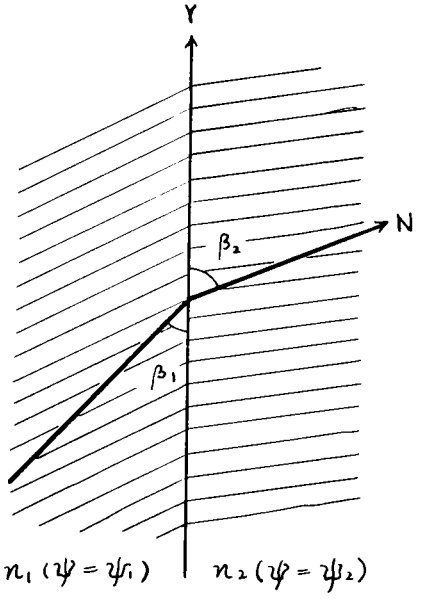


Fig. 1.11 Boundary in case "B" and refraction of the wave normal.

of ψ. Imagine a wave coming upon the boundary from the left with its wave normal having the angle $(\frac{1}{2}\pi - \varphi_1)$ with the direction perpendicular to the boundary, then the wave normal of the transmitted wave will be given by the angle $(\frac{1}{2}\pi - \varphi_2)$ which is determined by

$$n_2 \sin(\tfrac{1}{2}\pi - \varphi_2) = n_1 \sin(\tfrac{1}{2}\pi - \varphi_1) \,. \qquad (1.55)$$

This equation (1.55) is identical with (1.54).

Chapter 1.4 Ray tracing of the whistlers

1.4.1 Fundamental equations for ray tracing of the whistlers

A whistler wave generated by a lightning flash on the ground, will strike the ionosphere first of all. So it must be obtained in what path the wave is propagated in the ionosphere. However, considering the depth of the ionosphere to be very thin as compared with the whole path length, such calculation of the path in the ionosphere will not be so significant. Hence it will be sufficient to find only the direction which the ray will take at the top of the ionosphere (here we call the level of the maximum electron density, about 300Km above the ground, as the top of the ionosphere).

According to the discussion given in 1.3.2-1), a wave inside the plane ionosphere is governed by the Snell's law (1.42). In (1.42), refractive index n is assumed to be given by (1.7), although one may have a few objections such as the neglection of the collision frequency in the lower ionosphere. But near the top of the ionosphere, we can use (1.7). Then, by (1.42) and (1.7), we get

$$\frac{f_0}{\sqrt{f f_H}} \frac{\sin\beta}{\sqrt{\cos\Theta}} = \text{const.}, \qquad (1.56)$$

In the ionosphere, the electron density N increases very rapidly with increasing height, thus f_0 becomes very large at the top of the ionosphere compared with the value of the lower part, e.g. 10Mc/s at 300Km and 0.5Mc/s at the lower part of the E layer. Then

owing to the initial value of the right hand side constant in (1.56) at the lower part of the E layer, $\sin\beta$ in (1.56) becomes very small when f_0 becomes very large at the top. As a first approximation, it may be considered that β is zero at the top, regrdless the initial direction of the incident wave. By making β nil,

$$i - \alpha = 0 \quad \text{or} \quad i = \alpha, \tag{1.57}$$

$$\textcircled{H} + \psi = 0 \quad \text{or} \quad \textcircled{H} = -\psi. \tag{1.58}$$

And substituting (1.7) for n^2 of (1.8), we obtain

$$\tan\alpha = -\frac{1}{2}\tan\textcircled{H}. \tag{1.59}$$

From (1.57), (1.58) and (1.59), we finally get

$$\tan i_0 = \frac{1}{2}\tan\psi, \tag{1.60}$$

where i_0 is the value of i at the top. Thus we know that the ray at the top of the ionosphere depends only on the direction of the magnetic force of the latitude there, regardless the initial directions at the lower boundary of the ionosphere.

In the second place, let us consider the paths above the ionosphere. The paths will be obtained by using the method described in 1.3.2-2). As a refractive index, (1.3) is used. Then,

(Case "A") by (1.50)' and (1.3) ,

$$\frac{\gamma f_0}{\sqrt{f f_H}} \frac{\sin(\textcircled{H}+\psi)}{\sqrt{\cos\textcircled{H} + f/f_H}} = \text{const.} \tag{1.61}$$

(Case "B") by (1.54) and (1.3) ,

$$\frac{f_0}{\sqrt{f f_H}} \frac{\cos(\textcircled{H}+\psi)}{\sqrt{\cos\textcircled{H} - f/f_H}} = \text{const.}$$

or

$$\frac{\cos(\textcircled{H}+\psi)}{\sqrt{\cos\textcircled{H} - f/f_H}} = \text{const.} \tag{1.62}$$

because in case "B", $\frac{f_0}{\sqrt{ff_H}}$ has same value in both sides of the boundary.

1.4.2 Charts for calculation of ray paths

The value of

$$\frac{\sin(\textcircled{H}+\psi)}{\sqrt{\cos\textcircled{H}-f/f_H}} \tag{1.63}$$

is obtained from the given $\textcircled{H}$ with parameters ψ , and f/f_H. From (1.8) and (1.3) , we have

$$\tan\alpha = -\frac{1}{2}\frac{\sin\textcircled{H}}{\cos\textcircled{H}-f/f_H} \quad . \tag{1.64}$$

Consequently i is determined correspondingly from the given $\textcircled{H}$ with the parameters ψ and f/f_H. Therefore, we can produce charts of $\sin(\textcircled{H}+\psi)/\sqrt{\cos\textcircled{H}-f/f_H}$ vs. i with ψ and f/f_H as parameters.

The author has provided the charts of the cases

$\psi = 0^o, 10^o, 20^o, 30^o, 40^o, 50^o, 60^o, 70^o, 80^o, 90^o$

$f/f_H = 0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7.$

Namely, the charts in the Set-I (Fig. 1.12) are made with ψ as a parameter at each given value of f/f_H say, 0, 0.1, 0.2, -----etc. On the contrary, the charts in the Set-II (Fig.1.13) are made with f/f_H as a parameter at each given value of ψ . For the actual calculations, Set-II may be more convenient.

As to the curves of (1.63) corresponding to ψ in the range

$$90^\circ \le \psi \le 180^\circ ,$$

we find that the above produced charts are also usable, by such substitutions as shown here.

Consider the identity

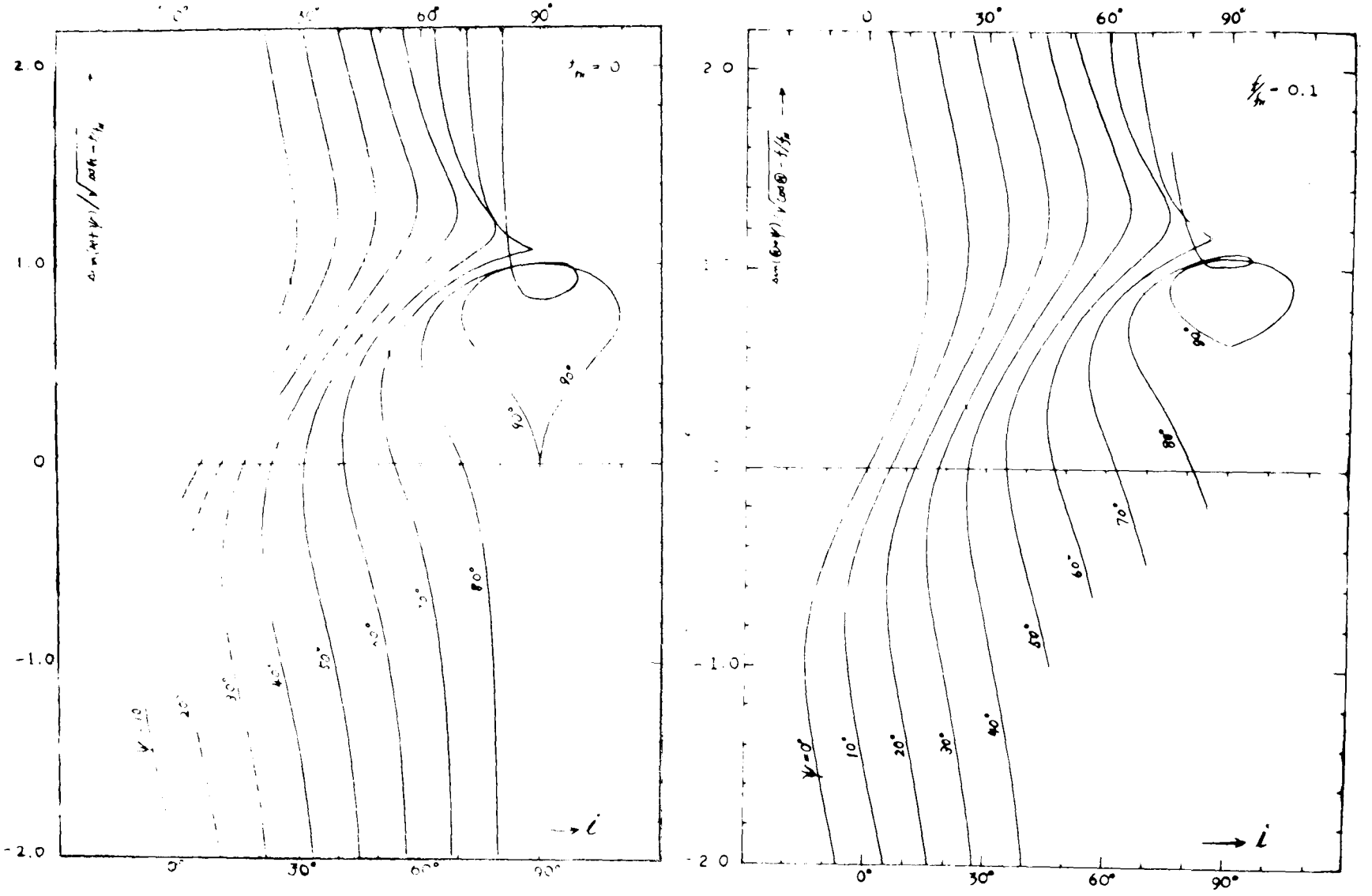


Fig. 1.12(a),(b) $(\sin\theta + y)/\sqrt{\cos\theta - f/f_H}$ against i for f/f_H = 0, 0.1.

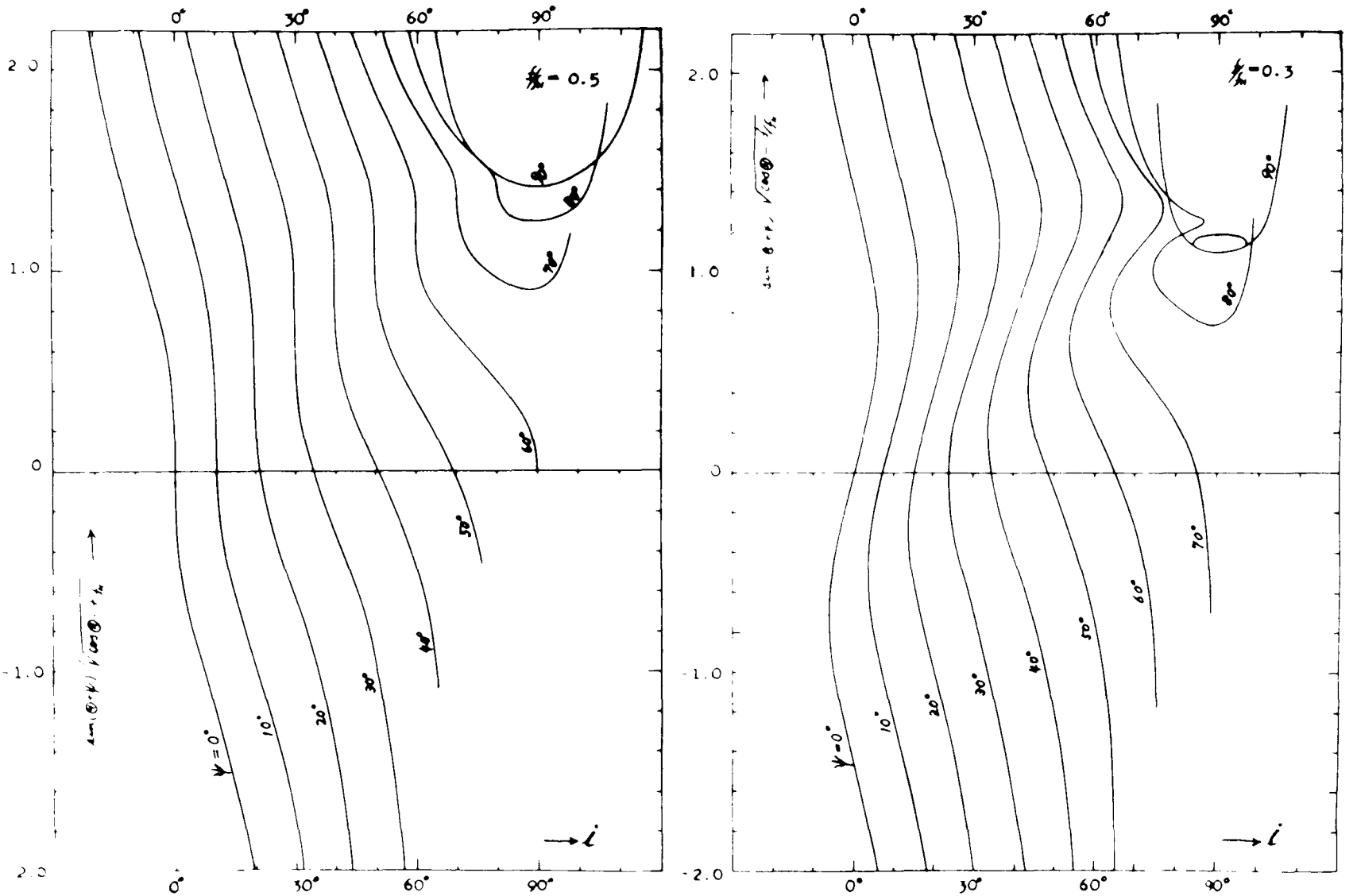


Fig. 1.12(c),(d) $\sin(\Theta - \mu)/\sqrt{\cos\Theta - f/f_H}$ against i for f/f_H = 0.3, 0.5.

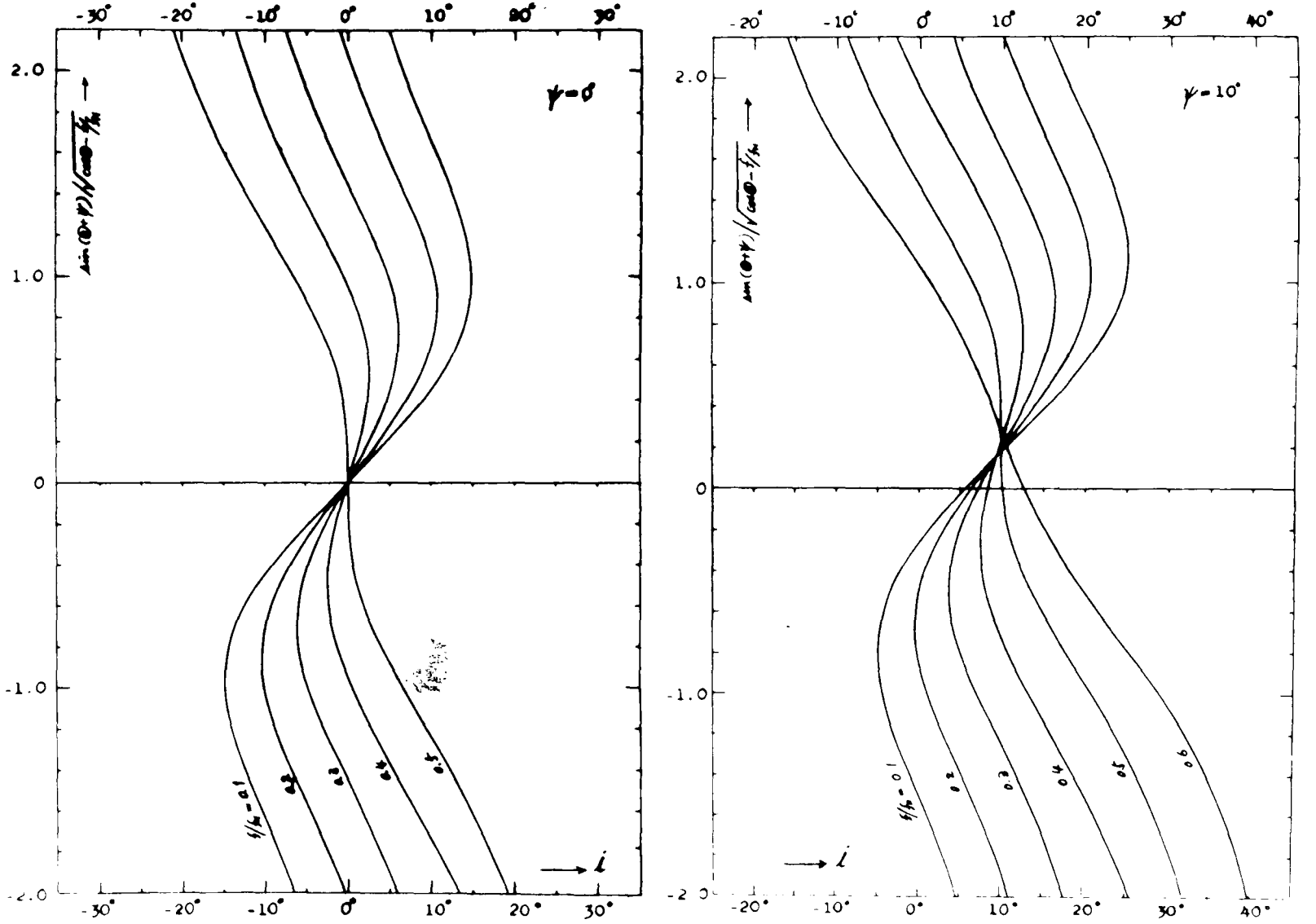


Fig.1.13(a),(b) $\sin(\Theta - \psi)/\sqrt{\cos\Theta - f/f_H}$ against i for $\psi = 0^\circ$, 10°.

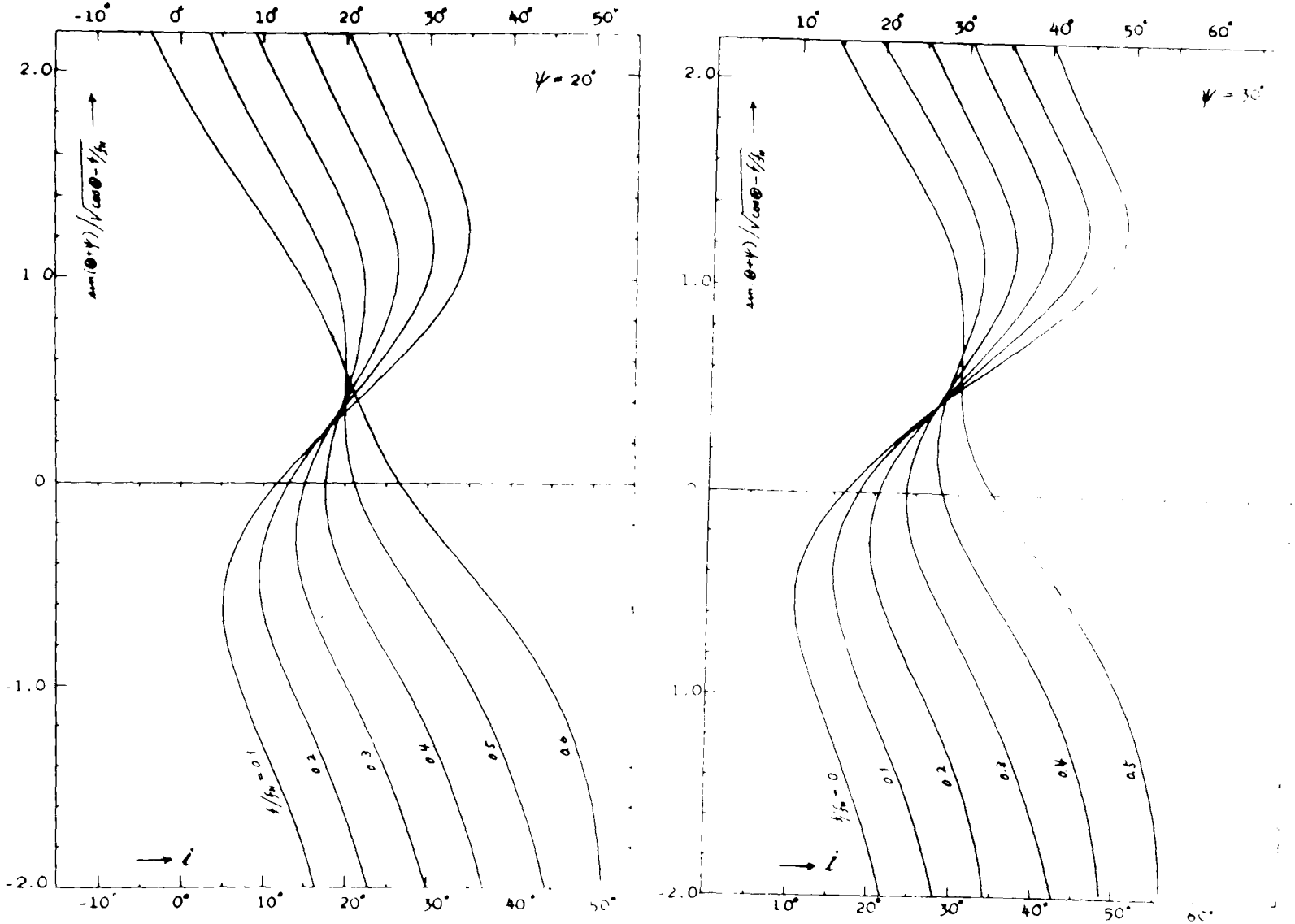


Fig. 1.13(c),(d) $\sin(\theta + \psi)/\sqrt{\cos\theta - f/f_N}$ against i for $\psi = 0^\circ$, 10°.

$$\frac{\sin(\textcircled{H}+\psi)}{\sqrt{\cos\textcircled{H}-f/f_H}} = \frac{\sin(-\textcircled{H}+\pi-\psi)}{\sqrt{\cos(-\textcircled{H})-f/f_H}} = \frac{\sin(\textcircled{H}'+\psi')}{\sqrt{\cos\textcircled{H}'-f/f_H}} \quad ,$$

where $-\textcircled{H} = \textcircled{H}'$ and $\pi - \psi = \psi'$.

By (1.64), $-\textcircled{H} = \textcircled{H}'$ will cause $-\alpha = \alpha'$. Then we have the relation

$$i = \textcircled{H} + \alpha + \psi = -\textcircled{H}' - \alpha' + \pi - \psi' = \pi - i'.$$

Hence it is clear that the curve of (1.63) vs. i corresponding to a value of $(90^\circ \leq \psi \leq 180^\circ)$ is just symmetrical to the curve corresponding to the ψ' $(= \pi - \psi)$ with respect to the line $i = 90^\circ$. For example, the value of (1.63) at $i = 130^\circ$ on the curve of $\psi = 120^\circ$ and $f/f_H = 0.2$ is obtained as the value at $i = 50^\circ$ $(= 180^\circ - 130^\circ)$ on the curve of $\psi = 60^\circ$ $(= 180^\circ - 120^\circ)$ and $f/f_H = 0.2$.

Regarding the boundary condition represented by (1.62), the same charts can be used with a little alteration in abscissa and parameter, namely the substitution

$$i \longrightarrow \frac{\pi}{2} - i \quad \text{in the abscissa}$$

and

$$\psi \longrightarrow \frac{\pi}{2} - \psi$$

in the parameter will do. For example, the value of (1.62) at $i = 20^\circ$ on the curve of $\psi = 30^\circ$ is obtained as the value at $i = 70^\circ$ on the curve of $\psi = 60^\circ$.

The curve of (1.62) corresponding to ψ in the range

$$90^\circ \leq \psi \leq 180^\circ$$

is found to be anti-symmetrical to that corresponding to the ψ' $(= \pi - \psi)$ with respect to the point $i = 90^\circ$ on the i-axis.

1.4.3 Geomagnetic line of force and gyro-frequency

The earth's magnetic field appearing in our ray calculation has been regarded as given by a magnetic dipole at the center of the

earth. A magnetic line of force $r(\theta)$ which starts from the colatitude θ_0 on the ground. and the field intensity $|H|$ arising from a magnetic dipole are obtained in the polar coordinates(r, θ) by the following formulae (Alfven, 1958). (See Fig. 1.14).

$$\left.\begin{aligned} r &= R\frac{\sin^2\theta}{\sin^2\theta_0}\ , \\ \tan\psi &= \tfrac{1}{2}\tan\theta\ , \end{aligned}\right\} \text{ magnetic line of force,} \qquad (1.65)$$

$$\left.\begin{aligned} H &= A\phi/r^3\ , \\ \phi &= \sqrt{1 + 3\cos^2\theta}\ , \end{aligned}\right\} \text{ field intensity,} \qquad (1.66)$$

$$A = 8.1\times10^{25}\ \text{gauss cm}^3\text{: for geomagnetism,}$$

where ψ is the complementary angle of magnetic dip, while θ is the colatitude and R is the earth's radius. According to these formulae, we find that the direction of magnetic field does not depend upon the radius r and is given in a very simple form as a function of θ.

Concerning the field intensity, it changes very slowly even in the same radius r according to the latitude. In our treatment the variation of $|H|$ with latitude is disregarded, that is, ϕ is considered as unity (the value on the equatorial plane).

Using this field intensity, the gyro-frequency(f_H) was calculated and f_H vs. height is illustrated in Fig. 1.15.

1.4.4 Electron density distributions assumed for our ray tracing

As was mentioned previously, the ray path and the distribution of the electron density in the exosphere are mutually correlated. Then we can not help taking the following way: Assume first various distributions of the electron density of the exosphere, and based

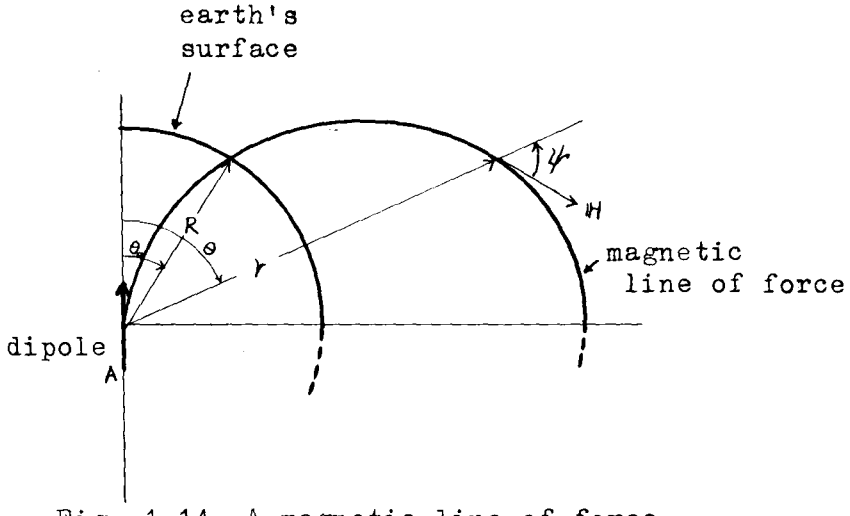


Fig. 1.14 A magnetic line of force from a dipole A in a magnetic meridian plane

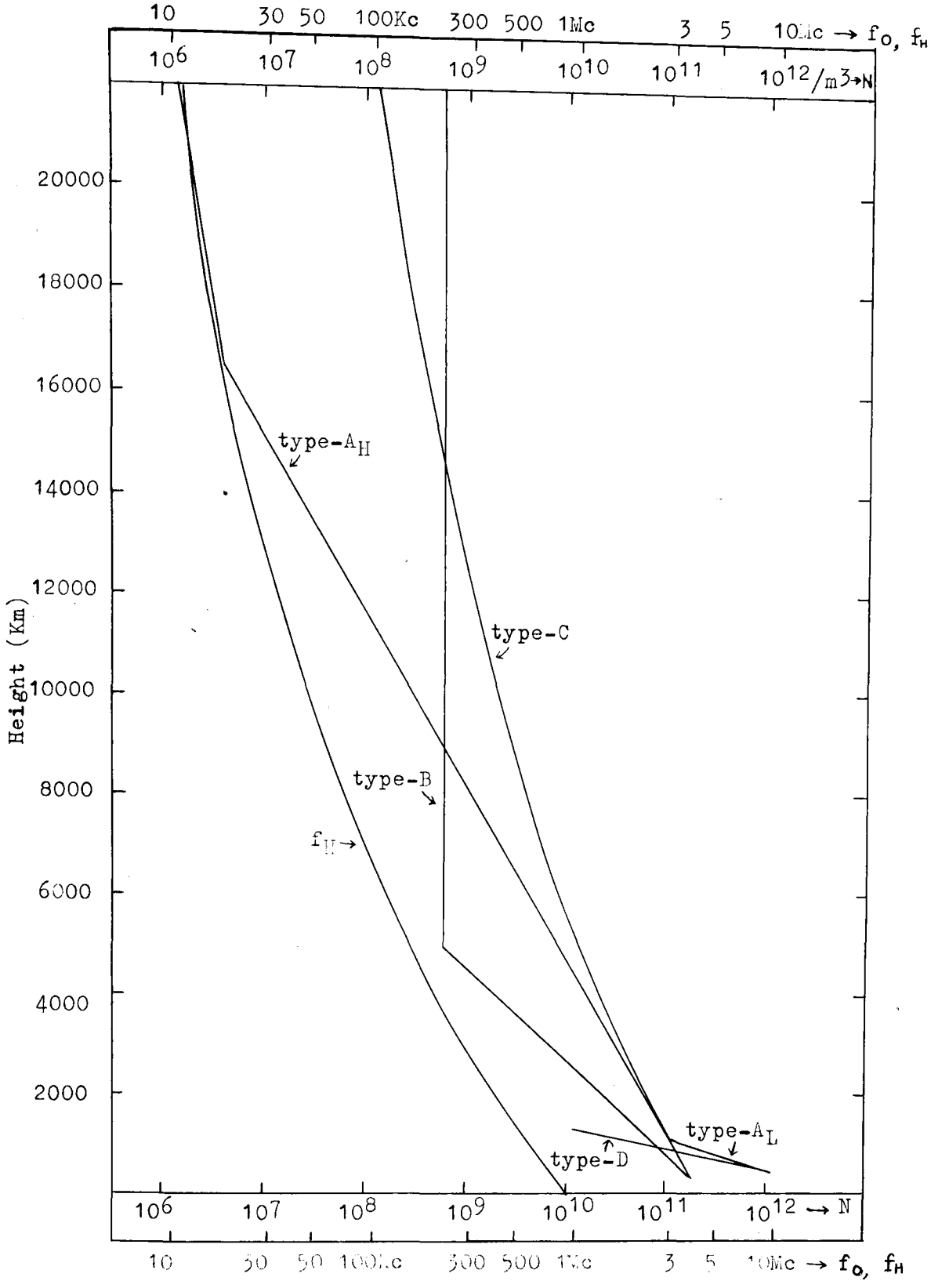


Fig. 1.15 Gyro-frequency and assumed electron density (per m^3) against height (km).

on the distribution calculate the ray path and dispersion D. Select the most appropriate one out of the assumed distributions, taking account of the properties of the observed whistlers.

It is however a difficult problem to assume a distribution of the electron density in the exosphere without any information thereof. From a measurement of the intensity and polarization of the zodiacal light, Siedentopf(1953) supposed the electron density to be approximately $6\times10^{8}m^{-3}$ at the distance of the earth's orbit from the sun. This value may be of the electron density a little distant from the earth, not the one in the proximity of the earth. It is because we can not imagine that the electron density decreases rapidly down to the above value in the outer space of the F layer.

After all the following distributions are tentatively assumed.

Type,	Height(km),	Number of electron(m^{-3})
A_H	$300 \leq h \leq 16500$,	$1.8\times10^{11}\exp\{-6.57\times10^{-4}(h-300)\}$
	$16500 \leq h$,	$4\times22870^{5}(h+6370)^{-5}$
A_L	$300 \leq h \leq 1000$,	$1.24\times10^{12}\exp\{-3.45\times10^{-3}(h-300)\}$
	$1000 \leq h$,	same as type A_H
B	$300 \leq h \leq 5000$,	$1.8\times10^{11}\exp\{-1.21\times10^{-3}(h-300)\}$
	$5000 \leq h$,	6×10^{8}(constant)
C	$300 \leq h$,	$1.8\times10^{11}\times 6670^{5}(h+6370)^{-5}$: nr = constant
D	$300 \leq h \leq 1000$,	$1.24\times10^{12}\exp\{-4.61\times10^{-3}(h-300)\}$

These distributions are also shown in Fig. 1.15.

In general the distribution will change not a little with latitudes. But it is not easy to take the latitude dependency into account, in calculating the ray path.

1.4.5 Some examples of the calculated ray path

Computation was made of ray paths in the space higher than 300Km (at the top of the ionosphere). Their initial ray direction were obtained from (1.60).

First, eq. (1.7) was used for n, that is the wave frequency is sufficiently small as compared with the plasma frequency f_o and gyro-frequency f_H. Ray paths calculated are illustrated in Fig. 1.16, using the distributions indicated in Table 1.2, namely the difference of the maximum electron density for high and low latitude is partly taken into consideration. The calculated dispersions and path lengths are listed in Table 1.2. Dotted lines on the graph are the geomagnetic lines of force.

Influence of the difference of the density distribution on the ray path is examined for the path starting from the latitude 40°, which is shown in Fig. 1.17. The ray paths under the distribution type C, are shown in Fig. 1.18.

Second, the case in which f/f_H is not neglected, is treated, then the refractive index is given by (1.3). Since the computation in this case is very trouble work, only for the whistlers starting from the latitude 40°, the paths are obtained for frequencies, 10, 20, 30Kc/s and under the distribution, type A_H.

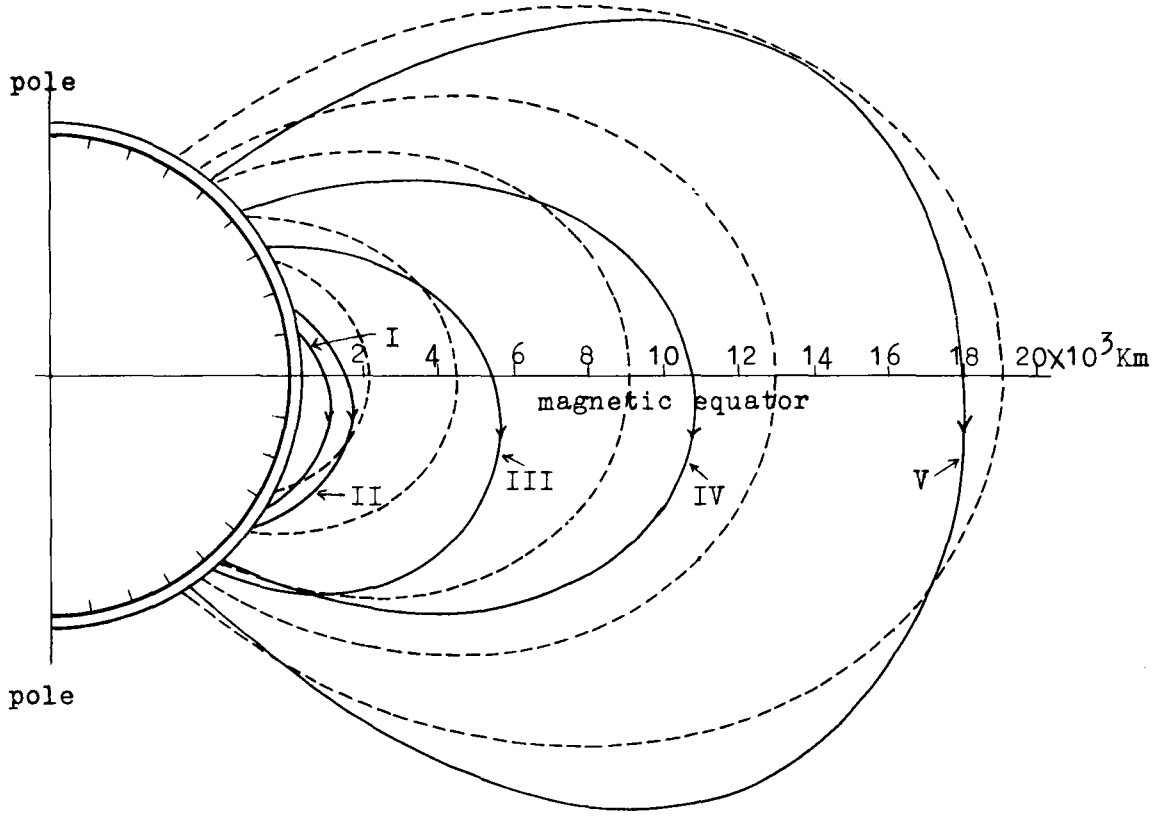


Fig. 1.16 Calculated ray paths (solid lines) for $f/f_H \ll 1$, under the electron density distribution A_L (for paths I, II, III) and A_H (for paths IV, V), and the geomagnetic lines of force (dotted lines).

Table 1.2 The results of calculations.

Curve No. in Fig. 1.16	Start θ	End θ	Dispersion $D(\sec^{\frac{1}{2}})$	Path length(Km)	Electron distribution
V	40°	180-33°	85	57,000	A_H
IV	50°	180-43°	90	31,400	A_H
III	60°	180-40°	85	19,200	A_L
II	75°	180-55°	62	7,700	A_L
I	80°	180-61°	53	5,500	A_L
	80°	180-76.5°	30	3,200	D

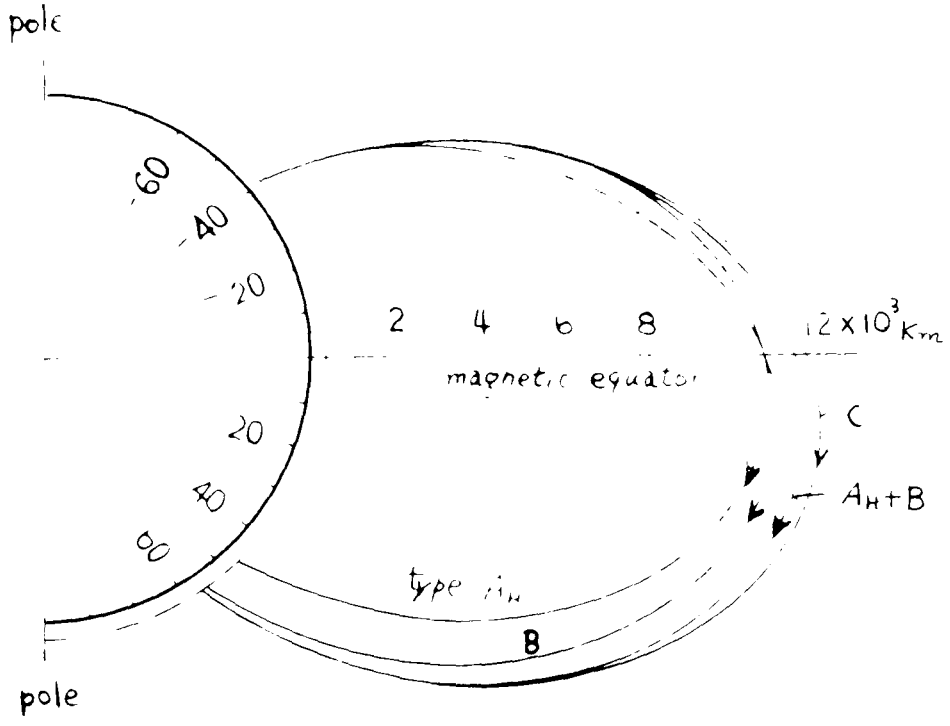


Fig. 1.17 Ray paths calculated under various distributions of the electron density for $f/f_H \ll 1$.

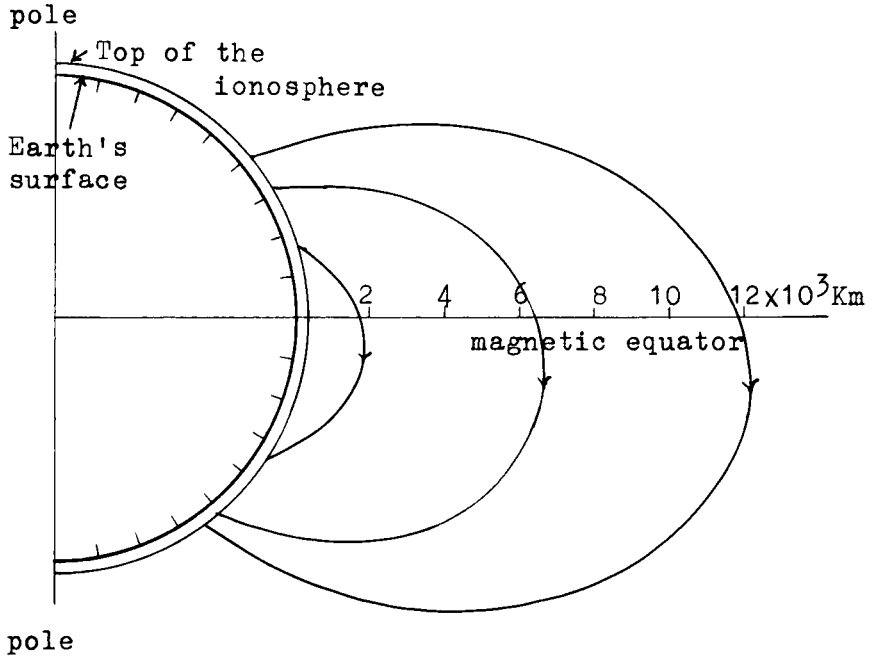


Fig. 1.18 Ray paths under the distribution (type-C) of electron density, f/f_H being assumed zero.

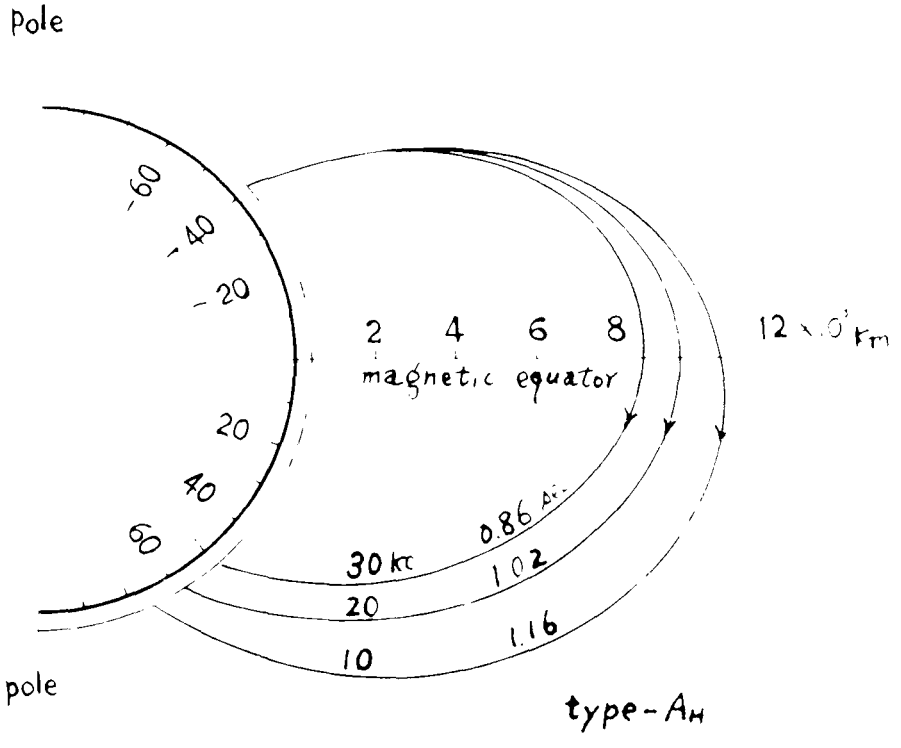


Fig. 1.19 Frequency dependence of the ray paths under the electron density distribution type-A_H

Chapter 1.5 Some considerations for the calculated results, as compared with the observed whistlers.

1.5.1 Asymmetric feature of the ray path

As Storey has shown theoretically, the ray is bunched around the direction of magnetic field. After the presentation of this theory, many theoretical investigations have been made with the assumption that the ray path coincides with a line of force of the geomagnetic field, for the sake of convenience.

As we see in Figs. 1.16 to 1.19, the calculated ray paths indicate asymmetric feature with respect to the magnetic equator, and do not follow the magnetic lines of force. These results teach us that the discrepancy between the actual path and the assumed one may be considerable.

However in Fig. 1.16, some ray path starting from higher latitudes seems to have more symmetrical feature than those from lower latitudes have. Under more strict consideration, ray path from higher latitudes in Fig. 1.16 may not be correct because especially for the path from higher latitudes the approximation $f/f_H \ll 1$ will not be satisfied. As illustrated in Fig. 1.19, the paths computed using more accurate expression of the refractive index show that the higher the frequency of the wave is , the more inner path the wave takes, so that the path from higher latitude becomes more symmetric.

One path was computed as a trial with the same condition as those shown in Fig. 1.16 besides a difference in the distribution of electron density, more sharply decreasing with height than the type-A_L. This path is listed at the bottom of the Table 1.2.

From this example, together with the results shown in Figs.1.17 and 1.18, it can be concluded that the rays take more inner paths by a more sharply decreasing distribution of electron density.

Very recently, a similar ray path calculation of the whistlers was made by Yabroff (1959), using an electronic computer. The results are almost identical with ours with a little discrepancy, which might be due to the error of our calculation. In order to compare the Yabroff's result with ours, Fig. 1.20 is given, where the final latitude of each path is plotted against the initial latitude of the paths under the distribution of the electron density type-A_E (dotted line by Yabroff and solid line by Kimura, the latter corresponding to Fig. 1.16). Therefore Yabroff's computation also indicates a considerable asymmetry especially for lower latitudes.

Helliwell and Gehrels (1958) carries out an experiment of the man-made whistler by using 15Kc/s wave intermittently transmitted at Annapolis. The signal received at Cape Horn (magnetically conjugate point of Annapolis) with average time delay of 0.7 second. This experiment (corresponding to a comparatively high latitude) sometimes showed that the ray path was not always symmetric and the shape of the ray path was greatly changeable.

1.5.2 Long whistler, short whistler and whistler train

Now that we have known for the ray path of the whistlers to be not always symmetric with respect to the magnetic equator, a question may arise about the mechanism of the long whistler and the whistler train, which have been considered as produced by twice or

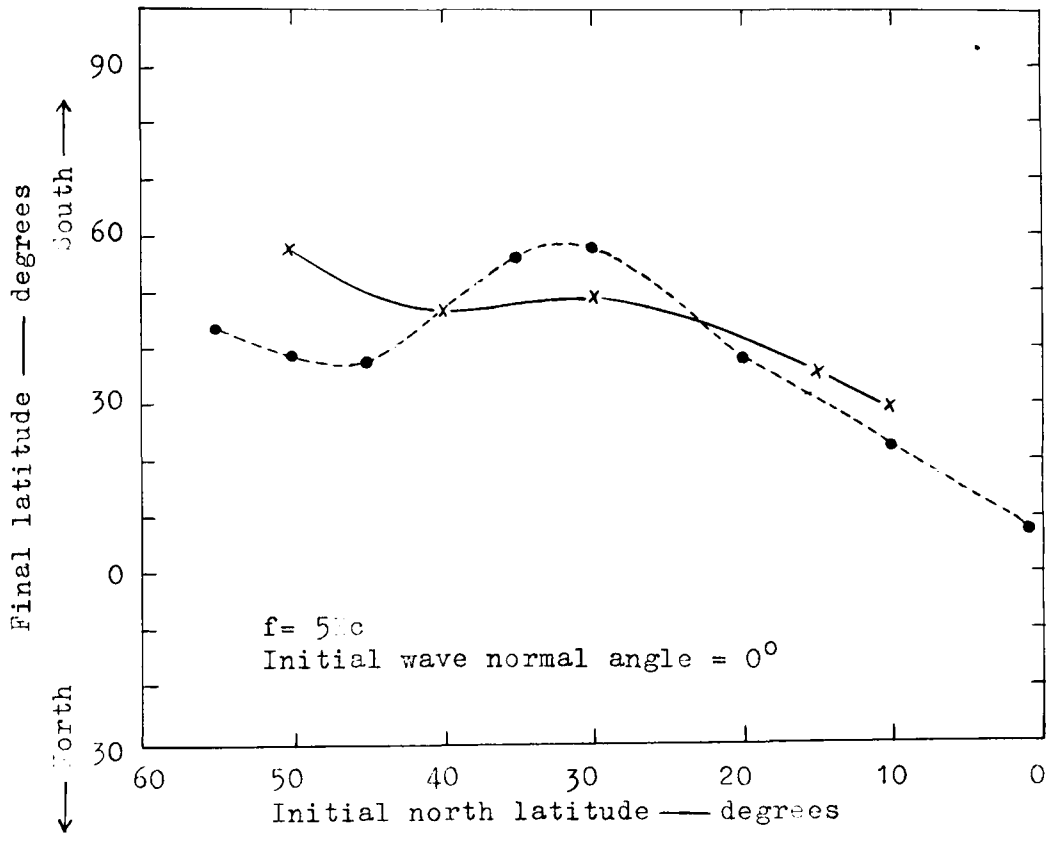


Fig. 1.20 Final latitude vs. initial latitude of the calculated ray path (solid line calculated by Kimura and dotted line by Yabroff).

more repetitions propagating along the same path.

Besides the case using very rapidly decreasing electron density with height, the asymmetry of the path will not be avoided, so that the apex shifts toward the other hemisphere, as shown in Fig. 1.16 etc. And this asymmetric feature is more intensified for the paths from lower latitudes and severer at lower frequency even for those from higher latitudes.

Hence it should be expected that in higher latitudes the long whistler may be detectable but in lower latitudes the long whistler can not be heard. This image appears to be almost satisfactory as we see that the long whistlers are heard at Wakkanai (geomagnetic latitude 35.5°), whereas only the short whistlers are observed at Toyokawa (24.5°).

Further, we can say that even in higher latitudes the distribution of electron density must be a certain special one for the long whistlers to be detectable. Moreover, for explaining the whistler trains, severer limitation on the distribution may be required.

As mentioned before, there is a kind of whistler train, the maximum loudness of which appears at the third or fourth repetition of the whistler. This phenomenon may be interpreted as follows. Suppose the case in which the ray path is nearly symmetric but with a little bit of asymmetric feature, and let a source be in the south from the receiving point of the northern hemisphere. Then the first retruned whistler (long whistler) will strike the earth a little bit north from the source due to a small amount of the asymmetry of the path. After a several repetitions the returned whistler will come to just the receiving point, and will be heard as the loud-

est whistler. Then they will pass over the receiving point to the north and again their loudness decreases. In this case the attenuation within the exosphere is assumed to be negligible.

Recently there was proposed a new thought for the ray path, to solve a question about the ray path asymmetry (Smith et al, 1959). Namely they have shown that the whistlers can be guided almost exactly along a magnetic line of force by a field aligned duct which may be produced by a local enhancement of the electron density along the magnetic line of force.

Chapter 1.6 Attenuation for VLF electromagnetic waves in the ionosphere

1.6.1 Preliminary remarks

In obtaining the ray path of the whistler, we have always neglected the mean frequency of collisions ν in the refractive index. But when we consider the penetration of the waves through the ionosphere, especially the lower part of the ionosphere, the collision of electron with neutral particles cannot be ignored. It will then contribute to the absorption of the wave.

At the bottom of the ionosphere, electron density, and therefore the critical frequency and the refractive index, increase very rapidly with height. Then there will be a partial reflection.

Really speaking, the diurnal variation of occurrence frequency of the whistlers (see Fig. 1.3) suggests that the ionosphere is in charge of the attenuation of the VLF electromagnetic waves in the daytime. Hence in this chapter a consideration on the attenuation within the ionosphere is presented.

Now we remind the refractive index n given by the formula (1.1). Under the quasi-longitudinal approximation, the refractive index for the extraordinary mode is simply reduced to

$$n^2 = 1 - \frac{X}{1 - Y_L - jZ}, \tag{1.67}$$

$$= 1 - \frac{f_0^2}{f - f_H \cos\theta - j\frac{\nu}{2\pi}} \frac{1}{f}. \tag{1.68}$$

Let n be given by the expression

$$n = \pm(\alpha - j\beta). \tag{1.69}$$

(Remark that in this expression, α is a phase constant and β is an absorption coefficient, and they are in quite another notation than those used earlier as angles). Then from (1.68) and (1.69), we get

$$\alpha = \sqrt{\frac{1}{2}\left\{1 - \xi + \sqrt{(1-\xi)^2 + \eta^2}\right\}} = \frac{\eta}{2\beta}, \tag{1.70}$$

$$\beta = \sqrt{\frac{1}{2}\left\{\xi - 1 + \sqrt{(\xi-1)^2 + \eta^2}\right\}} = \frac{\eta}{2\alpha}, \tag{1.71}$$

where

$$\left.\begin{aligned} \xi &= \frac{f_0^2}{f} \frac{f - f_H\cos\Theta}{(f - f_H\cos\Theta)^2 + \left(\frac{\nu}{2\pi}\right)^2}, \\ \eta &= \frac{f_0^2}{f} \frac{\nu}{(f - f_H\cos\Theta)^2 + \left(\frac{\nu}{2\pi}\right)^2}. \end{aligned}\right\} \tag{1.72}$$

Using the above defined notations, the amount of the attenuation by absorption suffered during the travel of a wave is given by the factor

$$\left.\begin{aligned} &e^{-\Gamma} \\ \text{where} \quad &\Gamma = \frac{2\pi f}{c}\int \beta\, ds, \end{aligned}\right\} \tag{1.73}$$

and ds is a path element (for the vertical incidence, ds is reduced to dh). The integration is performed along the path within the range under consideration.

Partial reflection coefficient (R) for vertical incidence is calculated by a complex quantity

$$R = \int \frac{1}{2\alpha}\frac{d\alpha}{dh} e^{-j\int \frac{4\pi}{\lambda}\alpha\, dh}\, dh, \tag{1.74}$$

where λ is the wave length (Namba and Maeda, 1939; Bremmer, 1951). The factor of attenuation caused by the partial reflection is defined by

$$1 - (\text{real part of } R). \tag{1.75}$$

1.6.2 Electron density and mean frequency of collision in the ionosphere

In the calculation of the attenuation, the following distribution of electron density in the ionosphere is used, which was proposed by Matumoto (1960). It is applicable to the season around the equinox of middle latitudes. That is:

1. For the F region: N(h) data calculated from h'-f curve at Cavendish laboratory is adopted.

2. For the E region: A modified Chapman layer is assumed, where the scale height is taken to be a linear function with respect to the altitude, and around the sunrise and sunset the curvature of the earth is taken into account.

3. For the D region: At midday, a model proposed by Nicolet (1960) is adopted.

Then the total distribution of the electron density is as shown in Fig. 1.21.

Mean frequency of collision of the electron with neutral particles at the altitude higher than about 120Km will vary diurnally, while in the E region or below, the frequency of collision is regarded as invariant of the time of day. In Fig. 1.22 the quantity is plotted.

1.6.3 Calculation of the attenuation in the ionosphere

In calculating the attenuation, the following assumptions were made.

1. The wave is assumed as incident vertically to the ionosphere. And the angle θ in eq.(1.72) is 30° irrespective of the altitude.

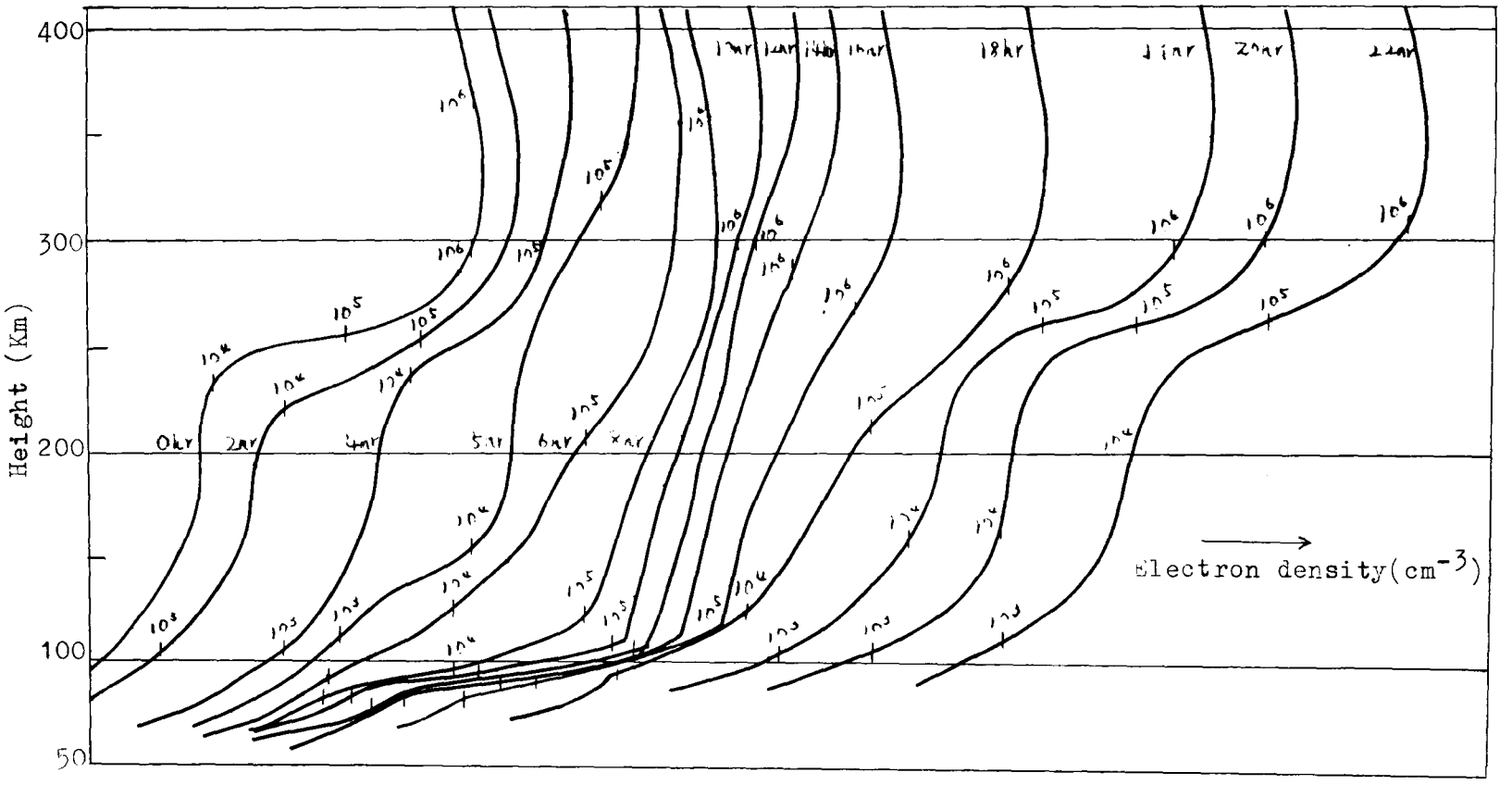


Fig. 1.21 Daily variation of N(h) curve of the ionosphere (for equinox and middle latitudes)(after Matsumoto)

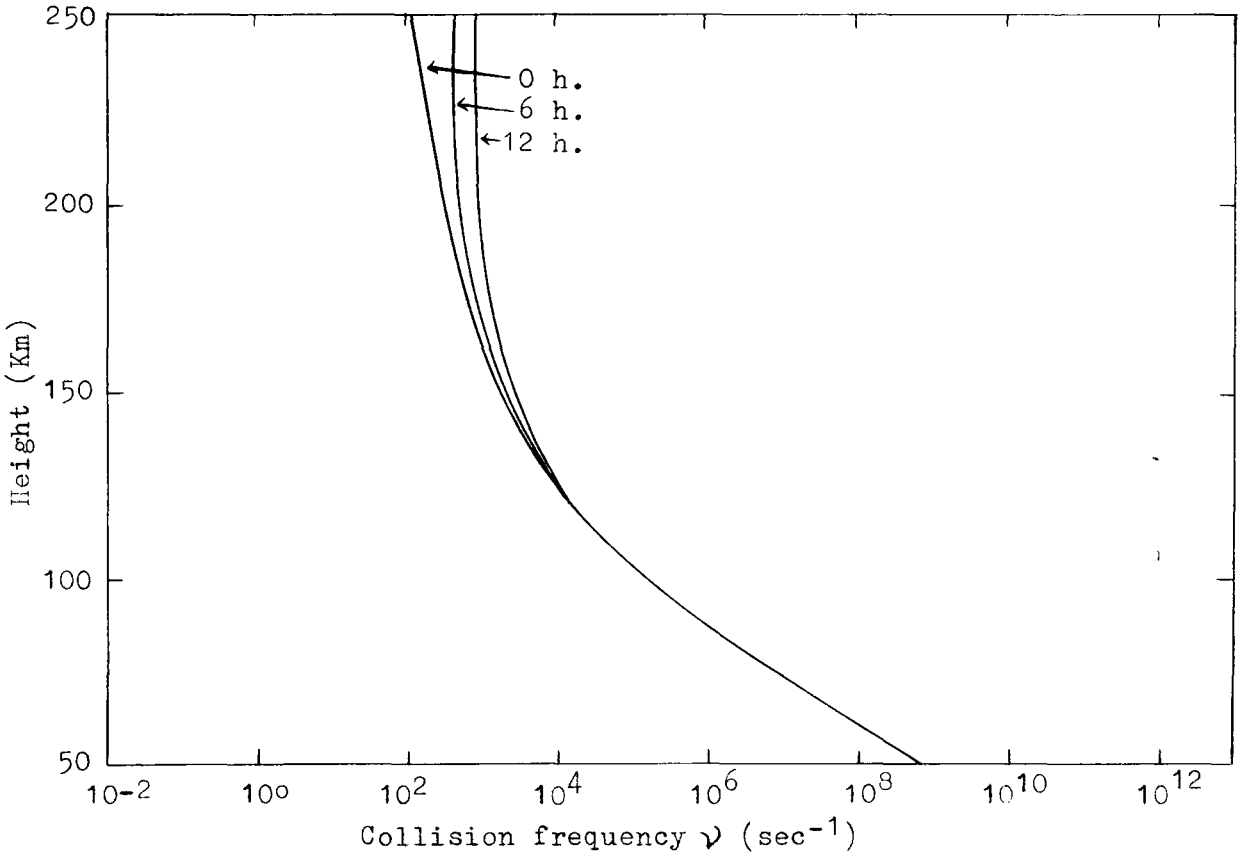


Fig. 1.22 Variation of the collision frequency ν with height.

2. Gyro-frequency is 1Mc/s.

Using the quantities previously supposed, (1.71) and then (1.73) were calculated. Diurnal variation of the attenuation (in db) given by (1.73) is shown in Fig. 1.23, for frequency 3Kc/s. Dependence of the attenuation upon frequency is as shown in Fig. 1.24. These graphs indicate that the attenuation is greater for higher frequency, and in the daytime. Fig. 1.23 will explain the diurnal variation of the intensity and the occurrence frequency of the whistlers and other VLF phenomena.

The partial reflection coefficient R given by (1.74) is about 0.6 at noon. Then the factor of attenuation caused by the partial reflection will be about 8 db.

In the exosphere, the frequency of collision is so low that the absorption due to the collosion will be almost negligible.

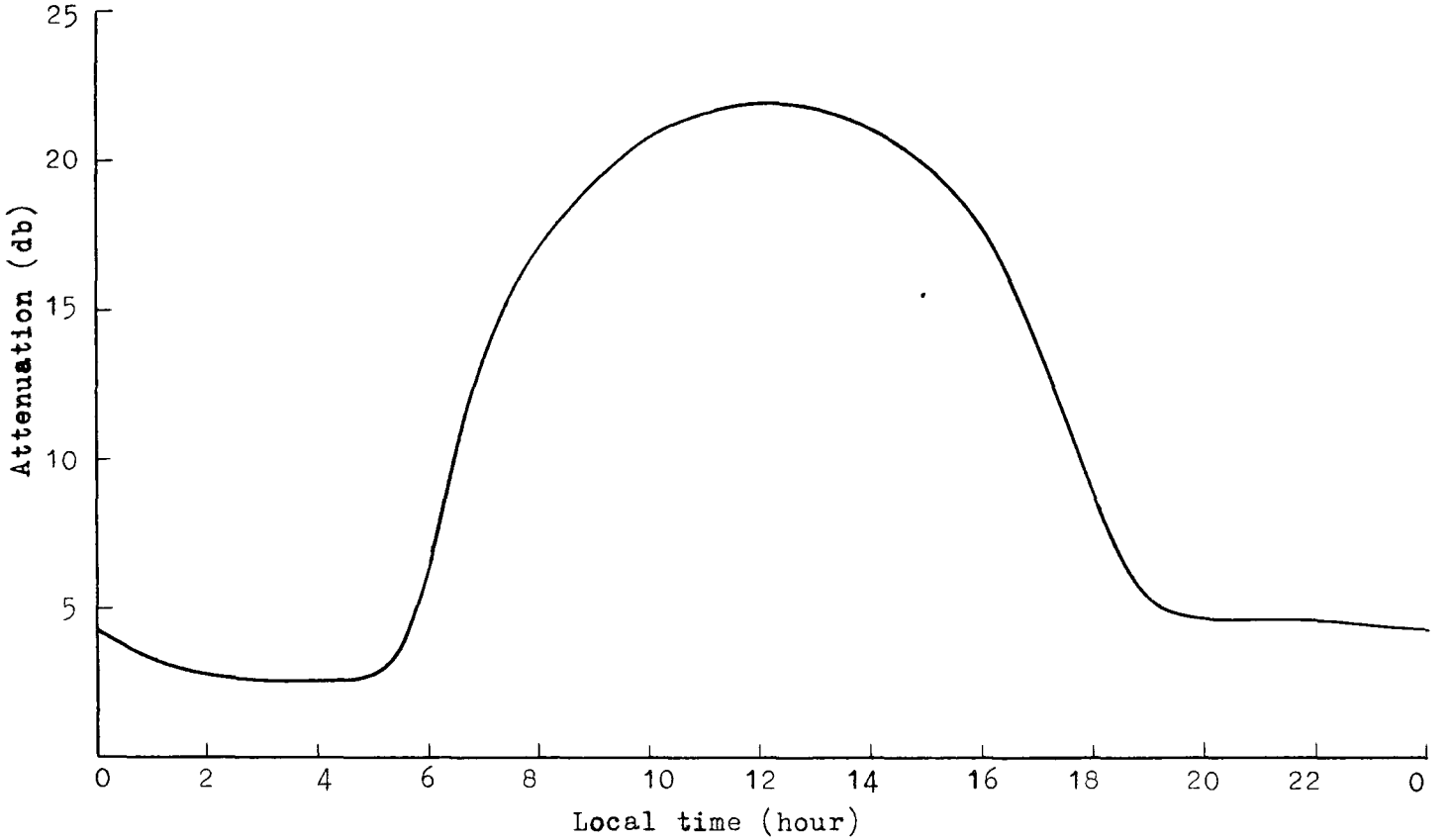
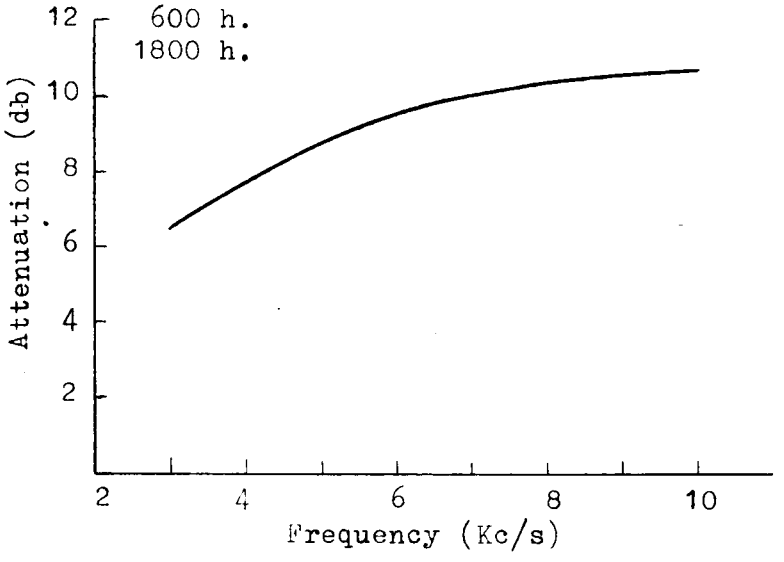


Fig. 1.23 Diurnal variation of the attenuation in the ionosphere for f=3Kc/s.



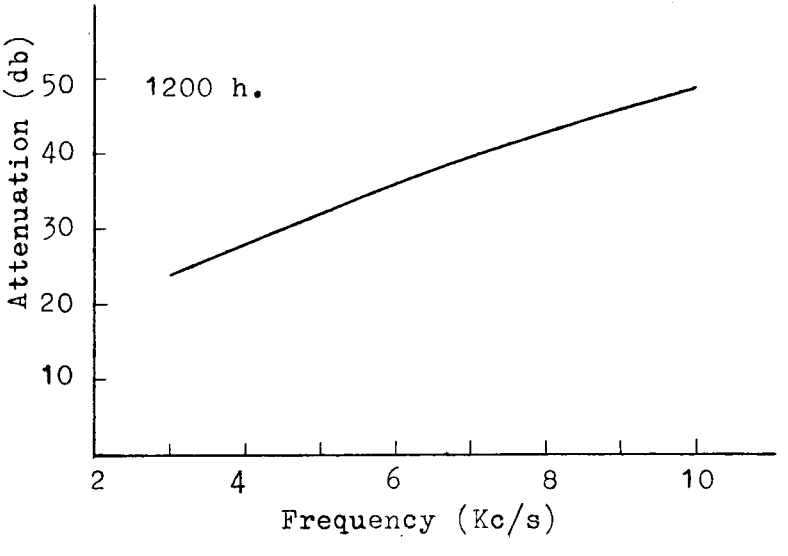


Fig. 1.24 Frequency dependence of the attenuation in the ionosphere.

Chapter 1.7 Concluding remarks

The problem of ray tracing of the VLF electromagnetic waves in the magneto-active plasma was treated. The author could prove that when the parameters of the medium change in one direction, the so-called Snell's law governs the behavior of the wave normal of the wave in the medium, which is even anisotropic, as well as in an isotropic medium. Then a simple numerically calculating method was devised for the ray tracing of the whistlers in the exosphere by the Snell's law.

The actual calculation of the ray path taught us many interesting informations, such as the asymmetry of the path, the dependency of the path on the distribution of the electron density etc., in spite that the calculation is an approximate method and naturally not so precise as that obtained by an electronic computer.

According to our calculations, it was concluded that even for the path starting at high latitude on the ground, the ray paths were not always symmetric with respect to the magnetic equator and actually some severe condition would be needed to explain the many times repeating echoes or train. Then as Smith has proposed, the field aligned ionization would help the propagation of the whistlers in some cases.

Attenuation of the VLF waves in the ionosphere was estimated. It amounted to 25 ~ 50db at noon, and 7 ~12db around 6h and 18h for frequency 3 ~ 10Kc/s. The attenuation due to the partial reflection is about 8db at noon.

As has not yet been mentioned, the direction of the wave normal

along the ray path is important, because according to the calculation the wave normal at the opposite end of the path is not always directed to the vertical of the ionosphere. It is associated with a very complicated problem of the penetration of the wave energy through the ionosphere. This problem must be treated wave theoretically in a spatially rapidly varying medium. Although the author did not deal with it at all, it may cause an additional attenuation within the ionosphere.

PART II

EFFECT OF A CHARGED PARTICLE BEAM ON THE PROPAGATION OF THE VLF ELECTROMAGNETIC WAVES IN THE IONOSPHERE AND EXOSPHERE

(Originating mechanism of the VLF emissions)

Chapter 2.1 Introduction

As described in the previous part of this paper, we have known the existence of the whistler mode of propagation of VLF electromagnetic waves in the ionosphere and exosphere owing to the atmospherics, called whistler. Hence, whatever VLF electromagnetic wave, VLF noise or emissions generated in the exosphere are propagated in the whistler mode. Of course, in this case, the wave is dispersed like whistlers due to the dispersive property of the medium.

Recently various VLF phenomena spontaneously originated in the exosphere, have drawn our attention as well as the whistlers. These phenomena are generally named "VLF emissions"(Gallet 1959) and are usually called dawn chorus, hiss etc. depending upon the frequency-time characteristics. The origin of the VLF emissions has been thought as existing in the exosphere rather than in the space below the ionosphere. The evidence is due to the fact that the VLF emissions correlate much with the magnetic activity of the sun and do not with terrestrial phenomena such as the lightnings.

We must, then, consider some mechanism of radiation taking place in the exosphere, that is , in a magneto-active plasma. Until now several theories have been proposed for the interpretation of the VLF emissions. These theories can be classified into two categories: The one is to consider directly mechanisms of electromagnetic wave emis-

sion, such as due to cyclotron radiation, Cherenkov radiation. The other is to consider a certain mechanism of VLF wave amplification, namely a weak noise, such as thermal noise, is amplified in the exosphere, so that it becomes powerful enough to be heard on the ground. The both classes of mechanism are, however, common on the point that they depend greatly on the charged particles injected into the exosphere with high speed.

The electromagnetic wave emissions caused by the previously described mechanisms, cyclotron radiation etc., are not so sufficiently intense as to explain the VLF emissions. Therefore the mechanism of amplification will be effective, if it is actually possible in the exosphere.

Gallet and Helliwell(1959) have suggested a possibility of TWT (traveling wave tube) like amplification by means of the incoming electron beam in the exosphere. Bell and Helliwell reported a theoretical calculation, the result of which supported the TWT theory.

In spite of the fact, the author has found that Bell's calculation was incorrect and TWT theory could not always be applicable to the VLF electromagnetic waves in the magneto-active plasma. In place of TWT mechanism, a new mechanism of amplification by means of the interaction of the wave with a proton cyclotron mode of the beam is found.

In what follows, the effect of a charged particle beam on the propagation of VLF electromagnetic waves in a magneto- active plasma is treated theoretically and then it is applied to reexamining the TWT mechanism of amplification in the exosphere and to verifying the possibility of a new mechanism of amplification by a proton beam.

Chapter 2.2 On the VLF emissions

2.2.1 Outstanding features of the VLF emissions

The VLF emissions can be heard with a receiver usually utized for the whistler observation. The frequency characteristics is also analysed by a sonagraph as is done on the analysis of the whistlers.

According to the frequency-time characteristics of the VLF emissions, they are distinctly classified into two groups (Gallet 1959). The first is "discrete type" of emissions and the second is "continuous type". The former includes "dawn chorus", "hooks", "risers", "falling tones", "quasi-horizontals" etc. (see Fig. 2.1). Such a phenomenon does appear for a short time duration of the order of 0.5 to 1 second, and the frequency of each phenomenon varies continuously in time with various ways within the frequency region from several kc/s up to several tens kc/s. At a moment each discrete emission has the frequency spectrum which is confined in a narrow frequency-time characteristics such as that taken by sonagraph is a continuous thin line.

The latter, continuous type, is called "hiss". This emission, on the other hand, is heard for more than one hour, and the occupied frequency band of the spectrum at a moment is as wide as several kc/s, and almost invariant in time (see Fig. 2.2).

The intensity of the strongest emissions detected at the ground is 10^{-13} to 10^{-14} $W/m^2(c/s)$ in power flux density (Allcock, 1957; Gallet and Helliwell, 1959).

There is another remarkable property of the emissions. The statistical study of the diurnal variation of the dawn chorus activity made by Allcock (1957) shows that the local time of maximum dawn chorus activity is linearly dependent upon geomagnetic latitude as

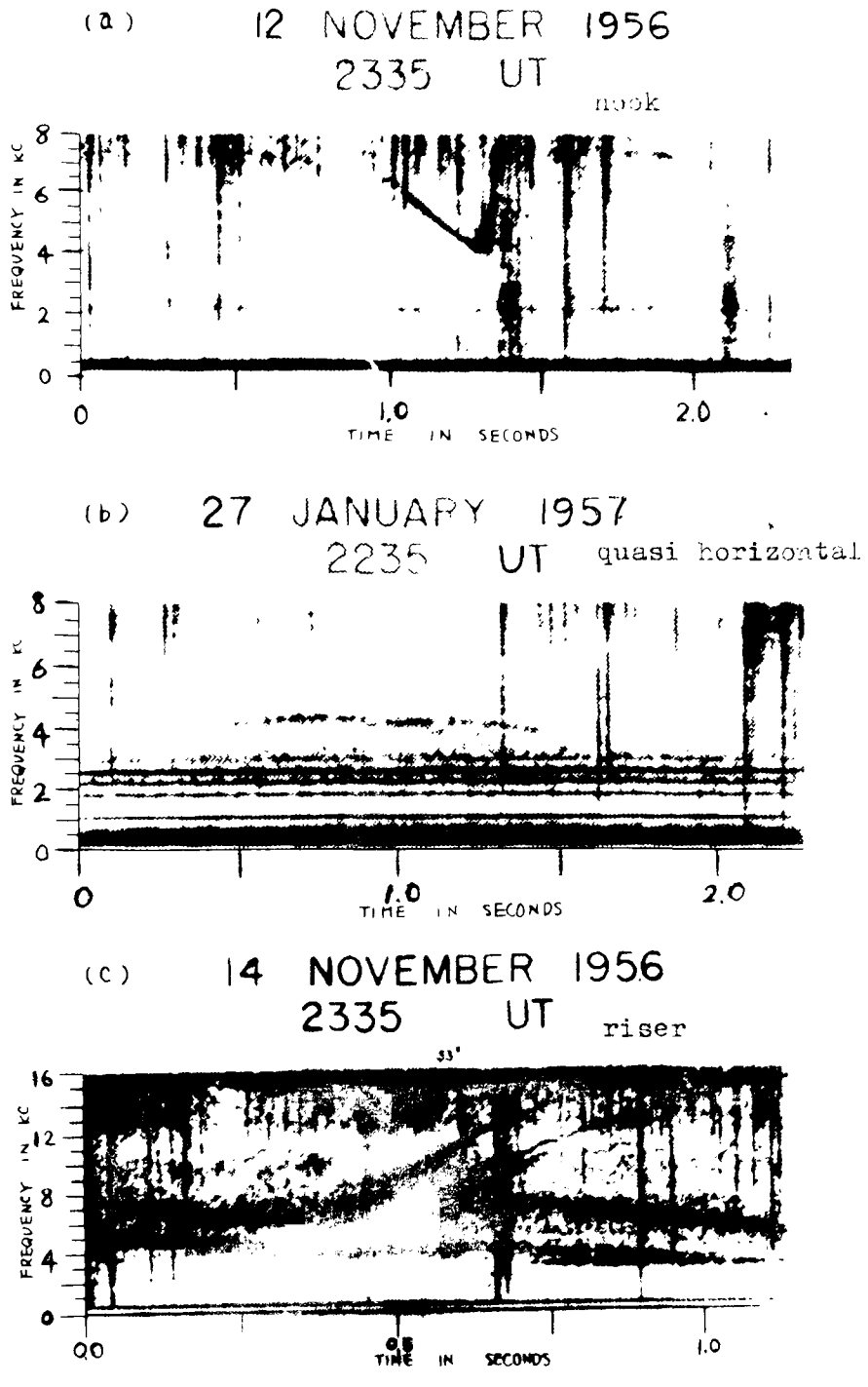


Fig. 4.1 VLF emissions--"discrete type"

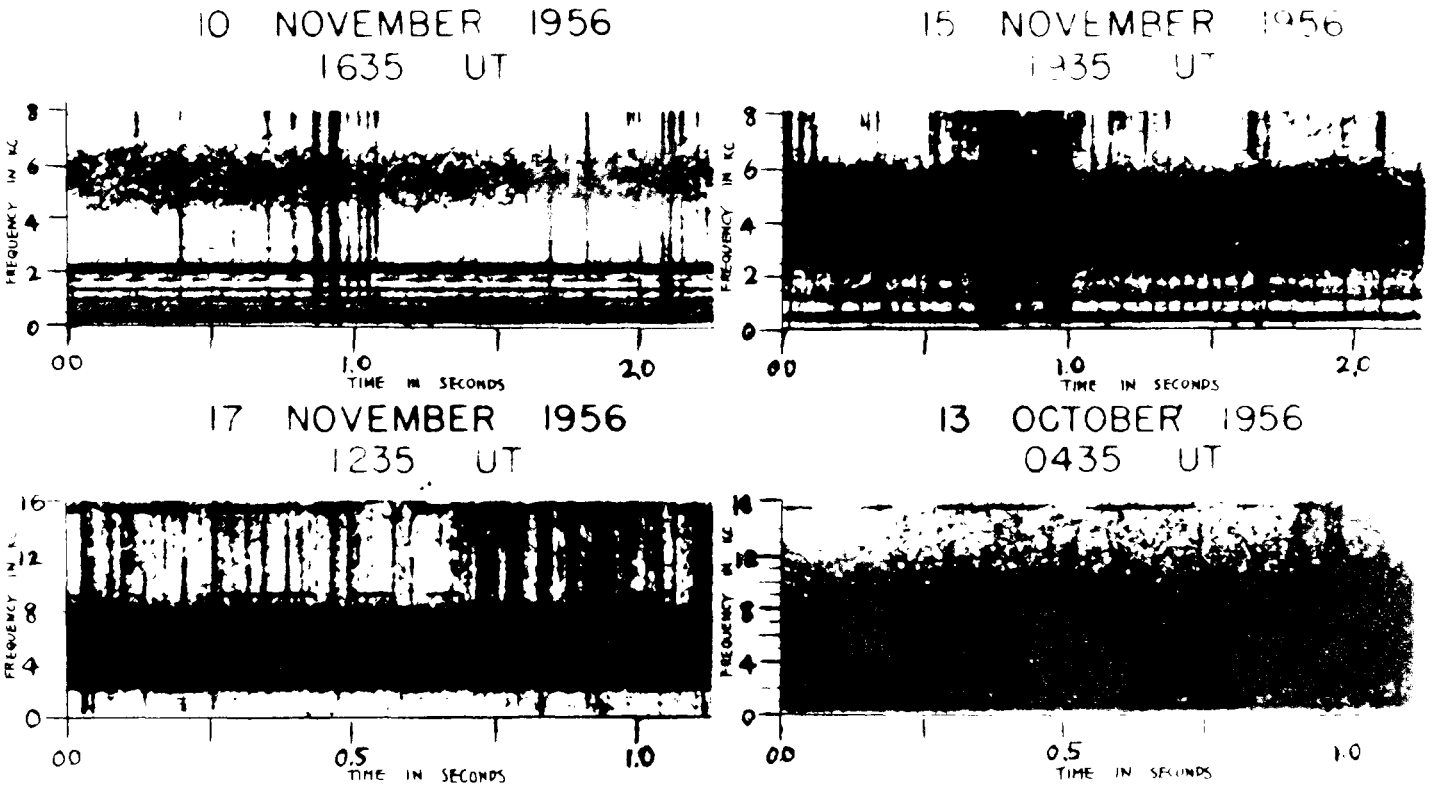


Fig. 2.2 VLF emissions--"continuous type"

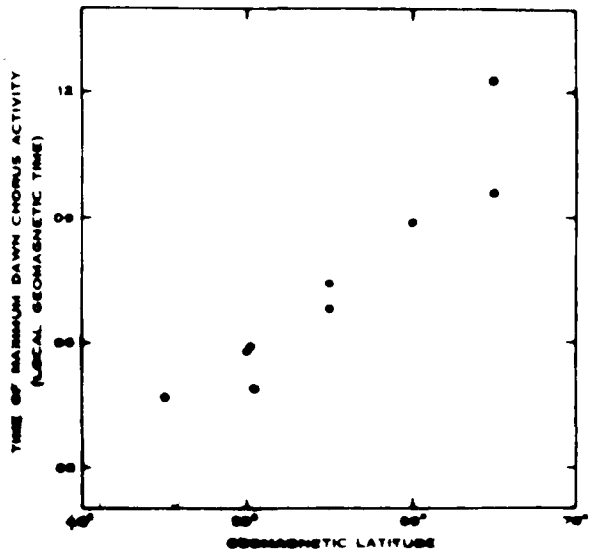


Fig. 2.3(a) Local time of maximum dawn chorus activity vs. geomagnetic latitude (after Allcock).

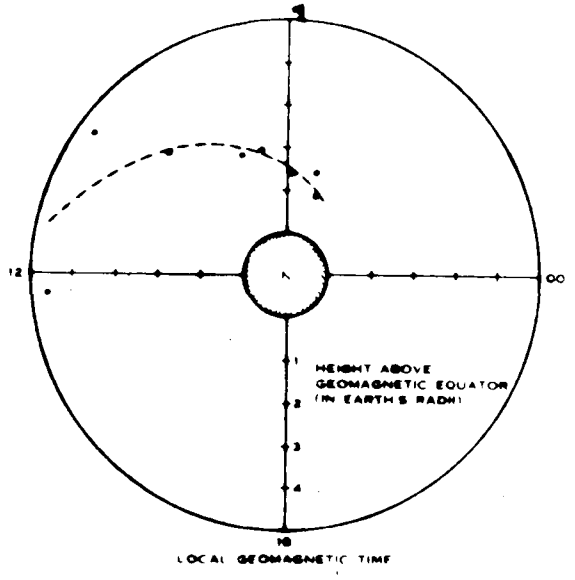


Fig. 2.3(b) Polar graph of local geomagnetic time against the height above the equator of the magnetic line of force passing through the observing station (after Allcock).

shown in Fig. 2.3(a), and when the data are replotted as a polar graph of local geomagnetic time against the height above the equator of the magnetic line of force passing through the observing station, the curve fitting the data is characteristics of the locus of a positively charged incoming particles being deflected by the earth's magnetic field (see Fig. 2.3(b)). Yoshida has obtained a similar result for the dawn chorus observed during geomagnetic storms (Yoshida, 1960).

Existence of a close association between aurorae and certain hiss has been reported (Martin, Helliwell, Marks, 1960). This evidence seems to imply also that the origins of VLF emissions are much related with incoming positive particles, especially protons.

2.2.2 Hypotheses of the originating mechanism of the VLF emissions

Originating mechanisms of the VLF emissions so far proposed are summarized as follows.

1. TWT (traveling wave tube) like mechanism of amplification (Gallet and Helliwell, 1959; Gallet, 1959; Bell and Helliwell, 1959): This hypothesis is based on the postulation that the longitudinal interaction of a whistler mode wave with a traveling plasma oscillation (or a space charge wave) of the incoming electron beam would bring about an amplification of the whistler mode wave provided that the average velocity U_o of the incoming beam approximates the phase velocity v_{ph} of the electromagnetic wave, that is $U_o \simeq v_{ph}$. According to Bell and Helliwell (1959) the gain of the amplification is expected to be about 2 db per wave length. The input signal is assumed to be provided by thermal radiation, whistler energy, or possibly Cherenkov radiation.

2. Cherenkov radiation (Ellis, 1957): Cherenkov radiation of a

frequency takes place if the velocity of a charged particle in a medium exceeds the phase velocity of the electromagnetic waves of the frequency in the medium. The required condition for the Cherenkov radiation is quite similar to that imposed on TWT mechanism (Ondoh, 1961).

3. Doppler-shifted cyclotron radiation of protons (Mac Arthur, 1959; Murcray and Pope, 1960): Gyro-frequency of protons in the exosphere is always very much smaller than that of the VLF emissions ordinarily heard. If the proton approaches with the velocity U_o to an observer, the observed frequency of the cyclotron radiation will be shifted up to several kc/s due to the Doppler dffect and as the result the radiation to be heard as a VLF emission. Then the condition of the proton gyro-frequency to be shifted up to a few kc/s is approximately given by $U_o \simeq v_{ph}$, being identical to the condition required for TWT mechanism.

Thus the frequency-time characteristics of the VLF emissions can be explained by means of any one of the previous three mechanisms. However the intensity of the Cherenkov radiation and proton cyclotron radiation possible in the exosphere are both too small to account for the observed strongest dawn chorus. Consequently the mechanism of amplification must work simultaneously with some of the electromagnetic wave radiation mechanisms or with some powerful seed.

Chpter 2.3 The effect of a charged particle beam on the propagation of electromagnetic waves in the magneto-active plasma

2.3.1 Preliminary consideration

Consider a region where a charged particle beam runs through an ambient plasma. The ambient plasma is assumed to be composed of electrons and protons and to be permeated by a static magnetic field B_o. The charged particle beam is also a plasma consisting of electrons and protons, and will flow in the average along lines of the static magnetic field, unless any constant electric field exists to provide the beam with a drift velocity perpendicular to the magnetic field. It is also assumed that the beam and the ambient plasma are uniform over a sufficiently large region of interest, far distant from boundaries.

In order to determine the propagation modes of VLF electromagnetic waves in such a coexisting region of the ambient plasma and the beam, the following equations are necessary and sufficient (Bailey, 1948; 1950).

i) Maxwell's field equations.

ii) The equations of conservation of electrons and positive ions constructing the beams.

iii) Maxwell's laws of the transfer of momentum in mixtures of different kinds of particles.

In applying the above equations to our problem, further assumptions will be made: Thermal motions of the individual particles in the ambient plasma and in the beam are neglected so that the particles are in zero temperature plasmas. Collisions between charged particles

and neutral particles are also neglected hereafter because in the practical medium that is in the exosphere the density of the neutral particles is very small. Physical quantities, such as magnetic field intensity, particle densities, particle velocities etc. are assumed generally to be made up of static (or constant) components and varying components (or perturbation), the latter being much smaller than the former. The static or constant quantities will be indicated if necessary by means of the suffix o, to discriminate them from the perturbations. Fundamental equations described above are generally non linear. To linearize them, products of perturbation quantities will be neglected.

The resulting equations for the perturbations yield plane wave solution of the form

$$Ae^{j(\omega t - kz)}$$

where ω is the angular frequency of the wave, k is the propagation constant and z axis in Cartesian coordinate system is taken as the direction of wave propagation. Then without loss of generality, the constant magnetic field B_o and constant beam velocity U_e, U_p of electrons and protons respectively are assumed to be directed so as to make an angle θ to z direction in the x-z plane.

After all our problem will be reduced to seeing if there exists any spatially growing mode of VLF electromagnetic wave under the influence of the charged particle beam.

2.3.2 Notations

To give easy understanding of the following equations, all symbols to be utilized will be listed up below in alphabetical order, where MKS units are used:

B_o, B; static and varying magnetic flux density vector ie. μH_o, μH

c; light velocity in free space.

$\nabla = \dot{\imath}(\partial/\partial x) + \dot{\jmath}(\partial/\partial y) + \dot{k}(\partial/\partial z)$

E; varying electric field vector.

e; magnitude of charge of an electron or a proton.

ε_o; dielectric constant of free space.

$f = \omega/2\pi$; wave frequency.

$f_{He} = \omega_{He}/2\pi$; electron gyro-frequency $= eB_o/2\pi m$.

$f_{Hp} = \omega_{Hp}/2\pi$; proton gyro-frequency $= eB_o/2\pi M$.

f_a; electron plasma frequency of ambient plasma $= \sqrt{\frac{e^2 N_a}{4\pi^2 m\varepsilon_o}}$

f_e; electron plasma frequency of the beam $= \sqrt{\frac{e^2 N_e}{4\pi^2 m\varepsilon_o}}$

f_p; proton plasma frequency of the beam $= \sqrt{\frac{e^2 N_p}{4\pi^2 M\varepsilon_o}}$

$\gamma = k_i/k_o$.

H_o, H; static and varying magnetic field vector.

J_a, J_e, J_p; convection current due to ambient electrons, moving electrons and moving protons respectively.

k, k_o; propagation constant in the medium and in free space.

k_r, k_i; real and imaginary part of the propagation constant k.

m, M; mass of an electron and a proton.

μ_o; permeability of free space.

n; refractive index of the medium $= k/k_o$.

N_a; average electron number density of the ambient plasma.

N_e, N_p; average electron and proton number density of the beam.

ρ_{oa}, ρ_a; average and varying electronic charge density of the ambient plasma.

ρ_{op}, ρ_p, ρ_{oe}, ρ_e; analogous meaning to the above quantities.

$\rho_{oa} = -eN_a$, $\rho_{oe} = -eN_e$, $\rho_{op} = eN_p$.

t; time

$\mathbb{U}_e$, $\mathbb{U}_p$; constant velocity vector of an electron beam and a proton beam.

u_a, u_e, u_p; perturbation terms of velocity of electrons in the ambient plasma, of those of electrons and protons in the beam.

$\mathbb{V}_e = \mathbb{U}_e/c$, $\mathbb{V}_p = \mathbb{U}_p/c$.

$X_a = f_a^2/f^2$, $X_e = f_e^2/f^2$, and $X_p = f_p^2/f^2$.

$\xi_e = nV_{ex}$, $\xi_p = nV_{px}$. $y_e = f_{He}/f$, $y_p = f_{Hp}/f$.

$\zeta_e = nV_{ex}$, $\zeta_p = nV_{px}$. $\zeta = n^2 - 1$.

x, y, z; Cartesian coordinates.

x, y, z as suffix imply respectively x, y, z components of the quantity.

θ; angle between z axis(wave normal) and vector $\mathbb{H}_0$ which is in the same direction as both $\mathbb{U}_e$, $\mathbb{U}_p$.

2.3.3 Fundamental equations

The fundamental equations to determine the modes of electromagnetic wave propagation in the magneto-active plasma under the influence of a charged particle beam, are Maxwell's field equations, conservation equations of charges and equations of transfer of momentum of the constituents

$$\nabla \times \mathbb{H} = \varepsilon_0 \frac{\partial \mathbb{E}}{\partial t} + \mathbb{J}_a + \mathbb{J}_p + \mathbb{J}_e , \tag{2.1}$$

$$\nabla \times \mathbb{E} = - \frac{\partial \mathbb{B}}{\partial t} , \tag{2.2}$$

$$\left.\begin{aligned} \nabla \mathbb{J}_e + \frac{\partial \rho_e}{\partial t} &= 0 , \\ \nabla \mathbb{J}_p + \frac{\partial \rho_p}{\partial t} &= 0 , \end{aligned}\right\} \tag{2.3}$$

$$\left.\begin{aligned} \mathbb{J}_a &= \mathbb{u}_a(\rho_{oa} + \rho_a), \\ \mathbb{J}_e &= (\mathbb{U}_e + \mathbb{u}_e)(\rho_{oe} + \rho_e), \\ \mathbb{J}_p &= (\mathbb{U}_p + \mathbb{u}_p)(\rho_{op} + \rho_p), \end{aligned}\right\} \tag{2.4}$$

$$\frac{\partial \mathbb{u}_a}{\partial t} = -\frac{e}{m}\left\{\mathbb{E} + \mathbb{u}_a \times (\mathbb{B}_0 + \mathbb{B})\right\}, \tag{2.5}$$

$$\frac{\partial \mathbb{u}_e}{\partial t} + \left\{(\mathbb{U}_e + \mathbb{u}_e)\nabla\right\}\mathbb{u}_e = -\frac{e}{m}\left\{\mathbb{E} + (\mathbb{U}_e + \mathbb{u}_e)\times(\mathbb{B}_0 + \mathbb{B})\right\}, \tag{2.6}$$

$$\frac{\partial \mathbb{u}_p}{\partial t} + \left\{(\mathbb{U}_p + \mathbb{u}_p)\nabla\right\}\mathbb{u}_p = \frac{e}{M}\left\{\mathbb{E} + (\mathbb{U}_p + \mathbb{u}_p)\times(\mathbb{B}_0 + \mathbb{B})\right\}, \tag{2.7}$$

where the effect of protons in the ambient plasma is neglected as is made in the magneto-ionic theory.

As the above equations are not always linear, in the following all products of two perturbations are neglected so as to linealize the equations. Static magnetic flux density B_0 in the equations (2.5) to (2.7) can be regarded as that of the applied magnetic field, if electrons and protons in the beam do not make any static current in the direction of the beam. Therefore in the following calculation, we assume that the number densities of electrons and protons in the beam are the same and moreover do the average velocities of electrons and protons, or assume that additional magnetic field due to a static convection current is negligibly small.

Chapter 2.4 Effect of an electron beam on the VLF wave propagation in the magneto-active plasma (Reexamination of the TWT mechanism in the exosphere)

The principle of TWT (traveling wave tube) like amplification is attributed to the energy exchange by the interaction between a longitudinal electron plasma wave and a longitudinal electric field of the electromagnetic wave. When it is applied to VLF electromagnetic wave propagating in the exosphere, the angle θ of the earth's magnetic field with z direction (wave direction) must not be zero, otherwise an electromagnetic wave propagating along the line of the earth's magnetic field ($\theta=0$) has no longitudinal electric field (Gallet and Helliwell, 1959; Gallet, 1959; Bell and Helliwell, 1959).

Now to verify the possibility of TWT like amplification in the exosphere, a quite general treatment must be made. Only a simplification is that we can neglect the effect of protons in the beam as well as that in the ambient stationary plasma.

By using the assumption of plane wave, all variables changing in $e^{j(\omega t-kz)}$, then the equations (2.1), (2.3) to (2.7) are rewritten in x, y, z components as follows, where varying magnetic field H is eliminated by means of (2.2),

$$\left(k^2-\frac{\omega^2}{c^2}\right)E_x+j\omega\mu_0\left(\rho_{a0}u_{ax}+\rho_{e0}u_{ex}+\rho_e U_{ex}\right)=0, \tag{2.8}$$

$$\left(k^2-\frac{\omega^2}{c^2}\right)E_y+j\omega\mu_0\left(\rho_{a0}u_{ay}+\rho_{e0}u_{ey}+\rho_e U_{ey}\right)=0, \tag{2.9}$$

$$-k^2E_z+j\omega\mu_0\left(\rho_{a0}u_{az}+\rho_{e0}u_{ez}+\rho_e U_{ez}\right)=0, \tag{2.10}$$

$$(\omega-kU_{ez})\rho_e-k\rho_{e0}u_{ez}=0, \tag{2.11}$$

$$j\omega u_{ax}+\omega_{Hz}u_{ay}+\frac{e}{m}E_x=0, \tag{2.12}$$

$$-\omega_{Hz}u_{ax}+j\omega u_{ay}+\omega_{Hx}u_{az}+\frac{e}{m}E_y=0, \tag{2.13}$$

$$-\omega_{Hx}u_{ay}+j\omega u_{az}+\frac{e}{m}E_z=0, \tag{2.14}$$

$$j(\omega - k U_{ez}) u_{ex} + \omega_{Hz} u_{ey} + \frac{e}{m} \frac{\omega - k U_{ez}}{\omega} E_x = 0, \tag{2.15}$$

$$-\omega_{Hz} u_{ex} + j(\omega - k U_{ez}) u_{ey} + \omega_{Hx} u_{ez} + \frac{e}{m} \frac{\omega - k U_{ez}}{\omega} E_y = 0, \tag{2.16}$$

$$-\omega_{Hx} u_{ey} + j(\omega - k U_{ez}) u_{ez} + \frac{e}{m} \left\{ \frac{k U_{ez}}{\omega} E_x + E_z \right\} = 0. \tag{2.17}$$

On eliminating $\mathbb{E}$, u_e, ρ_e from the above, we obtain

$$\left.\begin{aligned} C_{xx} u_{ax} + C_{xy} u_{ay} + C_{xz} u_{az} &= 0, \\ C_{yx} u_{ax} + C_{yy} u_{ay} + C_{yz} u_{az} &= 0, \\ C_{zx} u_{ax} + C_{zy} u_{ay} + C_{zz} u_{az} &= 0, \end{aligned}\right\} \tag{2.18}$$

where

$$\left.\begin{aligned} C_{xx} &= -\xi(\chi_a + \chi_e + \zeta) - \gamma_z^2 \xi, \\ C_{xy} &= j\{\gamma_z(\chi_a + \chi_e \xi + \zeta + \xi\xi) - \gamma_x \xi\eta\}, \\ C_{xz} &= \gamma_x \gamma_z \xi + \xi\eta(\chi_a - 1), \\ C_{yx} &= j\gamma_z(\chi_a + \chi_e \xi + \zeta + \xi\xi), \\ C_{yy} &= \xi(\chi_e + \chi_a + \zeta) + \gamma_z^2 \xi - \gamma_x^2, \\ C_{yz} &= -j\gamma_x(\chi_a - 1 + \xi\chi_e + \xi\xi), \\ C_{zx} &= \chi_e \eta - \gamma_x \gamma_z \xi, \\ C_{zy} &= j\gamma_x(\chi_a + \xi\chi_e + \zeta - \xi^2), \\ C_{zz} &= \chi_e + \xi^2(\chi_a - 1) + \gamma_x^2 \xi, \end{aligned}\right\} \tag{2.19}$$

with

$$\left.\begin{aligned} &\xi = 1 - n V_{ez}, \quad \eta = n V_{ex}, \quad \zeta = n^2 - 1, \quad n = k/k_0, \\ &V_{ez} = V_e \cos\theta, \quad V_{ex} = V_e \sin\theta, \quad \gamma = f_{He}/f, \quad \gamma_z = \gamma\cos\theta, \quad \gamma_x = \gamma\sin\theta, \end{aligned}\right\} \tag{2.20}$$

The necessary and sufficient condition for (2.18) to have non-zero solution is that the determinant formed by the coefficients of (2.18) equals zero.

$$\begin{vmatrix} C_{xx} & C_{xy} & C_{xz} \\ C_{yx} & C_{yy} & C_{yz} \\ C_{zx} & C_{zy} & C_{zz} \end{vmatrix} = 0, \tag{2.21}$$

This equation is generally called dispersion equation and is of the 8th degree of n (refractive index) in this case.

When $\theta = 0$, that is, the wave propagates in the direction of the earth's magnetic field vector, the dispersion equation yields

$$C_{xx}C_{yy} - C_{xy}C_{yx} = 0 , \tag{2.22}$$

$$C_{zz} = 0 , \tag{2.23}$$

Eq. (2.22) can be written as

$$n^3 - \frac{1}{V}(1-y)n^2 - \left(1 - X_e - \frac{X_a}{1-y}\right)n - \frac{1}{V}(X_a + X_e - 1 + y) = 0 , \tag{2.22}'$$

or

$$n^3 - \frac{1}{V}(1+y)n^2 - \left(1 - X_e - \frac{X_a}{1+y}\right)n - \frac{1}{V}(X_a + X_e - 1 - y) = 0 . \tag{2.22}''$$

The upper yields extraordinary electromagnetic wave mode and fast cyclotron mode of right handed polarization. The lower yields ordinary electromagnetic wave mode and slow cyclotron mode of left handed polarization.

(2.23) is reduced to

$$(1 - nV)^2 = \frac{X_e}{1 - X_a} . \tag{2.23}'$$

This equation provides space charge wave modes. Insofar as the practical TWT in very high frequency region is concerned, $X_a = 0$, so that (2.23)′ yields a fast space charge mode and slow space charge mode, that is,

$$n = \frac{1}{V}(1 - f_e/f), \tag{2.24}$$

$$n = \frac{1}{V}(1 + f_e/f). \tag{2.25}$$

The latter slow space charge wave mode causes an amplification of a slow circuit wave propagating along a spiral in a synchronous speed with the space charge wave due to the coupling between these two modes.

In our case, however, $X_a \neq 0$ and in addition $X_a \gg 1$ since the frequency of our problem is very low as compared with the electron plasma frequency of the ambient plasma, so that the root n of (2.23)' will become complex:

$$n = \frac{1}{V}\left(1 \pm j\sqrt{\frac{X_e}{X_a - 1}}\right) \qquad (2.26)$$

The above may be understood as the space charge wave itself being a growing wave (Sumi, 1959).

When $\theta \neq 0$, a coupling between the electromagnetic wave mode and the above mentioned growing space charge wave mode is expected. In general, (2.21) of the 8th degree can not be solved analytically. The roots were consequently obtained by means of the electronic computer KDC-1 of Kyoto University. In order to reexamine Bell's result (Bell and Helliwell, 1959) the following conditions of the medium and the beam are assumed.

$f_a = 1\text{Mc/s}$ $(N_a = 1.24 \times 10^{10}\text{m}^{-3})$,

$f_{He} = 1\text{Mc/s}$,

$N_e = 10^{-3}N_a = 1.24 \times 10^{7}\text{m}^{-3}$,

$U_e = 1.6 \times 10^{7}\text{m/s}$,

$f = 0.5 \sim 8\text{kc/s}$,

$\theta = 10^{\circ}, 30^{\circ}$.

In Fig. 2.4 and 2.5 the roots n of the dispersion equation (2.21) versus frequency f by inserting the above data are shown. To avoid complexity, three roots; forward propagaing extraordinary electromagnetic wave so called whistler mode and complex space charge wave modes (two conjugate roots) are merely plotted because other roots have no important effect on the coupling of the three modes of interest. These figures indicate that n of the whistler mode is always

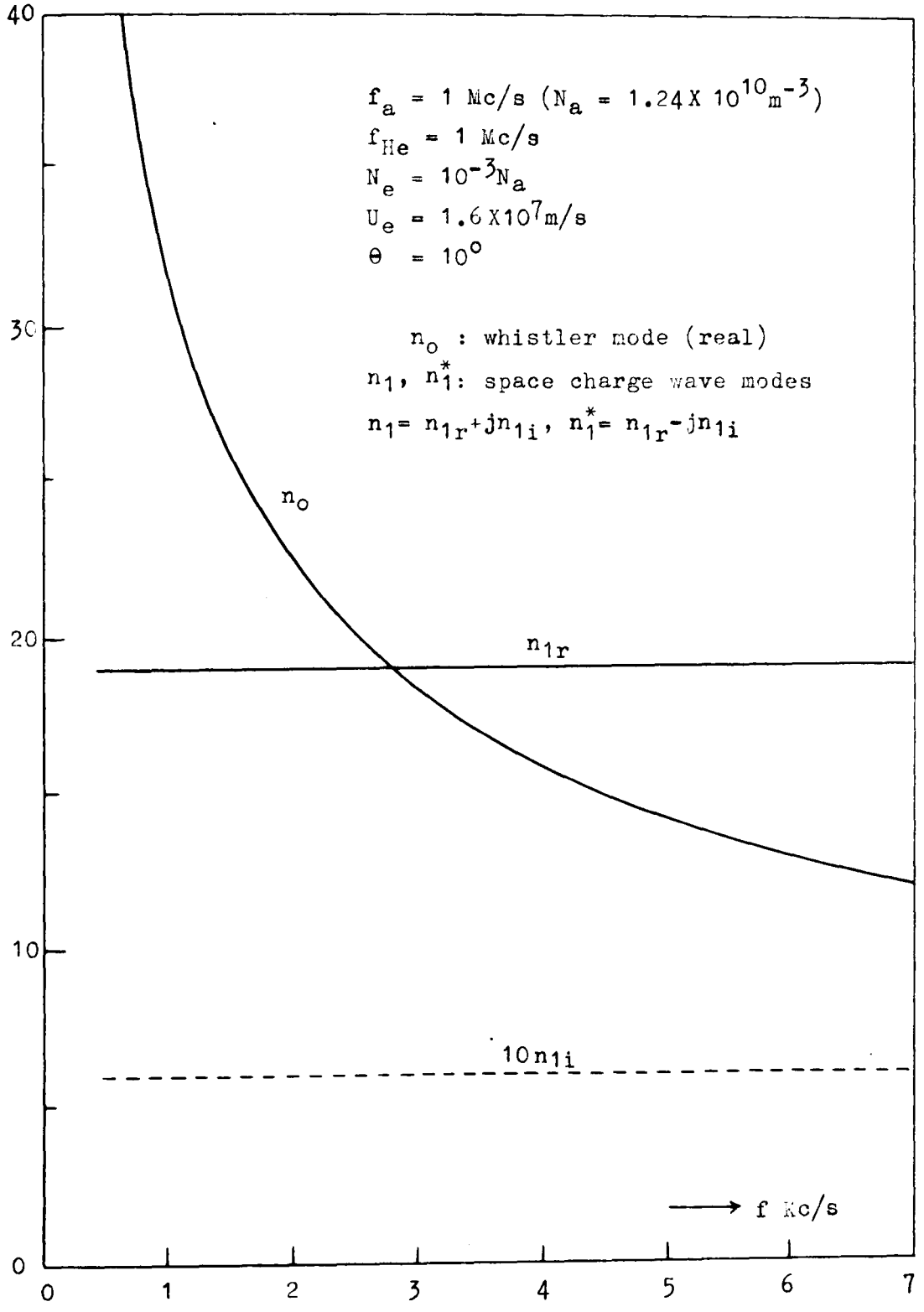


Fig.2.4 Refractive indices n of the coupled modes (θ =10^o)

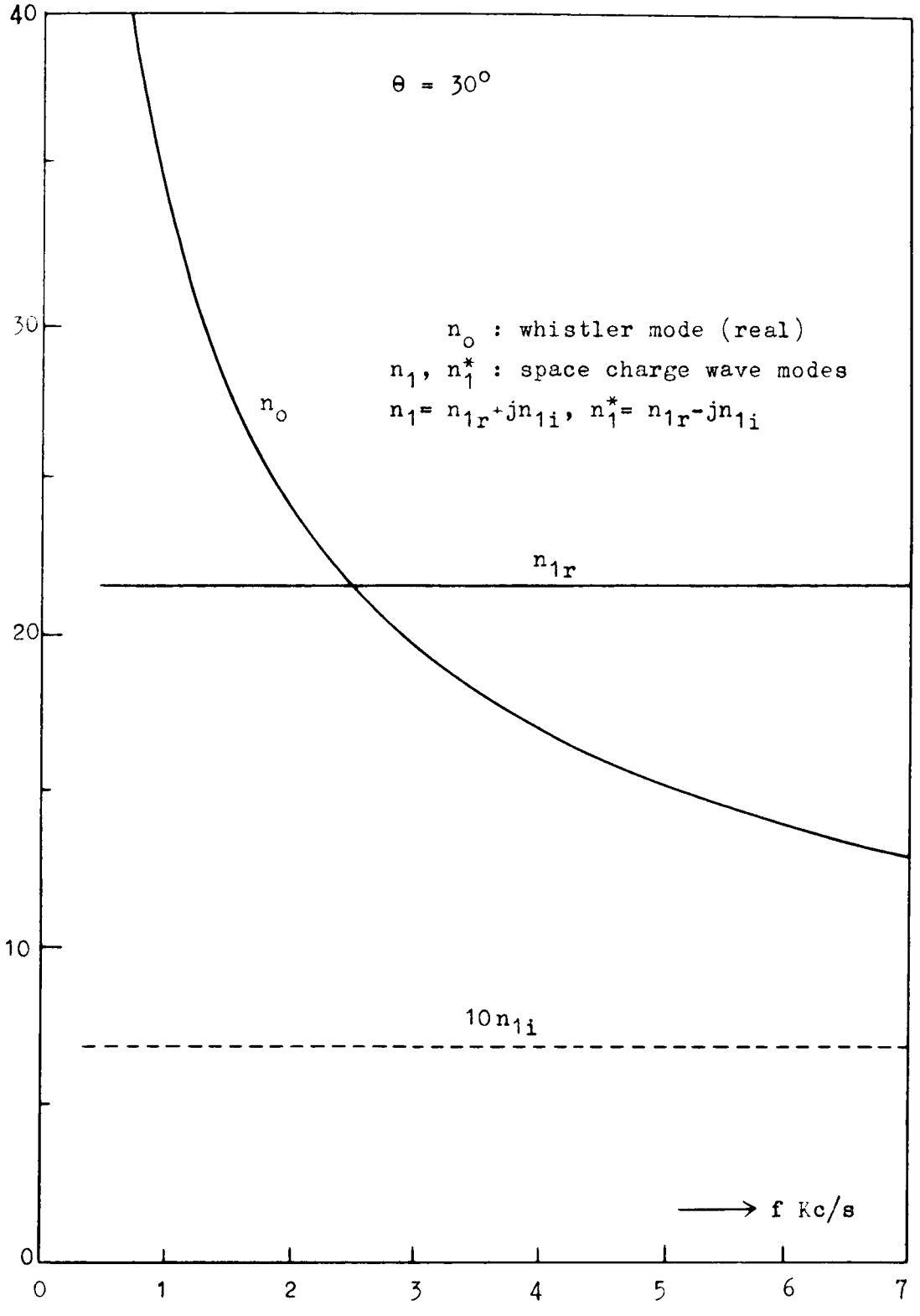


Fig.2.5 Refractive indices n of the coupled modes ($\theta = 30^\circ$)

real root over the coupling range of frequency and so an electromagnetic wave of the whistler mode can not give (or take) energy to (or from) space charge wave, even if the phase velocities of the both modes are in synchronism.

In concluding, the amplification of the VLF electromagnetic wave by means of the coupling with the electron beam in the exosphere can not be expected. Only a possibility is that a growing space charge wave will be partly converted to a wave of the whistler mode, if the space charge wave passes a certain boundary. In this case, however, there exists no frequency selectivity which is necessary to account for the queer frequency-time characteristics of the dawn chorus and the frequency spectrum of the hiss.

Chapter 2.5 Effect of a proton beam on the VLF wave propagation in the magneto-active plasma

2.5.1 Modes of electromagnetic wave propagation under the existence of a proton beam

In the preceding section we omitted the effect of protons. Nontheless the proton seems to have strong effect when the Doppler shifted proton gyro-frequency coincides the frequency of the electromagnetic wave of interest. In that case the transverse component of the electric field of the wave is important so that in the following only a perfect transverse case is treated; that is the case where the plane wave travels in the direction parallel to the earth's magnetic field ($\theta = 0$).

The fundamental equations are identical with (2.1) to (2.7) in chapter 2.3. After linealization and eliminating H by means of (2.2), x, y components of eqs. (2.1), (2.5), (2.6) and (2.7) are written as follows.

$$\left.\begin{aligned} \left(k^2 - \frac{\omega^2}{c^2}\right) E_x + j\omega\mu_0 (\rho_{0a} u_{ax} + \rho_{0e} u_{ex} + \rho_{0p} u_{px}) = 0, \\ \left(k^2 - \frac{\omega^2}{c^2}\right) E_y + j\omega\mu_0 (\rho_{0a} u_{ay} + \rho_{0e} u_{ey} + \rho_{0p} u_{py}) = 0. \end{aligned}\right\} \tag{2.27}$$

$$\left.\begin{aligned} j\omega u_{ax} + \omega_{He} u_{ay} + \frac{e}{m} E_x = 0, \\ -\omega_{He} u_{ax} + j\omega u_{ay} + \frac{e}{m} E_y = 0, \end{aligned}\right\} \tag{2.28}$$

$$\left.\begin{aligned} j(\omega - k U_e) u_{ex} + \omega_{He} u_{ey} + \frac{e}{m} \frac{\omega - k U_e}{\omega} E_x = 0, \\ -\omega_{He} u_{ex} + j(\omega - k U_e) u_{ey} + \frac{e}{m} \frac{\omega - k U_e}{\omega} E_y = 0, \end{aligned}\right\} \tag{2.29}$$

$$\left.\begin{aligned} j(\omega - k U_p) u_{px} - \omega_{Hp} u_{py} - \frac{e}{m} \frac{\omega - k U_p}{\omega} E_x = 0, \\ \omega_{Hp} u_{px} + j(\omega - k U_p) u_{py} - \frac{e}{m} \frac{\omega - k U_p}{\omega} E_y = 0, \end{aligned}\right\} \tag{2.30}$$

where (2.3) is unnecessary.

As is in the previous chapter, the dispersion equation is obtained, under the condition that $\mathbb{E}$, u_a, u_p, u_e must have non-zero solution. It becomes

$$\left\{(\xi_e+\gamma_e)(A+B)+\chi_e\xi_e(\xi_p-\gamma_p)(1+\gamma_e)\right\}\left\{(\xi_e-\gamma_e)(A-B)+\chi_e\xi_e(\xi_p+\gamma_p)(1-\gamma_e)\right\}=0, \tag{2.31}$$

$$\left.\begin{aligned} &\text{where} \quad A=(\chi_a+\chi_p+\zeta)\xi_p-\gamma_e\gamma_p\zeta, \\ &\qquad\quad B=(\chi_p+\zeta)\gamma_e\xi_p-(\chi_a+\zeta)\gamma_p, \\ &\text{with} \quad \xi_e=1-nV_e,\ \xi_p=1-nV_p,\ V_e=U_e/c,\ V_p=U_p/c, \\ &\qquad\quad \zeta=n^2-1,\ \chi_a=\frac{N_a e^2}{4\pi^2 m\varepsilon_0 f^2},\ \chi_e=\frac{N_e e^2}{4\pi^2 m\varepsilon_0 f^2},\ \chi_p=\frac{N_p e^2}{4\pi^2 m\varepsilon_0 f^2}, \end{aligned}\right\} \tag{2.32}$$

Eq. (2.31) is then rewritten as

$$\left\{nV_p-(1+\gamma_p)\right\}\left[\left\{nV_e-(1-\gamma_e)\right\}\left\{n^2-\left(1-\chi_e-\chi_p-\frac{\chi_a}{1-\gamma_e}\right)\right\}-\chi_e\gamma_e\right]+\chi_p\gamma_p\left\{nV_e-(1-\gamma_e)\right\}=0, \tag{2.33}$$

or

$$\left\{nV_p-(1-\gamma_p)\right\}\left[\left\{nV_e-(1+\gamma_e)\right\}\left\{n^2-\left(1-\chi_e-\chi_p-\frac{\chi_a}{1+\gamma_e}\right)\right\}+\chi_e\gamma_e\right]-\chi_p\gamma_p\left\{nV_e-(1+\gamma_e)\right\}=0. \tag{2.34}$$

Eq. (2.33) governs the right-handed polarized modes which include the extraordinary electromagnetic wave (whistler) mode and the slow cyclotron mode of protons. On the other hand (2.34) governs the left-handed polarized modes which include the ordinary electromagnetic wave mode and fast cyclotron mode of protons.

2.5.2 Spatially growing mode for VLF waves in the exosphere

Among the above two equations, it is clear that the one which possesses a possibility of the existence of spatially growing mode of the VLF electromagnetic wave is (2.33).

Dividing (2.33) by $nV_e-(1-\gamma_e)$, we have

$$\left\{nV_p-(1+\gamma_p)\right\}\left\{n^2-\left(1-\chi_e-\chi_p-\frac{\chi_a}{1-\gamma_e}+\frac{\chi_e\gamma_e}{nV_e-(1-\gamma_e)}\right)\right\}+\chi_p\gamma_p=0. \tag{2.35}$$

As each quantity of (2.35) has the following order of magnitude;

$$\chi_a \gg 1,\ \gamma_e \gg 1,\ \chi_e \simeq \chi_a/\gamma_e,\ \gamma_p \ll 1,\ \chi_p < 1,\ V_p \simeq V_e < 10^{-1},$$

$X_p y_p$ is so small that the roots of (2.35) will be obtained approximately from

$$n - \frac{1}{V_p}(1 + y_p) = 0, \tag{2.36}$$

$$n^2 - \left(1 - X_e - X_p - \frac{X_a}{1 - y_e} + \frac{X_e y_e}{nV_e - (1 - y_e)}\right) = 0. \tag{2.37}$$

Now we denote the root of (2.36) n_1, that is

$$n_1 = (1 + y_p)/V_p. \tag{2.38}$$

This mode is obviously a slow cyclotron mode of protons (Siegman, 1960). When one root of (2.37) is nearly the same as $|n_1|$ and $V_e \simeq V_p$,

$$nV_e - (1 - y_e) \simeq y_e,$$

so that (2.37) becomes

$$n^2 - \left(1 - X_p - \frac{X_a}{1 - y_e}\right) = 0.$$

Here we denote

$$n_o = \sqrt{1 - X_p - \frac{X_a}{1 - y_e}}, \tag{2.39}$$

this being undoubtedly the whistler mode of VLF electromagnetic wave.

Then (2.35) can be made as

$$(n - n_1)(n - n_o)(n + n_o) + X_p y_p / V_p = 0, \tag{2.40}$$

where $n_o$must be of the same order of n_1. In (2.40) as the quantity $X_p y_p / V_p$ is small, if we regard $n \simeq n_o \simeq n_1$, (2.40) may be approximately reduced to the following quadratic equation;

$$(n - n_o)(n - n_1) + X_p y_p / 2n_1 V_p = 0, \tag{2.41}$$

the roots of which are apparently,

$$n = \tfrac{1}{2}(n_o + n_1) \pm \sqrt{\tfrac{1}{4}(n_o - n_1)^2 - X_p y_p / 2n_1 V_p}. \tag{2.42}$$

n has complex roots only when

$$\tfrac{1}{4}(n_o - n_1)^2 < X_p y_p / 2n_1 V_p. \tag{2.43}$$

The maximum of the imaginary part of n occurs when

$$n_o = n_1, \tag{2.44}$$

and in this case roots n become

$$n = n_1 \pm j\gamma, \tag{2.45}$$

where

$$\gamma = \sqrt{X_p y_p / 2 n_1 V_p} = \sqrt{X_p y_p / 2(1 + y_p)}, \tag{2.46}$$

$$\cong \sqrt{X_p y_p / 2}. \tag{2.47}$$

These complex roots clearly mean that the whistler mode wave and proton cyclotron wave simultaneously grow or decay as they propagate in the positive z direction.

Now we can calculate the amplitude gain of the amplification. We denote the propagation constant k in real and imaginary parts,

$$k = k_r + jk_i,$$

then

$$k_r = k_o n_1 = \omega(1 + y_p)/U_p,$$

$$k_i = k_o \gamma \simeq (2\pi f/c)\sqrt{\tfrac{1}{2} X_p y_p} = 1.48 \times 10^{-8} \sqrt{f_p^2 f_{Hp}/f} \ (m^{-1}). \tag{2.48}$$

As it is convenenient to express f_p^2 and f_{Hp} in terms of N_p and f_{He}, (2.48) is rewritten as

$$k_i = 7.27 \times 10^{-11} \sqrt{N_p f_{He}/f} \quad (m^{-1}). \tag{2.49}$$

The amplitude gain Γ may be defined in the form;

$$\Gamma = e^{k_i \ell} \tag{2.50}$$

where ℓ is the length for which the waves propagate.

For one example, assume

$$f_{He}/f = 10^2, \quad N_p = 10^7 m^{-3}, \quad \ell = 10^6 m,$$

then

$$k_i = 2.3 \times 10^{-6} m^{-1},$$

$$\Gamma = e^{2.3} = 10 = 20 \text{ db}.$$

The condition necessary for amplification is given by (2.44). It can be understood as

$$v_{ph} = c/n_o = c/n_1 = U_p/(1 + y_p), \tag{2.5}$$

namely the phase velocity v_{ph} of the whistler mode wave must be nearly identical with the velocity of proton beam since $y_p \ll 1$.

Consequently this condition is equivalent to that proposed by Gallet and Helliwell (1959).

The frequency which satisfies (2.43) prescribes the frequency band to be simultaneously amplified. This band is approximately given by

$$\Delta f = 8 \frac{\gamma}{n_1} f , \qquad (2.52)$$

For the quantities $n_1 = 20$, $\gamma = 3.7 \times 10^{-2}$ ($N_p = 10^7$, $f_{He}/f = 10^2$),

$$\Delta f/f = 1.46 \times 10^{-2} = 1.46\% .$$

This value implies that the frequency band to be simultaneously amplified by a single velocity of the beam is quite narrow, being about 45c/s for the frequency region of about 3kc/s.

2.5.3 Physical interpretation of the spatially growing mode

Until now we have c nsidered only mathematically whether or not there were complex roots corresponding to the whistler mode wave. It is however necessary to prove that one of the complex roots really implies a growing mode in space. For that purpose, the energy consideration may be necessary and sufficient.

Now in speaking the conclusion in advance, the energy gained by the whistler mode wave and proton cyclotron wave is supplied only from the kinetic energy of the average velocity of the protons. Consequently the average velocity of the protons is slowed down in the direction of the earth's magnetic field, according as the whistler mode wave and the proton cyclotron wave are amplified. It is simply explained in the following way.

The power gain rate of the whistler mode wave during a traverse of dz is given by $\frac{\partial W_w}{\partial z} = \frac{1}{2} \Re \left\{ \frac{\partial}{\partial z} (\mathbb{E}_T \times \mathbb{H}_T^*) \right\}$,

$$= \frac{1}{2}\Re\{\nabla(\mathbb{E}_T \times \mathbb{H}_T^*)\}, \tag{2.53}$$

where $\Re$, *, suffix T imply "real part of", "complex conjugate" and "transverse component" respectively. By using (2.1) and (2.2), (2.53) is converted to

$$\begin{aligned}\frac{\partial W_w}{\partial z} &= -\frac{1}{2}\Re\left\{j\omega\frac{\varepsilon_0}{2}|E_T|^2 + j\omega\frac{\mu_0}{2}|H_T|^2 + \mathbb{E}_T \cdot \mathbb{J}_T^*\right\}, \\ &= -\frac{1}{2}\Re(\mathbb{E}_T \cdot \mathbb{J}_T^*), \\ &= \frac{1}{2}\Re\left\{eE_x(N_a u_{ax}^* + N_e u_{ex}^* - N_p u_{px}^*)\right. \\ &\quad \left. + eE_y(N_a u_{ay}^* + N_e u_{ey}^* - N_p u_{py}^*)\right\} \quad (W/m^2 \cdot m).\end{aligned} \tag{2.54}$$

From (2.30) u_{px}, u_{py} are determined in terms of E_x, E_y with k given by (2.48) for the maximum gain of amplification. Namely,

$$\left.\begin{aligned} u_{px} &= \frac{e}{M}\,\frac{\omega - kU_p}{\omega}\,\frac{j(\omega - kU_p)E_x + \omega_{Hp}E_y}{\omega_{Hp}^2 - (\omega - kU_p)^2}, \\ u_{py} &= \frac{e}{M}\,\frac{\omega - kU_p}{\omega}\,\frac{j(\omega - kU_p)E_y - \omega_{Hp}E_x}{\omega_{Hp}^2 - (\omega - kU_p)^2}, \\ \omega - kU_p &= -\omega_{Hp} - jk_iU_p, \end{aligned}\right\} \tag{2.55}$$

then

$$\left.\begin{aligned} u_{px} &\doteqdot -\frac{e}{M}\,\frac{\gamma_p}{2k_iU_p}(E_x + jE_y), \\ u_{py} &\doteqdot -\frac{e}{M}\,\frac{\gamma_p}{2k_iU_p}(E_y - jE_x), \end{aligned}\right\} \tag{2.56}$$

where k_iU_p is much smaller than ω_{Hp} so that k_iU_p was ignored against ω_{Hp}. In the same way u_{ex}, u_{ey}, u_{ax}, u_{ay} are given from (2.28) and (2.29):

$$\left.\begin{aligned} u_{ex} &\doteqdot \frac{e}{m}\,\frac{\omega_{Hp}}{\omega\,\omega_{He}^2}(-j\omega_{Hp}E_x - \omega_{He}E_y), \\ u_{ey} &\doteqdot \frac{e}{m}\,\frac{\omega_{Hp}}{\omega\,\omega_{He}^2}(-j\omega_{Hp}E_y + \omega_{He}E_x). \end{aligned}\right\} \tag{2.57}$$

$$u_{ax} = -\frac{e}{m}\frac{1}{\omega_{He}^2-\omega^2}(j\omega E_x - \omega_{He}E_y),$$
$$u_{ay} = -\frac{e}{m}\frac{1}{\omega_{He}^2-\omega^2}(j\omega E_y + \omega_{He}E_x). \qquad (2.58)$$

By using (2.56) (2.58), eq. (2.54) becomes simple, that is

$$\frac{\partial W_w}{\partial z} = \frac{1}{4}\frac{e^2N_p}{M}\frac{Y_p}{k_i U_p}|E_{oT}|^2 e^{2k_i z},$$
$$= \frac{1}{4}\frac{k_o^2 X_p Y_p}{\mu_o k_i U_p}|E_{oT}|^2 e^{2k_i z},$$
$$= \frac{1}{2}\frac{k_o^2 \eta_i \gamma}{\mu_o \omega}|E_{oT}|^2 e^{2k_i z} \quad (W/m^2m), \qquad (2.59)$$

where (2.38) and (2.46) are used. In this derivation it is apparent that the power gain of the electromagnetic wave is supplied from the work done by transverse proton velocity running against the transverse electric field. Notwithstanding the fact, the transverse velocity are not slowed down because it also acquires energy from the longitudinal velocity U_p, by means of the following mechanism.

The forces which act transversely on a proton are

$$F_1 = e\,\mathbb{U}_p \times \mathbb{B}_T, \qquad (2.60)$$

and

$$F_2 = e\,\mathbb{u}_{pT} \times \mathbb{B}_o. \qquad (2.61)$$

The force which acts longitudinally on a proton is

$$F_z = e(\mathbb{u}_{pT} \times \mathbb{B}_T). \qquad (2.62)$$

Physical works done by these forces per unit time, per unit cross section and per small axial length dz are respectively;

$$dW_1 = \tfrac{1}{2}e\,\Re(-u_{px}B_y^* + u_{py}B_x^*)U_pN_p dz, \qquad (2.63)$$
$$dW_2 = \tfrac{1}{2}e\,\Re(-u_{px}u_{py}^* + u_{py}u_{px}^*)B_oN_p dz = 0, \qquad (2.64)$$
$$dW_z = \tfrac{1}{2}e\,\Re(u_{px}B_y^* - u_{py}B_x^*)U_pN_p dz = -dW_1. \qquad (2.65)$$

By using eqs. (2.2), (2.56) and (2.65),

$$\frac{dW_z}{dz} = -\frac{dW_1}{dz} = -\frac{1}{2}e\Re\left\{\frac{k^*}{\omega}\frac{e}{M}\frac{\gamma_p N_p}{2k_i}(|E_x|^2 + jE_y\cdot E_x^* + |E_y|^2 - jE_xE_y^*)\right\},$$

$$= -\frac{1}{4}\frac{e^2}{M}\frac{N_p\gamma_p k_0 n_1}{k_i\omega}|E_{0T}|^2 e^{-2k_i z},$$

$$= -\frac{1}{4}\frac{\chi_p\gamma_p\omega k_0 n_1}{e^2\mu_0 k_i}|E_{0T}|^2 e^{2k_i z},$$

$$= -\frac{k_0^2 n_1\gamma(1+\gamma_p)}{2\mu_0\omega}|E_{0T}|^2 e^{2k_i z}, \qquad (2.66)$$

where (2.46) is inserted.

In addition, let $W_\perp$ be the transverse kinetic energy of the protons which flow across unit cross section in unit time, that is

$$W_\perp = \tfrac{1}{2}M\left|u_{pT}\right|^2 N_p U_p \ (W/m^2).$$

It will be easily proved that

$$\frac{\partial W_\perp}{\partial z} = \frac{1}{2}\frac{k_0^2 n_1\gamma}{\mu_0\omega}\gamma_p|E_{0T}|^2 e^{2k_i z} \ (W/m^2\cdot m). \qquad (2.67)$$

From the above results; (2.59), (2.66) and (2.67), it is proved that

$$-\frac{\partial W_z}{\partial z} = \frac{\partial W_1}{\partial z} = \frac{\partial W_w}{\partial z} + \frac{\partial W_\perp}{\partial z}. \qquad (2.68)$$

The identity (2.68) inevitably implies that the increased energy corresponding to the electromagnetic wave and the growing cyclotron wave is compensated by the depression of the longitudinal average velocity of the protons.

Chapter 2.6 Explanation of the VLF emissions in terms of the spatially growing mode by a proton beam

2.6.1 Frequency-time characteristics of the VLF emissions

According to the preceding chapters we obtained the conclusion that a VLF electromagnetic wave in extraordinary mode can be amplified by the transverse coupling with a synchronously running proton beam rather than by the longitudinal coupling with an electron beam as in TWT. But the condition required for amplification is identical with what has been used by Gallet and Helliwell(1959). Consequently the frequency versus time characteristics of the various VLF emissions may be accounted for from the condition if an appropriate distribution of the electron density in the exosphere is assumed as was shown in the cited paper.

Repeating here the condition, from (2.51)

$$v_{ph} = U_p, \tag{2.69}$$

where v_{ph} is the phase velocity of the whistler mode of electromagnetic wave and is given by

$$\begin{aligned} v_{ph} = c/n_o &= c\left[1 - \frac{f_p^2}{f^2} + \frac{f_a^2}{f(f_H - f)} \right]^{\frac{1}{2}}, \\ &\simeq c\left[1 + \frac{f_a^2}{f(f_H - f)} \right]^{\frac{1}{2}}, \end{aligned} \tag{2.70}$$

where (2.39) is used.

The identity (2.69) with (2.70) yields the appropriate frequencies to be amplified for the given velocity U_p, ambient electron plasma frequency f_a and electron gyro-frequency f_{He}, as

$$f_1 = \left(\frac{U_p}{c} \right)^2 \frac{f_a^2}{f_{He}}, \tag{2.71}$$

$$f_2 = f_{He}\left[1 - \left(\frac{U_p}{c} \frac{f_a}{f_{He}} \right)^2 \right], \tag{2.72}$$

in case U_p is much smaller than the velocity of light c. The VLF emissions are plausible to be explained by the frequency given by (2.71).

The frequency-time characteristics of the observed VLF emissions can, then, be interpreted as follows.

a) Discrete type of emissions; dawn chorus, hooks, quasi-horizontals....

The frequency to be amplified, f_1 given by (2.71), will spatially change depending upon the spatial distribution of ambient electron density N_a and gyro-frequency f_{He} and the velocity of proton beam. Now let a proton cloud the dimension of which is not so large, pass through the exosphere, then it appears that the cloud proceeds emitting the frequency f_1 determined by the space parameters of the position of the cloud at every moment. The emitted wave is propagated in the whistler mode down to the ground, so that the traveling time is given by the integration (1.11) in Part I. Hence the frequency time characteristics can be explained by a total characteristics, that is the frequency dependence of the traveling time plus the time variation of the emitted frequency f_1.

For example, let the beam velocity U_p be constant and the ambient electron density be of the form that

$$f_a^2/f_{He} = \text{constant},$$

then by (2.71) f_1 must be constant with time. Such a case is considered to be the case shown in Fig. 2.1(b), quasi-horizontal type. And in terms of an appropriate distribution of the ambient electron density, e.g. that given by a curve in Fig. 2.6, Gallet succeeded in showing the hook type of emissions.

In the same way any type of the discrete emissions can be

explained by an appropriate assumption of the electron density distribution. The pure frequency-time characteristics as shown in Fig. 2.1 suggests that the size of the incoming proton cloud must be less than about a hundred kilometers.

b) Continuous type of emissions; hiss

This type of emissions is easily understood. If the length of the proton beam is considerably large, waves emitted from a large region of the space reach the receiver simultaneously, then the energy spectrum becomes broad and continues for a long time.

2.6.2 Gain of the amplification

Next, to estimate the gain of amplification for the VLF emissions, the length of coupling between the electromagnetic wave and the beam must be known. From the frequency-time characteristics of the discrete type of emissions really observed, a hundred kilometers are supposed to be the axial (along a geomagnetic line of force) dimension L of the beam as described before. The length of coupling would be limited not only by the extent of spatial variation of the frequency to be amplified which is determined by (2.71), but also by such a dimension L of the beam, when the group velocity of the coupled mode is not the same as the average velocity of the protons.

According to (2.42), the refractive index corresponding to the phase velocity of the coupled modes is

$$n = \tfrac{1}{2}(n_0 + n_1),$$

for the frequency range where the amplification takes place. Then the group velocity v_g of the coupled modes is given by

$$v_g = c/n',$$

where $$n' = \frac{\partial(nf)}{\partial f} = \frac{1}{2}(n_0+n_1) + \frac{1}{2} f \frac{\partial(n_0+n_1)}{\partial f} .$$

In using (2.38) and (2.39), we obtain approximately

$$n' = 3n_0/4$$

for the strongest interaction. Then the group velocity of the waves, v_g is

$$v_g = 4c/3n_0 = (4/3)v_{ph} \simeq (4/3)U_p,$$

where v_{ph} is the phase velocity.

Consequently the wave energies of the whistler mode and the cyclotron mode travel with the velocity v_g, 4/3 times faster than that of the proton itself. It would appear that the possible length of coupling must be smaller than 4L, if the length of proton beam is L.

Now consider the quasi-horizontal type of emission. Then the maximum length of coupling will be 4L, that is several 100Km or so. In this case suppose

$$f_a^2/f_{He} = 8.55 \times 10^6 c/s$$

corresponding to f=3kc/s, U_p=5600Km/s and f_{He} on a geomagnetic line of force arising from the geomagnetic latitude 67^o on the ground, then k_i is calculated with N_p as parameter and plotted in Fig. 2.7 with height from 1000Km above the ground, for $N_p=10^5$, 10^6, $10^7 m^{-3}$ respectively. This figure indicates that about 20 db for 500Km coupling is expected in the lower region of the exosphere, if there exist $10^7 m^{-3}$ of protons in the beam.

For the hook type of emission, Gallet has postulated a distribution of N_a as shown in Fig. 2.6. Using the distribution of N_a and f_{He} (see also Fig. 2.6) explained previously, the frequencies f to be amplified are calculated for 4000, 6000, 8000, 10000Km/s of U_p as parameter (see solid curves in Fig. 2.8). According to Fig. 2.8,

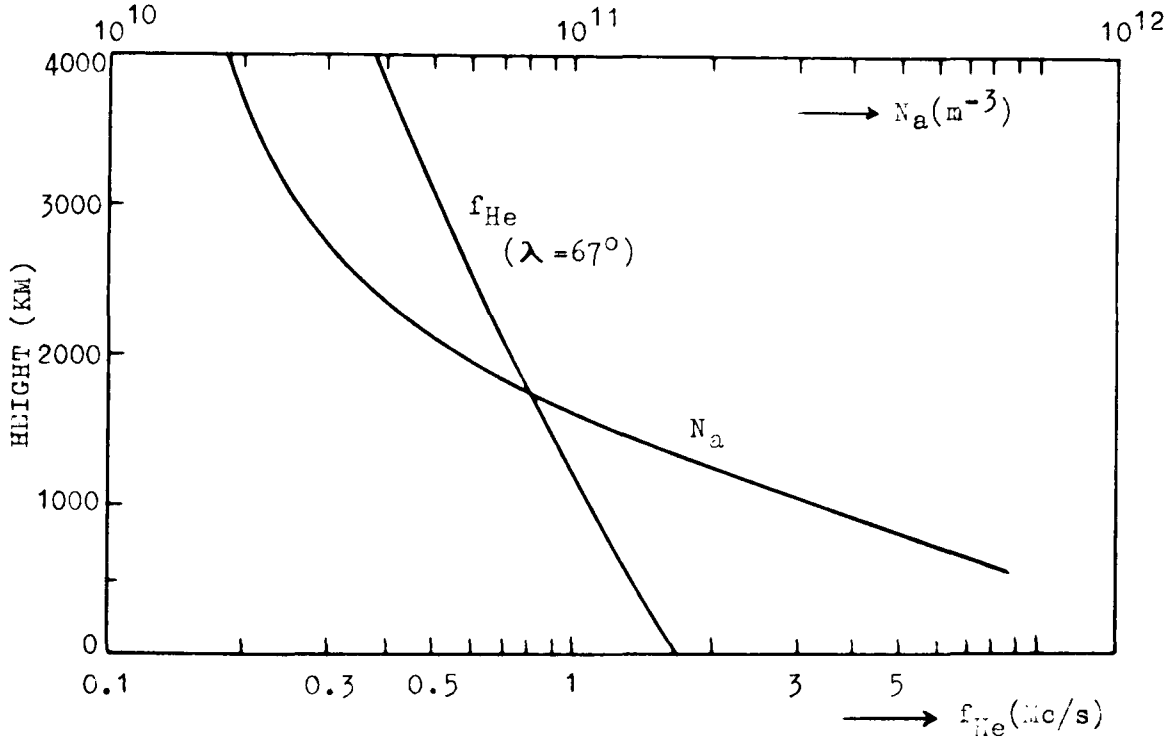


Fig. 2.6 Gyro-frequency f_{He} on the line of geomagnetic field arising from the geomagnetic latitude 67° on the ground, and assumed electron density N_a (after Gallet)

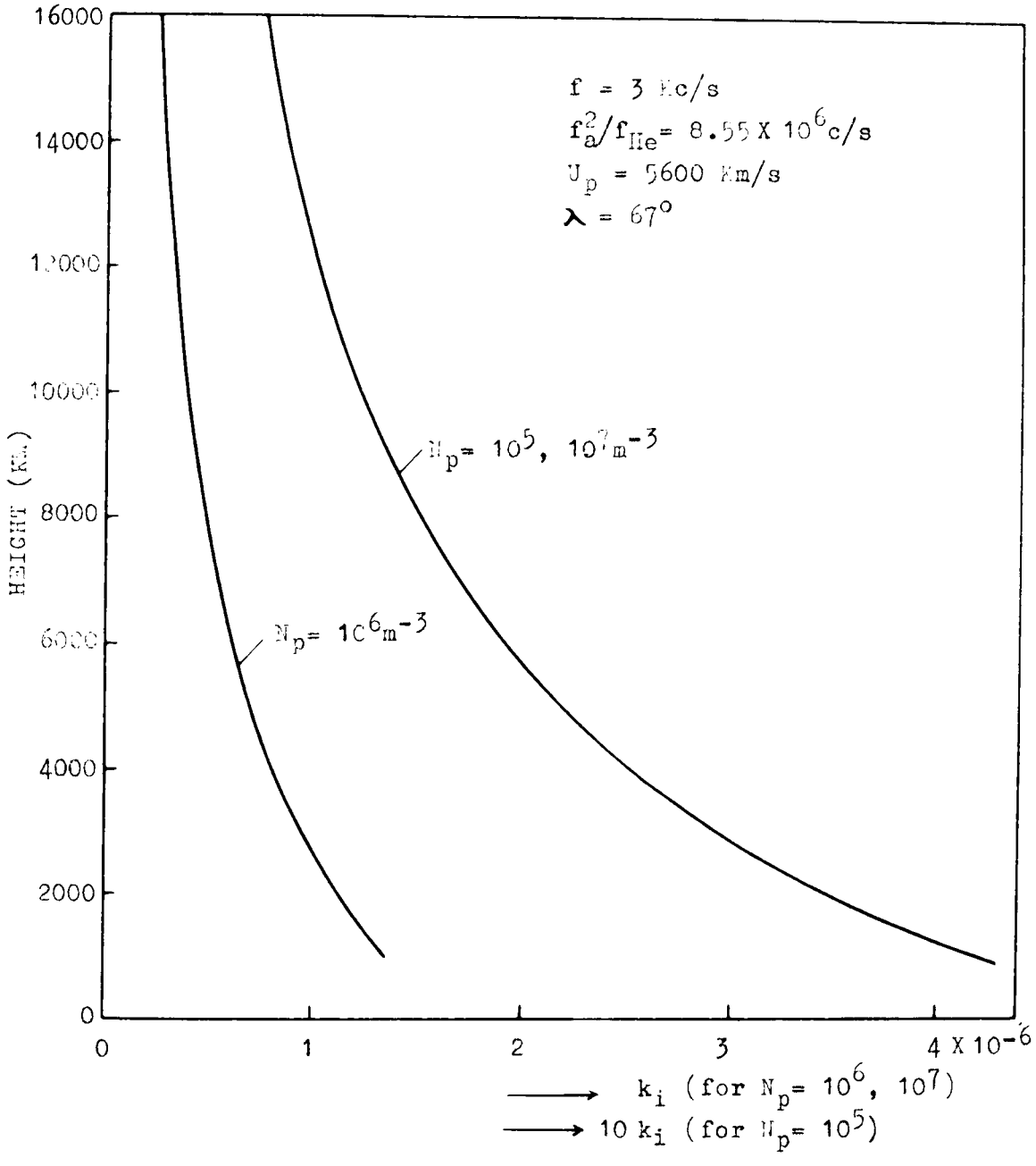


Fig.2.7 Imaginary part of the propagation constant k vs. height

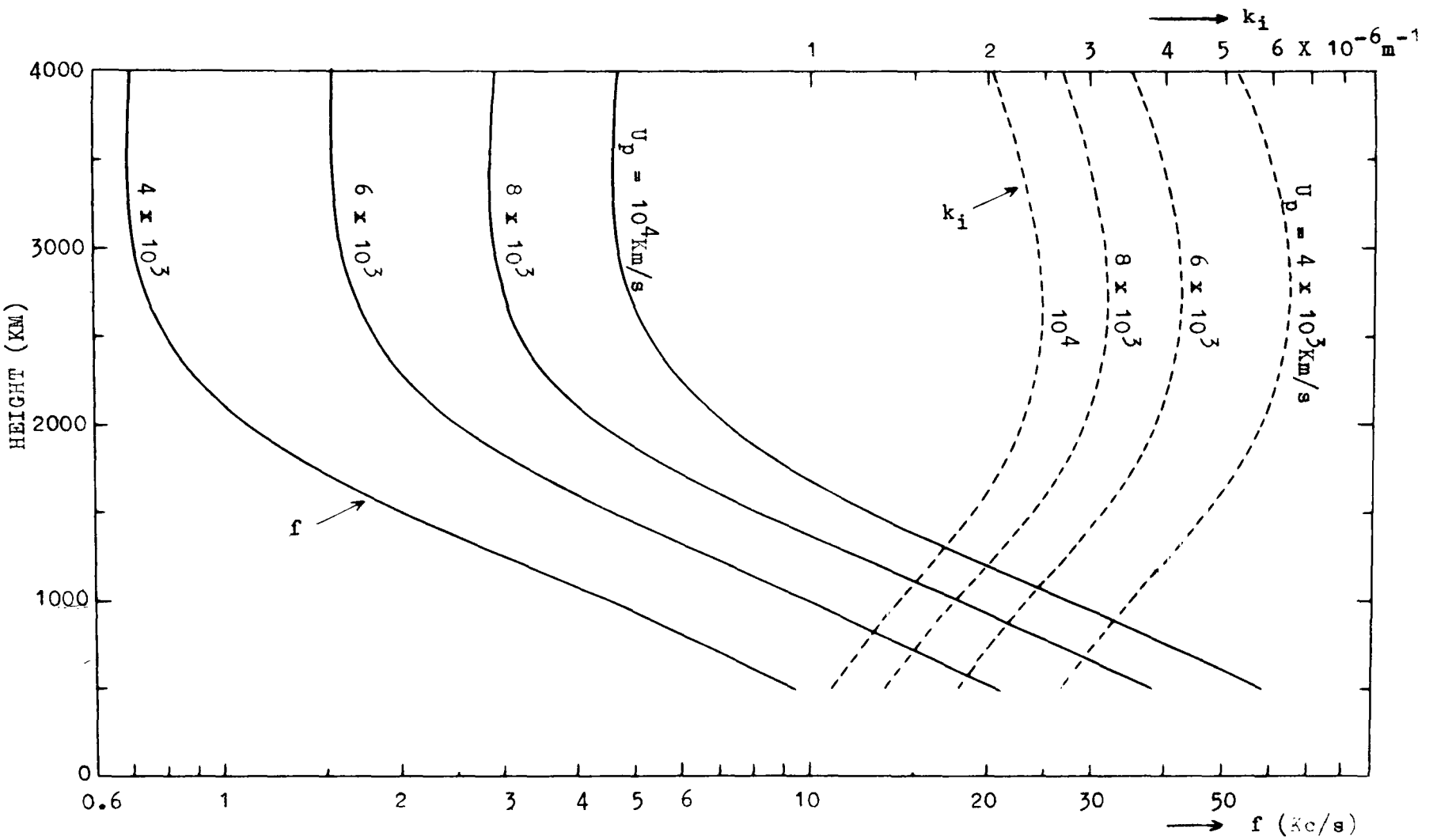


Fig.2.8 The frequencies to be amplified (solid curves) and corresponding k_i (dotted curves)

it is apparent that in the region where the frequency f changes rapidly, the possible length of coupling is so short, that we may not be able to expect even a few 100Km of coupling length. The dotted curves in the figure represent the corresponding k_i. In this case k_i has a maximum near 3000km above the ground.

Until now, we have treated single velocity theory. In practical case the velocity of protons might be distributed in a range. Such a velocity distribution widens the frequency band to be simultaneously amplified. As the result, the length of coupling would be lengthened a little. However the extent of velocity distribution must not be so wide, because the frequency-time characteristics of discrete type emissions are usually thin line on the sonagrams.

Although we have not yet known plausible data of the proton density in the beam, from the observation of the auroral proton, we can expect the proton density to be of the order of $10^5 m^{-3}$ in usual aurorae (Parker, 1959). Then $10^6 \sim 10^7 m^{-3}$ of proton density used in calculating the curves of Figs. 2.7 and 2.8 would not be impossible value, though it can not exceed much the value.

2.6.3 Consideration on the noise input (or seed) for amplification

Summarizing the above results, the gain of amplification expected by this proton mechanism is no more than 20db or so. Thus for the explanation of the powerful dawn chorus and hiss, a certain powerful noise input (or seed) is needed.

Possible seed or input might be thermal noise (for example, Bremsstrahlung), cyclotron radiation, Cherenkov radiation, or whistlers. Apart from the last one, other three are all rather weak, even Cherenkov radiation which is considered as most powerful in the exosphere

among the three, being of the order of $10^{-20}Wm^{-2}(c/s)^{-1}$ for $N_p=10^7m^{-3}$ (Ondoh, 1961). The whistler may be amplified by this proton mechanism, but it is improbable to regard that the whistlers are the peerless source of VLF emissions. Then another noise source must be sought for.

Now we have to remind the whistler mode of VLF electromagnetic wave propagation which provides a kind of radio windows in the ionosphere, so that VLF noise (or atmospherics) generated by lightnings or other discharges in the troposphere always permeates into the exosphere through the window. Electric field intensity of the VLF noise in the space between the ionosphere and the ground may be estimated from the charts of global intensity contour produced by C.C.I.R. (1959).

According to the charts, the electric field intensity of about 3kc/s under the ionosphere is as strong as $10^2\sim10^3\mu v/m$ for 1kc/s of bandwidth. Taking into account the absorption of the whistler mode wave within the ionosphere (about 5db at midnight as was shown in Fig. 1. in Part I), the power flux density of the noise in the exosphere during night time would easily become larger than 10^{-15} $Wm^{-2}(c/s)^{-1}$. These noises may have continuous frequency spectrum in the exosphere and the mechanism of amplification previously considered would be regarded to have a rôle of a band-pass filter which selects an appropriate narrow frequency band from the continuous noise spectrum and amplifies it by a little amount, say 20db necessary to contrast the frequency band with background noise level. Frequency characteristics of the band-pass filter may be varied due to the variation of the ambient electron density and earth's magnetic field

with altitudes according as the proton beam approaches the earth. Such a variation of characteristics would be detected at the ground as a dawn chorus or other kind of emissions, as was explained in 2.6.1.

From such a point of view, as the proton density of the beam the order of $10^{7}m^{-3}$ may be necessary.

It is worthwhile to examine whether or not the incoming proton beam has sufficient power density to amplify an ambient noise up to the power flux of about $10^{-14}Wm^{-2}(c/s)^{-1}$. As Gallet and Helliwell have checked (1959), here we calculate the power density in such a beam,

$$(\tfrac{1}{2}MU_p^2)(N_pU_p)\ Wm^{-2}.$$

Provided that the proton density is at least $10^{5}m^{-3}$ and the average velocity of protons U_p is $10^{7}m/s$, the power density is about 8.1×10^{-2} Wm^{-2}. This value implies that the amplification of the electromagnetic wave causes nothing but a tiny decrease of the average velocity of the protons.

Chapter 2.7 Concluding remarks

Our study has shown that in the magneto-active plasma, an electron beam did not give any special effect on the propagation of the VLF electromagnetic wave, but a proton beam had a great effect, that is , it could amplify the electromagnetic wave. The mechanism of amplification is due to a transverse coupling between the slow cyclotron mode of protons and the whistler mode of electromagnetic waves.

The condition of the amplification is that the phase velocity of the wave in the medium is nearly equal to the average velocity of the proton beam. The energy required for amplification of the waves is supplied from the kinetic energy of the moving protons. Consequently the average velocity of the protons in the direction of the applied magnetic field is slowed down.

This mechanism of amplification can be applied to the VLF emission which, we have supposed, may be generated in the exosphere. From the condition of amplification, the frequency-time characteristics can be explained by assuming an appropriate distribution of the electron density in the exosphere.

If this mechanism is really applicable to the VLF emissions, it will be consistent with the fact that the source of the VLF emissions appears to lie on the line of precipitation of positive particles in the exosphere (Allcock, 1957), and that a close association exists between aurorae and certain hiss (Martin et al, 1960).

However the gain of the amplification by such a transverse coupling with a proton beam is not so sufficiently large. The expected gain is no more than 20db by the coupling of several hundred km

along the earth's magnetic lines of force. Therefore the original noise seed (or input) for amplification must be already powerful enough in order to account for a powerful dawn chorus really observed.

In the upper ionosphere and in the exosphere there possibly exists the whistler mode of propagation for VLF electromagnetic waves. Thus, as far as the attenuation within the lower part of the ionosphere is small, intense atmospheric noises of continuous spectrum originating in the space below the ionosphere will be pervaded in the exosphere with the comparable intensity. This noise is quite so intense, that it can be the noise input for the amplification and then generates the VLF emissions.

Finally it must be noted that the general theory of Part II may also be applicable to the solar radio emission and some radiation in the micro wave region from the magneto-active plasma in laboratory experiments. It is because our arguments can be applied, as far as the wave frequency is much less than the electron plasma frequency of the ambient plasma, less than the electron gyro-frequency and larger than the proton gyro-frequency.

PART III

A NOVEL METHOD OF IONOSPHERE SOUNDING BY USING MIDDLE FREQUENCY BROADCAST WAVES

Chapter 3.1 Introduction

The observation of the ionosphere is usually made by the impulse sounding method which was first devised by G. Breit and M. A. Tuve. The apparatus and technique of the observation have already considerably advanced. Frequency used ordinarily ranges from 1Mc/s up to 20 Mc/s in Japan. The upper end of the frequency, 20Mc/s may be sufficient, as the critical frequency of the F layer will seldom exceed the value. However the lower end of the frequency 1Mc/s is not always appropriate, because during night time the critical frequency of the E layer often decreases until below 1Mc/s, and it is known that it becomes usually 600Kc/s$\sim$700Kc/s.

From these reason, the observation using lower frequency than 1Mc/s is needed. But this frequency region is in the middle frequency broadcast band that is 635Kc/s to 1605Kc/s, so we should desist from transmitting any impulse wave in this frequency band. This might be one reason why the observations using such low frequencies were not carried out. There seems to be another reason that at low frequencies the radio waves suffer a great deal of attenuation in the lower part of the ionosphere, so that the echoes clear enough could not be expected for determining the height of reflection, without powerful transmitter.

During night, the ionosphere does not receive solar radiation at all. Consequently it seems likely quite easy to treat theoretically

the time variation of the nocturnal electron density. But the treatment has not yet so much been done, due to the lack of the enough data. Only in the United States and in Australia, some observations have been carried on using powerful impulse transmitter (Watts and Brown, 1950; 1954; Watts, 1957; Mitra, 1957).

The electron density of the E layer is much associated with the absorption of the waves in the middle and low frequency band and especially it is needed to be known at night. Moreover it is apparently necessary to evaluate the conductivity of the night time E layer, which is used in the dynamo theory to explain the diurnal variation of the earth's magnetic field.

Under the above conditions and necessities, an ionospheric observation is wanted to be made without using the impulse method, in the middle frequency band, during night time. Only a method which satisfies such requirements may be the one which uses broadcast waves. A broadcast wave itself being quite powerful, will propagate partly along the ground surface, and be partly reflected from the ionosphere and reach a receiving point on the ground. The former is called the ground wave or direct wave and the latter is called the sky wave or reflected wave.

The difference of the traveling time of the both waves is naturally a function of the height of the reflecting layer. If we can extract only the sky wave against the ground wave, our object would be almost accomplished. For the purpose we have used a loop antenna, and had satisfactory results. In what follows, the experiment will be explained with some theoretical background.

Chapter 3.2 Principle

As one of the errors appearing in direction finding in the middle frequency band, we have experienced "night error". It is such a phenomenon as at night the electro-motive force of any loop antenna for a radio wave does not vanish in any direction of rotation of the antenna, so that we can not find the true direction of the transmitting station at night, while in the daytime the antenna can sense the true direction of the station.

This error is inevitable in all direction finders which utilize the loop antennae. It is because such a direction finder relies upon the ground wave, the magnetic field of which is horizontally polarized, however some elliptically polarized sky wave is included in the received wave at night.

In Fig. 3.1, the configulation of the electric field components of an incident wave is shown. The ground wave(denoted by E_G) has approximately vertical polarization due to the large conductivity of the ground, and is usually incoming horizontally to the antenna P along x axis. The sky wave(denoted by E_V, E_H) which is reflected at the ionosphere and incident upon P making an angle i with z, is usually left-handed quasi circularly polarized in ordinary mode and right-handed in extraordinary mode respectively in middle frequency band owing to the effect of the earth'smagnetic field (Appleton, Ratcliffe, 1927).

Let E_H be the horizontal component of the sky wave which is perpendicular to the incident plane, and E_V be the component on the incident plane which is perpendicular to both E_H and the incident direction. Then in general

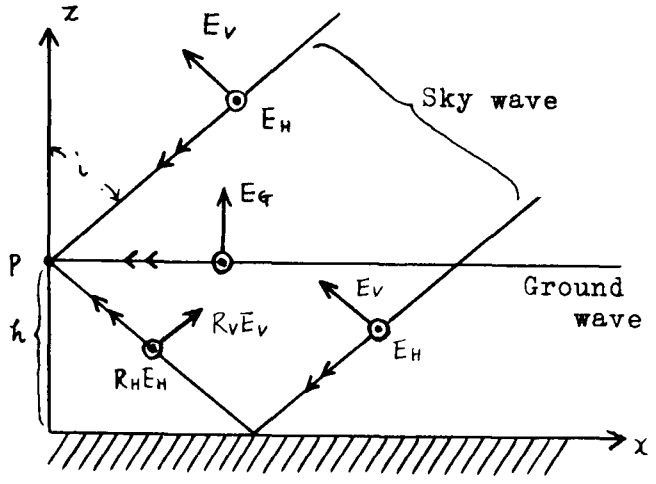


Fig. 3.1 Configuration of the electric fields of a incident electro-magnetic wave.

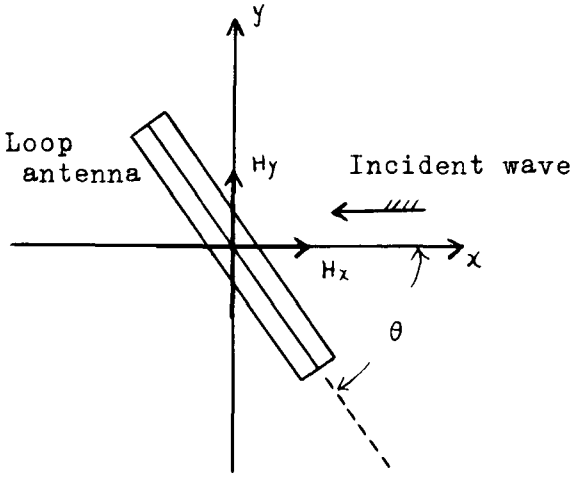


Fig. 3.2 Magnetic field components incident to a loop antenna.

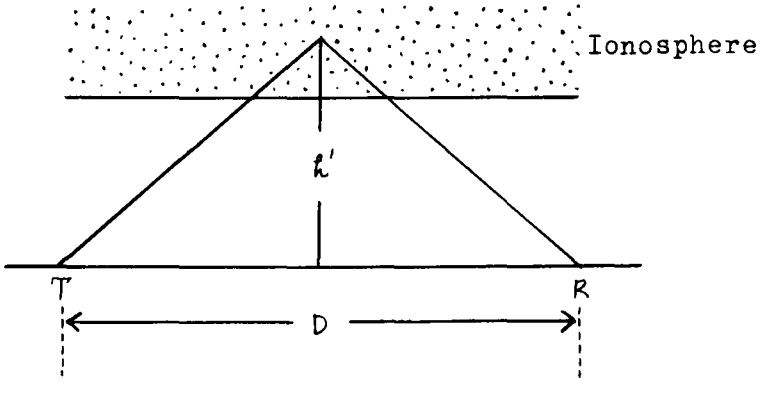


Fig. 3.3 Apparent height of ionospheric reflection.

$$E_V = \rho e^{-j\delta} E_H \ , \qquad (3.1)$$

holds, where ρ is the amplitude ratio nearly equal to unity and δ is the phase difference nearly equal to $\frac{1}{2}\pi$ in our case. Besides, as shown in Fig. 3.1, there is incident the ground-reflected wave of the sky wave, where R_H, R_V are the reflection coefficient of the ground for E_H, E_V respectively and are approximately -1, 1 respectively in case the ground is assumed to have infinite conductivity.

Now, x and y component of the magnetic field H in terms of E are given as follows.

$$H_x = \frac{1}{120\pi} (1 - R_H e^{-j\Delta}) E_H \cos i \ , \qquad (3.2)$$

$$H_y = -\frac{1}{120\pi} \left\{ E_G + (1 + R_V e^{-j\Delta}) E_V \right\} , \qquad (3.3)$$

where Δ is the phase difference due to the path difference between the direct component and the ground-reflected component of the sky wave and usually $e^{-j\Delta} \simeq 1$, when the antenna P is located sufficiently near the ground, compared with the wave length. Then, the loop antenna which is directed in such a way as shown in Fig. 3.2, will be penetrated by the magnetic flux ϕ given by

$$\phi = \mu_0 (H_y \cos\theta + H_x \sin\theta) NA \ , \qquad (3.4)$$

where A is the area and N is the number of turns of the loop antenna. Hence the electro-motive force e_l generated in the loop due to the change of ϕ with time will be

$$e_l = -j \frac{2\pi NA}{\lambda} \left[(E_G + 2E_V) \cos\theta - 2E_H \cos i \cdot \sin\theta \right] , \qquad (3.5)$$

where λ is the wave length.

It is naturally clear that if $E_V \neq 0$ and $E_H \neq 0$, there is no θ to make e_l zero. It is the cause of the night error.

On the contrary, if we have known the direction of the transmitting station, the angle θ can be adjusted as 90°. As the result the emf $e_{\ell N}$ of the loop is represented by

$$e_{\ell N} = j \frac{4\pi NA}{\lambda} E_H \cos i \ , \qquad (3.6)$$

namely the output of the loop antenna being generated only by the sky wave component. This adjustment will else be made in such a way that in the daytime when the sky wave component is very small due to the absorption in the ionosphere the output e_ℓ can be easily made zero by adjusting the angle θ (that is the condition $\theta = 90^{\circ}$), then this direction of the loop is held unchanged until the sky wave component become large at night.

After all the sky wave component can be extracted, in principle as far as it is not so small.

On the other hand, if we have a vertical antenna in the vicinity of P, the output of the antenna e_V will be

$$e_V = (E_G + 2E_V \sin i)\, h_V \ , \qquad (3.7)$$

where h_V is the effective height of the vertical antenna. Now the transmitting station being selected out of those located close to the receiving point, the following always holds;

$$|E_G| \gg |E_V| \ . \qquad (3.8)$$

Therefore we obtain even at night

$$e_V \simeq E_G h_V \qquad (3.9)$$

Then it is known that by using a loop antenna and a vertical antenna the sky wave and the ground wave can be separately received.

Now let D be the distance between the transmitting point and receiving point, and h' be the apparent height of the reflection

point (see Fig. 3.3), then the time lag τ of the sky wave(e_N) from the ground wave(e_V) is given by

$$\tau = (\sqrt{D^2 + 4h'^2} - D)/c, \tag{3.10}$$

c being the light velocity in free space.

For example, if we assume D=50Km, and h'=100Km, then τ=0.52msec. And if h'=300Km, $\tau \simeq$2msec. When the transmitted wave is a broadcast wave which is modulated by an audio frequency p(c/s), one cycle of 2000c/s of p corresponds 0.5msec. and one cycle of 500c/s does 2msec. It implies that if the modulation frequency p is smaller than 2000c/s, h' smaller than 100Km (corresponding to the E layer) can be determined from the time lag (or phase shift) of the modulating signal. For the F layer, the modulating frequency must be less than 500c/s.

In general, we have the following two theorems:

i) If a wave of the frequency f incident upon the ionosphere with an angle i_0 to the vertical is reflected at a height h_0, the frequency $f_\perp$ of another wave which is incident vertically and reflected at the same height h_0 must be

$$f_\perp = f \cos i_0. \tag{3.11}$$

ii) Whatever the incident angle i_0 is, the apparent height h' is invariant for any wave which is reflected at a constant height h_0(Martyn's theorem).

As the result, the observation using a wave of frequency f and incident with angle i_0 to the vertical is equivalent to the vertical sounding observation of the impulse method using the frequency

$$f_\perp = f \cos i_0.$$

Chapter 3.3 Description of the instrumentation

3.3.1 Requirements for the apparatus

To experiment our method, the following specifications should be taken into account.

1. A loop antenna is used which can be rotated arbitray to be able to extract the sky wave component only.
2. Phase characteristics for audio frequency of the two receivers and amplifiers for sky wave and ground wave must be equalized as good as possible.
3. A precise measurement is to be made of the time shift between the output (audio signal) wave forms of the sky wave receiver and ground wave receiver. Precision of the time measurement must be smaller than 30 μs which corresponds to 5Km in the precision of the height measurement.
4. When the broadcast wave is amplitude-modulated signals, the amplitude of the audio signals, by which the time lag is measured, may vary over a wide dynamic range. Consequently the phase comparator must respond exactly to the range of the amplitude.

According to the above specifications the apparatus is constructed as shown in Fig. 3.4.

3.3.2 Loop antenna system and receiver for sky wave extraction

In order to separate the sky wave from the ground wave, a gonioscope direction finder TD-B3a (Taiyo Musen Co.) was used. In place of turning a loop antenna, the same action is put in by means of crossed loop antennae and by turning the gonioscope in the receiver. The loop antennae are of 1m diameter and single-turned.

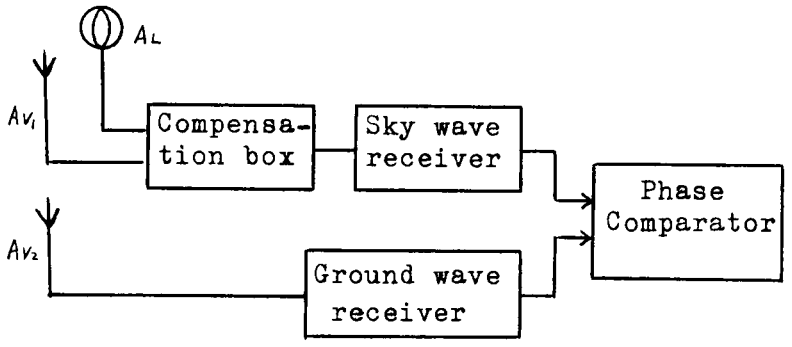


Fig. 3.4 General construction of the method.

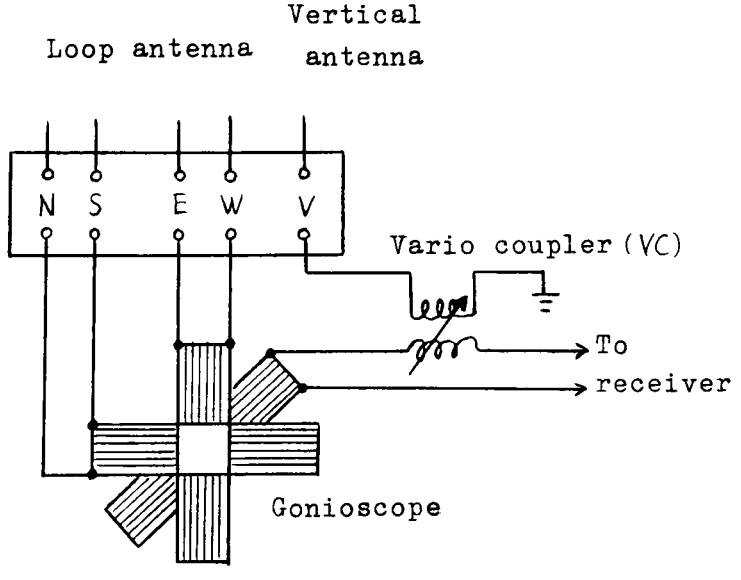


Fig. 3.5 Compensation box.

Perfect 8-type characteristics could not be obtained in the usual way, because of the vertical antenna effect. Therefore another vertical antenna is used so as to cancel the effect. It is carried out in the compensation box, shown in Fig. 3.5, that is, the emf of the vertical antenna whose effective height is about 10m, is added to the output of the gonioscope, through a vario-coupler VC. By adjusting this VC, the maximum to minimum ratio of the output of the antenna proper, hereafter abbreviated as "M-M ratio", was improved up to several thousands. This magnitude of the M-M ratio implies that if the sky wave intensity is larger than one thousandth of ground wave intensity, the sky wave can be separated. Most of the actual cases during nighttime, the sky wave is stronger than the above value. Then the M-M ratio of one thousand or two seems to be sufficient.

After the compensation box the signal is lead to a superheterodyne receiver which has one stage of radio frequency amplifier and two stages of intermediate frequency amplifier. The intermediate frequency is 107.5Kc/s and the bandwidth is about 3.5Kc/s.

The detector stage following the IF amplifiers must be made with attention on minimizing distortion. The circuit of the detector is shown in Fig. 3.6.

The receiver for ground wave is desirable to be the same circuit as the sky wave receiver. It is, however, not indispensable requirement, if the difference of the phase characteristics of the both receivers is known. In this case it is naturally favorable that the difference of the characteristics in audio frequencies is constant, especially in the range of some 100c/s ~ 2000c/s in order to simplify the time lag compensation.

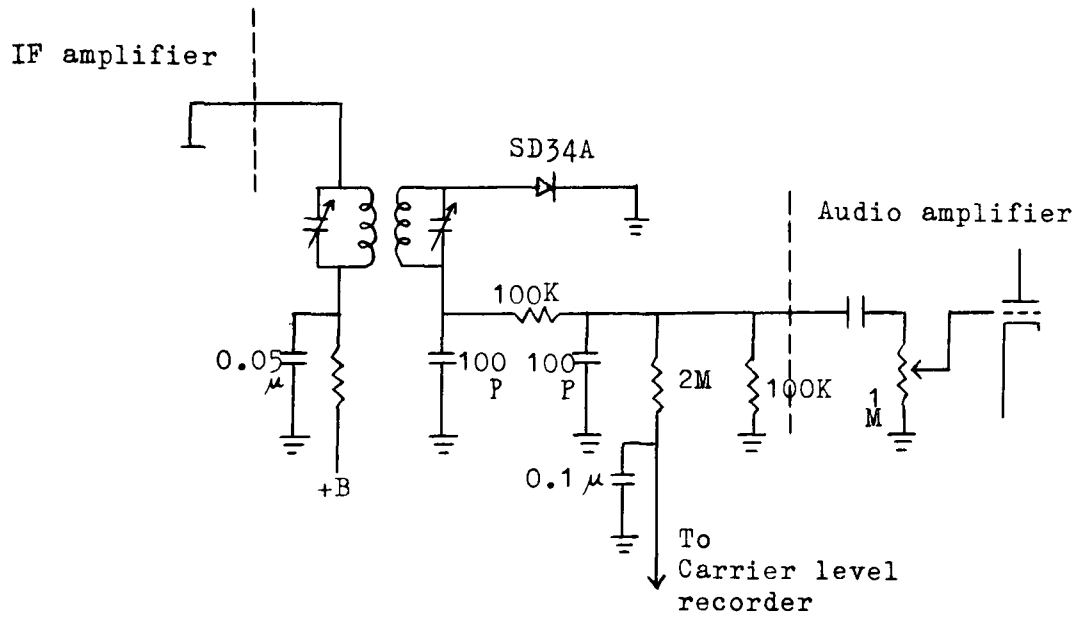


Fig. 3.6 Detector circuit.

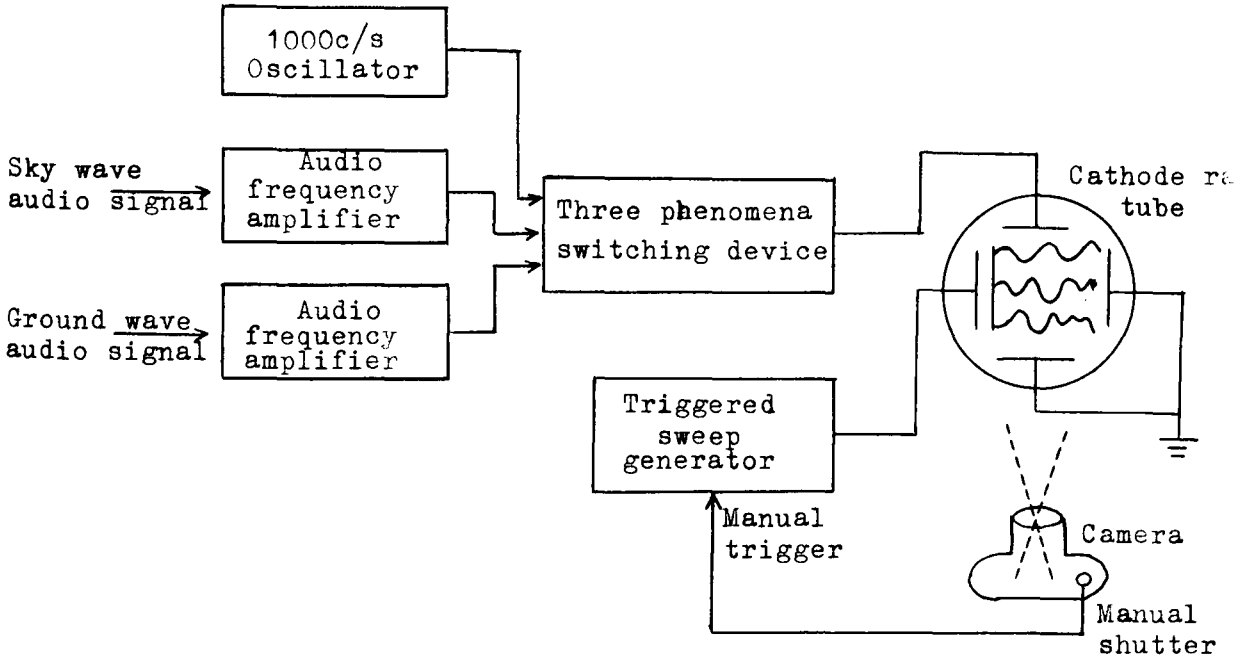


Fig. 3.7 Direct phase comparator

The carrier level of the received sky wave is very changeable, due to some kinds of fading. In our experiment, this carrier level was simultaneously recorded with a pen-recorder, since it is much associated with the error in measurement as described in chapter 3.5.

3.3.3 Phase comparator

a) Direct method

Measurement of the time lag of the sky wave with respect to the ground wave is reduced to the measurement of the phase difference within two channels for the same audio signal. The simplest one we have tried is as shown in Fig. 3.7. Namely the output signals of the two channels are displayed on a cathode ray tube with a manually triggered horizontal sweep. Such a display using a single beam tube is carried out by a "three phenomena switching device", and a standard signal say 1Kc/s is added on the display as a time scale. The view of these displays is recorded one by one on photographic film, simultaneously with the manual trigger of the horizontal sweep. Some examples of the record are shown in Fig. 3.8.

b) Pseudo impulse method

As favorable results were obtained for the experiment as was shown above, a rationalization of the method of phase comparison has been considered. A new method is indicated in the block diagram of Fig.3.9.

The audio signals after the detection stages are introduced to band pass filters which pass only the frequency range, 500c/s~650c/s. Such a filter has an action to make the wave form sinusoidal, otherwise the raw signal is of quite complicated wave form, because vocal and musical sounds consist of many higher harmonics besides a fundamental frequency. The frequency characteristics of the filter is

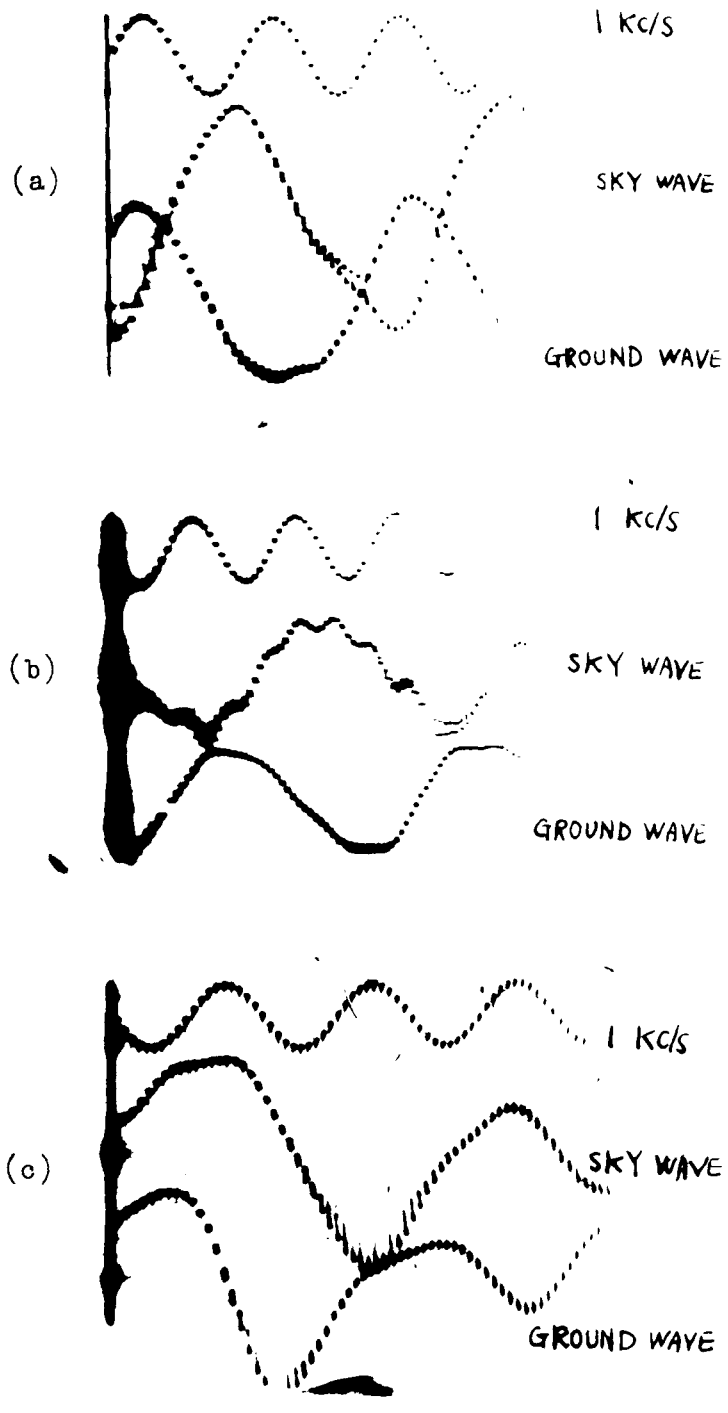


Fig. 3.8 Examles of records in the direct method.

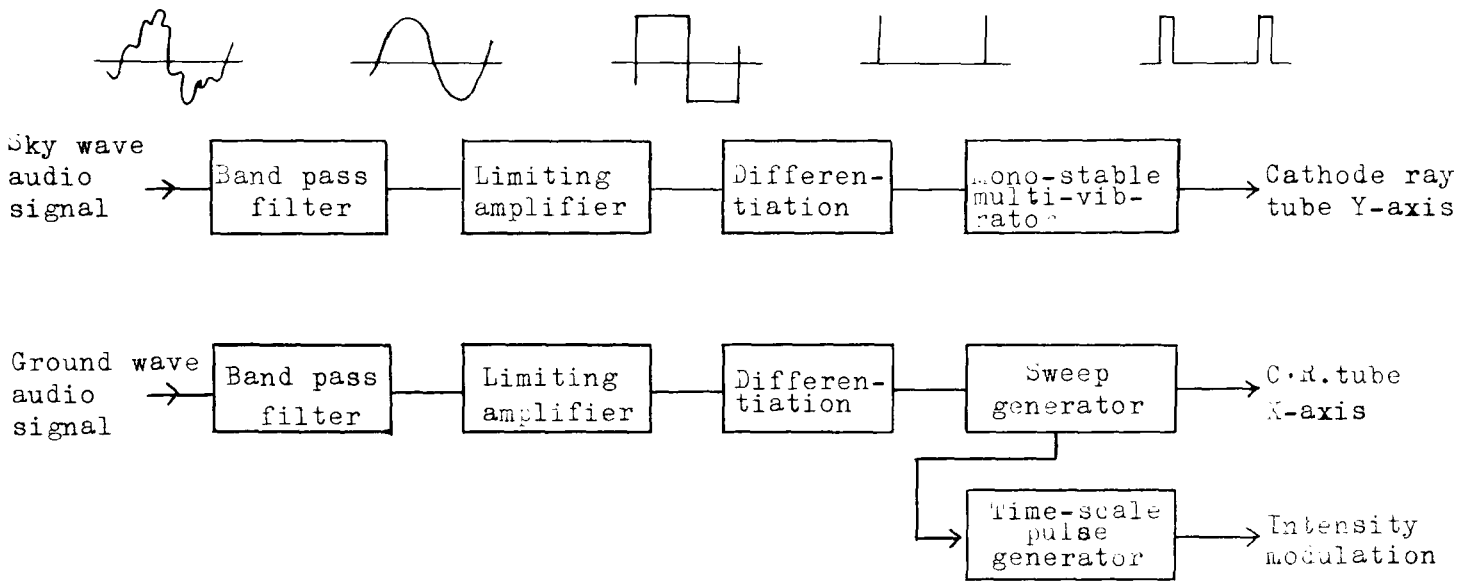


Fig. 3.9 Pseudo-impulse method of phase comparison.

not always unitary determinate, but was selected from the following point of view:

1. The lower the frequency is, up to the higher layer can be measured.
2. The energy spectrum of Japanese vocal sound peaks near 500c/s ~ 600c/s (Miura, 1955).
3. The band width should be narrow for the betterment of the wave form, but a narrow pass band will bitterly reduce the energy of the signal.

The filter used has a characteristics shown in Fig. 3.10.

The filtered signals are then amplified through double ended limiters to be shaped rectangular. In this case the phase of the original quasi-sinusoidal wave must be preserved. The limiting amplifiers are designed so as to satisfy the requirement. Namely as shown in Fig. 3.11 or by $V_3 \sim V_{11}$ in Fig. 3.12, in each stage of the limiting amplifiers the grid vlotage is clamped at ± 1.5V by two cells (Millman and Taub, 1956). Consequently a signal, the amplitude of which is smaller than 1.5V is amplified linearly and symmetrically in positive and negative sides, and larger than 1.5V will be shaped rectangular.

Now if there is some non-linearity in amplification, it results in the asymmetry of the wave form with respect to the zero line even for a sinusoidal input. Here we suppose the amount of the asymmetry being about 4% for one stage of amplification(it is a possible value for the tube 12AT7). Then because of the asymmetry and the level setting action of a coupling condenser C_1 in Fig. 3.11, the phase deviation of the zero crossing instant of a sinusoidal wave will be

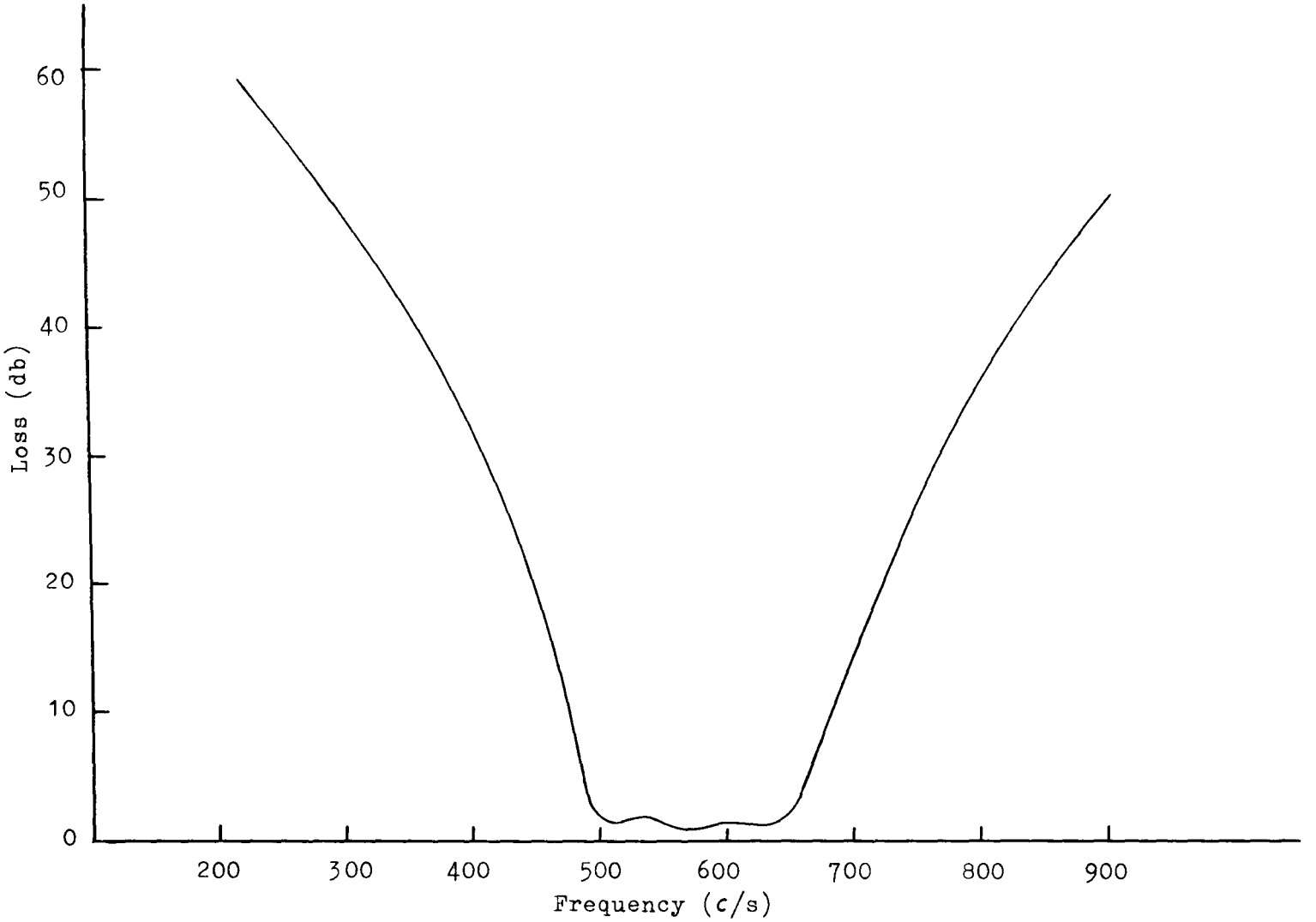


Fig. 3.10 Attenuation characteristics of the band pass filter.

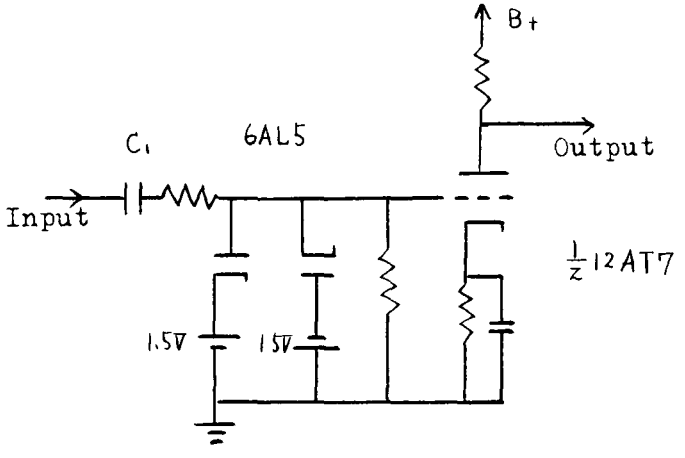


Fig. 3.11 Double-ended limiting amplifier.

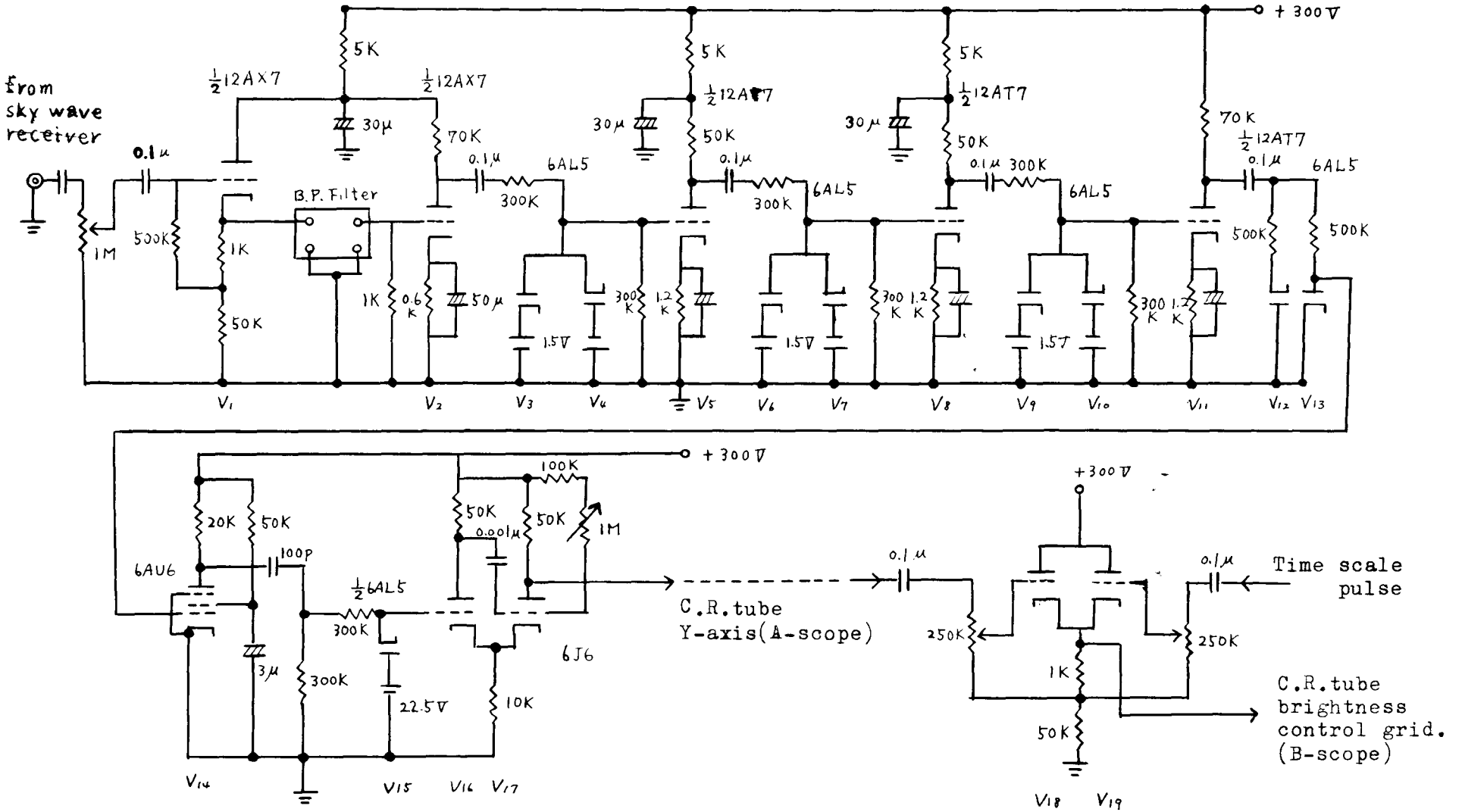


Fig. 5.12 Total circiuits of the sky wave impulse generator.

approximately 0.02 radian or 1.15°. It corresponds 0.32% of phase deviation. By three stages of such a limiting amplifier, the error which may arise from the transformation of the wave forms, does not amount even to 1%.

After being shaped rectangular, the signal is differentiated and varied to impulse series which occur at the zero level crossing (from negative to positive) instants of the rectangular wave.

As shown in Fig. 3.12, the differentiation is performed by 100pF and 300KΩ connected with the anode of V_{14}.

V_{15} is used as a clipper in order to prevent noise impulses from triggering the mono-stable multivibrator.

In the ground wave channel, these impulses are regarded as the direct wave impulses in the so-called impulse method of ionosphere sounder, and are used to trigger a sweep generator and time scale pulse generator.

The sweep generator was made so that the sweep speed and length are both variable (see Fig. 3.13). This sweep is applied to horizontal deflection plates of a cathode ray tube (ordinarily one sweep is about 2ms.).

The time scale pulse generator is a kind of pulsed Hartrey oscillator whose frequency is made 10Kc/s, and a Schmidt circuit. These time scale pulses (every 100μs) are applied to the brightness control grid of the cathod ray tube, so that the horizontal sweep is broken in every 100μs.

In the sky wave channel, the impulse series produced with an audio signal by the limiting amplifier, are used to trigger a mono-stable multivibrator which generates rectangular pulses of about

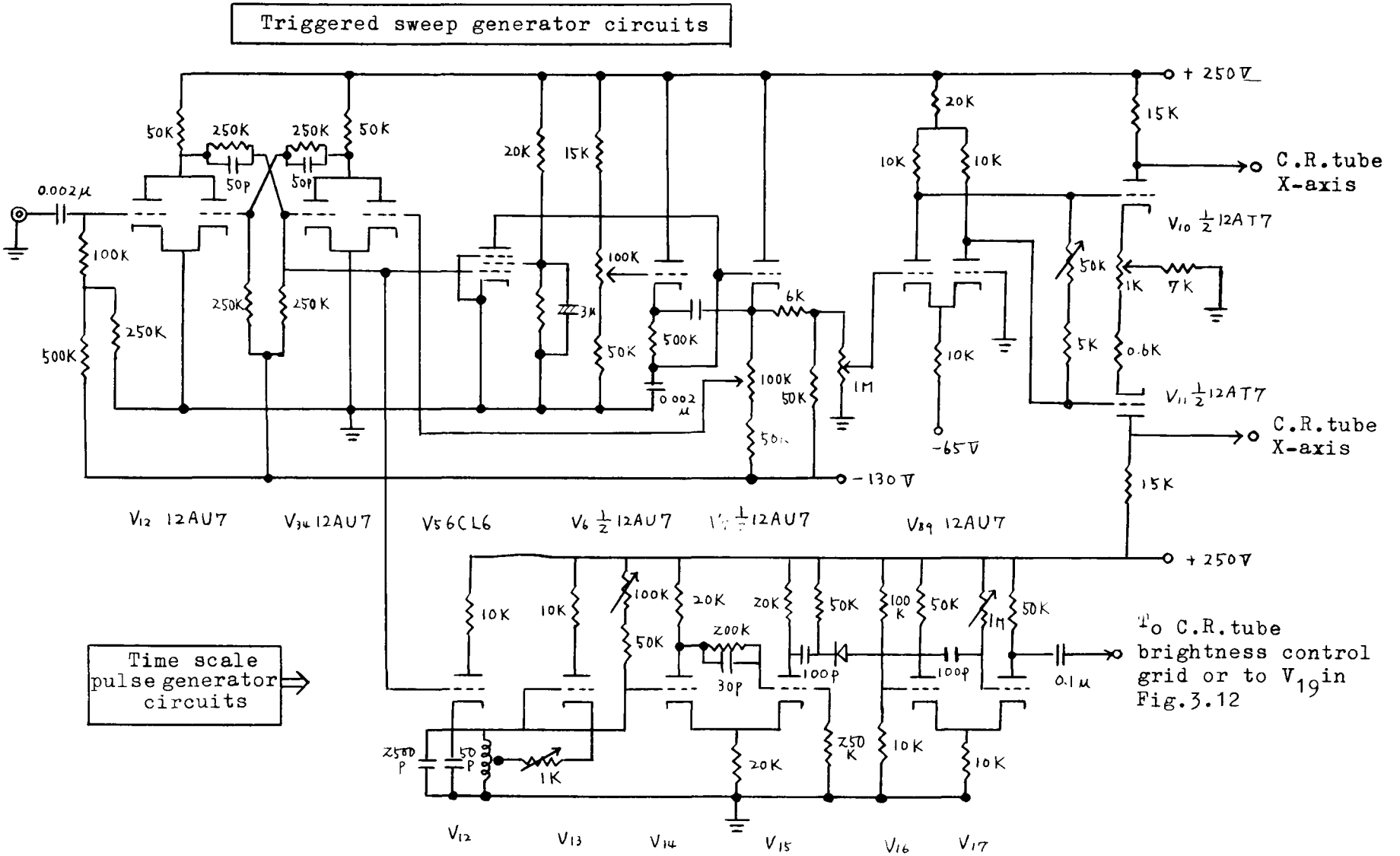


Fig. 3.13 Triggered sweep generator and time scale pulse generator.

100μs pulse width. These pulses are regarded as ionospheric echoes in the so-called impulse ionosphere sounding. Therefore these are applied to the vertical axis of the C.R.T. (it is here called A scope), or to the brightness control grid together with time scale pulses through a cathod follower for mixing (it is here called B scope).

To record these views on the C.R.T., a continuous photographic recorder was used, where photographic film runs in vertical direction with the speed of 2 & 4cm/min. In this case, for A scope, except the bright base line the fluorescent surface of the C.R.T. is masked and so the sky wave impulse appears as non bright part on the base line. This image is as shown in Fig. 3.14.

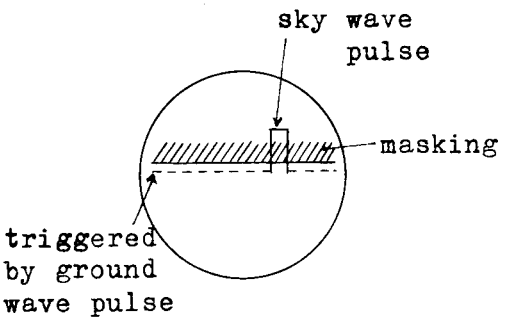


Fig. 3.14 Front view of the C.R.T.

The B scope may be seen as the negative of the A scope. But the former has a merit because it has an effect of an automatic noise limiter, because the noise whose amplitude is smaller than a certain limit does not appear on the film, while in the A scope, even small noise disturbs the image.

Chapter 3.4 Observation

3.4.1 Selection of the broadcast waves for the observation

Broadcast waves to be utilized for our experiment must satisfy the condition that at the receiving point in Kyoto University the ground wave should be much larger than the sky wave as explained in chapter 3.2. In consequence the broadcast station must be selected out of those located near the receiving point. On the other hand, most of the transmitting antennae now working are the anti-fading vertical antennae, so that their upward radiations are very weak as compared with their total radiations.

Then the distance between the transmission and reception must not be so proximate. Taking account of such points of view, transmission power, and frequency, four stations which are all located at Osaka and distant by approximately 55Km from the receiving point (Kyoto University) are selected. In Table 3.1 some remarks of the four stations are listed.

Table 3.1

Station	Frequency f	$f_\perp$ (*)	Power	Effective power for sky wave(*)	Distance
JOBK	670Kc/s	646Kc/s	100KW	390W	55Km
JOBB	830	800	100	300	55
JONR	1010	975	10	1.3	55
JOOR	1210	1168	10	114	55

(*) value corresponding to the apparent height of reflection, 100Km.

$f_\perp$ in this table indicates $f \cos i_0$, where i_0 is the incident angle to the ionosphere (see Fig. 3.3). However as i_0 is dependent upon

the apparent height h' for a given distance D, in the table i_o is assumed to be corresponding to h' of 100Km. For another h', i_o is depicted in Fig. 3.15.

The effective power for sky wave implies the radiating power in the direction making angle i_o to the vertical, which was calculated from the theoretical radiation characteristics of the antenna (*). In Fig. 3.15 this effective power for sky wave is also shown against the apparent height h'.

It was quite convenient for the analysis of data that the selected four stations were almost equally distant, 55Km, from the receiving point. For D=55Km, the time lag τ (ms) versus h' is as shown in Fig. 3.16.

3.4.2 Observed results

a) Direct method

From January to July in 1958, the direct method explained in 3.3.3 a) had been used for observation. From the films such as those demonstrated in Fig. 3.8, the time lag τ is measured by the time scale of 1Kc/s standard wave, then is transduced to the apparent height h' by using (3.10) or Fig. 3.16.

The apparent height of reflection for each broadcast wave is plotted in local time. Points of such plots generally disperse over a wide range in height, but if we draw an envelop of the plots, we will have a smooth line. Such a process is illustrated in Fig. 3.17. These envelops for the observed waves are collected in Fig. 3.18, 19, and 20.

The above process stems from the fact that the maximum time lag

(*) These data were given by courtesy of Kinki Radio Supervisory.

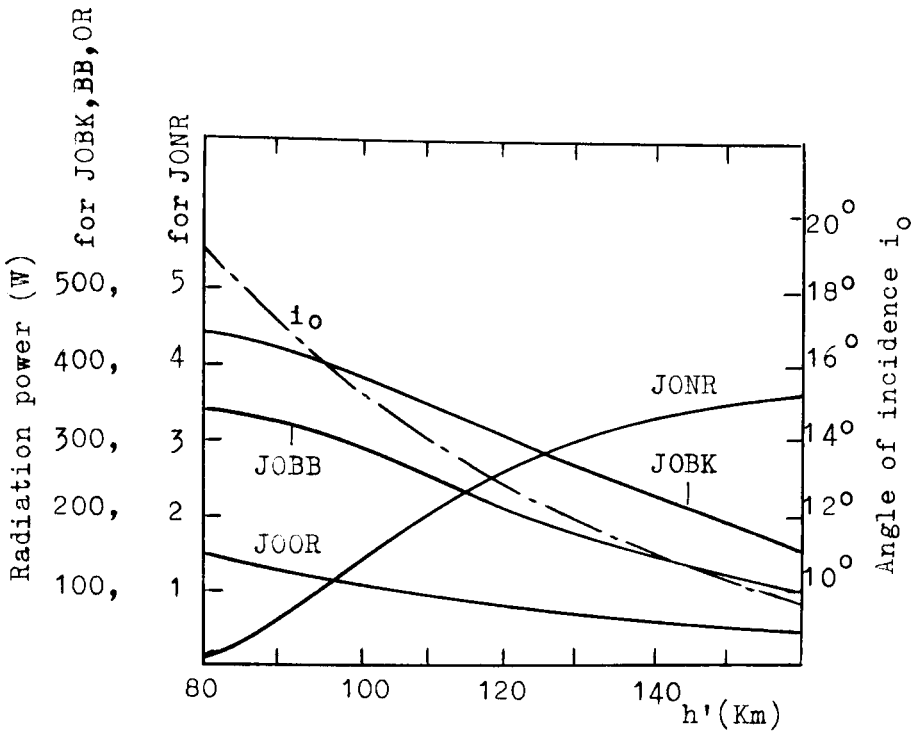


Fig. 3.15 Radiation power and angle of incidence as a function of the apparent height h'.

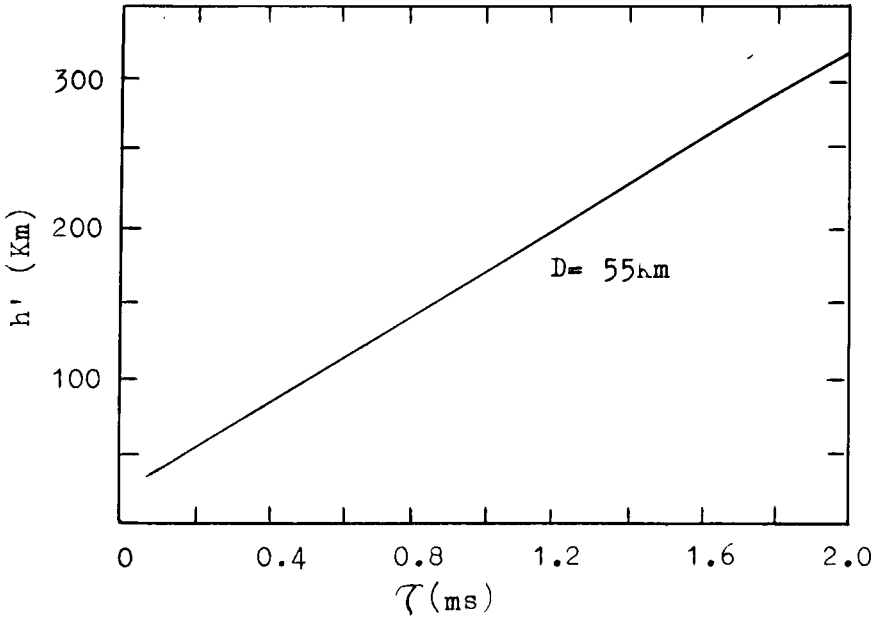


Fig. 3.16 Apparent height h' as a function of the time delay τ.

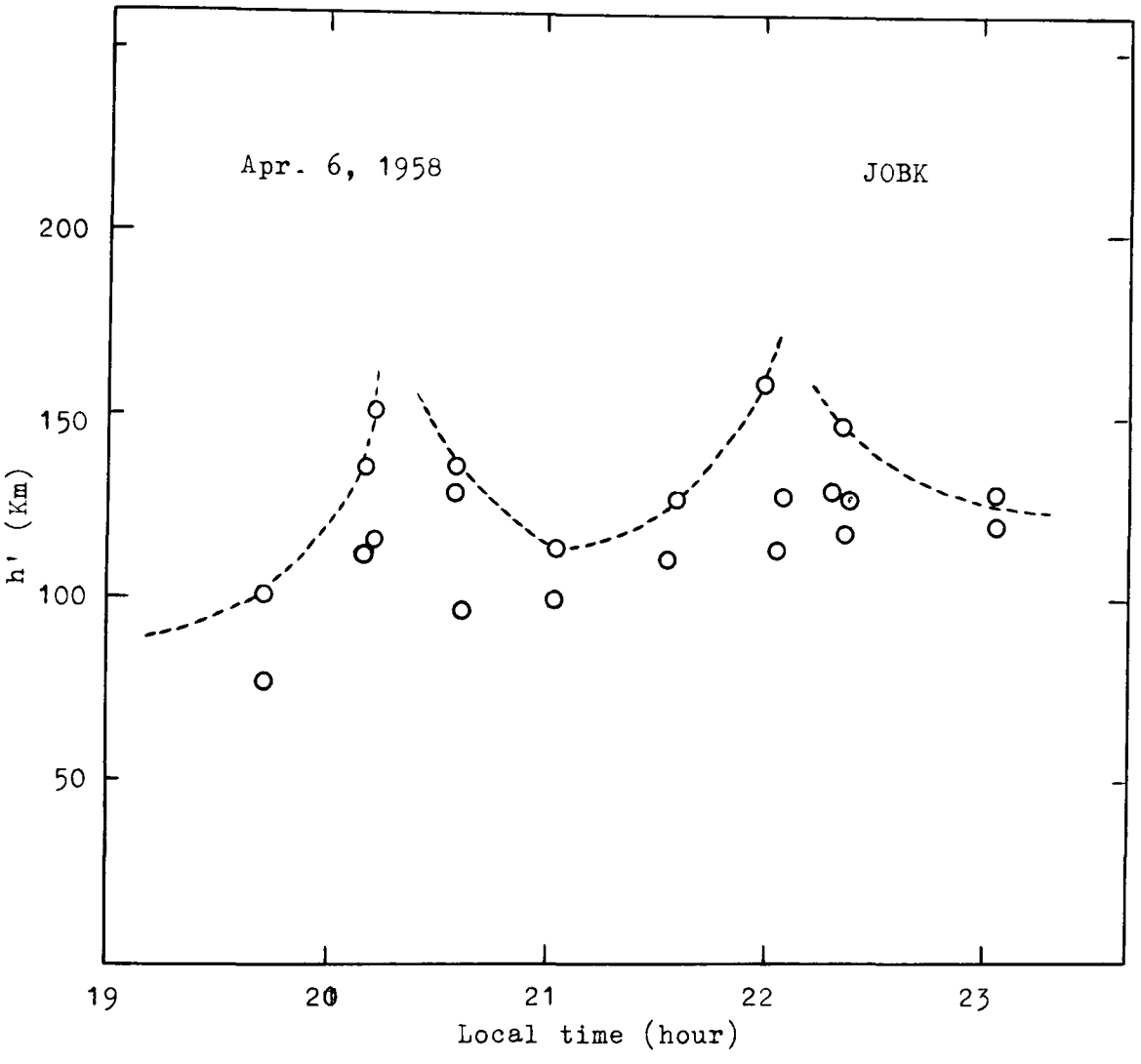


Fig. 3.17 Apparent height h' observed by the direct method. The envelope (dotted line) implies the variation of plausible h'.

may be the nearest value to the correct one, if the time lag is measured by the deviation of the peak position of the wave forms as illustrated in Fig 3.21, as will be explained in 3.5.2.

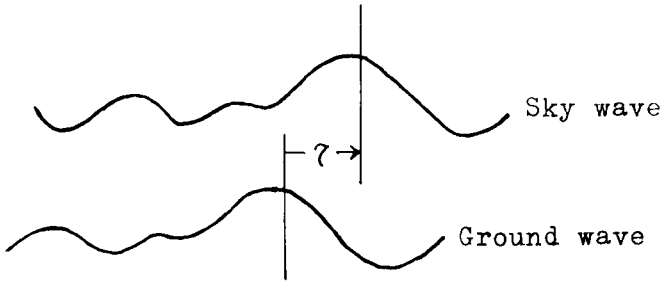


Fig. 3.21 Time lag measured by the deviation of local peak on the wave forms

Fig. 3.18 shows that each curve has one peak in h', which may imply that at this time the wave(frequency f) penetrates the E layer. In other words, the maximum critical frequency of the layer just coincides $f_L(=f\cos i_0)$at this time. On the right hand side of the peak, the curve goes down. This may be understood in such a way as the wave is reflected from the other upper layer and the retardation arising from the E layer is decreasing with time as the electron density of the E layer is falling off. The upper layer might be either another thick layer or a kind of sporadic E layer.

In Fig. 3.19 and 20, there appear two peaks in one frequency curve. The left (or earlier) peaks around 20h on April 20 might be the retardation at the time of penetration of the D layer and the right peaks around 21h might correspond to the penetration of the E layer. On April 6, the time of penetration of the same layer for two stations seems to be deviated a little bit as compared with that on April 20. Such a case might correspond to the decrease of the electron density being slow.

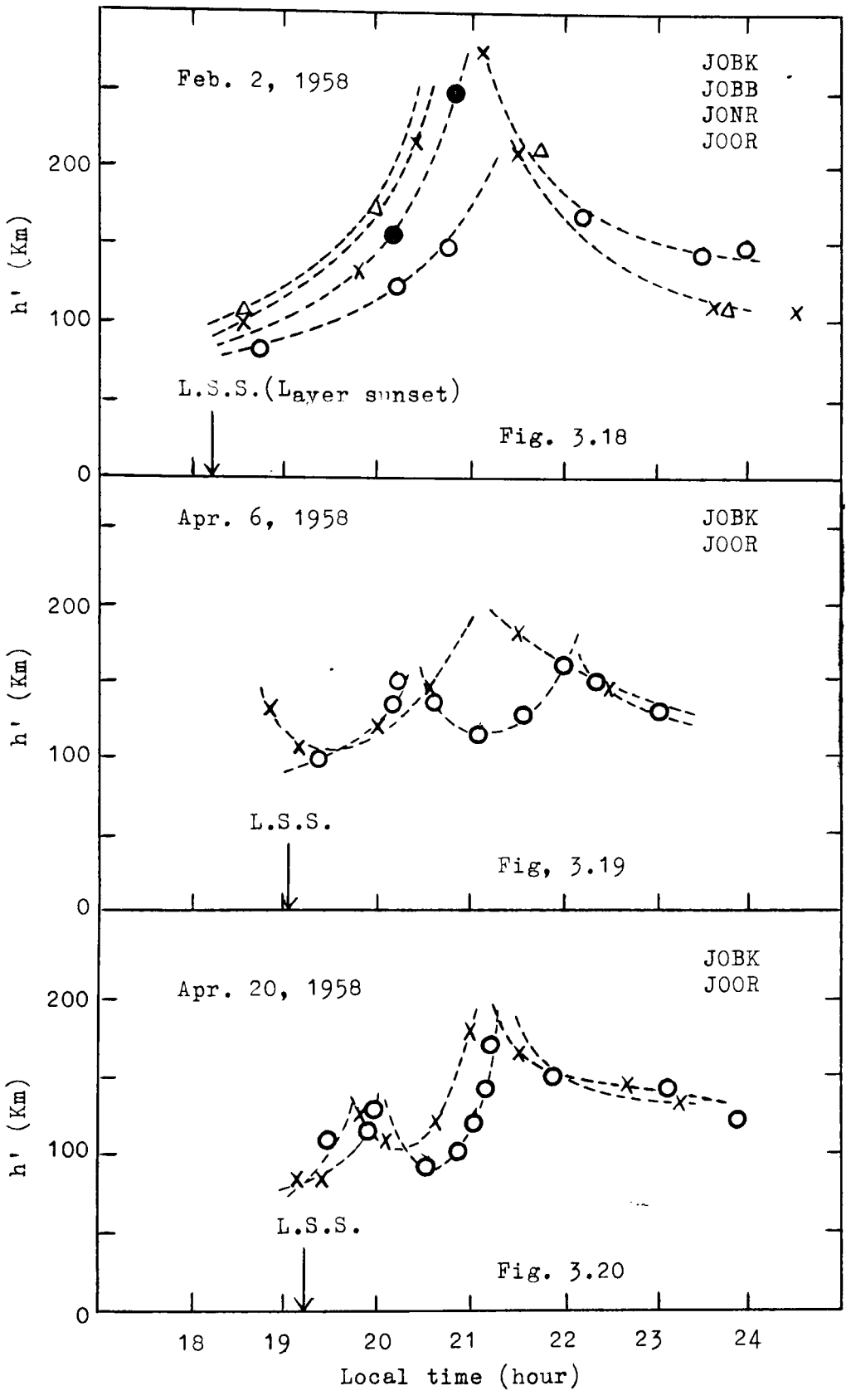


Fig. 3.18,19,20 Hourly variation of the apparent height h' for JOBK(o), JOBB(•), JONR(×) and JOOR(Δ).

From these hourly variation curves, the variation of the critical frequency of the D and E layers can be inferred.

Now we suppose the f_1 corresponding to the frequency f of a broadcast wave is approximately constant irrespective of the apparent height h'. Then, it can be concluded that at the instant of such a peak in a curve of frequency f, the critical frequency of the E or D layer is identical with the f_1 shown in Table 3.1. Discrimination of the D and E layers is self-evident, because the critical frequencies of each layer are considerably different.

From the above supposition, we can finally draw curves of the hourly variation of f_0E, as shown in Fig. 3.22. In this figure, the data of f_0E higher than 2Mc/s were borrowed from those of the impulse sounding made on the days at Kyoto University. The sunset at the altitude 100Km is taken as the origin of the time scale, which is about 40 minutes delayed from the ground sunset. Black circles and broken line indicate an anual average of the data at Watheroo in Australia (after Mitra, 1957) which can be compared with ours since Watheroo is located at a middle latitude and our data were obtained around the vernal equinox.

At a glance, the tendency of decrease of f_0E around the layer sunset is quite similar to that of Watheroo, while the minimum f_0E seems to be much less than that of Watheroo.

The chain lines in the same figure are those which are regarded as corresponding to the D layer. It is because those echoes appeared in April when the solar radiation arriving at the layer became intense, while they did not appear in February and generally in winter season.

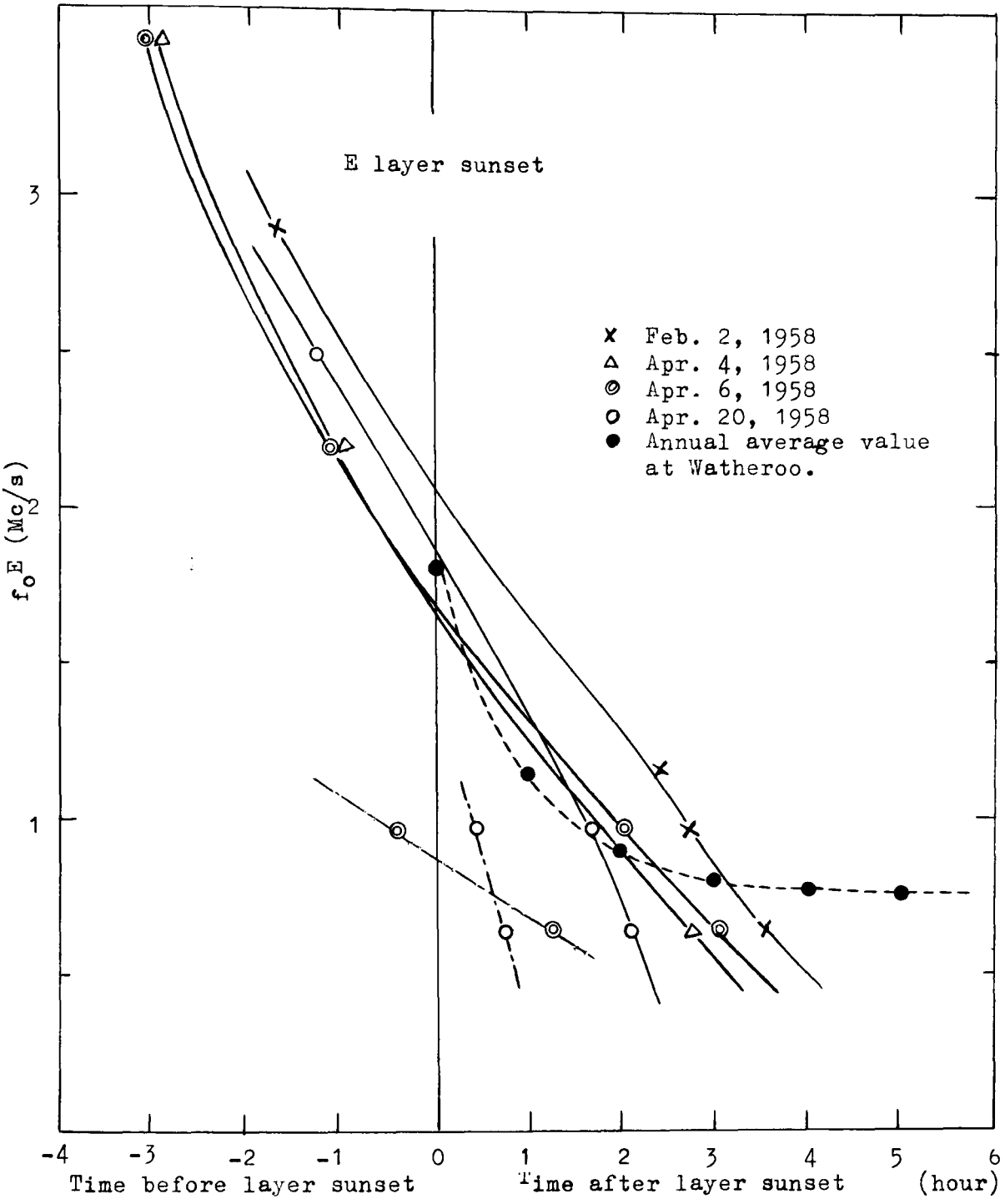


Fig. 3.22 Hourly variation of f_0E. The points which lie above 1.5 Mc/s line are corresponding to the data obtained by the impulse method.

Until now we have paid attention to only one mode of propagation, the ordinary mode. Actually in the ionosphere there exists another mode; the extraordinary mode. The latter mode of a frequency smaller than the gyro-frequency, generally not only suffers greater absorption, but also is reflected at higher altitude than the ordianry mode wave does. As the result, most of the echoes can be seen as those of the ordinary mode.

In general if we denote by f_o and f_x, the critical frequency of the ordinary and extraordinary mode respectively with respect to the same layer, we have a relation,

$$f_o^2 = f_x(f_x + f_H) \qquad (3.12)$$

in the range, $f_x < f_H$, where f_H is the gyro-frequency. It says that the penetration of the layer by the extraordinary wave occurs earlier in time than that by the ordinary wave, therefore the chain lines in Fig. 3.22 might be regarded as corresponding to the extraordinary mode. However it may not be so because even if the chain lines are assumed to be those of f_x, f_o and f_x at the same time do not satisfy (3.12).

b) Pseudo impulse method

Since October 1958, the new method explained in 3.3.3(b) has been used. Some observed films are illustrated in Fig. 3.23(a)~(d). In case of (a) and (b) of the figures, the separation of the sky wave is satisfactory and it surely indicates the actual time lag of the sky wave from the ground wave. On the other hand, in case of (c) and (d), the separation is not sufficient, so that the apparent time lag is only constant over a fraction of second, but is unsettled within several seconds.

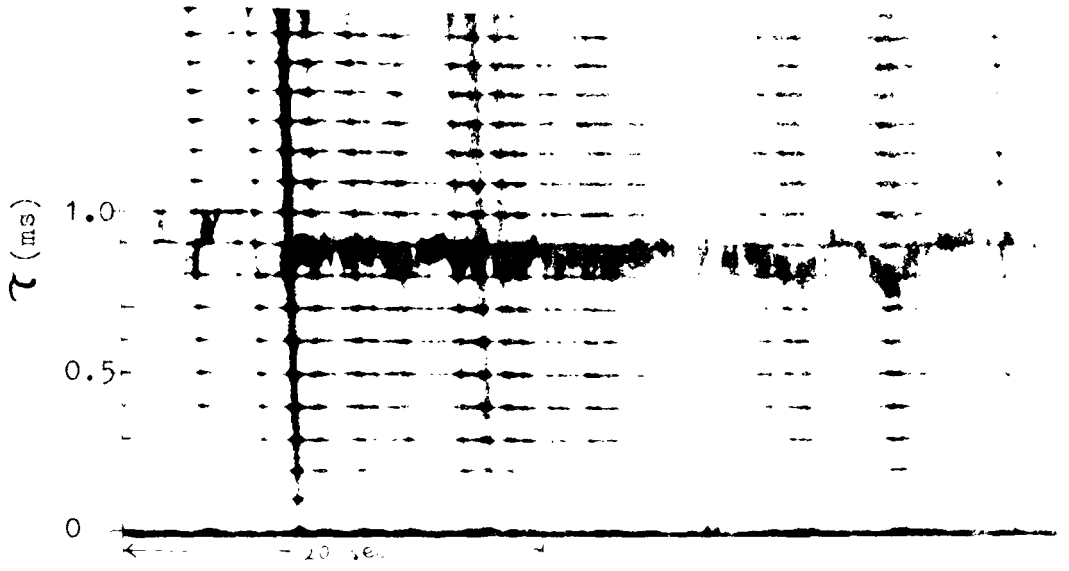
This phenomenon is explained in such a way that the quick variation of the carrier phase or amplitude of the sky wave due to a vertical shift of the reflecting layer or a fading results in the fluctuation of the apparent time lag, if some extent of the ground wave is left unerased in the sky wave channel.

Fig. 3.24 shows the relative time lag between the ground wave and sky wave channels, owing to the difference of characteristics of the two channels. It is the apparent time lag when a common signal is applied to the both channels. This amount of the relative time lag must be subtracted from the observed apparent time lag, when an actual ground wave and the corresponding sky wave are applied to the prescribed channels. It is apparent from Fig. 3.24 that in our experiment the relative time lag is almost constant, about 200μs, irrespective of frequency within the pass band of the filter.

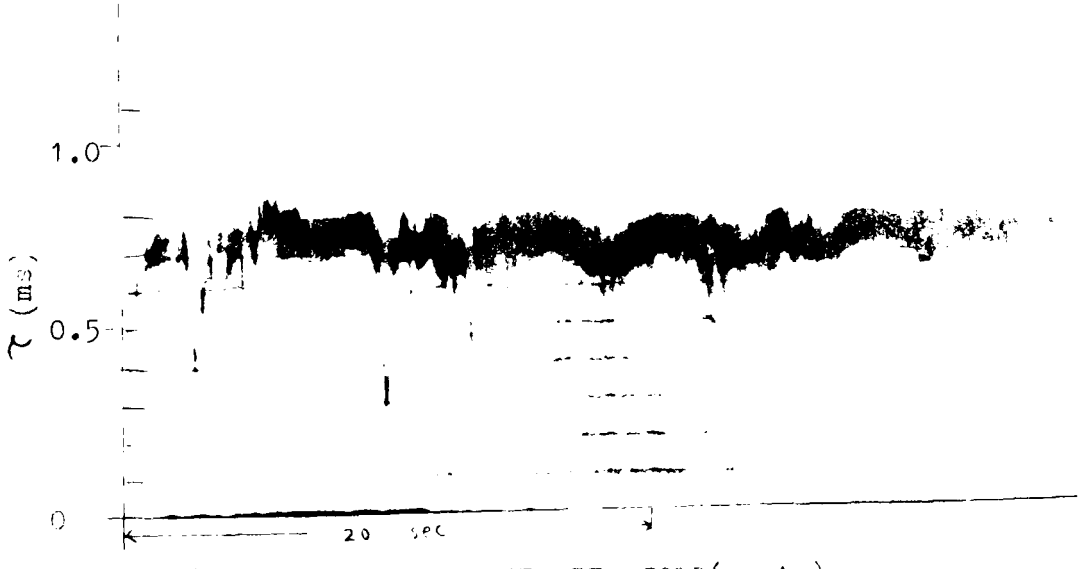
From such records as Fig. 3.23(a) and (b), hourly variation of the apparent height of reflection can be obtained for different stations. In fact, such a good record could not always be taken for all observing time. Most records were like Fig. 3.22(c) and (d), then we could not find the true apparent height.

To make good use of such records, a supplementary information was simultaneously recorded. Namely, the carrier level in the sky wave channel was recorded by a pen recorder as described at the end of 3.3.2. As will be explained in 3.5.3, with the aid of the level record we can often presume the actual time lag even from such a fluctuating record of time lag.

The observed results of the hourly variation of the apparent height in the evening of November to December, 1959 are shown in

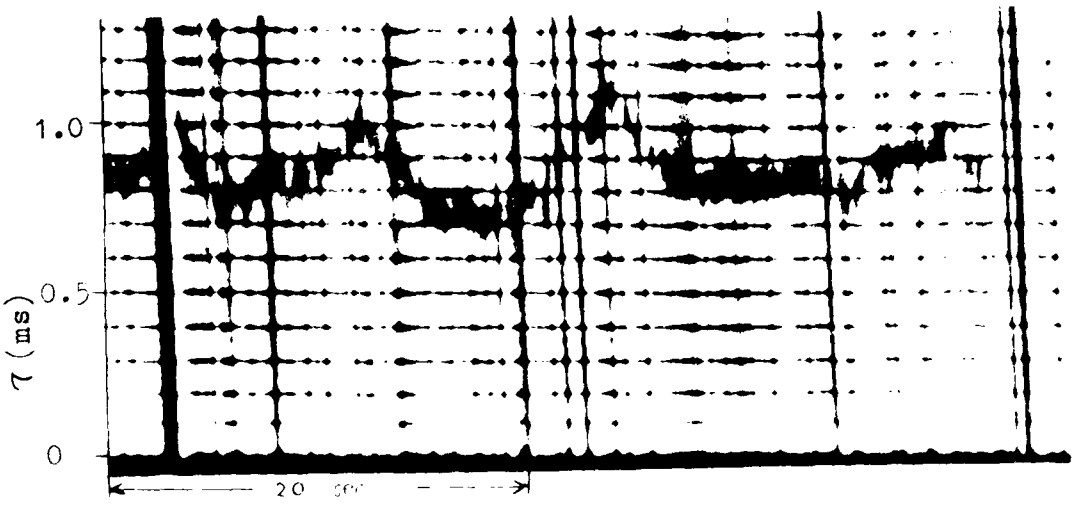


(a) 11 Jan. 1959 1920 JST JOBB(music)



(b) 15 Nov. 1959 2137 JST JOBB(music)

Fig. 3.23(a),(b) Records in the pseudo impulse method (sufficient separation).



(c) 11 Jan. 1959, 1850 JST JOBB(speech)

(d) 15 Nov. 1959, 2125 JST JOBB(music)

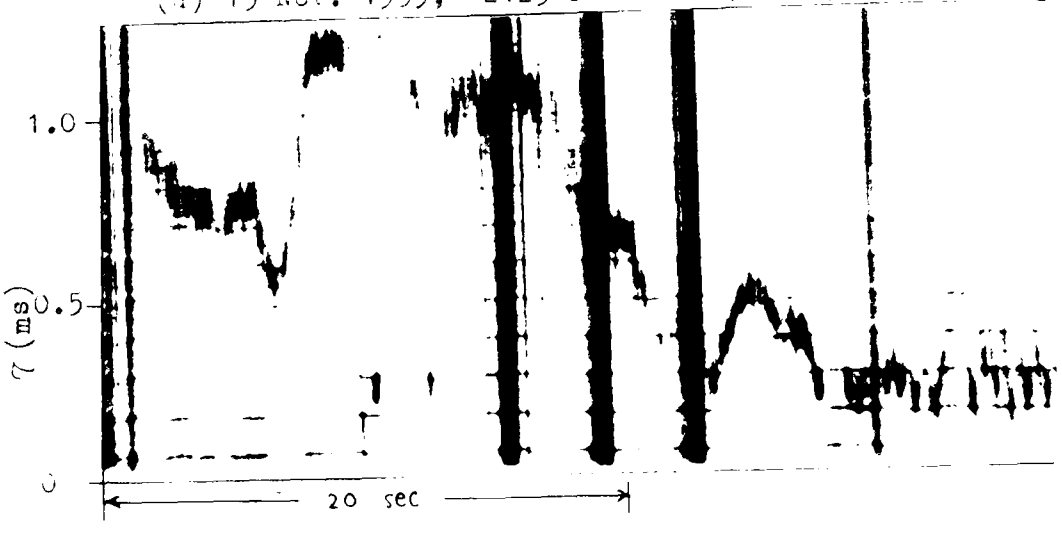
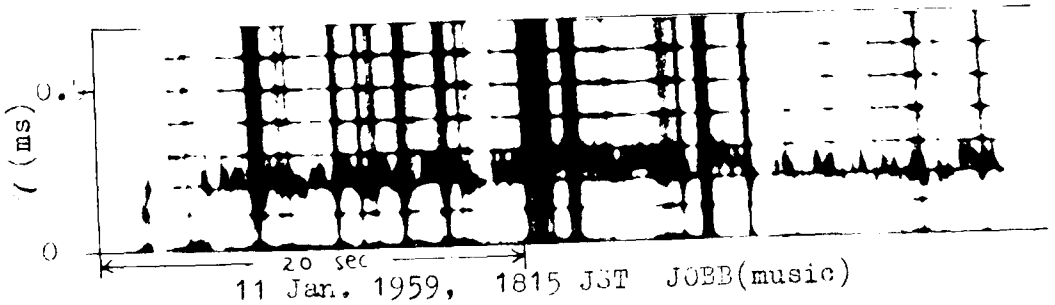


Fig. 3.23(c),(d) Records in the Pseudo impulse method (insufficient separation).



11 Jan. 1959, 1815 JST JOBB(music)

Fig. 3.24 Difference of the time delays through the sky wave and ground wave channels.

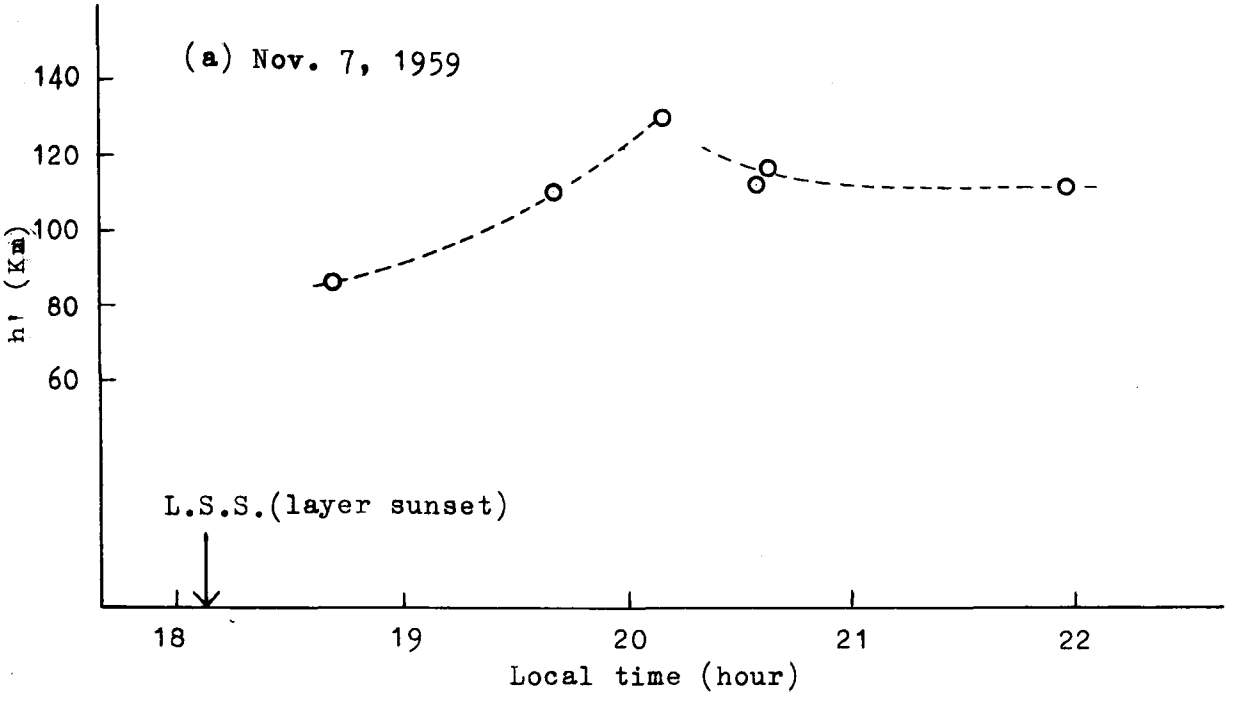
Fig. 3.25(a)~(h). According to these figures and the data observed at Kokubunji on the same days in the impulse method, we can have the following Fig.3.26 of the hourly variation of f_oE around the time of the layer sunset. It can be concluded that the critical frequency of the E layer decreases down to 650Kc/s or less around 20~21 h in local time about 2.5 hours after the layer sunset. In this case in terms of the layer sunset we imply the sunset at the E layer about 100Km above the ground.

Observations in the early morning were tried for four days around December 10, 1959. The hourly variations of the apparent height h' are shown in Fig. 3.27(a)~(d). It appears that the apparent height rapidly decreases just after or ahead of the layer sunrise. This phenomenon does not probably imply the retardation arising from a certain lower layer, but the real descending of the reflecting layer.

It is because the apparent heights of two broadcast waves different in frequency follow the same trace. There is another evidence of the rapid change of the layer height. Namely as seen in Fig. 3.28(a),(b), the apparent height of the lowest F region stratification, h'F indicates a similar hourly variation as those of the E layer around the layer sunrise.

The descending of the both layers can be interpreted as sunrise effect. Namely it is due to electron detachment from negative ions, taking account of the fact that the solar rays tangent to the ground strike the heigher region in the earlier time (Mitra 1952).

On the other hand we must notice a tendency of the layer ascending ahead of the descending. This phenomenon can not be explained by the above sunsise effect. Then it may lead to the conclusion that there



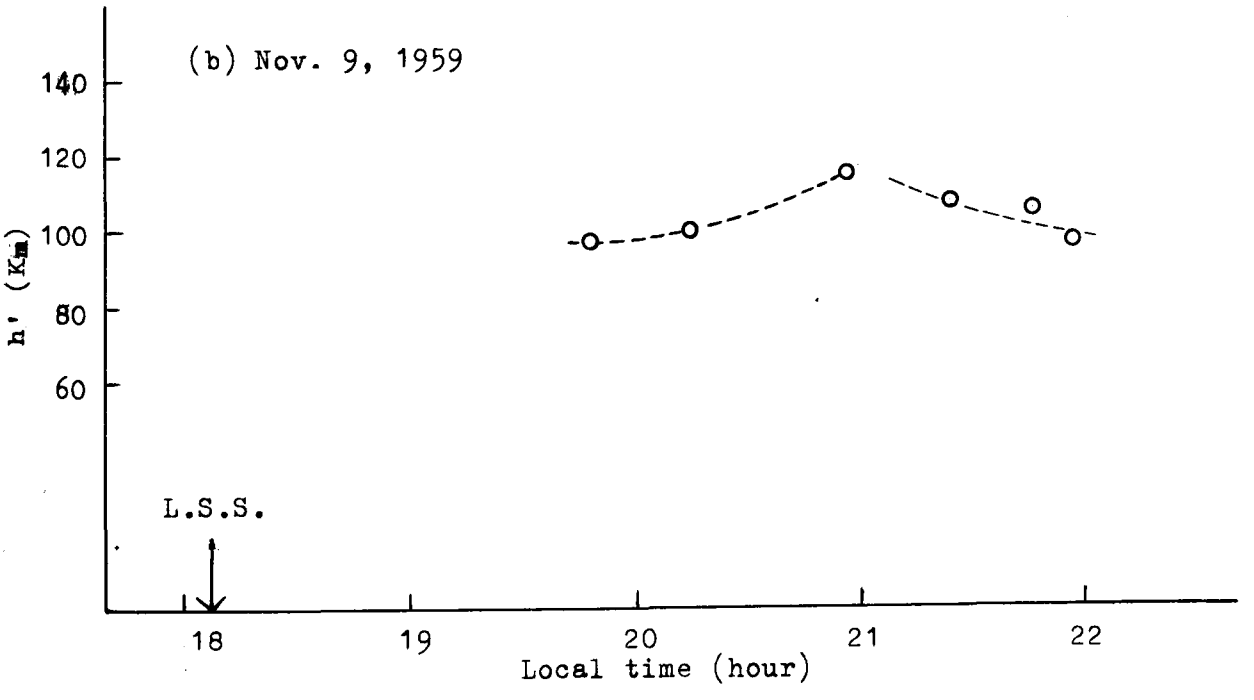
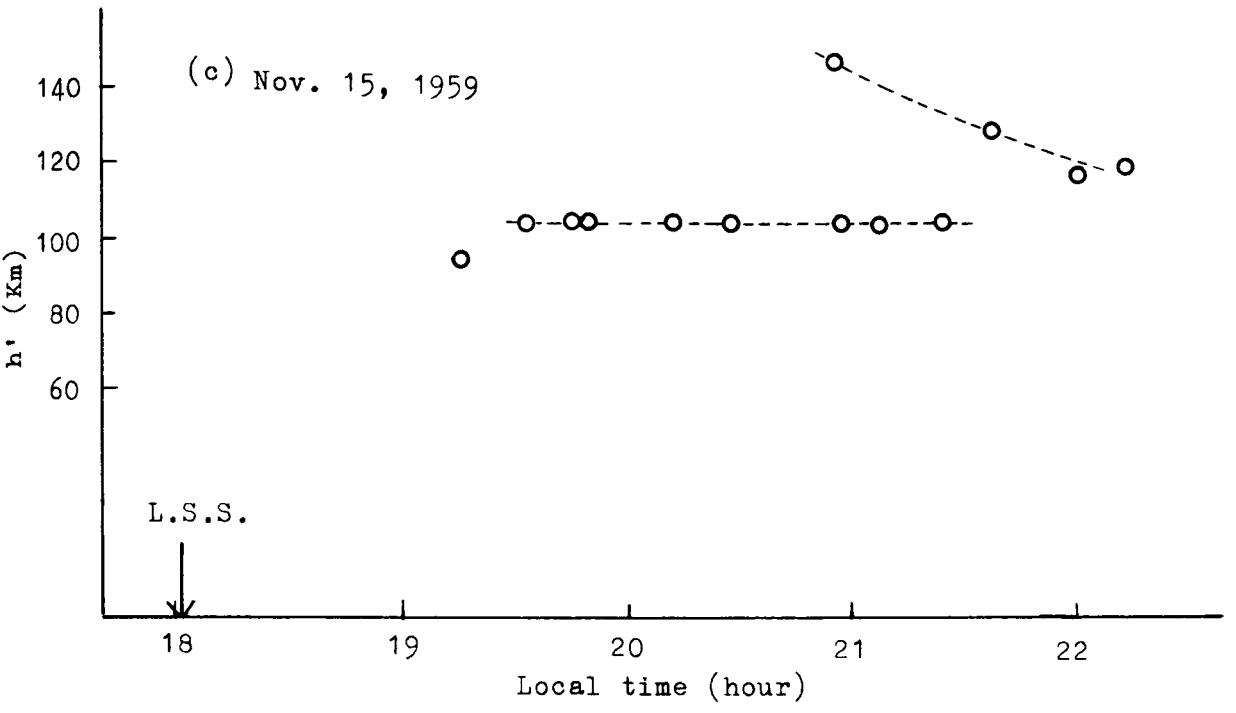


Fig. 3.25(a),(b) Hourly variation of the apparent height h' for JOBK.



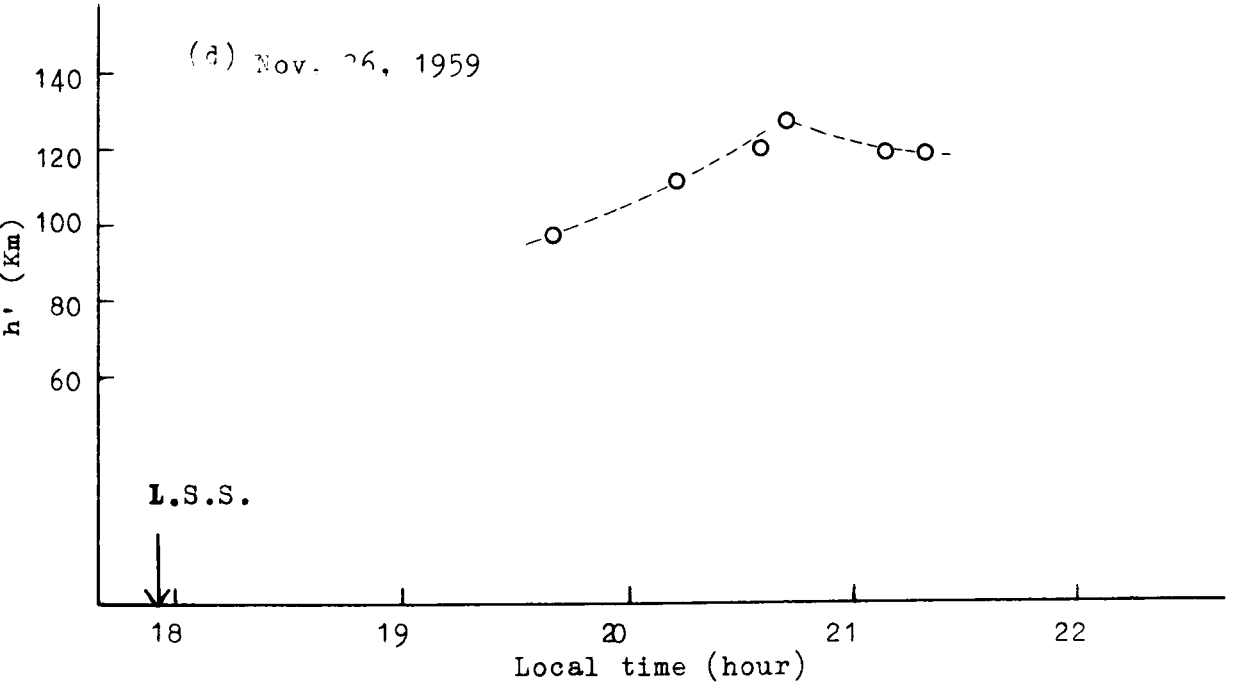
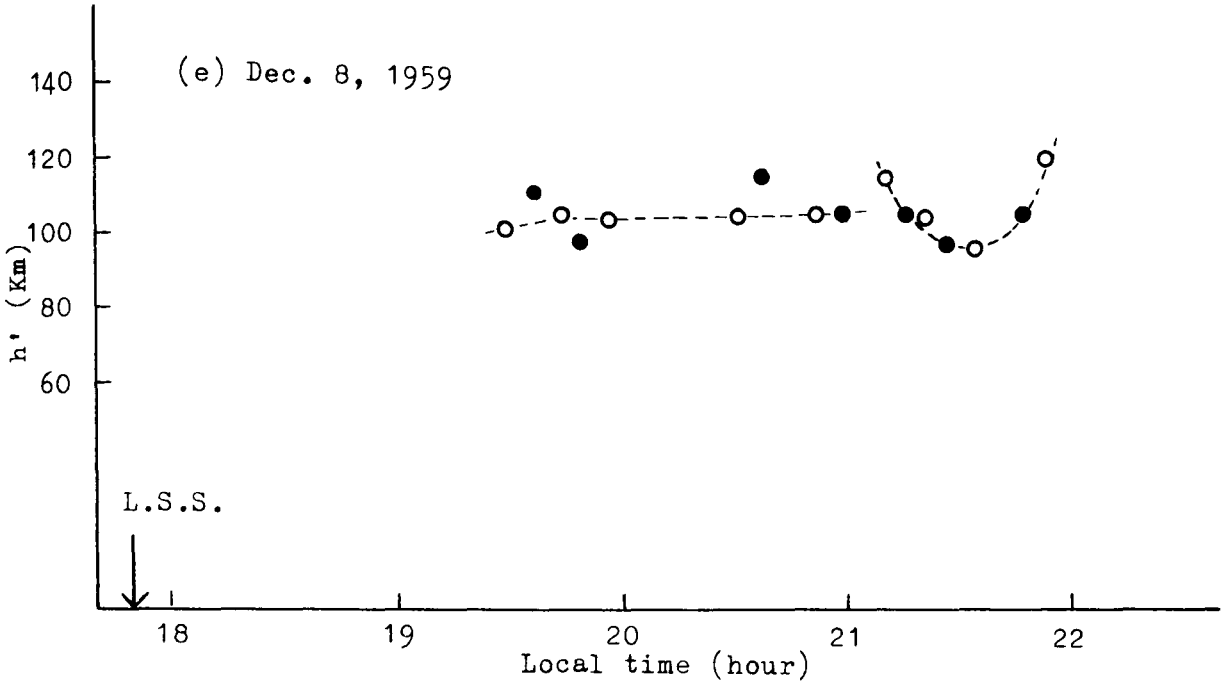


Fig. 3.25(c),(d) Hourly variation of apparent height h' for JOBK.



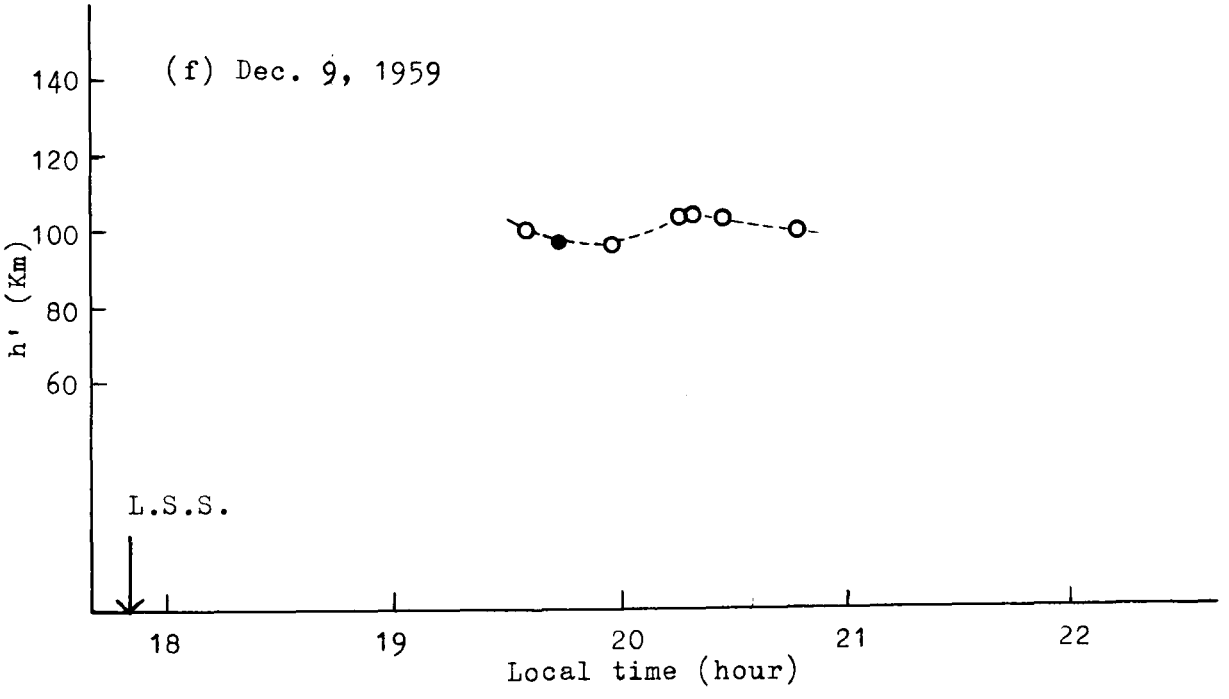
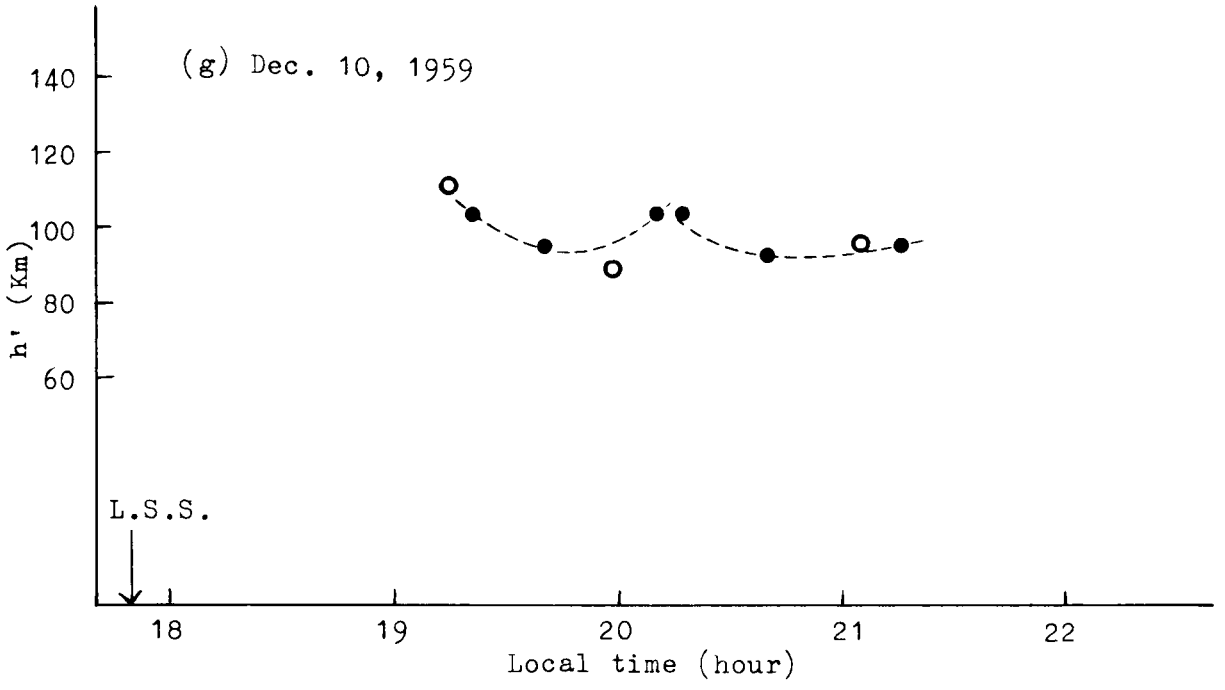


Fig. 3.25(e),(f) Hourly variation of apparent height h' for JOBK(o) and JOBB(●).



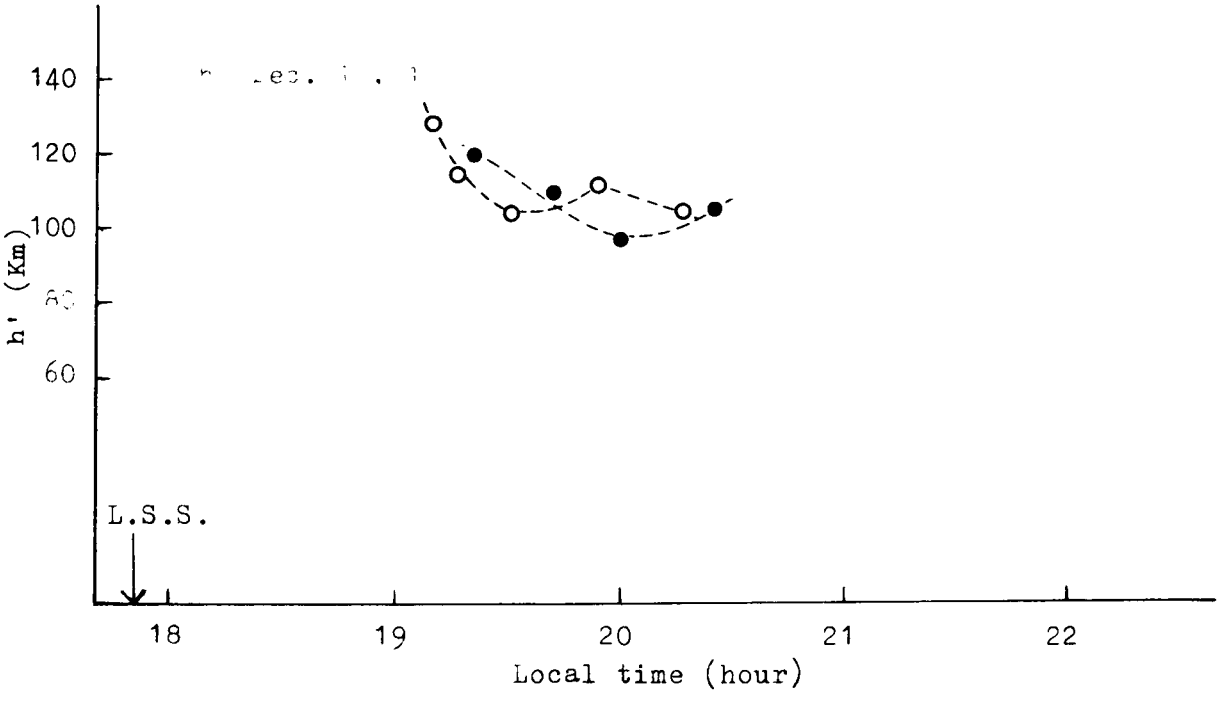


Fig. 3.25(g),(h) Hourly variation of apparent height h' for JOBK(o) and JOBB(●).

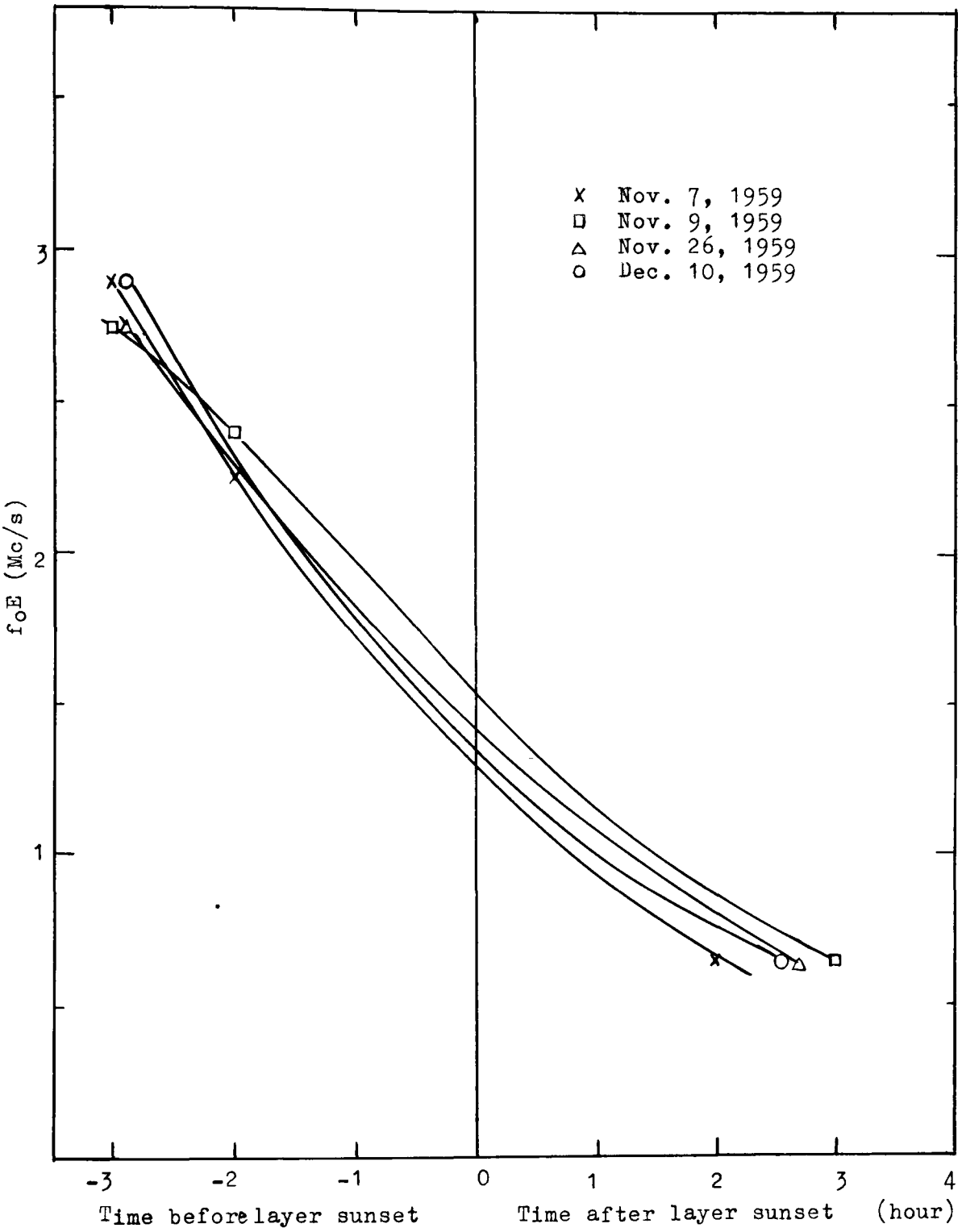


Fig. 3.26 Hourly variation of f_oE. The points which lie above 1.5 Mc/s line are corresponding to the data obtained by the impulse method.

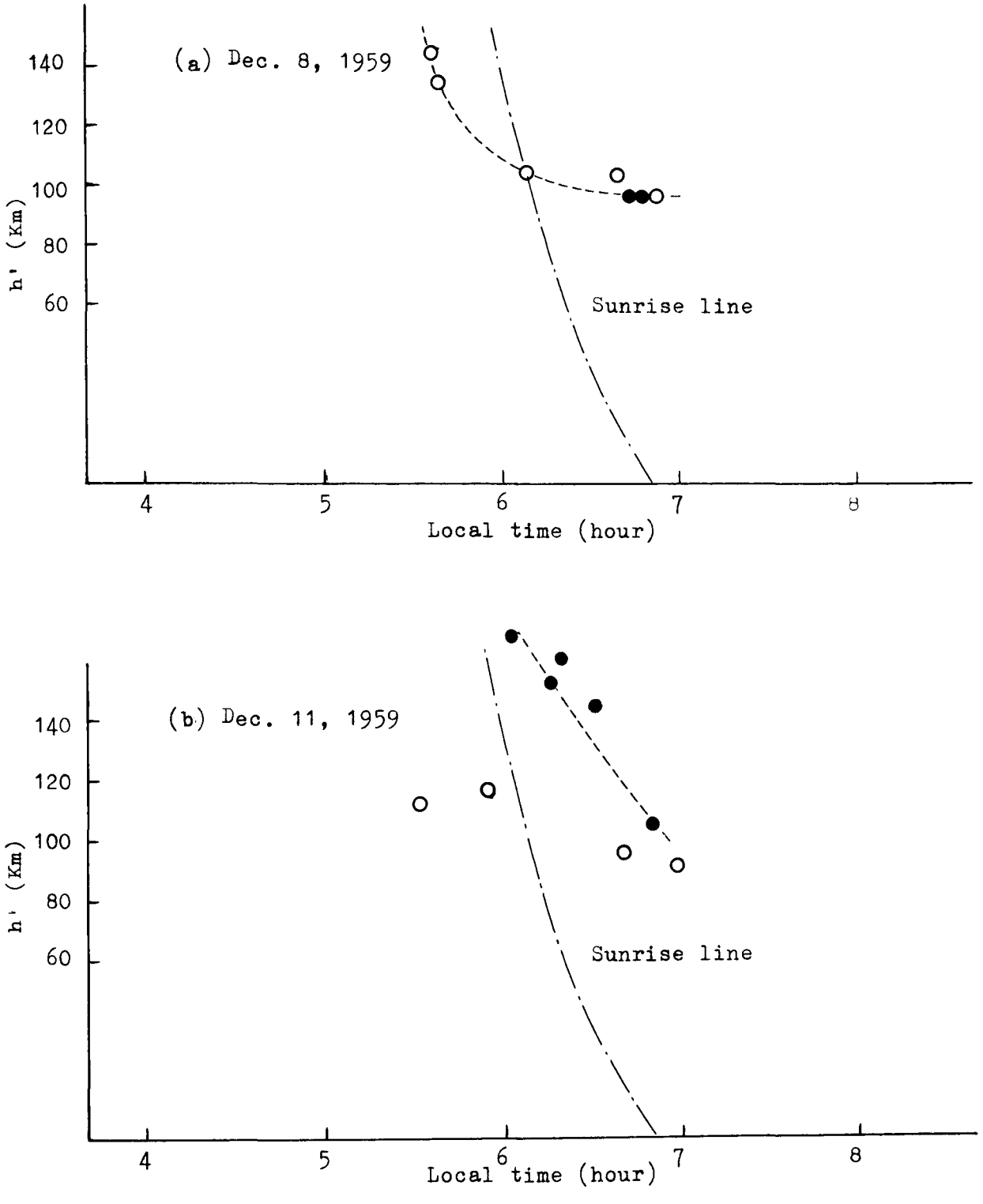
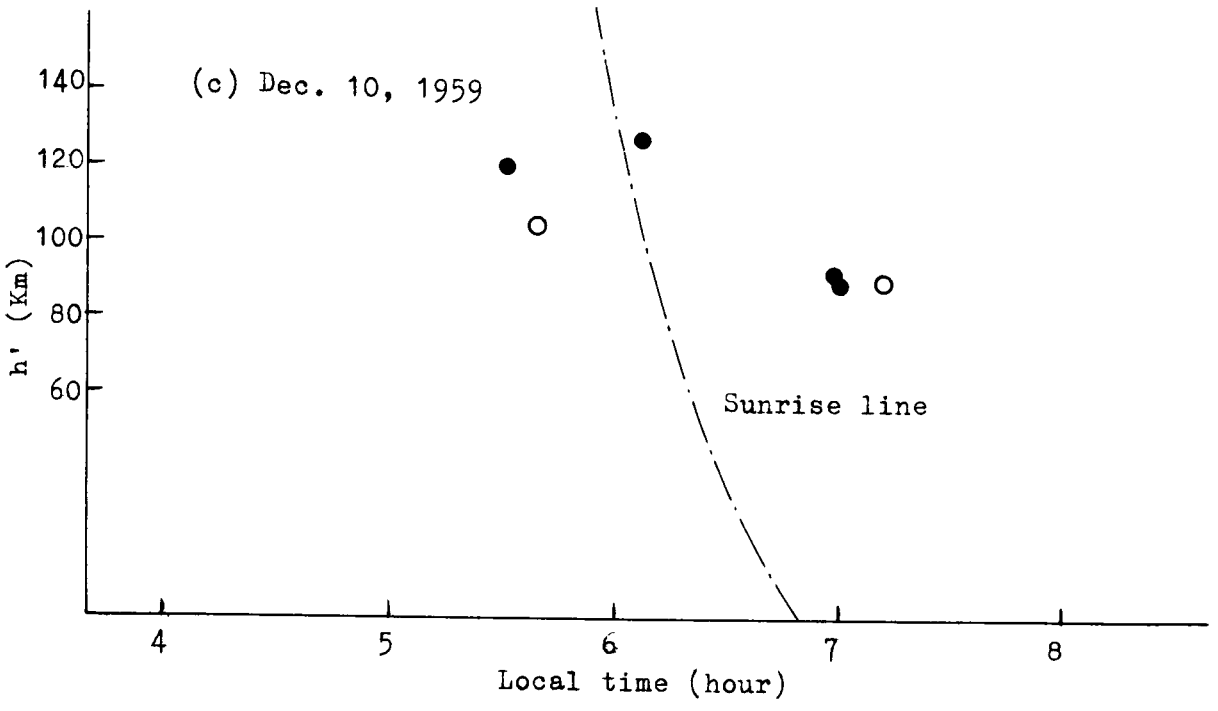


Fig. 3.27(a),(b) Hourly variation of apparent height h' for JOBK(o) and JOBB(●).



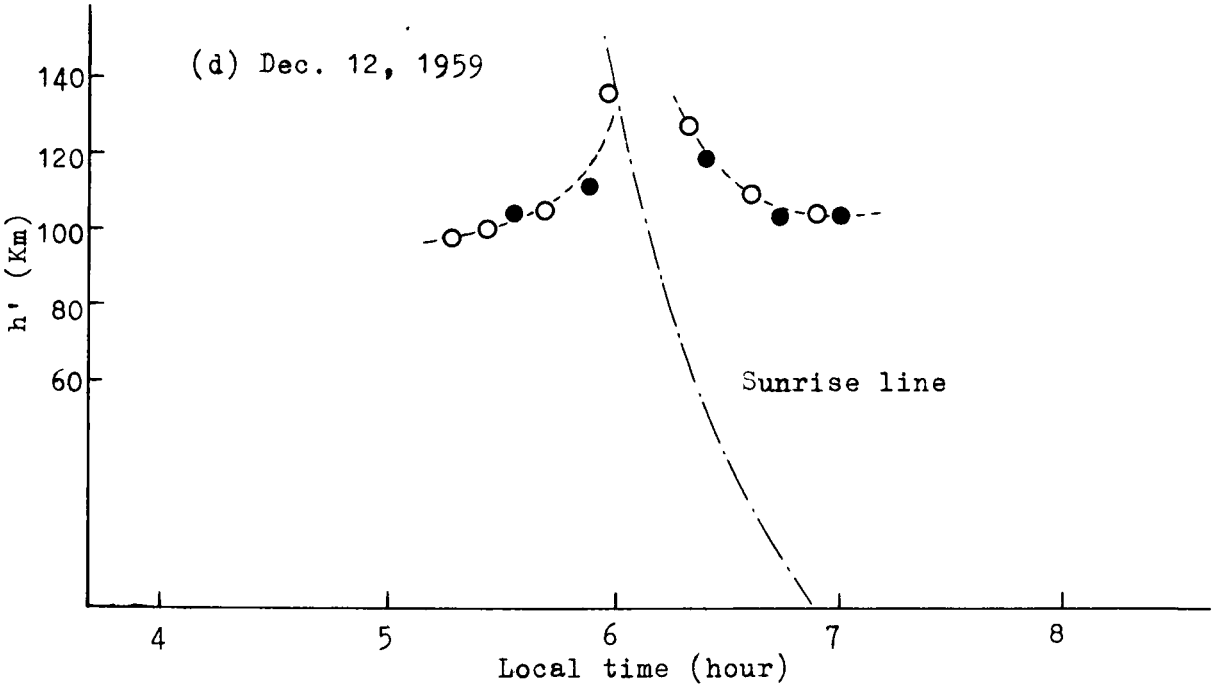


Fig. 3.27(c),(d) Hourly variation of apparent height h' for JOBK(o) and JOBB(●).

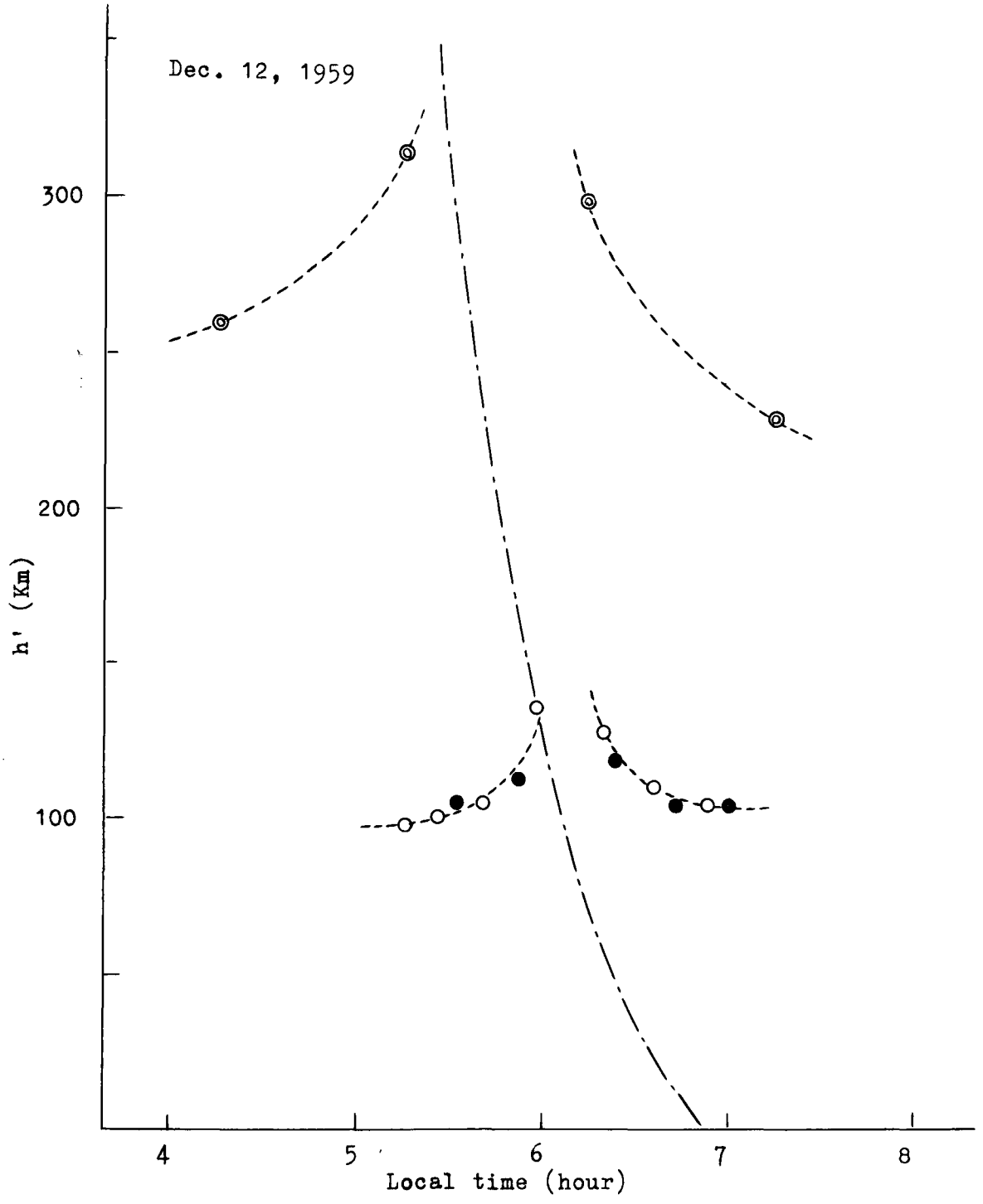


Fig. 3.28(a) Hourly variation of the apparent height h' for JOBK(o) and JOBB(●) in comparison with the variation of h'F at Kokubunji(◎) around the layer sunrise.

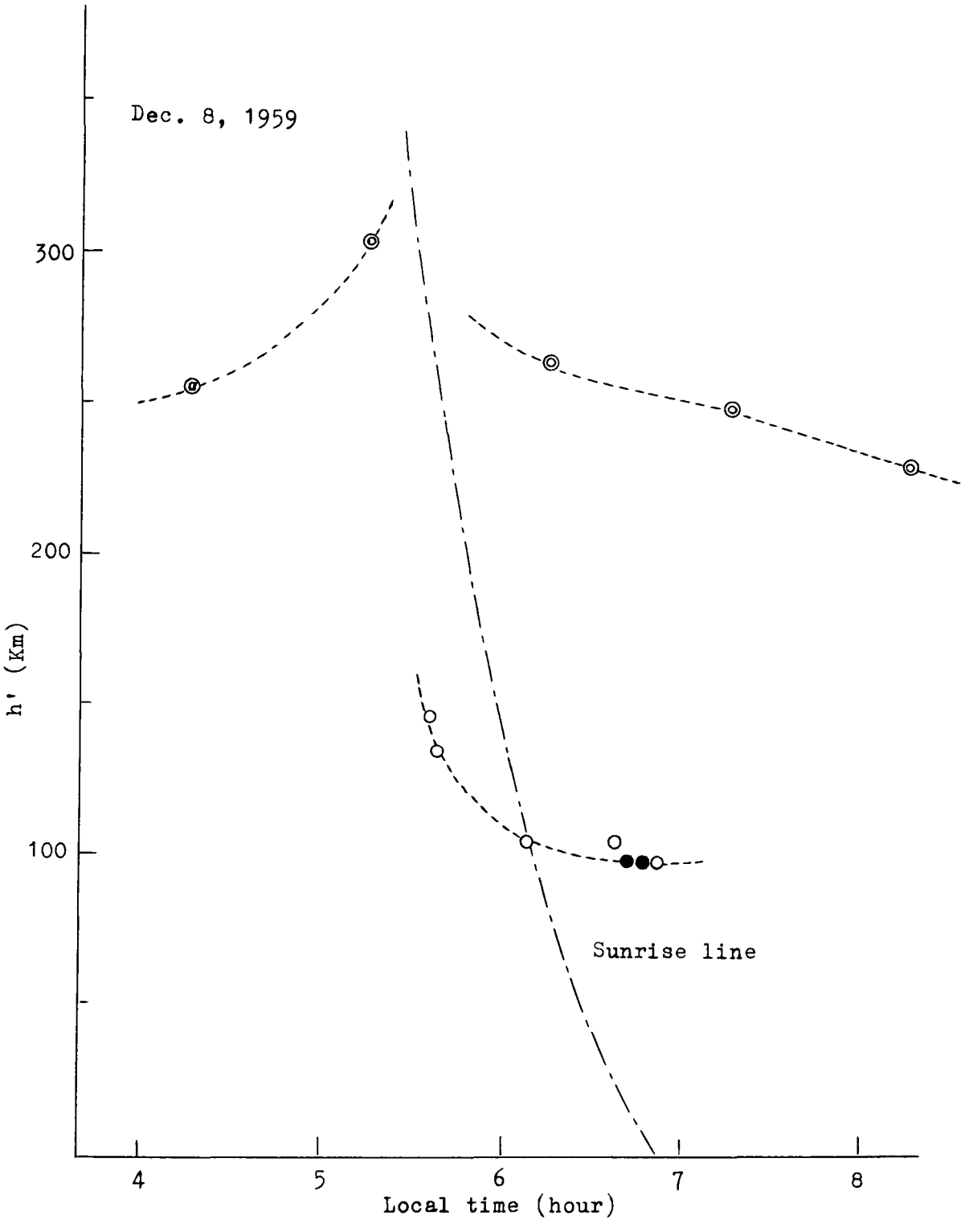


Fig. 3.28(b) Hourly variation of the apparent height h' for JOBK(o) and JOBB(●) in comparison with the variation of h'F at Kokubunji(◉) around the layer sunrise.

exist upward drifts both in the E layer and F layer, around the layer sunrise the phasesof which are different by an interval corresponding to the difference of the layer sunrise time for the E and F layers. Such vertical drifts of the layers are surely related with the dynamo currents in the layers. However these drifts can not be dealt with from dynamo-theoretically, taking the sunrise effect into consideration. Here we only point out the existence of the vertical drift in both layers around the layer sunrise.

3.4.3 Recombination coefficient of the E layer during night time

Daily variation of the electron density N in the E layer is generally determined by the equation

$$\frac{dN}{dt} = I - \alpha N^2, \qquad (3.13)$$

where α is the recombination coefficient and I is the ion production rate, and t is time. At night there is no ion production since the solar radiation does not attain to the region at all. Then the hourly variation of N at night will be given by

$$\frac{dN}{dt} = -\alpha N^2. \qquad (3.14)$$

The problem on the electron density during night time, therefore, seems quite simple. However in actual case, the electron density N does not follow such a simple equation with a constant recombination coefficient, but appears to follow the equation with variable α which depends on both altitude and time. According to A. P. Mitra(1957), α decreases after the layer sunset in such a way as shown in Fig. 3.29.

On the other hand our observed data is not so fine to deduce the variation of α. Only we can estimate an averaged recombination

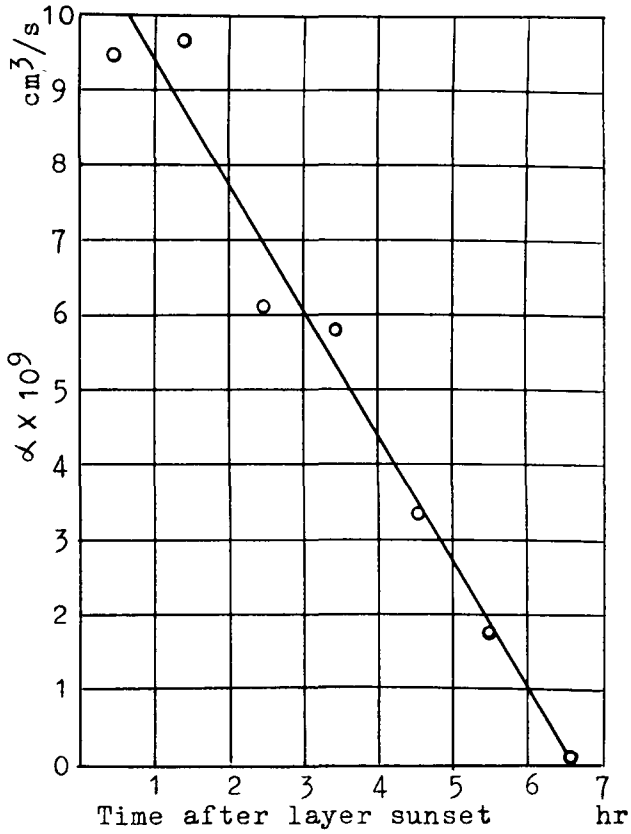


Fig. 3.29 Recombination coefficient at night obtained from f_0E observation at Watheroo. (after A.P.Mitra)

coefficient α_a for 2 or 3 hours after layer sunset. For the data on April 4 and 6 in 1959, we obtain $\alpha_a = 5 \times 10^{-8} cm^3/sec.$ for the altitude of the maximum density. For winter season, we can obtain the value $\alpha_a = 1.6 \times 10^{-8} cm^3/sec.$ on November 26 and December 10 in 1959. In addition, if we assume a tendency of decreasing α with time like in Fig. 3.26, the recombination coefficient at the layer sunset must be greater than those we have now estimated. From these data, it can be deduced that the recombination coefficient is usually greater than that obtained by A.P. Mitra and seems greater in spring than in winter. Naturally much data are further needed in order to obtain more plausible recombination coefficient.

Chapter 3.5 Estimation of the error

3.5.1 Comments on the error

Our method does clearly rely upon the excellency of separation of the sky wave by means of the loop antenna system. This excellency will depend upon not only the mechanical structure or electrostatic balancing of the antenna but also its location or environment. It results from the following reason: Owing to the secondary emission of conductors which are located near the antenna, apparent incident direction of a vertically polarized radio wave is shifted from the true direction by a little amomt. The conductors may usually be service wires for power supply. Then the impedance of the wire are fluctuated by users' switching at the ends of the wires, so that the intensity of the secondary emission will be unstable by the switchings. This effect results in some fluctuation of the incident direction of the radio wave. Consequently, in such a state the loop antenna or gonioscope must be adjusted at every moment so as to keep the best setting for sky wave extraction. Otherwise the intense ground wave compenent is apt to invade into the sky wave channel and cause a considerable error in observation.

In our case, batteries were used for the power source of the gonioscope receiver in order to avoid such switching effects as effective as possible. In spite of the consideration, it was difficult to prevent the effects completly.

From these points of view, an open field may be favorable for the location of the antenna. Not-with-standing, it was impossible to look for such a good place in the University, to try the experiment.

Another cause of error is sometimes due to the intermixing of waves reflected simultaneously from two layers. For example, when a sporadic E layer underlies the E layer, the Es layer is semi-transparent so that the both layers produce reflecting waves. In such a case, the sky wave at the receiving point, itself has been already distorted, then the observation gives us neither the height of Es layer nor that of the E layer. Our method is therefore powerless for these cases.

However in the former case, namely in case of insufficient separation of the ground wave and sky wave, observed data will be sometimes usable in estimating the apparent height h' of the reflecting layer, if in advance the effect of insufficient separation is evaluated. In the following, theorical arguments are given for the possibility.

3.5.2 Effect of the ground wave intermixing in the direct method

In the direct method explained in 3.3.3(a), audio wave forms are directly compared with each other. Therefore if the separation of the sky wave is unsatisfactory, the output audio signal of the sky wave receiver will be distorted from the original audio signal (that is the output of the ground wave receiver). Such records must be omitted in estimating the apparent height. However for the sake of necessity of ample records during short time, such records must also be used anyway.

Now we want to find the effect of the ground wave intermixing on the distortion of the wave form.

From eq.(3.5) the emf of a loop antenna is approximately

$$e_l = -j \frac{2\pi NA}{\lambda} \left[E_G \cos\theta - 2E_H \cos i \sin\theta \right] .$$

$$= K(E_{Ge} - E_{He}), \tag{3.15}$$

where

$$\left.\begin{aligned} K &= -j2\pi NA/\lambda, \\ E_{Ge} &= E_G \cos\theta, \\ E_{He} &= 2E_H \cos i \cdot \sin\theta. \end{aligned}\right\} \tag{3.16}$$

Assume that the original radio wave is of the frequency$(\omega/2\pi)$ and is modulated by a single audio frequency$(p/2\pi)$. Then

$$E_{Ge} = \{1 + k\sin pt\} E_{Go} e^{j\omega t}, \tag{3.17}$$

$$E_{He} = \{1 + k\sin p(t-\tau)\} E_{Ho} e^{j\omega t} e^{-j\omega\tau}, \tag{3.18}$$

where k is the degree of modulation and τ is the time lag of the sky wave with respect to the ground wave. Hence (3.15) becomes

$$e_\ell = KE_{Ho}\left[\rho(1+k\sin pt) - \{1+k\sin p(t-\tau)\} e^{-j\omega\tau}\right] e^{j\omega t}. \tag{3.19}$$

where $\rho = E_{Go}/E_{Ho}$ and indicating the degree of intermixing. After the detection stage of the receiver, the output signal is represented by the modulus of the ( )-bracketed part of (3.19);

$$\rho(1+k\sin pt) - \{1+k\sin p(t-\tau)\} e^{-j\omega\tau}. \tag{3.20}$$

This signal will be generally distorted from the original sinusoidal wave of the angular frequency p.

Now we suppose that the amount of shifting of the sky wave from the ground wave is measured by the deviation of a certain local peak in the wave form, as was shown in Fig. 3.21. The phase of the positive peak point in one period is calculated under a given $\omega\tau$, and is shown in Fig. 3.30, under the assumption; k=1, $p/2\pi$=500c/s, τ=500μs, ρ=0.5 and 1.0. In this graph, phase deviation of a peak of the compound wave form is represented in %, where 0% implies no error in estimating the time lag of the sky wave and positive %

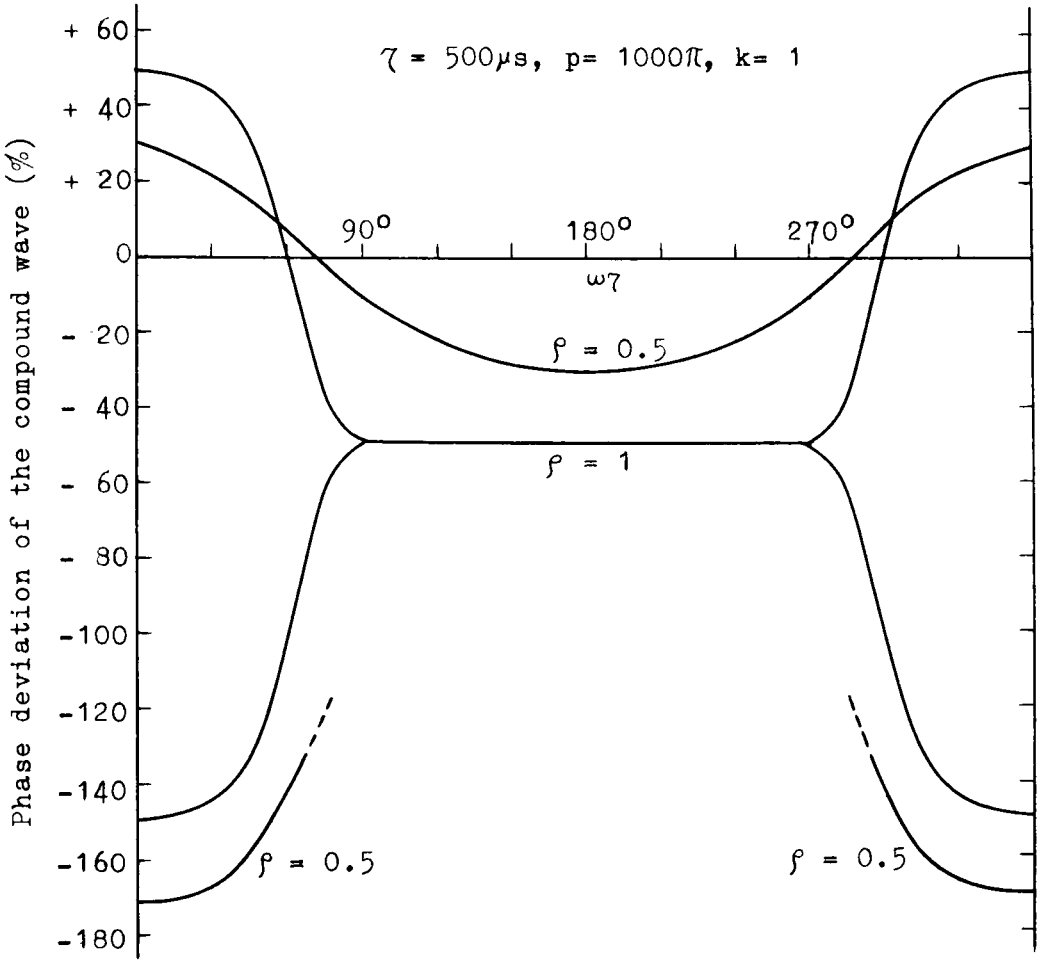


Fig. 3.30 Phase deviation of the positive peak of the compound wave with respect to that of the sky wave.

implies a larger time lag than the true one. Two branches around $\omega\tau = 0^{\circ}$(or 360°) represents the existence of two positive peaks in one period, that is the appearance of the second harmonics due to distortion. Then it may be concluded that if we omit such data as those having two peaks in one period, apparent time lag of the compound wave form from the ground wave is clearly less than that of the true sky wave.

In consequence, in case the wave form on the film is a little distorted when it is compared with the ground wave, the evaluated apparent height must be smaller than the actual apparent height. This conclusion has given the basic way in treating the observed data, as was carried out to produce Fig. 3.18, 19, 20.

3.5.3 Error in the pseudo impulse method

Secondly we are going to investigate an error which will arise from the pseudo impulse method. The difference from the previous case is that in this case a band pass filter is used to pick up only the signal of frequencies ranged from 500 to 650c/s. Then a signal represented by (3.20) whose fundamental frequency is in the above range, loses the higher harmonic components by the action of filter. Therefore the phase of only the fundamental frequency of (3.20) is important. This phase is calculated by the numerical harmonic analysis.

In Fig. 3.31(a), (b), (c), the difference between the phase of the fundamental frequency($p/2\pi$) component obtained by the above analysis and the phase of the ground wave signal of frequency $p/2\pi$, is indicated in % against $\tau = 500\mu s$ for various $\omega\tau$, with the assumption; $p/2\pi = 500c/s$, $k=0.5$ and $\rho = 0.1$, 0.2, 0.5 respectively. This

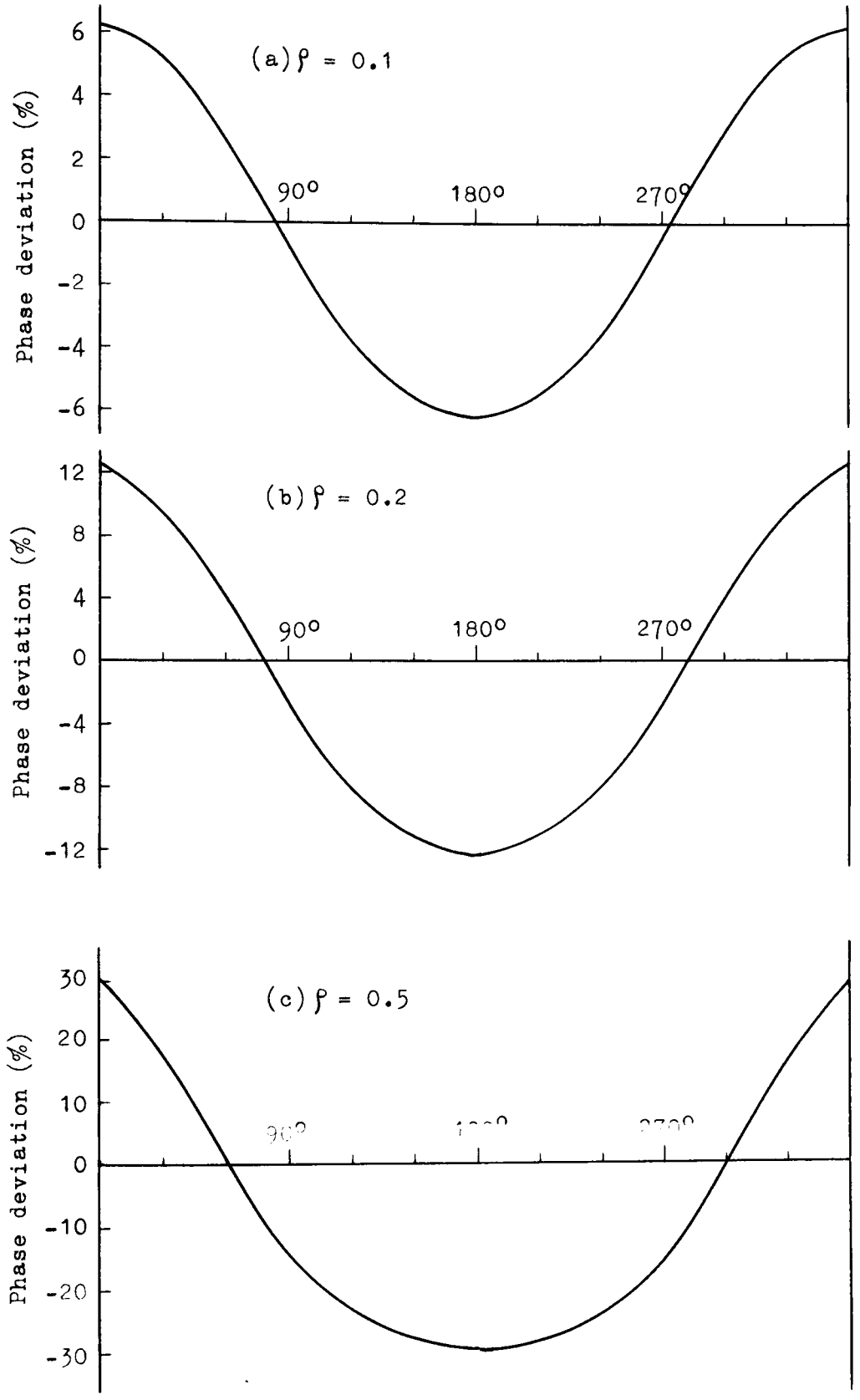


Fig. 3.31(a),(b),(c) Phase deviation of the signal of frequency 500c/s due to the intermixing.

phase deviation can be regarded as an error in obtaining the apparent height h'. According to the figures, the maximum deviation takes place at 0 or 180° of $\omega\tau$ and the magnitude is approximately linearly dependent upon ρ as shown in Fig.3.32.

In Fig. 3.33 the magnitude of the maximum deviation for various modulation frequency p within the pass band of the filter is shown, with $\rho=0.1$ and $\omega\tau=180^{\circ}$, 210°, 240° as parameter.

The deviation for different k is shown in Fig. 3.34 with $\rho=0.1$, $\tau=500\mu s$, $p/2\pi=500c/s$ and $\omega\tau=180^{\circ}$, 210°, 240° as parameter.

According to Fig. 3.31~3.34, we can conclude that:

1. The magnitude of deviation depends upon the fraction of the ground wave component inermixing into the sky wave channel. In addition it is greatly dependent upon the phase lag $\omega\tau$ of the carrier of the sky wave from that of the ground wave. In case $\omega\tau$ is around zero, the time lag appears to be larger than the true time lag, while in case $\omega\tau$ is around 180°, it appears smaller. Since ω is of the order of 10^6 in our case, $\omega\tau$ will change from 0° to 360° according to a small change of the time lag τ, say $10\mu s$, such as due to a bit of drift of the reflecting layer.

2. If $\omega\tau$ is assumed to be constant, the error in time lag is almost invariant irrespective of the modulation frequency p and the degree of modulation k.

Consequently, when the intermixing of a certain fraction of the ground wave component into the sky wave channel can not be prevented, the time lag of the sky wave looks varying so much in time as shown in Fig. 3.23(c),(d). In turn, even if the time lag appears to be constant for a short time, it does not always imply the correct time

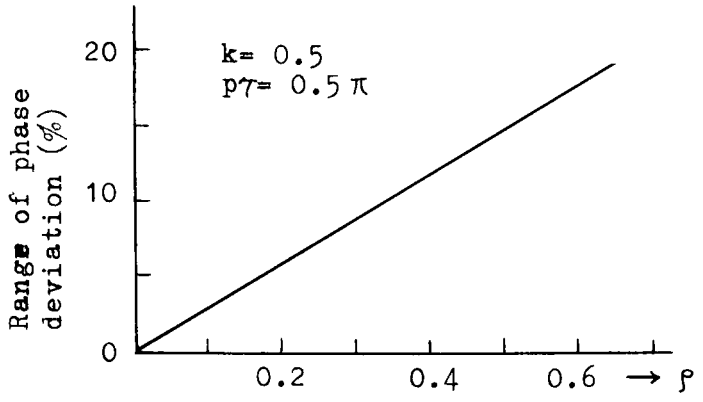


Fig. 3.32 Range of phase deviation as a function of the intermixing ratio ρ.

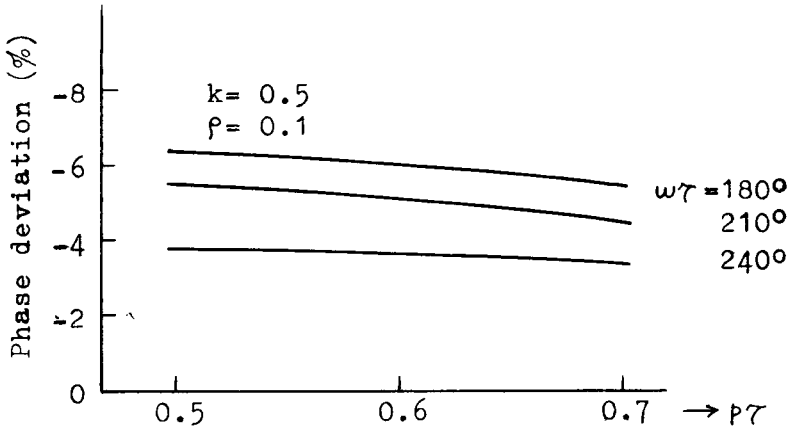


Fig. 3.33 Phase deviation as a function of the modulation frequency p. ($\tau = 500\mu s$)

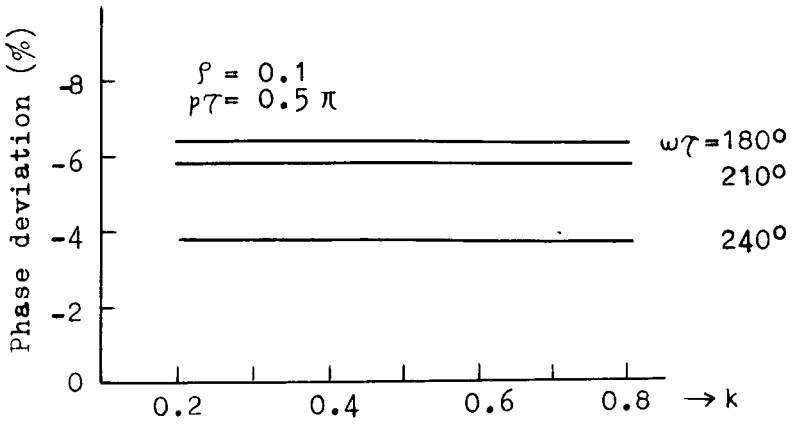


Fig. 3.34 Phase deviation as a function of the index of modulation k.

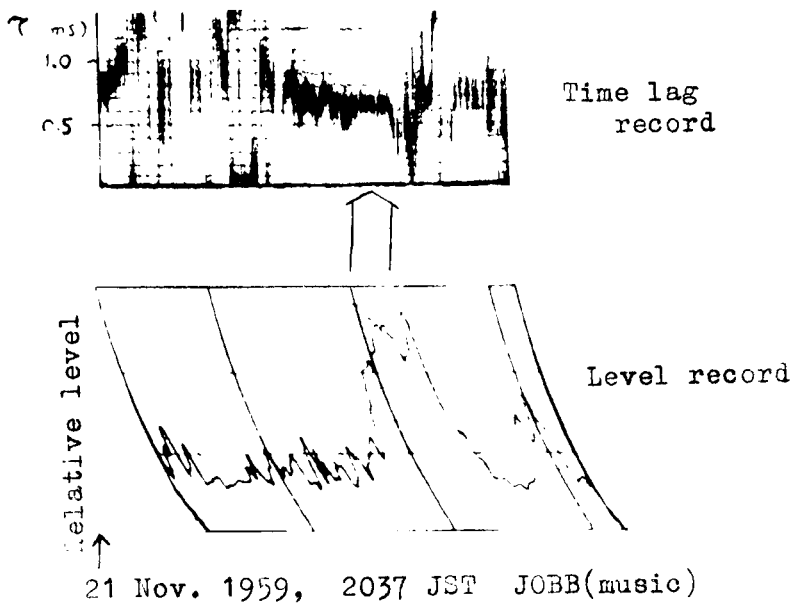
lag, because it is possible for γ and ρ to be sufficiently invariant for a while.

In this stage we must consider how to reduce the error in evaluating the apparent height h'.

Under the most satisfactory condition in eliminating the ground wave by the loop antenna system, the observed time lag must be invariant for a period of the order of minute. Yet in case some extent of invasion of the ground wave into the sky wave channel exists, the observed time lag fluctuates according to the fluctuation of the sky wave intensity or the fluctuation of the phase of the sky wave carrier. Therefore it is sometimes effective to compare the record of the time lag with the record of the carrier level of the sky wave channel. If the time lag varies in association with the variation in carrier intensity of the sky wave, at the maximum of the intensity, the time lag will become correct value, since the ground wave component invaded into the sky wave channel may be so stable in intensity that at the maximum of the carrier level recorded, the intermixing ratio ρ will become minimum. Such examples of the correlation are shown in Fig.3.35.

However the maximum of the carrier level does not always result in the correct time lag, but the minimum of the carrier level sometimes indicates the true time lag τ. This case may be due to the variation of the incident direction of the ground wave with time.

In case of the carrier phase changing with time, we can see a view that the apparent time lag τ changes largely, in spite that the carrier level does not so largely change. In such a case, the true time lag τ may be the mean value of the fluctuating time lag. Such data however, could not actually be found in the records.



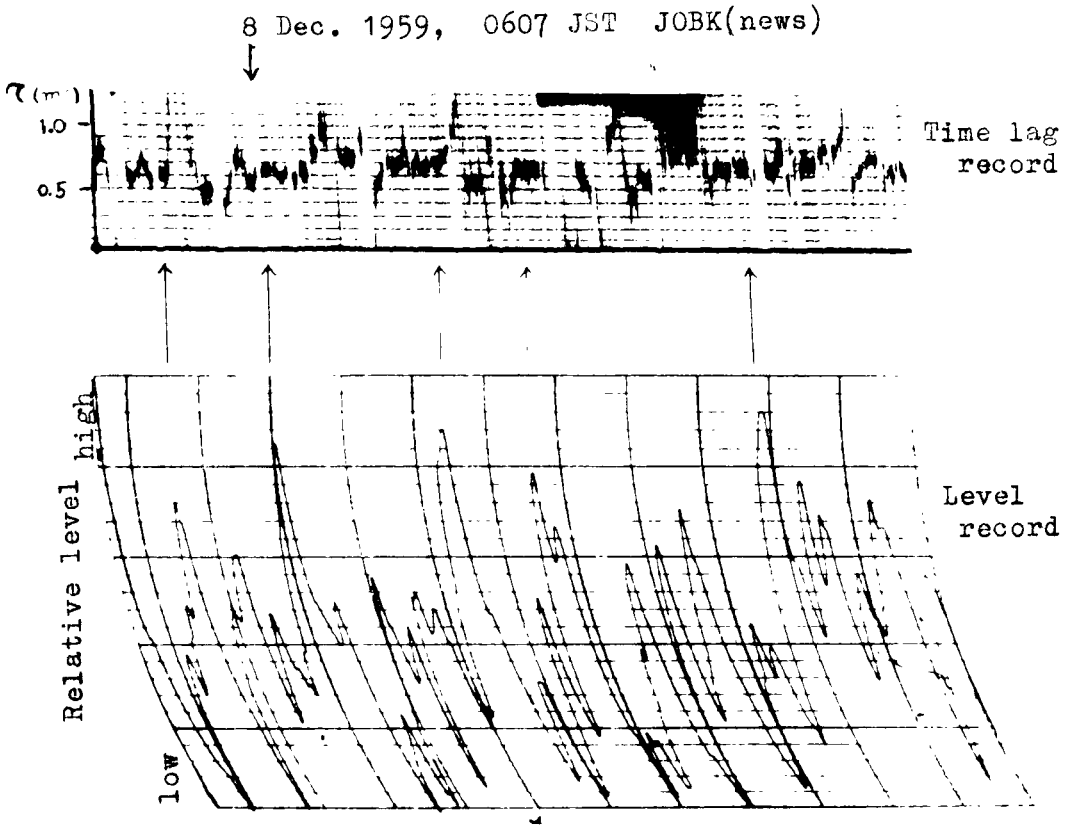


Fig. 3.35 Correlation of the time lag τ and the level in the sky wave channel.

Chapter 3.6 Concluding remarks

Although an idea of observation of the ionosphere by means of obliquely propagating broadcast waves may not be so wild, the experiment to ascertain its possibility was quite difficult. It is because first a satisfactory separation of the weak sky wave from the strong ground wave is highly troublesome, and second it is difficult to keep the best condition of the separation for a long time against the instability of various parameters such as the effect of secondary emissions by neibouring conductors, or may-be an actual variation of the incident direction of the ground wave.

In spite of the difficulty, the result of our experiment seems to be satisfactory. We could obtain the data of the critical frequency of the E layer during nighttime, though the shape of the layer could not be determined because of deficiency of the data and yet the inefficient technique of observation.

The critical frequency of the E layer probably decreases down to or even less than 650Kc/s, 2 or 3 hours after the layer sunset. Then the recombination constant must be as great as $2\times10^{-8}cm^3/s$.

Observations made early morning have shown rapid ascending and descending of the E layer around layer sunrise. This phenomenon is similar to those appearing in the height variation of the F layer around the layer sunrise. It may prove the existence of a vertical drift around the layer sunrise in the both layers.

The effect of unsatisfactory separation of the ground wave and sky wave is investigated and as the result even if some fraction of the ground wave component invades into the sky wave channel, the

correct h' often can be deduced.

Then if a favorable location is selected where the extraction of the sky wave can be always satisfactorily expected, this method of ionospheric sounding will play a powerful rôle during nighttime.

CONCLUSION

A purpose of the author's studies was to clarify unknown problems in the field of the electromagnetic wave propagation through several topics concerning the ionosphere and exosphere. It is because the theoretical and experimental investigations on such problems will enable us to acquire not only the solution of the problems of the electromagnetic waves proper, but also some information on the unknown quantities of the media.

Accordingly as has been described in this paper, the author has studied mainly the VLF phenomena associated with the ionosphere and exosphere, and also investigated a method of sounding the nighttime E layer.

In what follows, the summary of this paper and the problems to be developed in future are given.

In part I, the problem of ray tracing of the VLF electromagnetic waves in an anisotropic and inhomogeneous medium, like in the exosphere and ionosphere was mainly treated. Although for the ray tracing in isotropic medium, The Snell's law is usable, the author found that this law could be applied to the anisotropic media with respect to the wave normal.

Then an approximate method of numerical calculation of the ray path was devised, which is applicable to the whistlers. By using this method , ray paths of the whistlers were calculated under some assumption of the distribution of the electron density.

The calculated results indicates that the ray paths are neither necessarily identical with the lines of force of the geomagnetic field, nor symmetrical with respect to the magnetic equator. The asymmetry is conspicuous for the paths starting from the low latitude.

Attenuation in the ionosphere for the VLF waves was also calculated. The attenuation by absorption amounts to 35~50db in the daytime, but it is only 5db at midnight. This quantity can explain the diurnal variation of the occurrence frequency of the whistlers and VLF emissions.

In part II, a possibility of amplification of the VLF electromagnetic waves by a charged particle beam running through a magnetoactive plasma was examined, in order to find a mechanism of the VLF emissions in the exosphere and ionosphere. The problem of amplification by means of a beam is not so fresh problem, because originary such a problem has been developed in the field of microwave tube and in the field of solar radiation. Accordingly a possibility of the TWT like amplification was first reminded in connection with the VLF emissions.

However the author has found that in the exosphere, such longitudinal coupling mechanism does not yield any spatially growing mode for the whistler mode wave. In turn, he examined the transverse coupling of the proton cyclotron mode of the beam and the whistler mode. And this coupling was proved to yield a spatially growing mode of the proton cyclotron wave and the whistler mode waves. The energy for these two growing mode is supplied from the average velocity of the proton beam.

Since the gain of amplification is not so large, there needs some powerful noise to be amplified for the explanation of the VLF emissions. It was also pointed out that the atmospheric noise generated in the troposphere can always invades into the exospheric

space in the whistler mode, so that these noise, sufficiently intense, can be a noise source which is amplified to generate the VLF emissions.

This mechanism of amplification may be applicable to the problems of radiation from a magneto-active plasma in microwave region and the problem of the solar radiation.

In part III, a quite new method of ionosphere sounder was explained. We have not yet known much about the nighttime E layer. This is because the observation by using middle frequency band has not been performed at all in Japan.

The author has experimented a new method which utilizes the middle frequency broadcast waves. Separation of the sky wave and the ground wave was succeeded by using a loop antenna. In spite that the experiment was not so much satisfactory, the results suggest us the possibility that it can be used as one of the ionosondes, during nighttime, provided that an appropriate location is selected.

As to the futere problems, some remarks are finally added.

In part I, a wave theoretical study on the VLF waves returned to the ionosphere should be made. For more plausible estimation of the electron density in the exosphere, a systematic experiment of man-made whistler will be necessary.

In part II, an argument taking account of the velocity distribution of charged particles in the beam and in the ambient plasma, is left untouched. This effect may decrease the gain and widen the frequency band to be amplified. Although the effect of nonlinearity may not be so important, our treatment would be inperfect,

without any examination of the effect. Three dimensional coupling problem seems to be interesting. An incoming beam may have an action to change the wave normal of a propagating whistler mode wave.

In part III; though our experiment was satisfactory in a sense, some difficulties are left in the use of the method for a routine observation. One is the instability of the extraction of the sky wave component. Second is the difficulty of observing many waves simultaneously. The first will be solved to a certain extent by selecting an appropriate location of the experiment. The second is partly a technical problem and partly an economical problem.

ACKNOWLEDGEMENT

The author wishes to express his sincere thanks to Prof. K Maeda for his kind guidance and encouragement throughout this study. Many thanks are due to Prof. T. Sakai and Dr. S. Kato for their valuable discussion and encouragement in promoting his work.

He is also indebted to Prof. Y. Takeya of Osaka City University, Assistant Prof. H. Matsumoto of Kobe University, Messrs S. Yajima and T. Tsuda for their discussion and interests in these problems.

Further he expresses his gratitude to Dr. S. Kato for his assistance in using KDC-1 electronic computer and to the staffs of KDC-1 who prepared the library subroutines for the computor. He must also express his thanks to Mr. J. Otsu of Nagoya University and Mr. T. Ondoh for their valuable criticism on the studies of the VLF phenomena, and Dr. S. Kawazu of Electrical Comunication Laboratory for his suggestion on the experiment of the ionosphere sounder.

He is grateful to Messrs Y. Kawamura, A. Ezuka and H. Nobukuni for their enthusiastic co-operation in the experiment and calculation during their graduate studies. Finally he appreciates the kind assistance of Mr. S. Takeda and other staffs of Maeda research laboratory in preparing this paper.

REFERENCES

PART I

Alfven, H.; Cosmical Electrodynamics. Oxford University Press, London New York, 1950, p. 4.

Appleton, E. V.(1932); J. Inst. Elec. Eng. 71, 642.

Booker, H. G.(1935); Proc. Roy. Soc. A. 150, 267.

Bremmer, H.; Terrestrial radio waves. Elsevier Pub. Co., New York-Amsterdam-London-Brussels, 1949, p. 301.

Bremmer, H.(1951); Theory of electromagnetic waves-A Simposium, p.169.

Crary, J. H., Helliwell, R. A., and Chase, R. F.(1956); J. Geophys. Res. 61, 35.

Dinger, H. H.(1956); Navy Res. Rep. No. 4825.

Helliwell, R.A., Crary, J. H., Pope, J. H., and Smith, R. L.(1956); J. Geophys. Res. 61, 139.

Helliwell, R. A., and Gehrels, E. (1958); Proc. IRE. 46, 785.

Helliwell, R. A., and Morgan, M.G.(1959); Proc. IRE. 47, 200.

Iwai, A., Otsu, J., Murata, Y., and Kato, T.(1957); Bull. Res. Inst. Atmosph. Nagoya University. 7, 26. (in Japanese)

Iwai, A. and Otsu, J.(1958); Proc. Res. Inst. Atmos., Nagoya University. 5, 50.

Kimpara, A.(1960); Proc. Res. Inst. Atmos. Nagoya University. 7, 40.

Kimura, I.(1957); Master Dessertation.

Koster, J. R., and Storey, L. R. O.(1955); Nature, London, 175, 36.

Lepechinsky, D.(1956); J. Atmos. Terr. Phys. 8, 297.

Maeda, K., and Kimura, I.(1956); Rep. Ionos. Res. Japan. 10, 105.

Maeda, K., and Kimura, I.(1956); Denki-Gakkai General Meeting, Report No. 137 (in Japanese).

Maeda, K., and Kimura, I.(1957); Denki-Yon-Gakkai General Meeting, Report No. 623 (in Japanese).

Maeda, K., and Kimura, I.(1959); J. Atmos. Terr. Phys. 15, 58.

Matsumoto, H.(1960); Report at the Committee on Radio Wave Propagation Dec. 8 1960. (in Japanese)

Namba, S., and Maeda, K.; Radio wave propagation. Corona Book Co., 1939. p. 11. (in Japanese)

Nicolet, M. and Aikin, (1960); J. Geophys. Res. 65, 1469.
Siedentopf, H., Behr, A., and Elsassen, H.(1953); Nature, London. 171, 1066.
Smith, R. L., Helliwell, R. A., and Yabroff, I. W.(1959); Technical Report No. 3. Dec. 31, 1959, Stanford Electronics Laboratories.
Storey, L. R. O.(1953); Phil. Trans. Roy. Soc. A. 246, 113.
Yabroff, I.(1959); Final letter report, Stanford Research Institute. AFOSR-TN-60-71, December 1959.

PART II

Allcock, C. Mck.(1957); Australian J. Phys. 10, 286.
Bailey, V. A.(1948); Australian J. Sci. Res. A. 1, 351.
Bailey, V. A.(1950); Phys. Rev. 78, 428.
Barrington R. E.(1959); Proceedings of the Symposium on Physical Process in the Sun Earth Environment. Defence Research Board, Canada, p. 223.
Bell, T. F., and Helliwell, R. A.(1959); ibid, p. 215.
Brice, N. M.(1960); Technical Report No. 7, Stanford Electronics Laboratories.
C. C. I. R.(1959); Document of the IXth plenary assembly, Los Angeles, III, p. 223.
Ellis G. R.(1957); J. Atmos. Terr. Phys. 10, 302.
Gallet, R. M. and Helliwell, R. L.(1959); J. Res. Nat. Bur. Stand. 63D, 21.
Gallet, R. M.(1959); Proc. IRE. 47, 221.
Kimura, I.(1961); Rep. Ionos. Space Res. Japan. 15, No. 2 (in press).
Mac Arthur,J. W.(1959); Phys. Rev. Letters. 2, 491.
Martin, L. H., Helliwell, R. A., and Marks, K. R.(1960); Technical Report No. 1, Stanford Electronic Laboratories.
Murcray, W. B., and Pope, J. H.(1960); Phys. Rev. Letters 4, 5.
Ondoh, T.(1961); J. Geomag. Geoelec. 12, 77.
Parker, E. N.(1959); Proc. IRE. 47, 239.
Siegman, A. E.(1960); J. App. Phys. 31, 17.
Sumi, M.(1959); J. Phys. Soc. Japan 14, 653.
Yoshida, S.(1960); Uchusen-Kenkyu 5, 397 (Journal of Cosmic Research in Japanese).

PART III

Appleton, E. V., and Ratcliffe, J. A.(1927); Proc. Phys. Soc. 117, 576.

Maeda, K., Kato, S., and Kimura, I.(1958); Denki-Yon-Gakkai General Meeting, Report No. 824. (in Japanese)

Maeda, K, and Kimura, I.(1958); Report at the Committee on Radio Wave Propagation, Dec. 4 1958. (in Japanese)

Maeda, K., and Kimura, I.(1959); Denki-Yon Gakkai General Meeting, Report No. 1205. (in Japanese)

Millman, J., and Taub, H.; Pulse and Digital Circuit. Mc Graw Hill Book Co., 1956.

Mitra, S. K.; Upper Atmosphere. The Asiatic Society, Calcutta, 1952.

Mitra, A. P.(1957); J. Atmos. Terr. Phys. 10, 140

Watts, J. M., and Brown, J. N.(1950); J. Geophys. Res. 55, 179.

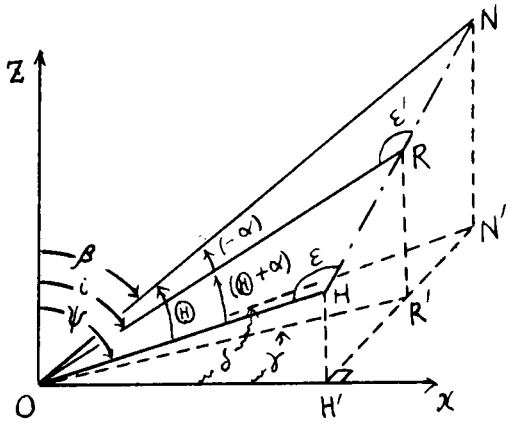
Watts, J. M.(1957); J. Geophys. Res. 62, 484.

Miura, T.; Tsuwa Hin-shitsu(Transmission quality). Electrical Communication Engineering Series. Kyoritsu Pub. Co., 1955.

Appendix 1.1

(The deduction of the equations (1.39) and (1.40) from (1.37) and (1.38))

The relative relation of the angles is illustrated in a diagram as below. $\vec{OH}$, $\vec{ON}$, $\vec{OR}$, indicate magnetic field, wave normal and ray direction respectively and let the points H, R, N be in a plane perpendicular to the x axis. Then, these points are all on one straight line. Their project upon xy-plane are the points H', R', N'.



(i) By using a principle of trigonometry, we get the relations

$$\frac{\overline{HR}}{\sin(Ⓗ+\alpha)} = \frac{\overline{OR}}{\sin\varepsilon},$$

$$\frac{\overline{HN}}{\sin Ⓗ} = \frac{\overline{ON}}{\sin\varepsilon}.$$

These identities yield the relation

$$\frac{\overline{HR}}{\overline{HN}} = \frac{\overline{OR}}{\overline{ON}} \frac{\sin(Ⓗ+\alpha)}{\sin Ⓗ}$$

Further we have the identity $\dfrac{\overline{HR}}{\overline{HN}} = \dfrac{\overline{OR}\sin i \sin\gamma}{\overline{ON}\sin\beta\sin\delta}$

Accordingly, $$\frac{\sin i \sin\gamma \sin Ⓗ}{\sin(Ⓗ+\alpha)} = \sin\beta\sin\delta \qquad \text{(A-1)}$$

holds. By means of the above relation, (1.38) becomes

$$n\sin\beta\sin\delta = const \longrightarrow (1.40)$$

(ii) According to the trigonometry, there hold the next identities

$$\frac{\overline{ON}}{\sin\varepsilon} = \frac{\overline{HR}}{\sin(-\alpha')} = -\frac{\overline{HR}}{\sin\alpha},$$

$$\frac{\overline{OH}}{\sin(\pi-\varepsilon)} = \frac{\overline{OH}}{\sin\varepsilon} = \frac{\overline{RN}}{\sin(\textcircled{H}+\alpha)}$$

From these, we have $$\frac{\overline{ON}}{\overline{OH}} = -\frac{\overline{HR}}{\overline{RN}}\frac{\sin(\textcircled{H}+\alpha)}{\sin\alpha}.$$

Then $$1 - \frac{\overline{ON}}{\overline{OH}}\frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)} = 1 + \frac{\overline{HR}}{\overline{RN}} = \frac{\overline{HN}}{\overline{RN}} = \frac{\overline{ON}}{\overline{OH}}\frac{\sin\beta\sin\delta}{\sin\psi\tan\gamma}$$

or $$1 = \frac{\overline{ON}}{\overline{OH}}\left\{\frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)} + \frac{\sin\beta\sin\delta}{\sin\psi\tan\gamma}\right\} \qquad \text{(A-2)}$$

Now we have another identity, namely

$$\overline{ON}\sin\beta\cos\delta = \overline{OH}\sin\psi \quad \text{or} \quad \frac{\overline{ON}}{\overline{OH}} = \frac{\sin\psi}{\sin\beta\cdot\cos\delta} \qquad \text{(A-3)}$$

And from (A-1)

$$\sin\beta\sin\delta = \frac{\sin\gamma\sin i\sin\textcircled{M}}{\sin(\textcircled{H}+\alpha)} = \sin\gamma\sin i\{\cos\alpha - \sin\alpha\cot(\textcircled{H}+\alpha)\}$$

$$= \sin\gamma\sin i\left\{\cos\alpha - \frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)}\cos(\textcircled{H}+\alpha)\right\}$$

$$= \sin\gamma\sin i\left\{\cos\alpha - \frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)}(\sin\psi\sin i\cos\gamma + \cos\psi\cos i)\right\} \qquad \text{(A-4)}$$

where the relation (1.27) was used. Applying (A-3) and (A-4) to (A-2), then we have

$$1 = \frac{\sin\psi}{\sin\beta\cos\delta}\left[\frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)} + \frac{\sin\gamma\sin i}{\sin\psi\tan\gamma}\left\{\cos\alpha - \frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)}(\sin\psi\sin i\cos\gamma + \cos\psi\cos i)\right\}\right]$$

Consequently,

$$\sin\beta\cos\delta = \left[\frac{\sin\alpha\sin\psi}{\sin(\textcircled{H}+\alpha)} + \sin i\cos\gamma\left\{\cos\alpha - \frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)}(\sin\psi\sin i\cos\gamma + \cos\psi\cos i)\right\}\right]$$

$$= \sin i\cos\alpha\cos\gamma + \frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)}\left[\sin\psi - \sin^2 i\cos^2\gamma\sin\psi - \sin i\cos\psi\cos i\cos\gamma\right]$$

$$= \sin i\cos\alpha\cos\gamma + \frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)}\left[\sin\psi(1-\sin^2 i\cos^2\gamma) - \sin i\cos\psi\cos i\cos\gamma\right]$$

$$= \sin i \cos\alpha \cos\gamma + \frac{\sin\alpha}{\sin(\textcircled{H}+\alpha)} \left[\sin\psi \sin^2\gamma + \sin\psi \cos^2 i \cos^2\gamma - \sin i \cos\psi \cos i \cos\gamma \right]$$

$$= \cos\gamma \cos\alpha \left[\sin i - \frac{\tan\alpha}{\sin(\textcircled{H}+\alpha)} \left\{ (\sin i \cos\psi - \sin\psi \cos i \cos\gamma) \cos i - \sin\psi \sin\gamma \tan\gamma \right\} \right]$$

$$= \frac{1}{n} \quad \text{eq. (1.37)}$$

Thus we obtain,

$$n \sin\beta \cos\delta = \text{const.} \longrightarrow (1.39)$$