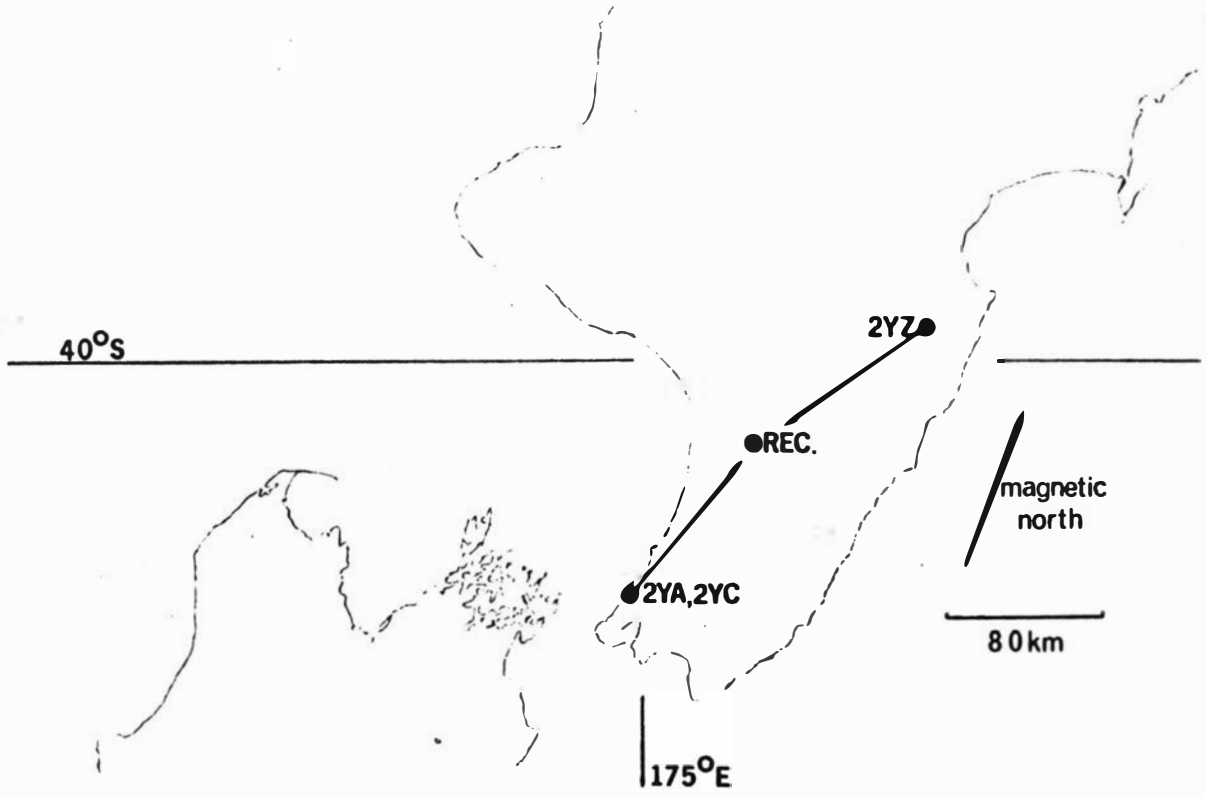
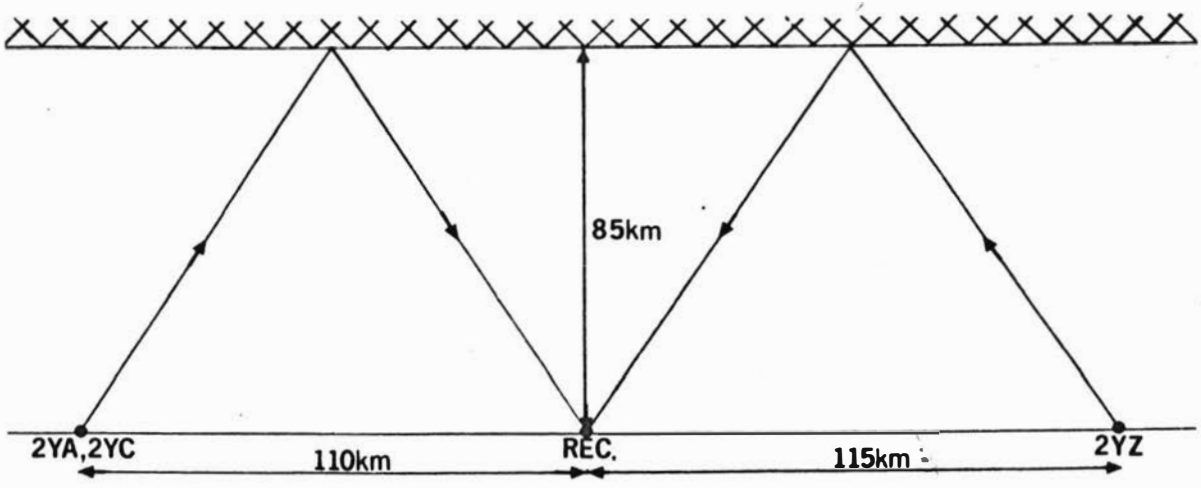
IONOSPHERIC REFLECTIONS AT MEDIUM FREQUENCIES

A thesis presented in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy in Physics at
Massey University

RALPH ERNEST HENDTLASS

1973





frontispiece

Abstract

This thesis examines a possible explanation for some of the time variations of the night-time sky wave at distances between 80 and 160 km from low-medium-frequency transmitters. In the postulated model this variation is explained in terms of interference between the signals travelling along many paths from the transmitter to the receiver via a lower ionosphere disturbed by the passage of a ducted acoustic wave. A rigorous solution to this model is not possible, but by using reasonable approximations a solution is obtained which is suitable for analysing experimental records of sky wave variations. These experimental records were obtained by using an interferometer technique to separate the sky wave from the ground wave. Some extremely good agreement was found between theory and experiment showing that this could be a useful technique for studying ducted acoustic waves. Certainly some, at least, of the variations commonly observed in the night-time sky wave signal are caused in this way.

ACKNOWLEDGEMENTS

I am deeply grateful to Mr. E.R. Hodgson who not only suggested this project, but also gave freely of his time, advice and encouragement. Also I would like to thank Professor N.F. Barber of Victoria University of Wellington, for his valuable assistance.

I also wish to express my appreciation to all the technical staff under Mr. D.M. Williams, and especially Mr. L.G. Cranfield, for the use of their skills and time. The University Grants Committee also assisted with grant 70/143.

Finally I thank Mrs. D.G. Stichbury for her patience and ability to turn my illegible writing into orderly typescript.

Index

Frontispiece			ii
Abstract			iii
Acknowledgements			iv
Index			v, vi, vii
List of figures.			viii, ix
SECTION ONE:		INTRODUCTION TO THIS WORK AND A REVIEW OF EARLIER WORK	1
1-1	:	This experiment, origin and outline.	1
1-2	:	A brief survey of experimental techniques for observing the ionosphere between 50 and 100 km.	3
1-2-1	:	H.F. vertical incidence ground based sounding techniques.	4
1-2-2	:	H.F. oblique incidence ground based sounding techniques	5
1-2-3	:	L.F. ground based sounding techniques	6
1-3	:	The lower night-time ionosphere at middle latitudes.	7
1-4	:	A physical process and a postulated mathematical model.	12
SECTION TWO:		MATHEMATICAL SOLUTION OF THE IDEALIZED MODEL.	16
2-1	:	The deformation of the ionization boundary by the passage of an acoustic wave.	16
2-2	:	Factors affecting the received sky wave due to the perturbed boundary.	22
2-2-1	:	Interference terms.	22
2-2-2	:	The validity of the approximation used in the above section.	27
2-2-3	:	The reflection coefficient.	30
2-2-4	:	The reflecting area.	32

2-3 : Factors affecting the received sky wave due to the experimental arrangement. 37

2-4 : The range over which receivable group reflections can occur. 39

2-5 : The extention to a continuous curve of expected sky wave amplitude versus group position. 40

2-6 : The extention of the proceeding work in this section to an arbitary length acoustic group. 42

SECTION THREE: A DESCRIPTION OF EXPERIMENTAL METHOD. 43

3-1 : The aerial system. 43

3-2 : Rejection of the ground wave. 45

3-3 : Rejection of the sky wave. 45

3-4 : Limitations of the automatic sky wave rejection circuitary. 49

3-5 : An outline of practical circuit details. 50

3-6 : Experimental routine. 52

SECTION FOUR: RESULTS, CONCLUSIONS AND SUGGESTIONS. 54

4-1 : Results of early work. 54

4-2 : The determination of parameters from records made using a high chart speed. 58

4-3-1 : Recognition of significant events from high speed records. 61

4-3-2 : Results obtained from high speed records. 64

4-4 : Comparison with other experiments. 69

4-5 : Conclusion and suggestions for future research. 70

APPENDIX ONE: THE PREDICTION OF THE SKY WAVE SIGNAL TO BE EXPECTED FROM AN ACOUSTIC WAVE GROUP OF ARBITARY LENGTH. 74

APPENDIX TWO: SIMPLIFIED CIRCUIT DIAGRAMS 76

APPENDIX THREE: FLOW CHART FOR CALCULATING THE EXPECTED SKY WAVE AMPLITUDE USING THE THEORY OF SECTION TWO. 81

REFERENCES: 82

List of Figures

Figure i : Experimental, mid-latitude, night-time electron density distributions. 10

Figure ii : Mid-latitude, night-time electron density distribution as a function of solar zenith angle. 10

Figure iii : The geometry of the idealized model. 23

Figure iv : 28

Figure v : 28

Figure vi : 28

Figure vii : The ratio of the specular reflection area to any one of the non-specular reflection areas plotted as functions of the distance of the wave group from the transmitter and the total aerial extent. 36

Figure viii : The value of B required for signals reflected from position x to arrive at the receiver plotted as a function of x, and the possible range of reflection points as a function of L/A. 39

Figure ix : A block diagram of the system used in this experiment. 44

Figure x : The polar response of the aerial system when set up for ground wave rejection. 47

Figure xi : The polar response of the aerial system when set to reject signals with an angle of arrival of 60^{o}. 47

Figure xii : The additions required to the system shown in figure ix to enable that system to hunt for the sky wave. 48

Figure xiii : A typical night's record. 55

Figure xiv : Records showing examples of false indication of angle of arrival and examples of the monotonic variation in the angle of arrival. 57

Figure xv : A plot of L vs Δx. 59

Figure xvi : A set of theoretical curves of sky wave amplitude vs group position. 62

Figure xvi : A histogram of the occurrence of detected values of L in a total population of 82. 64

Figure xviii : An example of the good agreement between theory and experiment obtained by using successive approximation to get a best fit. 65

Figure xix : The angle of arrival of an observed sky wave compared with the theoretical values obtained from a computed curve. 66

Figure xx : Theoretical sky wave amplitude vs position curves for both 570 kHz and 660 kHz . 68

Figure xxi : Simplified circuit diagrams. 79

Simplified circuit diagrams, continued. 80

SECTION ONE : INTRODUCTION TO THIS EXPERIMENT AND A REVIEW OF EARLIER WORK.

A short description of this experiment and the early conclusions to be made from it are presented first. This is followed by a brief review of the main experimental techniques used to study the lower ionosphere and a picture of the night-time ionosphere as revealed by experiments using these techniques. The final part of this section is devoted to explaining a postulated physical process which might cause some of the observed effects, and in attempting to reduce this to a mathematically tractable model.

1-1: This experiment, origin and outline.

It has been known since the 1920's that both sky and ground waves are present in the night-time zone of fading which lies between 80 and 160 km from any medium frequency broadcast transmitter. Considering the relative path lengths, and accepting that the reflection coefficient of the ionosphere will probably be of the order of 0.1, it is clear that the sky wave amplitude will always be less than the amplitude of the ground wave unless focussing has occurred. Since the composite signal at a receiving aerial is at times observed to fade away totally, some form of focussing must be occurring.

Massey University (40° 21' S, 175° 34' E) happens to be approximately equidistant from three medium frequency broadcast stations of the New Zealand Broadcasting Corporation. These are stations 2YA Wellington (41° 06' S, 174° 50' E, 570 kHz), 2YC Wellington (41° 06' S, 174° 50' E, 660 kHz) and 2YZ Napier (39° 48' S, 176° 40' E, 630 kHz), and are shown in the frontispiece. The situation seemed to offer an opportunity for

an inexpensive study of night-time fading. The frontispiece also illustrates the proposal as it stood at the time work was begun. The sky waves arriving at the receiver (REC) would be recorded and it was hoped that the time variation of the received sky wave amplitude would give evidence about the focussing mechanisms.

One receiver and recorder was made and has been in use for about two years. A second receiver and recorder came into operation about one year ago. Two suitably spaced aerials at the receiver had their outputs combined so that the signal due to the direct ground wave was as small as possible, whereas the sky wave, coming down quite steeply, gives a large signal. The amplitude of the received signal (smoothed to a time constant of ten seconds) was recorded continuously usually from sunset to sunrise for it is only at night that the absorption from the D region (at a height of approximately 50 km) becomes small enough that an appreciable sky wave amplitude is received at the ground. Further details are given in section three.

Recordings made showed that the skywave was continually varying, occasionally reaching an amplitude many times that of the ground wave (twelve times the ground wave amplitude being observed on one occasion). Initial records on slow running chart paper (1 inch per hour) revealed that both the sky wave amplitude and the variation in this amplitude tended to reach a maximum at approximately local midnight. Comparisons of records made over a period of a year suggested that high sky wave levels occur more often in winter than in summer. It was noticed that even a single receiver sometimes gives records that suggest a special type

of moving disturbance. Recordings made with a faster chart speed (12 inches per hour) occasionally showed a distinct periodicity of sky wave maxima: a group of five or six regularly spaced peaks occurring within a total of ten to fifteen minutes. It is mainly these single station records and their possible interpretation that are considered in this work.

It is the suggestion of the writer that these peaks of intensity are due to various orders diffracted by a short packet of acoustic waves travelling through the lower ionosphere. In justification of this suggestion, it is shown in section four that if one assumes an acoustic wave packet to be travelling in a 'sound channel' (at a height of 85 km where there is a local minimum of sound speed), one can pick the acoustic wavelength and the number of waves in the packet in such a way as to account well for the relative magnitudes of the successive peaks in the record as well as for the interval between them.

1-2: A brief survey of experimental techniques for observing the ionosphere between 50 and 100 km

The region of the ionosphere lying between heights of 50 and 100 km has proved the most difficult region to investigate by means of direct measurement techniques; the atmosphere is too dense for satellites to remain in orbit and only a limited number of rocket borne measurements can be made. Consequently most observations have been made using ground-based techniques. For completeness, and because the relationship between the results of these techniques and the results obtained here are referred to in section 4-4, a brief survey of the experimental techniques for observing the ionosphere between 50 and 100 km

is presented below. This is followed in section 1-3 by a short description of the lower night-time ionosphere at middle latitudes, including the previously mentioned 'sound channel'. The main development of this work is returned to in section 1-4.

1-2-1: H.F. vertical incidence, ground based sounding techniques.

The conventional swept-frequency ionsonde has been used for several decades to record the heights of the various ionospheric layers and the variations in their electron concentration. Provided the range of observing frequencies extends as low as about 0.25 MHz investigations can be made of the lower D layer, although it is difficult to make observations in the same detail as those which can be made for greater heights. This is due to the low reflection coefficients and the inability of the usual N(h) reduction techniques to provide electron concentration distributions. Several new types of ionosonde are being built which are expected to overcome current limitations by providing information in digital form (BIBL and OLSON, 1967) or by using pulse-coded pulse compression techniques to improve signal to background discrimination (COLL and STOREY, 1964,1965). Continuous wave ionosondes were proprosed by FENWICK and BARRY in 1967 and J.W. WRIGHT at Boulder U.S.A. is developing the dynsonde and kinesonde techniques, a brief description of which may be found in ECCLES and KING, (1970).

The partial reflection of radio waves in the approximate frequency range from 2 to 6 MHz by the lower ionosphere is the basis of another valuable technique. The basic experimental measurement is the ratio of the relative amplitudes of extraordinary and ordinary waves

as a function of height. From this the electron collision frequency may be deduced for heights at which the differential absorption is negligible and the electron density may be deduced for heights at which the differential absorption is measureable. Measuring the phase between the two components rather than the amplitude has recently been employed by VAN BIEL et.al.(1969) and FRASER (1965,1968)and FRASER and VINCENT (1970) have attempted to use the partially reflected echoes to study D region and neutral atmospheric motions.

1-2-2: H.F. oblique incidence, ground based sounding techniques.

Four different techniques for probing the ionosphere use pulsed waves at oblique incidence. Two of these involve radar installations; the study of radio aurorae by means of auroral radar (transmitter and receiver are at a common site) and the study by meteor radar of meteor trails and their movements. A third technique is the use of high frequency ground backscatter in which radio waves reflected from the ionosphere are backscattered at the ground and return via the same paths to the transmitter. Swept frequency oblique sounding is the basis of the final technique. The auroral radar is used primarily to provide information on the nature of the ionosphere irregularities associated with aurorae while the meteor radar provides averaged information on atmospheric winds and turbulence and the electron and ion diffusion coefficients in the height range from 80-110 km (the results are usually assumed to apply to an average height of about 95 km). High frequency ground backscatter is used to determine the skip distance of obliquely propagated radio waves and the oblique ionosondes are used to

provide a real-time determination of the range of frequencies which can propagate over specific paths.

Techniques that use continuous waves at oblique incidence can be put into three main classes. These use the doppler shifts caused either by vertical ionospheric movements or by ionization changes, the forward scatter of VHF waves from the ionosphere, and the study of aurorae by systems in which the widely spaced transmitter and receiver are both on the equator side of the auroral irregulatities. The doppler technique is used for studying solar EUV radiation bursts and F2 layer drift velocities; forward scatter is used for the study of sporadic-E ionization and spread-F irregularities; and the radio aurorae are studied to determine their frequency and extent and possibly provide information about the auroral irregularities themselves.

1-2-3: L.F. ground based sounding techniques.

Probing techniques at L.F. and V.L.F. frequencies involve measuring one or more of the parameters such as amplitude, phase, time of flight and polarization of a radio signal as received at locations either near or far from a radiation source. Much of the present knowledge about the lower ionosphere has been deduced from observations of this kind.

Single frequency steep incidence measurements have been made in many parts of the world since the classic observations made at the Cavendish Laboratory were reported in 1951. Parameters studied include the apparent reflection height, h'(f), and the magnitude and phase of the various ionospheric conversion coefficients. DEEKS (1966) used the data from the Cavendish experiments to obtain a model

of the electron density distribution in the lower ionosphere. The Naval Electronics Laboratory Centre, San Diego, uses steep incidence pulses on three frequencies simultaneously to deduce a best fit electron density profile.

Oblique incidence methods range from simple single frequency single path observations to multifrequency multipath observations. The simplest observations can only identify changes in the ionosphere. The more complicated observations can give more information in principle, but for maximum use the data should be absolute rather than relative. The object is to be able to define the ionospheric structure along the path taken by the signal.

The International Union of Radio Science set up an ad hoc Working Group on Electromagnetic Probing of the Upper Atmosphere in 1968, to review all of the radio techniques used to study geophysical phenomena in the ionosphere. Their report, to which the reader is referred for more detailed information than could be given here, is contained in a special issue of the Journal of Atmospheric and Terrestial Physics (Volume 32, No. 4, April, 1970).

1-3: The lower night-time ionosphere at middle latitudes.

While more attention has been paid to the F region, a considerable number of attempts have been made to measure directly or indirectly parameters in the D and E regions. Ground based low frequency measurements have provided most of the information but have been supplemented more recently by techniques based on partial reflection of high frequency signals and by measurements made during rocket flights. As the number of experimental results increased, more and more work has been devoted to the establishment of theoretical

models of the D and E regions capable of explaining these results. This has required attention to the chemistry of these regions; the D region can no longer be considered to be solely the result of solar photo-ionization of NO and the importance of atomic oxygen, ozone and water vapour and of ionization by cosmic rays must be recognised (DOHERTY, 1970).

The published work on the D and E regions has concentrated largely on the day-time conditions; only a fraction consider the night-time regions at mid-latitudes. As mentioned before, DEEKS (1966) used the data from the classic Cavendish observations to deduce experimental electron density distributions, some of which were for night-time mid-latitude conditions. Other experimental electron density distributions have been produced by GARDNER and PAWSEY (1953) and MECHTLY and SMITH (1968). RADICELLA (1968) has produced theoretical models of electron and ion densities in these regions for both low and medium solar activity. He compared electron density results with the three references cited above and ion density results with the results of HALE et al (1967) and SAGALYN et al (1967). The electron density distribution in the D region during the night and pre-sunrise period at mid latitudes and a fairly high level of solar activity has been deduced by THOMAS and HARRISON (1970). They used a full wave integration method to compute values from a model electron density distribution and then varied this model to achieve a best fit to the radio propagation measurement made by BAIN et al (1952), STRAKER (1955), BRACEWELL et al (1954) and WEEKS and STUART (1952 a,b.).

Given that the ionosphere parameters vary with both time and position and that electron density distribution

cannot be deduced directly from ground based measurements it is not surprizing that absolute agreement between theory and experiment has not been found. The mid-latitude, night-time electron density is shown in figure i as a function of height and in figure ii as a function of both height and solar zenith angle. A common feature of all the distributions in figure i is the change of about two orders of magnitude in the electron density between altitudes of 80 and 90 km, followed by a small electron density gradient for the next 5-10 km. Above this the electron density again increases more rapidly. Figure ii suggests that the two order increase in electron density forms quickly after sunset and disappears equally quickly at sun-rise. The reason for this sudden increase is not clearly understood but it has been suggested that it may be due to the transport of nitric oxide into this region (GEISLER and DICKINSON, 1968).

Both spatial and temporal small-scale variations of electron density are known to exist in the ionosphere and this stratification during both day and night has been the subject of considerable work. HELLIWELL (1949) and HELLIWELL et al (1951), observing at 100 kHz and 325 kHz, report a multiplicity of strata between 90 and 130 km. A review paper on the methods and results of a number of workers up to about 1956 reaches the conclusion that one stratum at about 85 km is observed consistently (ELLYETT and WATTS, 1959).

Fewer workers have reported observations of irregularities at heights of about 85 km. FRASER and VINCENT (1970) using a vertical pulse experiment at 2.4 MHz have studied D region strata and irregularities at times

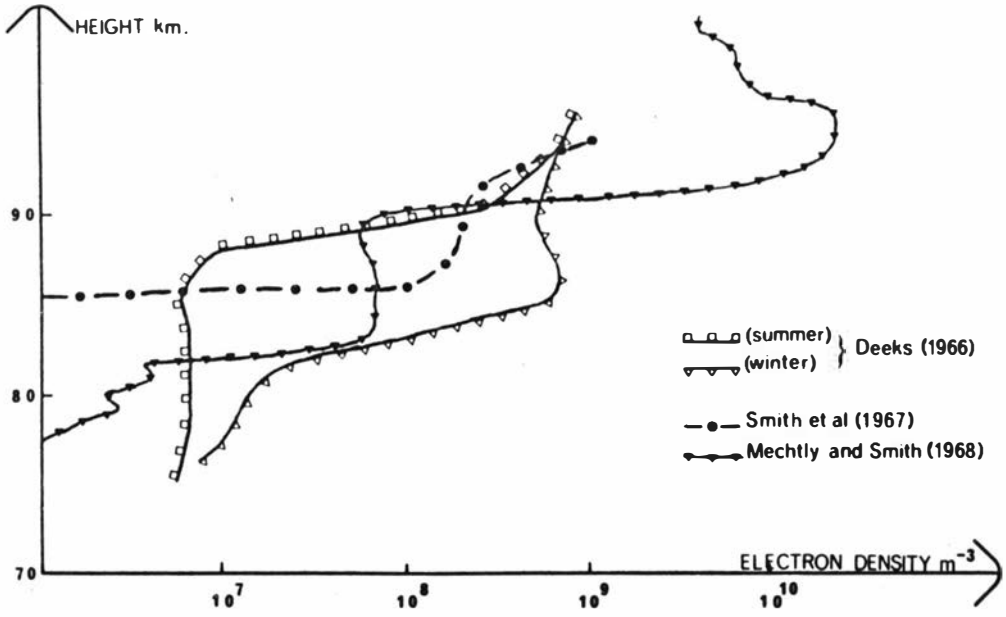
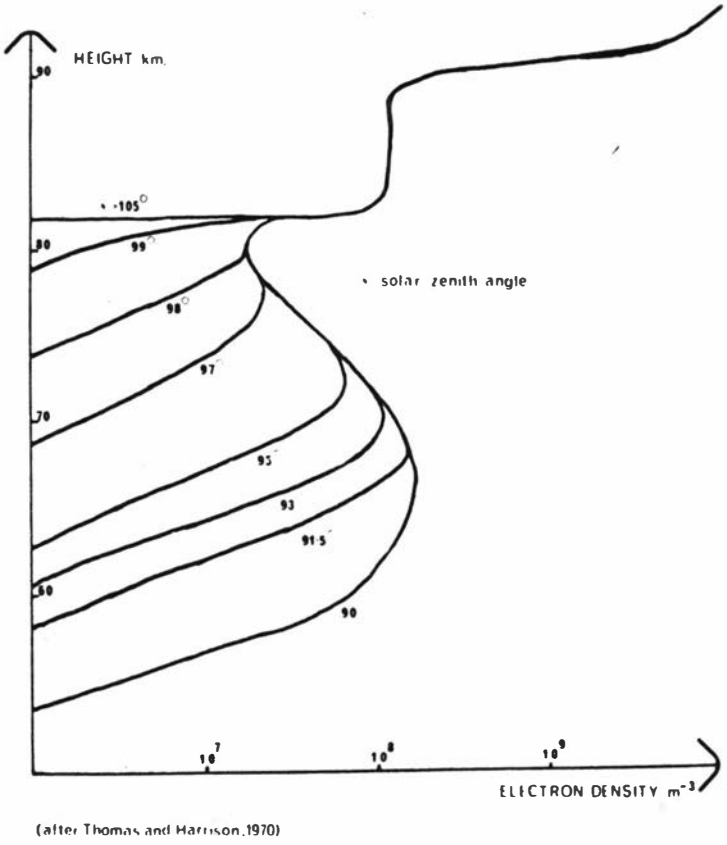


Figure i: Experimental, mid-latitude, night-time electron density distributions.



(after Thomas and Harrison,1970)

Figure ii: Mid-latitude, night-time electron density distribution as a function of the solar zenith angle.

close to local noon. They report observing coherent reflections from moving irregularities at heights of approximately 80 km. GOSSARD and PAULSON (1968), also using pulsed transmission, report periodic fading from the D region. Most of this they attribute to interference from an off-path reflector, but one case at about local midnight they interpret as evidence of a first mode gravity wave whose kinetic energy resides mainly in the upper sound channel. FELGATE and GOLLEY (1971) report some observations of D region irregularities using a pulsed transmitter at 1.98 MHz and a large (178 dipole) receiving array. They do not state the time at which measurements were made or report detailed results, stating only that the scale (ground distance between pattern maxima) of the D region patterns observed was considerably smaller than the scale of E-region patterns. The only case for which figures were quoted had a scale of about 200 m. They classified all D region patterns as random.

A further feature of the night-time ionosphere is the local minimum of temperature, and hence of sound speed, at about 85 km (GEORGES, 1967, page 30). It is well known that acoustic waves can be reflected and ducted in regions of minimum sound speed. The same author gives, in his section 3.9.1, methods for calculating the absorption of an acoustic wave in the atmosphere. A wave of given frequency reaches maximum amplitude at a height at which the absorption begins to dominate over the exponential increase in amplitude due to the decreasing air density. An inspection and extrapolation of the curves given suggests that at a height of 85 km waves with periods of the order of 5 seconds may reach their maximum

amplitude. As GEORGES points out, wave reflection, ducting and effects of ionospheric winds have been ignored and thus the above figure of five seconds must be considered as only indicating the order of magnitude.

In conclusion it might be stated that, whereas a medium frequency signal propagating obliquely upwards during the day-time is heavily absorbed by the D region, at night-time the same signal travels upwards through a region of low electron density and is then incident on a fairly sharp boundary across which the electron density changes by about two orders of magnitude. At approximately the same height there is a local minimum of sound speed, and the possibility of reflection and ducting of acoustic waves. At this height, also, waves with a period of the order of five seconds may reach their maximum amplitude.

1-4: A physical process and a postualted mathematical model.

In order to explain the periodic events noticed, a process must be postualted which is physically reasonable in light of the known facts about the ionosphere. The high levels of sky wave are presumably due to reflections from many points arriving at the receiver. These points must be inclined to the horizontal at an angle appropriate to their spatial coordinates, and thus the smooth ideal mirror ionosphere must be disturbed. Since the sky wave amplitude changes relatively slowly and consistent sequences of events can be observed over a period of several minutes these perturbations must show significant stability. Random perturbations of a scale size much less than the wavelength of the probing radio signal are extremely unlikely to produce sky wave amplitudes of the magnitude observed. The periodic nature of some

events suggests interference is occurring between different reflected rays and this requires a deformation of the ionospheric mirror coherent over a finite spatial extent.

The physical process postulated here for the periodic variation of the ionospheric reflection height is as follows. A microbarum, assumed to occur as a short burst a few wavelengths long, originates some two or three thousand kilometers away from a transmitter radiating signals in the low medium frequency range (around 0.6 MHz). Examination of a typical spectrum of microbarums (DOWN, 1967) shows that this wave will have a period of approximately 4 secs. It rises at a shallow angle into the ionosphere and is trapped by the local temperature minimum at about the same height as the sudden increase in electron density (approximately 85 km). Accepting a velocity of sound of 250 ms^{-1} at this height, the wavelength of the microbarum will be about 1 km. Due to ionospheric winds the path of the microbarum from the source to the vicinity of the experiment would not be a straight line if projected onto the surface of the earth. However the microbarum is assumed to travel in a straight line over the short distance between the transmitter and the receiver.

The sudden increase in electron density mentioned before is affected by the passage of the ducted microbarum which distorts this boundary from plane. The pressure wave causes changes in the ionization concentration and thus modifies the reflection coefficient. Radio waves from the transmitter incident obliquely onto this boundary are reflected with significant reflection coefficients. In the vicinity of the periodically distorted boundary there are a number of reflection points via which signals can travel from transmitter

to receiver. The vector sum of the amplitudes of these sky waves gives the received skywave amplitude which will thus vary as the microbarum moves. Since the disturbing phenomenon (the microbarum) is not random the variations in the received sky wave may be expected to be themselves non-random and therefore relatively easily identified.

In order to express this physical description in a mathematical form suitable for investigation certain simplifications must be made. The first of these is to use ray optics. Since the wavelength of the probing signal is only a few times smaller than the wavelength of the acoustic wave, clearly the use of ray optics under these conditions is not above suspicion. There are various criteria for the point at which ray optics may be said to break down, one simple one being that ray optics are not valid if the number of Fresnel zones in an area is less than the number of reflection points in that area. Accepting this criterion and using values appropriate to this experiment, ray optics is (at worst) not valid when the acoustic wave group is between 45 and 65 km from the transmitter. This, as can be seen from the frontispiece, is also the region for specular reflection. There are, however, no definite boundaries; the approximation involved in the use of ray optics becoming poorer the more the above condition is violated. Provided the condition is not grossly violated it may be hoped that the coarse features predicted when the acoustic group is in this region will still be recognisable. For the above reasons and in the interests of mathematical tractibility ray optics have been used at this stage of the investigation.

Three other approximations will be introduced now,

others will be introduced, and justified, as needed later. In order to simplify the arithmetic a flat earth is assumed. This will hardly introduce a significant error when the greatest distances considered are approximately 100 km. All ionization at heights below the sudden night-time increase is ignored. Although there will be partial reflections from this ionization, the reflection coefficients will be extremely small, as will the amount of bending of the radio waves. The short acoustic wave group is idealized to consist of a whole number of wavelengths with a constant amplitude. This obviously will not correspond closely to practical conditions but is adopted for mathematical simplicity.

SECTION TWO: MATHEMATICAL SOLUTION OF THE IDEALIZED MODEL.

In this section equations are derived from which an approximate, but reasonable, continuous curve of theoretically predicted sky wave amplitude versus group position may be produced. Firstly the effect of the passage of an acoustic wave along the ionization boundary is calculated. This completes the derivation of the idealized model and is followed by the derivation of the many factors which affect the received sky wave amplitude. For convenience these factors are divided into two types; factors due to the perturbed boundary and factors due to the experimental arrangement. The former are calculated first and consideration is given to the approximations involved in these calculations before the later class of factors are calculated. The range over which receivable reflections can occur is then considered briefly. This is followed by a discussion of the extension of the above calculation to give a continuous curve of expected sky wave versus group position. Finally brief mention is made of the extension of the above to perturbing groups of arbitary length.

2-1: The deformation of the ionization boundary by the passage of an acoustic wave.

The effect of the trapped acoustic wave on the ionization boundary must be determined before ray optics can be used. This is done in two stages: the particle velocities induced by the acoustic wave are calculated using a method similar to that used by KANTOR and PIERCE (1968). The effect of these induced velocities on the previously plane boundary is then considered.

Consider an acoustic wave propagating in the x direction through a homogenous weakly ionized gas. The wave

has angular frequency ω, wave number k and neutral particle velocity $\underline{v}_n$. The wave is assumed to be mechanical in nature, but introduces electron and ion velocities ($\underline{V}_e$ and $\underline{V}_i$) via collisions. The earth's magnetic field $\underline{B}_o$, making an angle σ with the x axis and lying in the x-z plane, will affect the ion and electron motion, as will the induced electric field $\underline{E}$.

By solving Maxwell's equation for $\underline{E}$ the induced field may be expressed in terms of $\underline{V}_e$ and $\underline{V}_i$. Assume that the current density is $eN_p(\underline{V}_i - \underline{V}_e)$, when N_p is the ambient charged particle number density (assumed equal for both ions and electrons) and that all first order quantities associated with the disturbance vary with t and x as $e^{j(\omega t - kx)}$. Using Maxwell's equations

$$\nabla \times \underline{B} = j\omega\varepsilon_o\mu_o\underline{E} + \mu_o\underline{J}$$

and $$\nabla \times \underline{E} = -j\omega\underline{B}$$

it follows that:

$$\nabla \times (\nabla \times \underline{E}) = \omega^2\varepsilon_o\mu_o\underline{E} - j\omega\mu_o\underline{J} .$$

Since $\underline{E} = \hat{x}\,E_x\,e^{j(\omega t - kx)} + \hat{y}\,E_y\,e^{j(\omega t - kx)} + \hat{z}E_z\,e^{j(\omega t - kx)}$

$$\nabla \times (\nabla \times \underline{E}) = k^2 E_y\, e^{j(\omega t - kx)}\hat{y} + k^2 E_z\, e^{j(\omega t - kz)}\hat{z}$$

Hence, solving for E_x, E_y and E_z gives:-

$$E_x = \left[j eN_p\,(V_{ix} - V_{ex})\right] / (\omega\varepsilon_o) \qquad (1)$$

$$E_y = \left[-j\omega\mu_o e\,N_p\,(V_{iy} - V_{ey})\right] / \left(k^2 - (\omega/c)^2\right) \qquad (2)$$

and $$E_z = \left[-j\omega\mu_o\,eN_p\,(V_{iz} - V_{ez})\right] / \left(k^2 - (\omega/c)^2\right) \qquad (3)$$

where $c = (\mu_o\varepsilon_o)^{-\frac{1}{2}}$ is the speed of light.

The equations governing the motion of electrons, ions and neutrals are taken to be those for a three component

gas. Using representative orders of magnitude one may easily show that the pressure forces due to ions and electrons are insignificant, and these are therefore neglected. The first order equations of motion for the charged particles are therefore

$$\nu'_{in} (\underline{V}_i - \underline{V}_n) = (e/m_i) (\underline{E} + \underline{V}_i \times \underline{B}_o) \qquad (4)$$

and $$\nu'_{en} (\underline{V}_e - \underline{V}_n) = -(e/m_e) (\underline{E} + \underline{V}_e \times \underline{B}_o) \qquad (5)$$

where ν'_{in} and ν'_{en} are effective ion-neutral and electron-ion frictional collison frequencies defined by the momentum equation

$$m_i \frac{\partial \underline{V}_i}{\partial t} = m_i \nu'_{ij} (\underline{u}_j - \underline{u}_i)$$

As a consequence of this $\nu'_{ij} = \frac{m_j}{m_i} \nu'_{ji}$

and

$\nu'_{ij} = \frac{m_j}{m_i + m_j} \nu_{ij}$ where ν_{ij} is the (normal) average collison frequency.

Since the ion and neutral particle masses are nearly equal ν'_{in} is approximately half the average collision frequency of an individual ion with neutrals. However, ν'_{en} may be interpreted directly as an average collision frequency.

Simplifications can be introduced. Acceleration terms and electron-ion collisions are neglected in the equations (4) and (5) since, for moderate frequencies propagating in the lower ionosphere at night, ω is very much less than both ν'_{en} and ν'_{in} and the number of ions (assumed singly positively charged) is small compared to the number of neutrals. Further, because of the small

density ratio of ions to neutrals, the effect of the induced ion and electron motions on the neutral motions is negligible, and we may consider $\underline{V}_n$ to be the particle velocity associated with an acoustic wave propagating in the x direction. Since the charged particles should have negligible influence on the speed of the wave we take $K = \omega/a$, where a is the speed of sound. The term $(e/m_i)\,\underline{V}_i \times \underline{B}_o$ in equation (4) can be neglected in comparison with the ion neutral term as the ion-neutral collision frequency is much greater than the ion cyclotron frequency at altitudes around about 85 km.

Then using $k^2 = \omega^2/a^2$ and $a^2 << c^2$ equations (1), (2), (3) and (4) can be solved for the ratios V_{iz}/V_{ez} and V_{iy}/V_{ey} to give:-

$$V_{iz}/V_{ez} = V_{iy}/V_{ey} = j\alpha / (\nu'_{in} + j\alpha)$$

where $\alpha = e^2 \mu_o N_p a^2/(m_i \omega)$

α/ν'_{in} is the order of 10^{-4} for 1 Hz at 85 km and we may conclude that any current at right angles to the wave normal will be carried almost entirely by electrons.

In the z and y component equations the ratio of the term $(e/m_e)\underline{E}$ to the term $\nu'_{en}\underline{V}_e$ has magnitude $e^2 \mu_o N_p a^2/(m_e \omega \nu'_{en})$. For 1 Hz this ratio is 10^{-3} at 100 km and 10^{-4} at 90 km, and thus the terms E_z and E_y in equation (4) can be neglected. However E_x cannot be neglected and remains the sole significant mechanism for coupling of the electron and ion motions.

Equation (5) may be solved to find:

$$V_{ez}/V_{ey} = \omega_{eL}/\nu'_{en} \tag{6}$$

$$V_{ey}/V_{ex} = -\nu'_{en}\,\omega_{eT}/(\nu'^{2}_{en} + \omega_{eL}^{2}) \tag{7}$$

$$\frac{V_{ex}}{V_{nx}} = \frac{\left[R\,\nu'_{en}(1+j\delta) + \nu'_{in}\right](\nu'^{2}_{en} + \omega_{eL}^{2})}{\nu'_{in}(\nu'^{2}_{en} + \omega_{eL}^{2}) + R\,\nu'_{en}(1+j\delta)(\omega_{e}^{2} + \nu'^{2}_{en})} \tag{8}$$

and $$\frac{V_{ix} - V_{ex}}{V_{nx}} = \frac{jR\delta\,\nu'_{en}\,\omega_{eT}^{2}}{\nu'_{in}\,(\nu'^{2}_{en} + \omega_{eL}^{2}) + R\nu'_{en}\,(1+j\delta)(\omega_{e}^{2}+\nu'^{2}_{en})}$$

where $R = m_e/m_i$ and $\delta = \omega m_i \nu'_{in}\,\varepsilon_o/(N_p e^2)$

$\lesssim 10^{-4}$ for 1 Hz at 85 km,

and thus $V_{ix} - V_{ex} \ll V_{nx}$

Physically this means that the x components of the electron and ion displacement are approximately equal in order to reduce the magnitude of the induced field E_x.

At 85 km $\nu'_{en} \ll \omega_e$ and $R\,\nu'_{en} \ll \nu'_{in} \ll \nu'_{en}$

Provided cos σ is not small then from equations (6), (7) and (8)

$V_{ez}/V_{nx} \approx \tan\sigma$

$V_{ey}/V_{nx} \approx 0$ and $V_{ex}/V_{nx} \lesssim 1$.

Hence we see that under these conditions the electrons will tend to move along the magnetic field lines. This might be termed a quasi-longitudinal approximation. Plotting these results shows them to be valid at heights of 85 km for values of σ at least up to 70°.

Now consider the passage of an acoustic wave, of large vertical extent, travelling in the $+x$ direction along a boundary at a height $z=H$ when the electron density concentration per unit volume is N above the boundary and zero below. In the absence of the wave these electron densities are uniform in the x and y directions.

The displacement of a neutral particle within the wave may be described by

$$\Delta x = A' \sin(\omega t - kx).$$

Hence

$$V_{nx} = V_{ex} = A' \omega \cos(\omega t - kx)$$

and

$$V_{ez} = A' \omega \tan \sigma \cos(\omega t - kx).$$

The vertical displacement of a particle whose "rest" coordinates are x, y, H is therefore

$$\Delta z = A' \tan \sigma \sin(\omega t - kx).$$

The previously plane boundary has therefore been distorted into a sinusoidal form, the sinusoidal perturbation having the same wavelength and velocity as the acoustic wave, but whose amplitude A is $= A' \tan \sigma$. The electron density along the boundary is no longer uniform, but is a function of x. For plane waves propagating in the x direction the continuity equation for electrons has the form (BOOKER and DYCE, 1965)

$$\Delta N/N = (k/\omega) V_{ex}$$

$$= k A' \cos(\omega t - kx) \qquad (9)$$

where $\Delta N/N$ is the fractional change in the electron density concentration.

No physical interpretation can be made if $\Delta N/N < -1$ which can happen if the amplitude of the acoustic wave is too high. In terms of the wavelength to amplitude ratio of boundary perturbation (L/A) this occurs for values of L/A less than approximately 2.5 .

The effect of this variation in electron density will be considered later.

2-2: Factors affecting the received sky wave due to the perturbed boundary.

The expected received field strength of the sky wave component depends on many factors. The factors to be calculated immediately are those due to the sinusoidally perturbed boundary itself.

Figure iii illustrates the geometry of the idealized model and the quantities labelled on it are defined as follows:-

H = height of unperturbed boundary above the ground.

D = distance from transmitter to receiver.

L = wavelength of sinusoidal perturbation.

A = amplitude of sinusoidal perturbation.

x = horizontal coordinate, measured from the transmitter.

y = vertical coordinate, measured from the ground.

θ = launch angle of a particular ray.

β = gradient of the boundary at the point of reflection of this ray.

Ne = unperturbed electron density above the boundary.

λ = wavelength of probing radio wave.

2-2-1: Interference terms.

It has been shown that the effect of an acoustic wave travelling horizontally along a sharp discontinuity in ionization produces a sinusoidal variation in the reflection

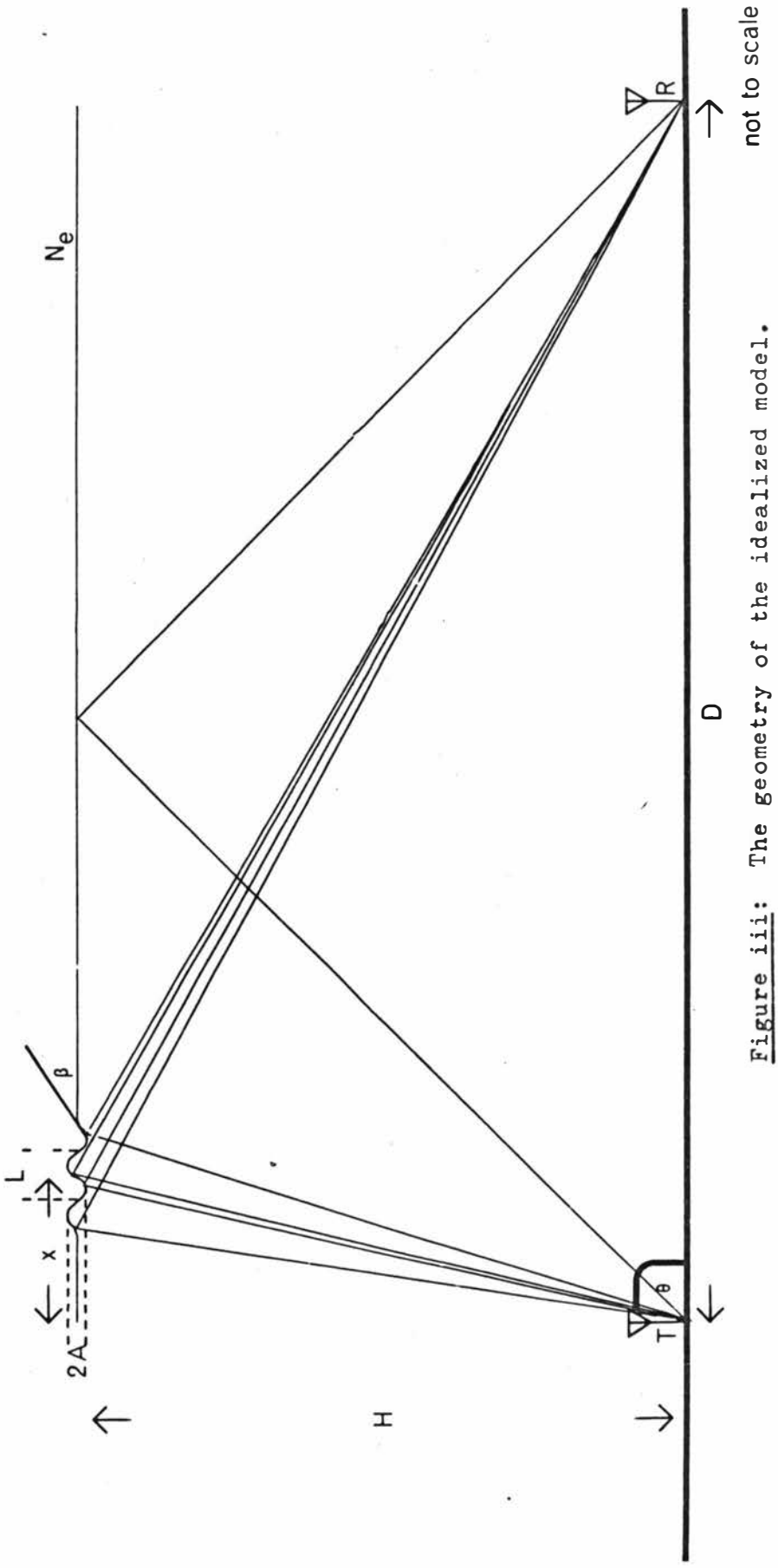


Figure iii: The geometry of the idealized model.

height. The night-time ledge of ionization which exists at a height of approximately 85 km is considered to cause a sharp discontinuity in the refractive index and the passage of an acoustic wave travelling horizontally along this discontinuity has been shown to cause a sinusoidal variation in the reflection height.

This perturbation may be written at $t = o$ as,

$$y = H + A \sin(2\pi x/L) . \qquad (10)$$

Then a ray incident at an angle Θ to the horizontal will be reflected from a general point (x_1 , y_1) at an angle $\Theta - 2\beta$ from the horizontal, where

$$\tan \beta = dy/dx = (2\pi A/L) \cos(2\pi x/L) . \qquad (11)$$

If the observing point on the ground is a long way from the reflecting surface, then all rays arriving at that point from a small range of x can be considered parallel. These will have come from points within this range of x for which β is the same. These points can be found for a given value of β by solving equation (11). The validity of this approximation will be considered later.

Let $2\pi x/L = 2m\pi + \delta$ where m is some integer, then the general solution of equation (11) is

$$2\pi x/L = 2n\pi \pm \delta \quad \text{where } n \text{ is any integer.} \qquad (12)$$

Equation 10 gives as solutions for x two series, one of which can be labelled + and corresponds to the $+\delta$ series, and the other which can be labelled - corresponding to the $-\delta$ series. The horizontal spacing between x_n^+ and x_{n+1}^+ or x_n^- and x_{n+1}^- is L, and the horizontal spacing between x_n^+ and x_n^- is $\delta L/\pi$. Since all the rays in the

group considered come from points for which both β and θ are considered constant, the reflection coefficient has the same (generally complex) value for each reflection. The contribution from a pair of points x_n^+ and x_n^- can be combined to give a composite disturbance from these two points, and then the total disturbance due to some number n of these paired points can be found, each of the paired points being displaced from the next paired point by a horizontal separation L .

See figure iv . Let all phases be measured relative to the phase of the disturbance from x_1^- .

The extra path length to x_1^+ is

$$\left\{(x_1^+ - x_1^-)^2 + (y_1^+ - y_1^-)^2\right\}^{\frac{1}{2}} \sin(\alpha + \theta)$$

where $\tan\alpha = (x_1^+ - x_1^-)/(y_1^+ - y_1^-)$

The extra path length from x_1^+ compared to the path from x_1^- is

$$-\left\{(x_1^+ - x_1^-)^2 + (y_1^+ - y_1^-)^2\right\}^{\frac{1}{2}} \sin(\alpha - \theta + 2\beta)$$

Combining the above two results, the path to and from x_1^+ is longer than the path to and from x_1^- by

$$2\left\{(x_1^+ - x_1^-)^2 + (y_1^+ - y_1^-)^2\right\}^{\frac{1}{2}} \cos(\alpha + \beta) \sin(\theta - \beta)$$

Since

$$x_1^+ - x_1^- = \frac{\delta L}{\pi}$$

and (using equation (12) and equation (10))

$$y_1^+ - y_1^- = 2A \sin\delta$$

hence

$$\tan\alpha = (\delta L/2\pi A \sin\delta) .$$

Substituting the above relations and manipulating, it follows that the signal received via reflection from x_1^+ lags the signal reflected via point x_1^- by an angle ϕ given by

$$\phi = \frac{4}{\lambda}\left[2A\pi \sin\delta \cos\beta - \delta L \sin\beta\right]\sin(\theta - \beta). \qquad (13)$$

For reflecting points within the same series, $\Delta x = L$, and $\delta = \pi$. Thus substituting into (13) the signal from $x_{(n-1)}$ lags the signal from x_n by 2γ where

$$2\gamma = \frac{-4\pi}{\lambda} L \sin\beta \sin(\theta - \beta) \qquad (14)$$

Using equations (11) and (12) it follows that

$$\tan\beta = 2\pi A \cos\delta / L \qquad (15)$$

and substituting (14 and (15) into (13), ϕ can be expressed in terms of δ and γ giving:-

$$\phi = 2(\delta - \tan\delta)\gamma/\pi \qquad (16)$$

Using the standard result for the interference between two slits, the composite disturbance from the paired point x_n due to the contributions x_n^+ and x_n^- can be expressed as being proportional to

$$\sin\phi/\sin(\phi/2) \qquad (17)$$

The composite disturbance from all the x_n in a perturbation group can similarly be expressed as being proportional to

$$\sin n\gamma/\sin\gamma \qquad (18)$$

There will be one x_n per wavelength of perturbation and thus n represents the number of wavelengths in the perturbation group.

2-2-2: The validity of the approximation used in the above section.

In the above section it has been assumed that β is a constant for all reflection points within a perturbation wave group. It is now important to consider the limitations imposed by this approximation in order that the maximum group size consistant with this approximation can be determined, and in order that the analysis can be extended beyond a single group.

In general there are two reflection points per wavelength. Consider the two wavelengths shown in figure v , where x marks the centre position of the group for which the value of β has been evaluated from equation (29), (see 2-3). The value of β evaluated at x occurs at positions a, b, c and d as shown. However the true reflection points are either slightly greater values of β (cases a and b, shown as a' and b') or at a slightly lower value of β (cases c and d , shown as c' and d'). The variation required in β is shown highly exaggerated and will in practice be only a very small fraction of a degree. Consider the region between a and b (or c' and d'). The picture can be considered symmetrical about the point of inflexion since the required β is essentially a linear function of x over the range of x used.

The error in assuming that the rays from a' and b' are parallel will still be negligible when the error in assuming that the rays from a' and d' are parallel is starting to become significant. Since the reflection coefficient „ R „ is a slowly varying function of the angle of incidence, it may be considered constant over a very small range of

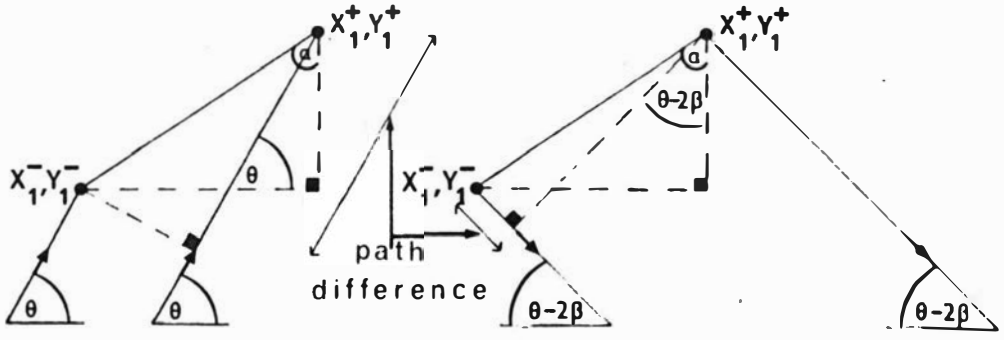


figure iv

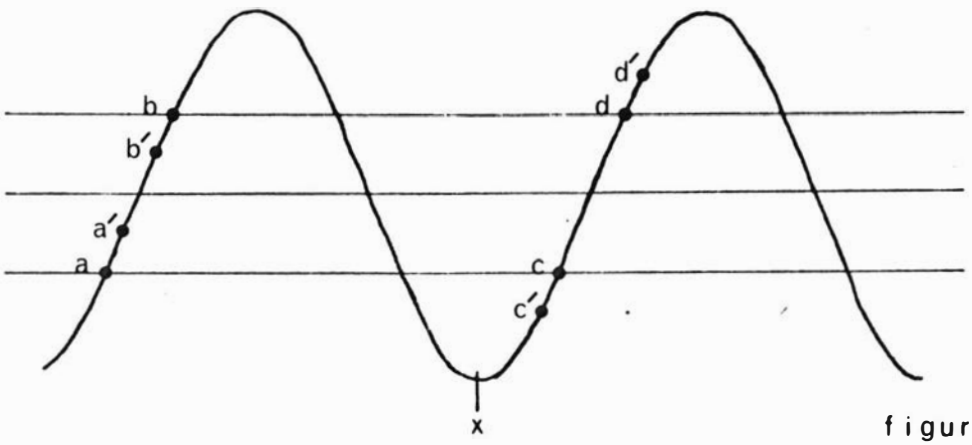


figure v

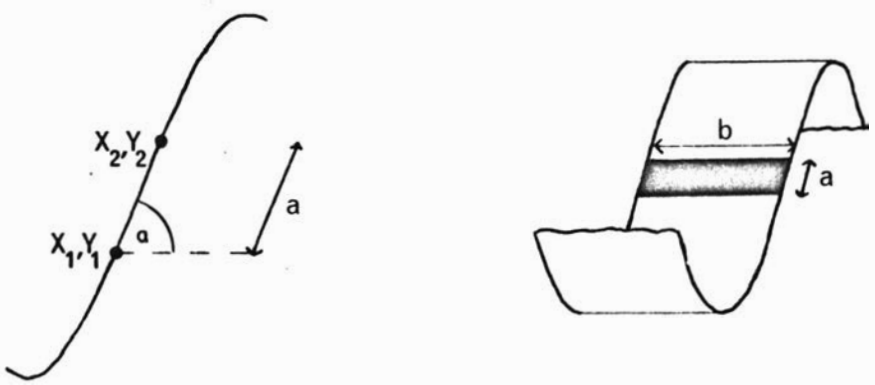


figure vi

angle. It may be considered linear over a larger range of angles where it may be no longer be assumed constant.

Thus, on adding the two rays from a' and b', the phase error introduced by the different path lengths is cancelled out. However, although the values of β at a' and b' will only vary very slightly from one another, they may both differ rather more from the value of β for position x . In this case, although there would be no significant phase error on combining the rays from a' and b' , there would be an error in the amplitude. This is likely to become more serious as the number of wavelengths in the group is increased. Never-the-less it follows that any amplitude error on combining the rays from a' and b' will, to a good approximation, be equal and opposite to the amplitude error on combining the rays from a' and d' . Hence neglecting the extra inverse distance factor (approximate since we are considering amplitude) the total amplitude and phase errors over a group containing an even number of wavelengths cancel out.

For the geometry of this problem, D = 110 km and H = 85 km , a calculation shows for a perturbation wavelength of 1 km and a perturbation amplitude of 0.1 km the error is of the order of 10^{-4} of a wavelength at 570 kHz .

Thus, provided the group considered is symmetrical about the point x for which β is worked out, and the group contains an even number of wavelengths, little error is introduced by the above analysis even for groups containing up to, say, ten wavelengths. The values θ and β corresponding to the central position x of the group should be used when calculating „ R„ .

2-2-3: The reflection coefficient.

Since the transmitter in the experiment described transmits a vertically polarized signal and the receiving aerials used are only sensitive to vertical polarization, only one of the four possible reflection coefficients is of interest. This is $_{\prime\prime}R_{\prime\prime}$ and depends, among other things, on the values of θ and β and also on the electron density at the point of reflection. In general this reflection coefficient is complex. JONES (1970) has pointed out that the ray approach may be extended down to frequencies of about 100 kHz by the consideration of the complex part of the reflection coefficient (due to the complex part of the refractive index).

Consideration of the collision frequency (approximately 10^5 sec^{-1} at 85 km at night (JONES, 1970)) shows that this has an insignificant effect on the value of the refractive index. The effect of the earth's magnetic field cannot, however, be neglected a priori. Theory shows that the reflection coefficient $_{\prime\prime}R_{\perp}$ will not be zero if the effect of the earth's magnetic field is significant. This has been experimentally investigated using an aerial sensitive only to horizontal polarization. Tests carried out using signals at 570 kHz showed only small values of horizontal signal strength, typically a few percent of the vertically polarized reflected skywave signal. This is consistent with an extrapolation of the results obtained from the rigorous application of the magneto-ionic theory to the calculation of the reflection coefficients for a sharply bounded ionospheric model by JOHLER and WALTERS (1960). Thus it may reasonably be assumed that the effect

of the earth's magnetic field may be ignored in this experiment for the purposes of calculating the refractive index of the ionosphere at the height of interest.

Since the reflection is occuring (in the model) at a sharp discontinuity in the refractive index, the reflection coefficient is simply that predicted by the Fresnel formula for oblique incidence and no magnetic field. Idealizing the boundary so that the refractive index above the boundary is n and the refractive index below is unity, this reflection coefficient is (see, for example, Budden "Radio Waves in the Ionosphere", chapter 8),

$$"R" = \frac{n^2 \cos\mu - (n^2 - \sin^2\mu)^{\frac{1}{2}}}{n^2 \cos\mu + (n^2 - \sin^2\mu)^{\frac{1}{2}}} \qquad (19)$$

where μ is the angle of incidence. In terms of the geometry outlined at the beginning of this section

$$\mu = \pi/2 - (\theta - \beta) \qquad (20)$$

and thus

$$"R" = \frac{n^2 \sin(\theta - \beta) - (n^2 - \cos^2(\theta - \beta))^{\frac{1}{2}}}{n^2 \sin(\theta - \beta) + (n^2 - \cos^2(\theta - \beta))^{\frac{1}{2}}} \qquad (21)$$

It remains to calculate the refractive index n above the boundary. This depends on the electron density and it has been shown that the electron density along the boundary is modified by the passage of an acoustic wave in an x direction along the boundary.

The variation in electron density is given by (equation (9))

$$\Delta N = N k A' \cos(\omega t - kx)$$

which can be written at time t=0 as

$$\Delta N = N k A' \cos(kx).$$

Using equation (11) and recalling that $A = A' \tan\sigma$

$$\Delta N = N \tan\beta / \tan\sigma .$$

Let N_e be the (quiescent) electron density above the boundary before the passage of the acoustic wave and N_e^* the electron density during this passage.

Then $N_e^* = N_e(1 + \tan\beta/\tan\sigma)$.

Let n be the quiescent refractive index at the boundary. This may be defined by

$$n^2 = 1 - X_1 = 1 - (f_p/f)^2$$

when $f_p^2 = K N_e$ and f is the radio wave frequency. Then n^*, the refractive index above the boundary during the passage of the acoustic wave, may be defined in a similar way, as

$$(n^*)^2 = 1 - \frac{N_1 K(1 + \tan\beta/\tan\sigma)}{f^2}$$

$$= n^2 - X_1 \tan\beta/\tan\sigma \qquad (22)$$

Thus from an assumed value of N_e, the quiescent value of refractive index n may be calculated as K is known (K is approximately $81\ m^3$ if N_e is expressed as electrons per cubic metre). Thus using equation (22) the value of the perturbed refractive index n^* may be calculated, which when substituted into equation (21) give the reflective coefficient "R".

2-2-4: <u>The reflecting area</u>.

In the preceeding part of this section reflection from points has been considered. Consideration is now

given to reflection from finite areas. In order to get an idea of the magnitude of the correction term required, the reflecting area is calculated in an approximate way. The sinusoidal perturbation of the reflecting layer is considered to be of infinite extent in the z direction. The size of the reflecting area depends on the height a, width b, and inclination α of the area as shown in figure vi. The approximate method used to calculate the factor a is as follows. It is assumed that the receiver has finite extent in the x direction of magnitude 2R, symmetrical about the point a distance S from the transmitter. Since the distances from the transmitter to the centre of the receiver (S) and to the furthest point of the receiver (S+R) are not the same, different values of β will be required for transmissions to these two points. Let these be β_1 and β_2. Let the sinusoidal perturbation be centred at $x = x_1$. Knowing the description of the perturbation at t=0, the principal value of x (called x_{p1}) corresponding to this value of β_1 can be calculated. A close approximation to the x coordinate of the reflection point is therefore $x_1 + x_{p1}$. Similarly a close approximation to the x coordinate of the reflection point for rays reaching the furthest extreme of the receiver is $x_1 + x_{p2}$.

Substitution of these values of x into the defining equation of the perturbation gives the corresponding values of y. Neither of these coordinates accurately describes the true reflecting points since the values of x and y differ from those used when calculating β. However the discrepancies will be small and the errors will have the same sign unless d^2y/dx^2 changes sign between

the two reflection points calculated. The situation is avoided by using one half the receiver dimension in the x direction (reception at the two points x =S and x= S+R). Thus even when subtracting the x and y coordinates of the two calculated reflection points as a first step to obtaining the distance between them, the error introduced is likely to be small. If further accuracy were desired an iterative process based on the above could be pursued, but, in view of the previous assumptions made in this work, such a process is thought to be of little value here.

Writing the above in terms of equations derived before:-

From (29) (see 2-3), the two values of β are given by

$$\tan 2\beta_1 = H(S - 2x_1)/(S_{x1} - x_1^2 + H^2)$$

and

$$\tan 2\beta_2 = H(S + R - 2x_1)/((S + R)x_1 - x_1^2 + H^2)$$

From (11) the x coordinates of the reflection points are given by

$$x_1 + x_{pn} = x_1 + (L/2\pi) \cos^{-1} (L \tan \beta_n/2\pi A) \quad \text{where}$$

n = 1 or 2 and the corresponding y coordinates are given by substitution into equation (10).

Hence for the first approximation the width of the slit is given by

$$((x_{p2} - x_{p1})^2 + (y_2 - y_1)^2)^{\frac{1}{2}} \qquad (23)$$

Since the slit is inclined at an angle to the perpendicular to the incident beam, a correction must be made which consists of multiplying by the cosine of this angle. Let α be the angle between the horizontal and

the line joining (x_1, y_1) to (x_2, y_2), measured positive clockwise. Then, with Θ defined as before, the correction is equivalent to multiplying by

$$\sin(\Theta - \alpha) \tag{24}$$

The dimension b is more easily calculated. Since the perturbation is considered to be of infinite extent in the z direction, reflection has no effect on motion along the z axis if the sine wave is moving along the x axis. The effective z dimension of the slit, b, is therefore related to the z dimension of the receiver R_1 by

$$b = (A/(A+B))R_1$$

where A is the distance from transmitter to reflection point and B is the distance from the receiver to reflection point. Thus b is proportional to

$$(x^2 + y^2)^{\frac{1}{2}}/((x^2 + y^2)^{\frac{1}{2}} + ((S-x)^2 + y^2)^{\frac{1}{2}}) \tag{25}$$

The product of equations (23), (24) and (25) gives the correction factor for effective slit size.

The reflecting area calculation above is only applicable to calculating the amplitudes of the various signals reflected from the sinusoidal part of the boundary to within a constant. Obviously a similar, simpler, calculation may be carried out to obtain, to within the same constant, a reflecting area for the reflection from the unperturbed point of symmetry. This reflecting area is given by

$$\text{area} = (R R_1 H)/(2\sqrt{4H^2 + \Delta^2}) \tag{26}$$

The absolute magnitudes of the amplitudes of the different skywaves depend on the assumed dimension of the receiving aerial, linearly for the case described by equation (26) and non linearly for the cases described by equations (23), (24) and (25). Since all these sky waves have to be summed in order to obtain the total received sky wave it is important to know the ratio of the 'flat' reflecting area to one of the reflecting areas on the 'wave' boundary. This ratio is plotted as a function of total aerial extent (AE) (i.e., 2R), and also as a function of the position of the sinusoidal section of the boundary in figure vii.

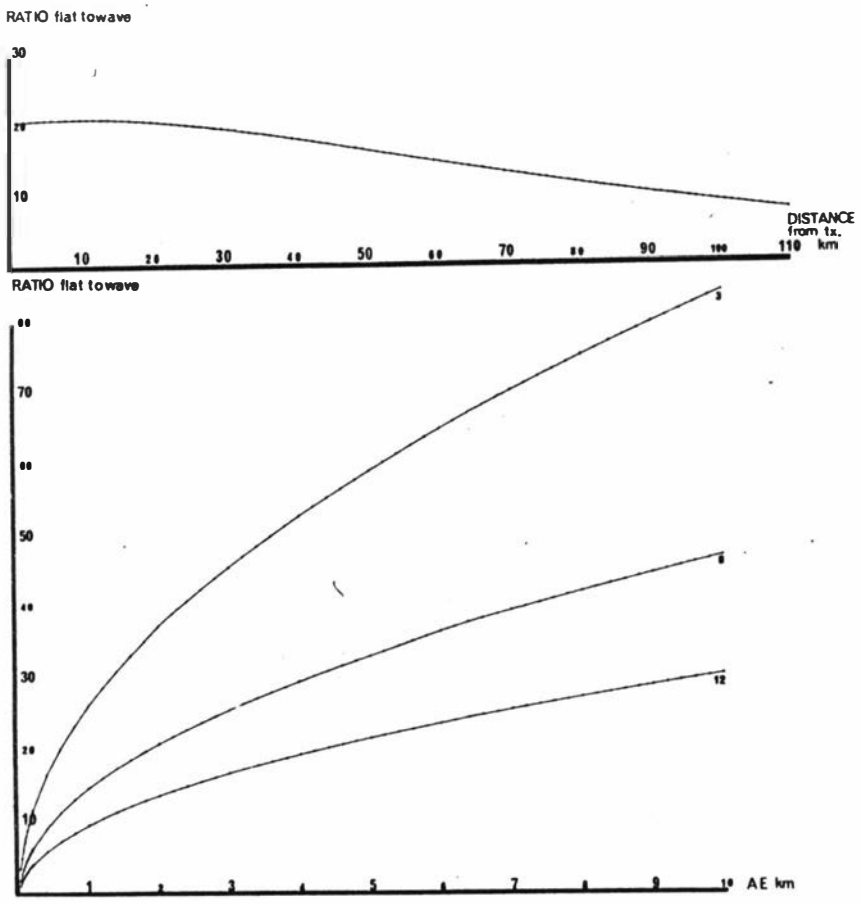


Figure vii: The ratio of the specular reflection area to any one of the non-specular reflection areas plotted as a function of the distance of the wave group from the transmitter (top) and of the total aerial extend (AE) (bottom).

Top: -L=1km, L/A=8, AE=2km. Bottom: -L=1km, wave group centred 10 km from the transmitter, L/A=3, 8 or 12 as shown.

As the aerial extent decreases the ratio approaches unity. Although the analysis on which these ratios are based is extremely approximate, since the dimension (2R) of one of the individual aerials used was a few centimetres the ratio of the reflecting areas for the two types of reflection is taken to be strictly unity.

As a consequence of this assumption no absolute calibration can be provided for the amplitude axis of theoretically calculated curves.

2-3: Factors affecting the received sky wave due to the experimental arrangement.

Previously in this section the sinusoidally perturbed region of the boundary has been treated in isolation, the only stipulation made being that it be at a considerable distance from either the transmitter or the receiver. Factors due to the experimental arrangement are now considered and in order to relate these to the position x of the group centre it is convenient to express θ and β in terms of D, H and x.

From figure iii it is clear that

$$\text{Tan}\,\theta = H/x \tag{27}$$

$$\text{Tan}\,(\theta - 2\beta) = H/(S-x) \tag{28}$$

and hence

$$\text{Tan}\,2\beta = H(S-2x)/(Sx - x^2 + H^2)\,. \tag{29}$$

Up to now the factors considered only affect the amplitude of the received signal. The next three factors, however, affect both the amplitude and the phase. The first factor dependent on the geometry of the experiment is the inverse total path distance factor. The path distance from transmitter to receiver via the central point

of the perturbation group is given by:-

$$S = (x^2 + H^2)^{\frac{1}{2}} + ((S - x)^2 + H^2)^{\frac{1}{2}} . \qquad (30)$$

Therefore the amplitude factor is $1/S$ and the phase factor introducted by the path length is

$$\Delta \phi_1 = 2\pi S/\lambda \qquad (31)$$

The second and third factors to be considered in this section are the polar radiation pattern of the transmitter and the polar response pattern of the receiving aerial. These are peculiar to the experiment described.

The transmitter (station 2YA, 570 kHz) used a $5/8\,\lambda$ vertical aerial earthed at the lower end. The radiation pattern from a vertically grounded wire, assuming a perfect earth, is (see for example TERMAN, Radio Engineering, Chapter 15)

$$E \propto (\cos(2\pi L/\lambda) - \cos((2\pi L \sin\theta)/\lambda))(\cos\theta)^{-1}$$

which substituting for the length becomes

$$E \propto (\cos(5\pi/4) - \cos((5\pi \sin\theta)/4))(\cos\theta)^{-1} . \qquad (32)$$

For the receiving aerial system used (see section 3-2) the system is adjusted so that the ground waves from the two aerials are of equal amplitude and in phase. Then for rays incident at an angle $(\theta - 2\beta)$ to the horizontal the signal from aerial two leads the signal from aerial one by $\phi = 2\pi D(1 - \cos(\theta - 2\beta))/\lambda$. From elementary optics the amplitude of the received signal after the difference amplifier is:-

$$\alpha(2(1 - \cos\phi))^{\frac{1}{2}} \qquad (33)$$

and the phase shift is proportional to

$$\sin\phi/(\cos\phi - 1) . \qquad (34)$$

2-4: The range over which receivable group reflections can occur.

In order that reflections from the group occur and arrive at the receiver, the value of β required for those paths (from equation (29)) must be equal to or less than maximum value of β available from equation (11). That is

$$\tan^{-1} 2\pi A/L \geq \tfrac{1}{2} \tan^{-1} (H(S - 2x)/(Sx - x^2 + H^2)).$$

The left hand side of equation (35) gives the maximum

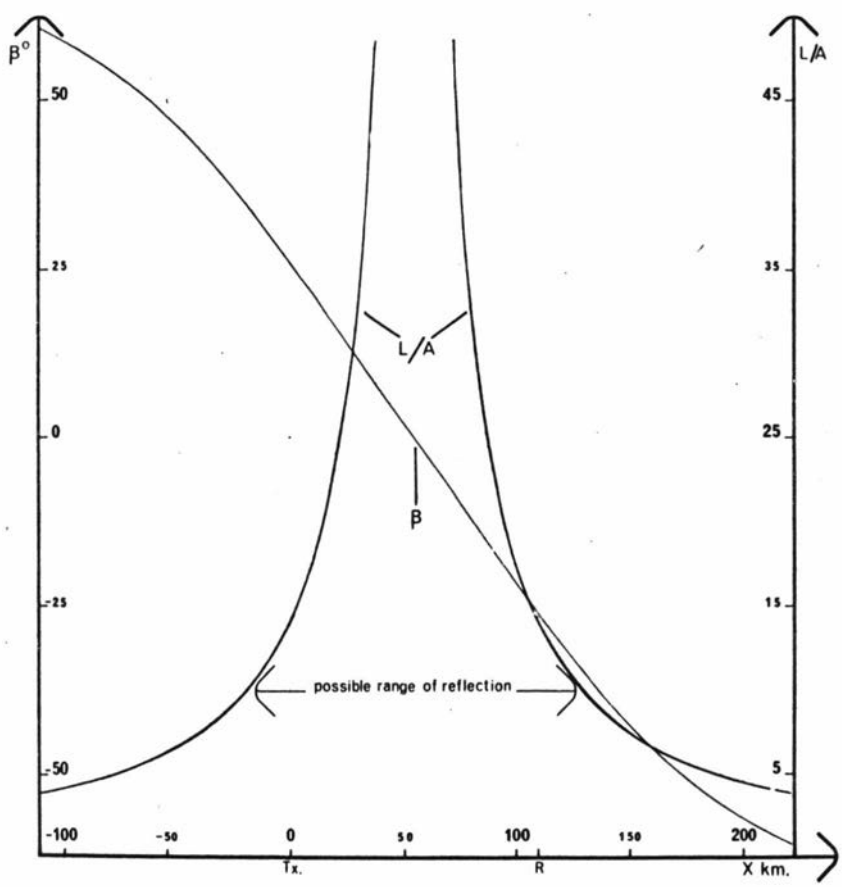


Figure viii: The value of B required in order that signals reflected from position x arrive at the receiver plotted as a function of x . Also shown is the possible range of reflection points deduced from this shown as a function of L/A (calculated for S = 110 km & H = 85 km).

slope of the reflecting surface (β_{max}), and solving the full equation gives the range of x over which reflections of concern can occur. For each numeric value of β_{max} equation (35) may be solved in its equality form to give two values of x. These correspond to the two possible signs for β_{max}, which is positive for $x < S/2$, negative for $x > S/2$ and symmetrical about $x = S/2$ (see figure viii). The two values of x calculated are the maximum and minimum values and hence the range over which receivable group reflections can occur is established. This is plotted as a function of L/λ in figure viii.

2-5: The extension to a continuous curve of expected sky wave amplitude versus group position.

The expected value of the sky wave reflected from a wave group centred at the position x is proportional to the product of equations (17, (18), (21), (32) and (33) and inversly proportional to equation (30). Equation (22) will also be needed for the evaluation of equation (21).

To obtain the total expected sky wave the component above must be combined with the specular component reflected from the unperturbed point of symmetry midway between the transmitter and the receiver. In combining these two components due regard must be paid to their relative phases. Care must be exercised when the position occupied by the wave group encompasses the point of symmetry.

When the position of the group centre is displaced slightly from the point of symmetry the values of β required for the specular and group reflections differ slightly. The greater the difference in the required values of β, the

less the overlap between the two reflecting areas. When the group centre is at the point of symmetry the specular reflection point is identical to one of the group reflection points.

As the wave group moves over the point of symmetry the specular reflection slowly merges with one of the group reflection points. Since absolute reflecting areas are not known, it is not possible to determine the exact rate at which the specular contribution should be disregarded in order not to count it twice. For reasons outlined below it is probably advisable to discount the specular contribution whenever the wave group encompasses the point of symmetry.

The above calculation obviously cannot be carried out for every possible value of x. A continuous curve must be built up by straight line interpolation between points calculated for a chosen series of positions x. The increment of x must be chosen to be sufficiently small to avoid excessive distortion due to the linear interpolation. The fine detail due to the interference between the group sky wave and the specular sky wave is unlikely to be observed in practice.

The disregarding of the specular contribution when the position of the wave group includes the point of symmetry has been suggested as this enables an easy, reasonable approximation to the correct shape of the curve. The position on the calculated curve when the wave group first encompasses the point of symmetry will be indicated by a sudden change in magnitude. The point just before the drop is correct, as is the calculated value corresponding to the wave group being

centred on the point of symmetry. Using the curve between these two points as a model the two points can be joined providing an approximate curve in this region. Since this is also the region in which the approximation of ray optics is suspect (section 1-4), there is little to be gained by improving this method.

2-6: The extension of the proceeding work in this section to an arbitrary length acoustic group.

The preceeding method may be extended in order to investigate the sky wave signal to be expected from an acoustic group whose length is too great for the approximation of section 2-2-3 to be considered reasonable. The long group is subdivided into any number of groups, provided that each individual group is small enough for the approximation mentioned above to hold. The effect of each group is calculated individually and these are summed before being combined with the specular reflection as outlined above.

As this form of analysis has not been needed in this work the details of this extension are not given here but in appendix one.

SECTION THREE: A DESCRIPTION OF THE EXPERIMENTAL METHOD.

An outline of the experiment has already been given in section 1-1; in this section the methods used to carry out the experiment are described. This is done in a block diagram form without great detail as to how each block performs its allotted function. Brief mention is made of some of the more general practical circuit requirements; if more detail is required the reader is referred to the simplified circuit schematic and explanatory notes in appendix two. Finally in this section, mention is made of the experimental routine and of the methods used to try to identify circuit malfunction.

3-1: The aerial system.

In order to observe either the ground wave or the sky wave a system must be used which rejects the unwanted component. This rejection must be of a high order. Since at these frequencies it is not possible to obtain sufficient directivity with a single aerial some form of array is required. As the directivity property can be obtained from the array itself maximum flexibility is obtained by choosing each element to have a uniform vertical polar response.

The array used consisted of only two elements, the common axis of which was orientated towards the transmitter. Each element consisted of a horizontal ferrite rod whose long axis was placed broadside to the transmitter. This had the incidental advantage that with the north-south transmitter-receiver orientation any signals from Australia fell into the natural null of the ferrite. Each

ferrite was very broadly tuned and coupled to a pre-amplifier which first matched it to the co-axial line and second provided a trap to reject the local radio station (940 kHz). In addition the ferrites were provided with electrostatic screening to minimise any local pick-up.

The use of two aerials provides the necessary means by which signals of the same frequency might be identified by their angle of arrival. For any one signal arriving the phase difference between the signals as available from the two aerials is a measure of the angle of arrival of the signal. No one signal can be selected in preference to all others, but any one signal may be rejected. This requires phase shifting, amplitude matching and a comparison amplifier. A block diagram of the actual system used in this experiment is shown in figure ix.

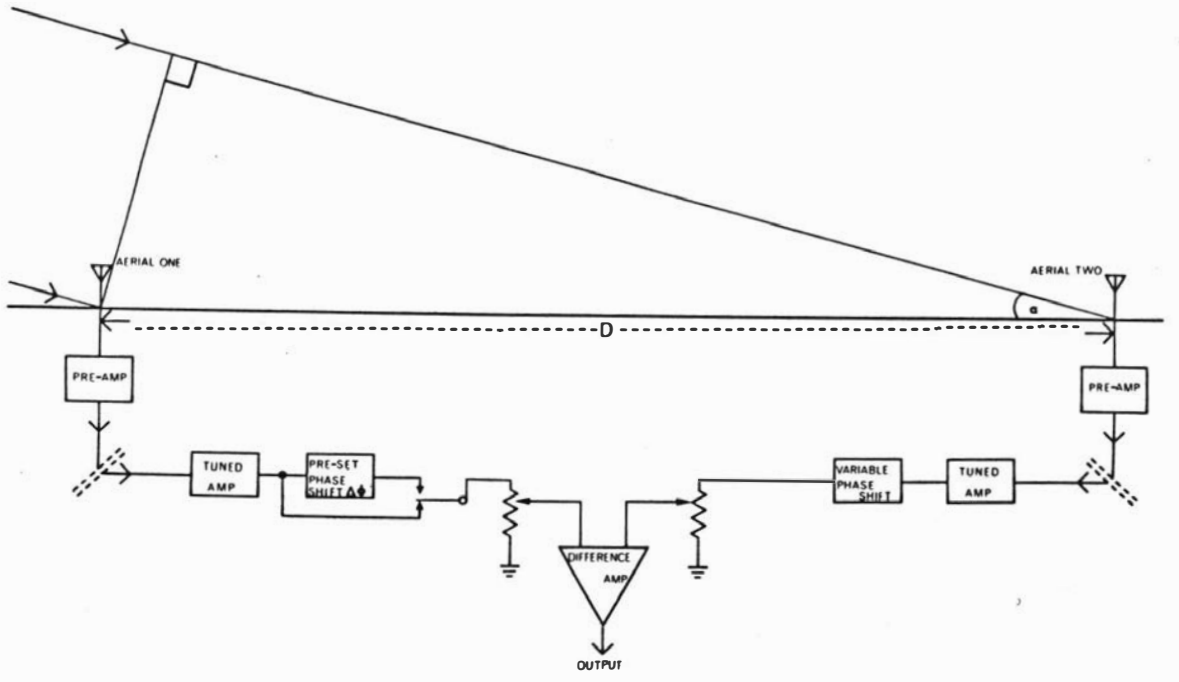


Figure ix: A block diagram of the system used in this experiment.

3-2: Rejection of the ground wave.

During the day time when only the ground wave is present in any strength the preset phase shift is switched out and the variable phase shift and the variable attenuators are adjusted until the output from the difference amplifier is (ideally) zero. In practice the ground wave is reduced by 40 dB from the signal available at one aerial element. The horizontal ground wave signal arriving at aerial two lags the signal at aerial one by $360D/\lambda$ degrees. The variable phase shift has advanced the aerial two signal by this amount and also compensated for any difference between the two channels up to the difference amplifier. The response of the aerial system is now of the form shown in figure x. This is basically the method used to obtain the sky wave. The inter-aerial spacing used was 0.11 km. As the direction of the ground wave is constant day and night any signal output from the amplifier at night is assumed to be due to sky wave and will have been amplified by the aerial system by an amount which is a function of the angle of arrival. An obvious, but useful, check on the system performance is to ensure that the output remains zero all day and returns to zero the next day following the fluctuations at night.

3-3: Rejection of the sky wave.

When, following adjusting the system as outlined above, the present unity gain phase lag $\Delta\emptyset$ is introduced into the lead from aerial one, it can easily be shown that total cancellation will occur from a signal arriving at an angle α to the horizontal, given by

$$\Delta\emptyset = (360D/\lambda)(1-\cos\alpha) \text{ degrees.}$$

The value of $\Delta\emptyset$ needed for any reasonable value of α is quite small and within the range of a simple filter. For this reason a pre-set phase lag in lead one has been used rather than a preset phase advance in lead two as a low pass filter is inherently less susceptible to noise than a high pass filter. The total aerial response pattern is now as shown in figure xi. The pattern must be symmetrical about the horizontal axis since the aerials cannot distinguish between a signal from the sky and a signal at a similar angle from underground.

The first attempts to obtain a relatively uncontaminated night-time ground wave were made using this approach. The value of α was altered in an effort to identify the angle of arrival of the skywave. Success would have been indicated by a constant ground wave signal and a varying skywave signal. This would only be achieved for one unique angle of arrival, and the lack of success recorded is hardly surprising. In an effort to overcome this a variable 'pre-set phase lag $\Delta\emptyset$' was employed as shown in the block diagram in figure xii.

A voltage controlled phase shift was inserted in series with the pre-set phase lag $\Delta\emptyset$. The voltage controlled phase shift must have a constant gain and is driven by a ramp generator which is free running provided the input to it is high. During the day time, when the sky wave is absent, the output of the voltage

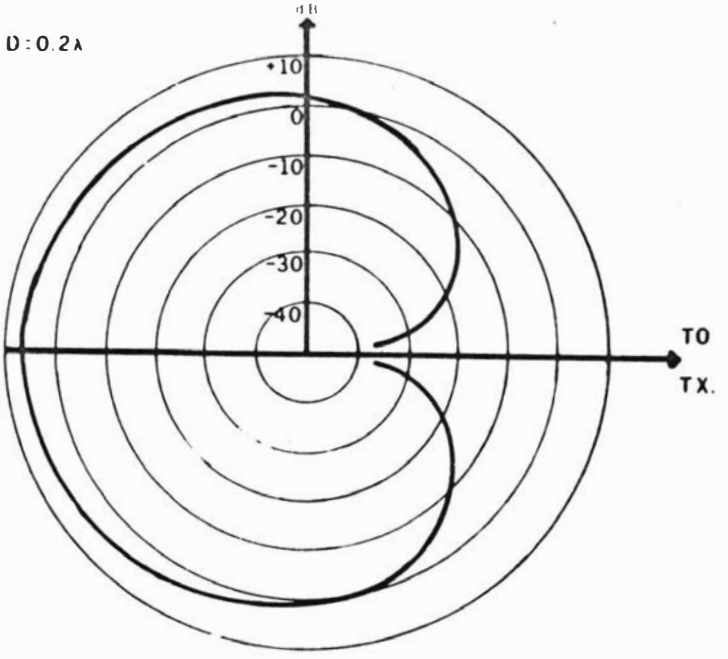


Figure x: The polar response of the aerial system when set up for ground wave rejection.

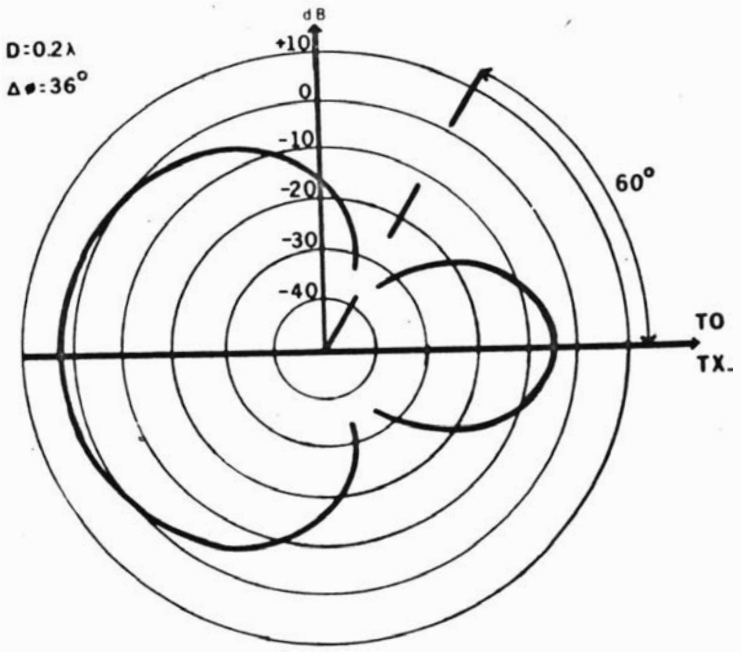


Figure xi: The polar response of the aerial system when set to reject signals with an angle of arrival of 60^{o}.

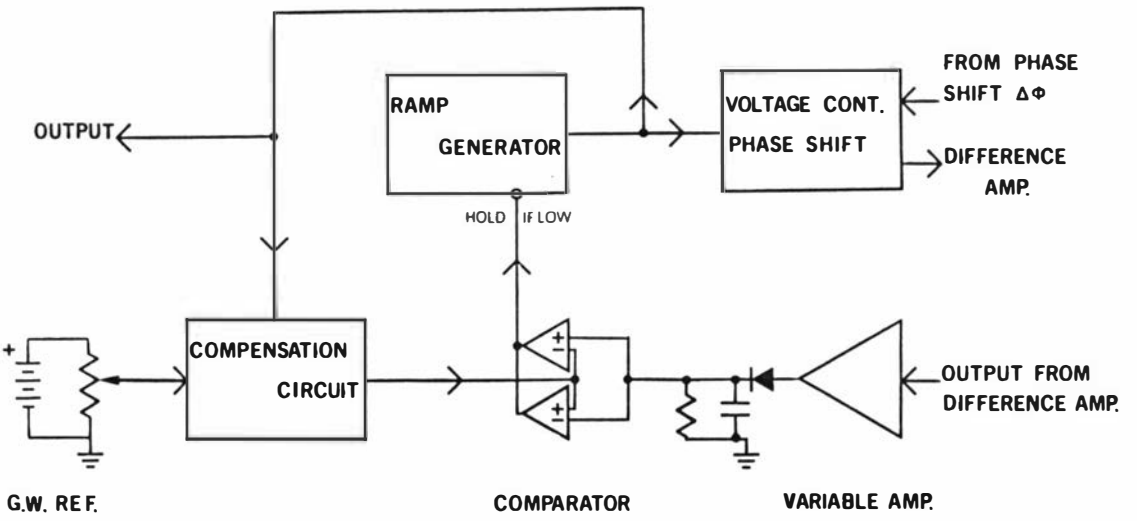


Figure xii: The additions required to the system shown in figure ix to enable that system to hunt for the sky wave.

controlled phase shift is held at zero. The circuit is then adjusted as described above to produce a null at an angle α less than the minimum expected angle of arrival of the sky wave. The output from the difference amplifier is the signal due to the ground wave and this is amplified, rectified and measured. An equal voltage is used as the ground wave D.C. reference.

At night the output from the difference amplifier is not necessarily pure ground wave but may be contaminated by sky wave. This output is amplified, rectified and compared with the day time ground wave reference. The comparator used provides a positive output if the two signals differ by more than a pre-determined fraction, and this output is used to provide the input to the voltage controlled phase shift. Since the magnitude

of the pure ground wave signal expected is a function of α a compensating circuit is included so that the comparator is always presented with a reference signal appropriate to the elevation angle of the response pattern null.

When the null in the aerial response pattern is at the same angle as the arriving sky wave, the output from the comparator goes low and the circuit ceases to hunt, remaining stable until the ground wave is again contaminated. The value of the angle α at which the sky wave is arriving can be determined by measuring the input to the voltage controlled phase shift.

3-4: Limitations of the automatic sky wave rejection circuitary.

Since the normal specular reflection would be expected to exist at all times during the night, the fluctuating night-time sky wave must be expected to consist of contributions arriving at at least two angles. Because of this it might at first seem impossible to use the method outlined above to hunt down and reject the sky wave. Whilst it is not possible to track weak non-specular reflections in this way, the non-specular reflections of concern in this experiment must have a greater amplitude than the specular reflection in order to produce the high sky wave levels observed. It is possible to track a strong reflection in the presence of a relatively weak specular reflection by providing a dead band in the comparator. This decreases the accuracy with which the angle of arrival can be measured. Small events will not

cause the circuit output to change and the usefulness of this type of circuit is thus limited to large or what might be termed major events whose non-specular contribution occurs solely from one limited area. A major event consisting of strong reflections from two or more separate areas will cause the circuit to continue hunting, being unable to find any one position to reject all the large sky waves. A further limitation of the sky wave rejection circuitry may become apparent following a period of hunting. When the received sky wave, having been due to two or more main reflections, becomes due to one reflection alone, the circuit will cease to hunt and will record the angle of arrival of this signal. When, however, the last two (or more) major reflections cease together (or within one sweep of the circuit) a false indication may occur. When the signal input into the comparator is due only to ground wave, the circuit is stable at any angle due to the ground wave reference correction circuit. Therefore hunting will cease at the disappearance of all major sky wave components and the output, which normally corresponds to the angle of arrival, will remain at whatever value it had at the time of disappearance. Events of this kind can be detected by carefully examining the sky wave record to establish whether a major peak was occurring at the time hunting ceased. Examples of this can be found in figure xiv(a).

3-5: An outline of practical circuit details.

The operation of the radio frequency circuitary has been outlined above in terms of block diagrams. A

detailed description of the contents of each block will not add materially to an understanding of this work; if desired simplified schematic circuits may be found in appendix two. However brief mention is made of particular requirements placed on various circuit blocks in the order a signal would meet them.

The aerial units have already been described. They are connected to the rest of the circuit by coaxial cable, which also serves to carry power up to drive the circuits in them. This coaxial cable has to be fairly carefully matched at either end to prevent pick up of local signals. The amplifiers which receive the signal from the aerials are conventional with double tuning between stages. Provided a common local oscillator were used for the two channels, superhetrodyne receivers could have been used, but the simple receivers provided sufficient selectivity in this experiment. The difference amplifiers were integrated circuits chosen for high common mode rejection. As the signal levels at the outputs of these amplifiers were normally lower than either of their two inputs the risk of signal degradation by noise pick up was possible. To guard against this the final output stage of the difference amplifier was a further single tuned stage, whose gain was, for convenience, made variable.

Following separation the signals were converted to D.C. by a simple peak reading detector. The rectifying diode was, of course, non linear when only slightly forward biased but approximately linear later. Since large signals are the main area of interest, this was felt to be

adequate. The small time constant built into the circuit attenuated the signal fluctuation due to the modulation, and, initially, the type of recorder used was relied upon to complete the removal, as far as possible, of the remaining 'high' frequency signal fluctuation. The recorder used originally was of the dot printing, moving coil type with a natural cut-off at about one hertz. With the later availability of a multichannel U.V. recorder whose galvanometers had a 3dB point of 300 hertz, an amplifier with a variable time constant was used for each signal to be recorded. The time constant of the amplifier was variable from ten seconds to about one thousand seconds.

It was originally intended to attempt to measure phase shifts in the received sky wave, utilising the pure ground wave as a reference. Since even the automatic hunting circuit for sky wave rejection had the severe limitation of only being able to deal with one strong sky wave signal at once, it was not possible to ensure an uncontaminated ground wave signal at night. During moderate and highly disturbed sky wave conditions the proportion of time during which the sky wave rejection circuit was hunting was sufficiently high to preclude the use of a phase reference oscillator flywheel synchronised to the time averaged ground wave.

3-6: Experimental routine.

The calibration procedure described above was carried out each afternoon, recording being automatically carried out from shortly before dusk to shortly after dawn. Since the main station recorded, 2YA, normally transmits

twenty-four hours a day, the normal check against a failure in equipment or calibration was to make sure that the sky wave signal returned to zero the next day. The one occasion per fortnight when the transmitter was shut down for normal maintenance was particularly useful for checking for evidence of interference from outside sources. No case of outside interference was detected. During the routine maintenance at the transmitter it would occasionally be keyed on for a series of short intervals. Use was made of this to check the operation of the circuits for automatic sky wave hunting and rejection. An example of this is shown in figure xiv(a).

SECTION FOUR: Results, conclusions and suggestions.

For convenience the results presented in this section are divided into two parts. First the results obtained from the earlier work using slow chart speeds are presented. However, before the later results using higher chart speeds are presented, an account is given of the selection and analysis procedures followed for these results.

No experiment is ideal nor provides all the results that could be required in a form that is easy to digest. Therefore the final part of this section is devoted to a review of some of the weaknesses of this experiment, with suggestions given for improvement and extension in the hope that they will be of use in future research.

4-1: Results of early work.

In the initial stages recordings were made on chart paper moving at a speed of one inch per hour and a maximum of two channels could be recorded simultaneously. The slow chart speed prevented any detailed study being made of the structure of the night-time sky wave. However the highly variable strength of the sky wave signal became apparent as did the surprizingly high level reached on occasions. A section of a typical nights recording is shown in figure xiii.

Section 3-3 discusses the difficulty of measuring the angle of arrival of the sky wave by varying the position of the aerial pattern null on a trial and error basis. However inspection of recordings made over a

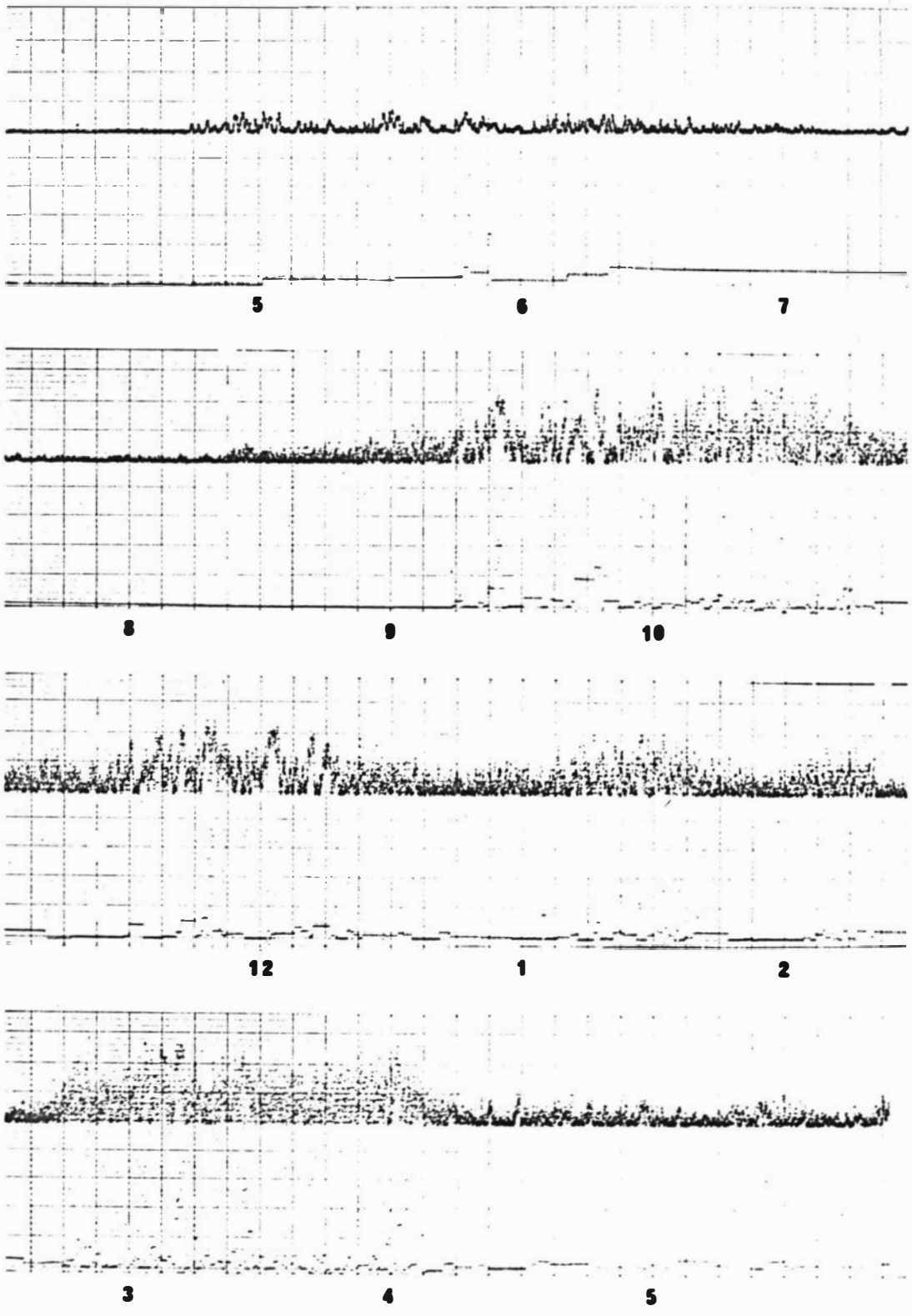


Figure xiii: A typical night's record. Local time as marked, top trace sky wave amplitude, bottom trace angle of arrival ($\alpha_{basic} = 26^{o}$).

period of some weeks with different null positions showed that for any reasonable value of α (as defined in 3-3) there would be occasions on which particular peaks would be suppressed. Events of this type were noticed by comparing the records with another set made simultaneously using rejection of the ground wave.

There was without doubt no unique angle of arrival; the range of skyward null angles which produced noticeable peak suppression was very large (30 degrees to 80 degrees elevation). Since one peak would be suppressed whilst neighbouring ones would not, and considering the effective angular width of a null (approximately 10 degrees for a minimum of 30 dB amplitude suppression), it was felt that successive maxima often came from quite different areas of the ionosphere. The theory that small random fluctuations around the specular reflection point gave rise to the high sky wave levels observed did not, therefore, seem to fit the general case, although it may be true for some events.

In order to further examine the arrival angles of the sky wave the automatic sky wave rejection circuit described in section 3-3 was developed and used. As might be expected in view of the limitations of this circuit, an appreciable percentage of the time was spent in continual hunting (see figure xiii). When not hunting most of the time is spent jumping from one arrival angle to another with little obvious pattern. On occasions the arrival angle changes in a monotonic way (see figure xiv). This suggests a steadily moving reflection

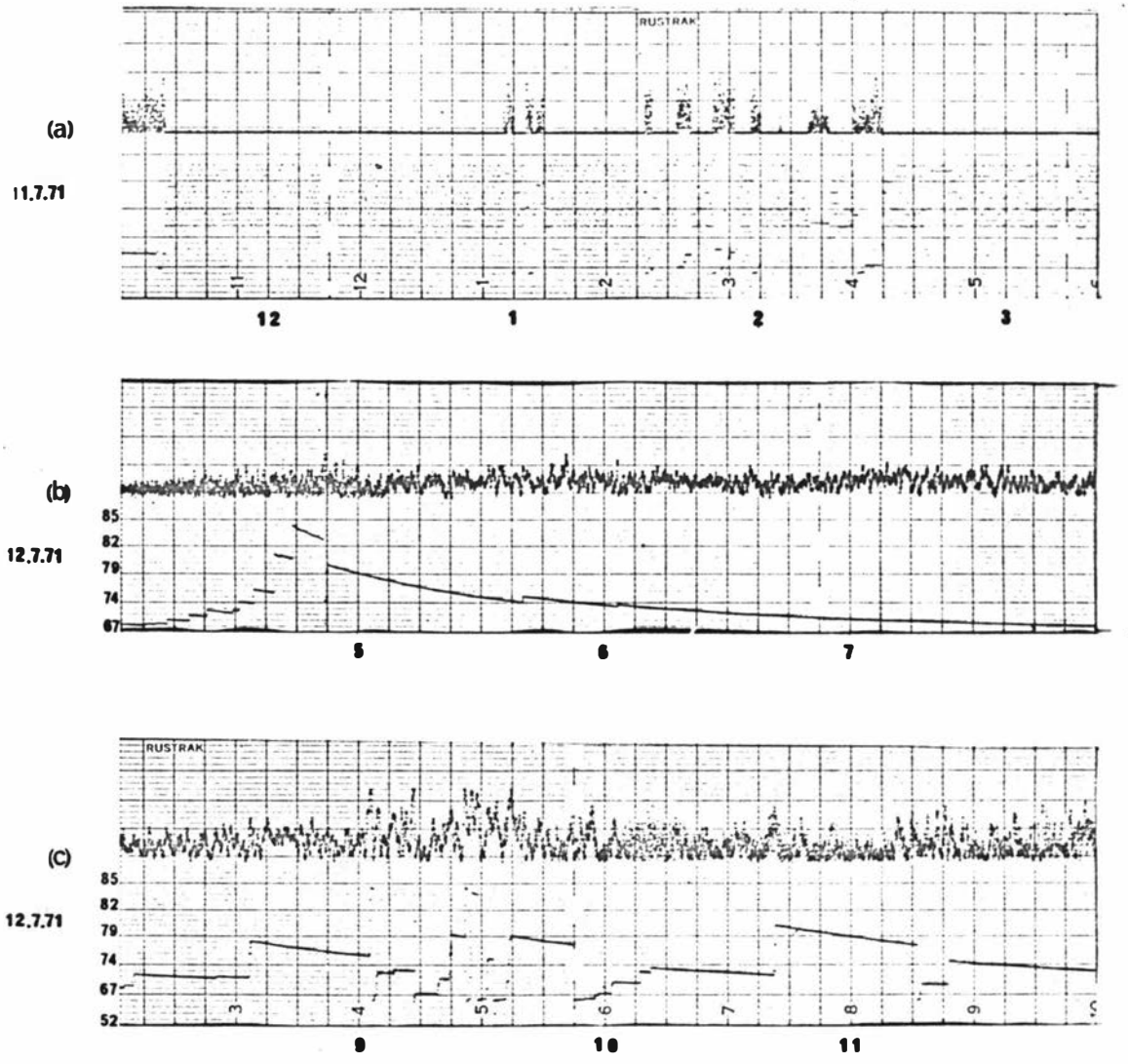


Figure xiv: Records, top trace sky wave amplitude, bottom trace angle of arrival, local time as marked. (a) shows examples of false indication of angle of arrival and (b) and (c) show examples of the monotonic variation in the angle of arrival which first suggested a steadily moving source. The example in (b) is rare and occurs over a longer time period than can be explained in terms of this work.

zone, the signal from which varies in strength as the zone moves. These could be the types of reflection considered in this work. With the advent of a high speed chart recorder it was hoped to study in more detail the variation with time of the angle of arrival in this type of event. However, when sufficient sensitivity is obtained to measure the angle of arrival of each of the peaks of a sequence, the circuit also becomes sensitive to other reflections. For this reason the record is often spoilt by reflection other than one from the moving reflection zone.

The main test as to whether an observed event can be described successfully by the theory of section two is taken to be the production of a theoretical curve to match the practical curve with physically reasonable values for all parameters.

4-2: The determination of parameters from records made using a high chart speed.

There are five parameters required before an attempt can be made to fit the theoretical amplitude vs. distance curves to the experimentally obtained amplitude vs. time curves. These are the perturbation height (H), the perturbation velocity (V), the perturbation wavelength (L), the effective number of perturbation waves in the group (N) and the ratio of perturbation wavelength to amplitude ($^{L}/A$). These are not independent and two must be known in order that the other three can be determined. The two about which most is known are the height H and velocity V and these are therefore chosen to be the independent variables. By recording at the higher chart

speed it is possible to make measurements which will enable two of the three remaining parameters to be calculated.

An inspection of the equations which define the theoretical curves shows that the term which changes most for a given change in group position is $\sin N\gamma / \sin \gamma$. γ will always be zero at $x = \frac{1}{2} S$, leading to a maximum of this term at this position (or time) which may be referred to as the central maximum. The next maximum of this term will occur when $\gamma = \pm \pi$, which occurs when the perturbation group is at a position $\frac{1}{2} S \pm \Delta x_1$. By using the defining equation of γ(14), a curve of Δx_1 vs. L can be plotted, as in figure xv.

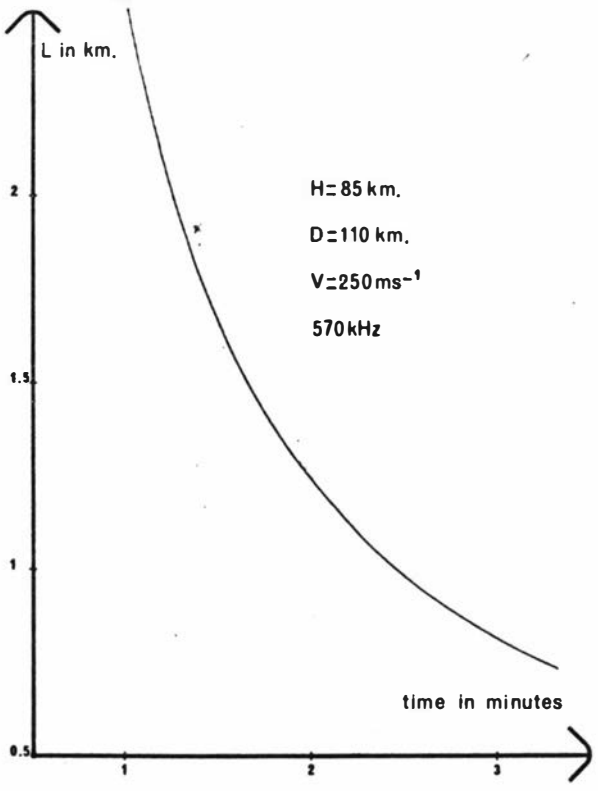


Figure xv: A plot of L vs. Δx, the latter axis being calibrated in time on the assumption that the velocity of sound at a height of 85 km is 250 ms^{-1}.

Using this, a measurement of the time between the central maximum of the pattern and either of the adjacent maxima yields at once the wavelength of the perturbation. Since the times involved are of the order of minutes L can be measured to reasonable precision provided V and H are assumed accurate. As will be shown, γ is an approximately linear function of position for small values of Δx, and this is used in the estimation of N.

To express γ as a function of Δx equations (14), (27) and (28) are used. Expressing these in terms of Δx, they become,

$$\gamma = (-2\pi L/\lambda) \sin\beta \sin(\theta - \beta),$$

$$\tan\theta = H/(\tfrac{1}{2}S + \Delta x),$$

and

$$\tan(\theta - 2\beta) = H/(\tfrac{1}{2}S - \Delta x).$$

Since $\cos A - \cos B = 2\sin\tfrac{1}{2}(A+B)\sin\tfrac{1}{2}(B-A)$,

$$2\sin\beta\sin(\theta-\beta) = \cos(\theta - 2\beta) - \cos\theta .$$

Using $\cos A = 1/(1+\tan^2 A)^{\frac{1}{2}}$, substituting and manipulating yields

$$2\sin\beta\sin(\theta-\beta) = \left\{ \frac{H(\tfrac{1}{2}S-\Delta x)\left(1+\left(\frac{\frac{1}{2}S-\Delta x}{H}\right)\right)^2 - H(\tfrac{1}{2}S+\Delta x)\left(1+\left(\frac{\frac{1}{2}S-\Delta x}{H}\right)\right)^2}{((\tfrac{1}{2}S)^2 + H^2 - (\Delta x)^2)^2} \right\}^{\frac{1}{2}}$$

Provided Δx is small both $(\tfrac{1}{2}S + \Delta x)/H$ and $(\tfrac{1}{2}S - \Delta x)/H$ are small compared to unity, and without incurring an unreasonable error only the first two terms of the expansion

$$(1 + x^2)^{\frac{1}{2}} = 1 + \tfrac{1}{2}x^2 - \tfrac{1}{8}x^4 + \ldots\ldots$$

need be considered.

Substituting and simplifying gives

$$\sin\beta\sin(\theta-\beta) = \frac{-H\Delta x}{(\tfrac{1}{2}S)^2 + H^2 - (\Delta x)^2}$$

Considering the values of S (=110 km) and H (=85 km) appropriate to this experiment, the $(\Delta x)^2$ term in the bottom line of this expression may be neglected to first order for small values of Δx. Hence:

$$\gamma = \frac{2\pi\ L\ H\ \Delta x/\lambda}{(\frac{1}{2}S)^2 + H^2}$$

This may not be accurate enough for determining L from the separation of the orders but is adequate for estimating the value of N by the method below, since for an ideal perturbation group N can only have integer values.

For Δx_1 as defined above, $\gamma = \pm \pi$, and then approximately

$$L = \lambda((\tfrac{1}{2}S)^2 + H^2)/(2\,H\,\Delta x_1)$$

Let Δx_2 be the half width of the central maximum at $x = \frac{1}{2}S$.

These minima occur when $N\gamma = \pm \pi$, and therefore approximately

$$N = \lambda((\tfrac{1}{2}S)^2 + H^2)/(2\ L\ H\ \Delta x_2)$$

Substituting for L gives

$$N = \Delta x_1/\Delta x_2$$

This clearly provides a simple method of estimating N.

4-3-1: Recognition of significant events from high speed records.

The full night's record is inspected by eye and sections with peaks at regular intervals are selected. These are then more closely examined and compared with a set of calculated curves, such as the set shown in figure xvi.

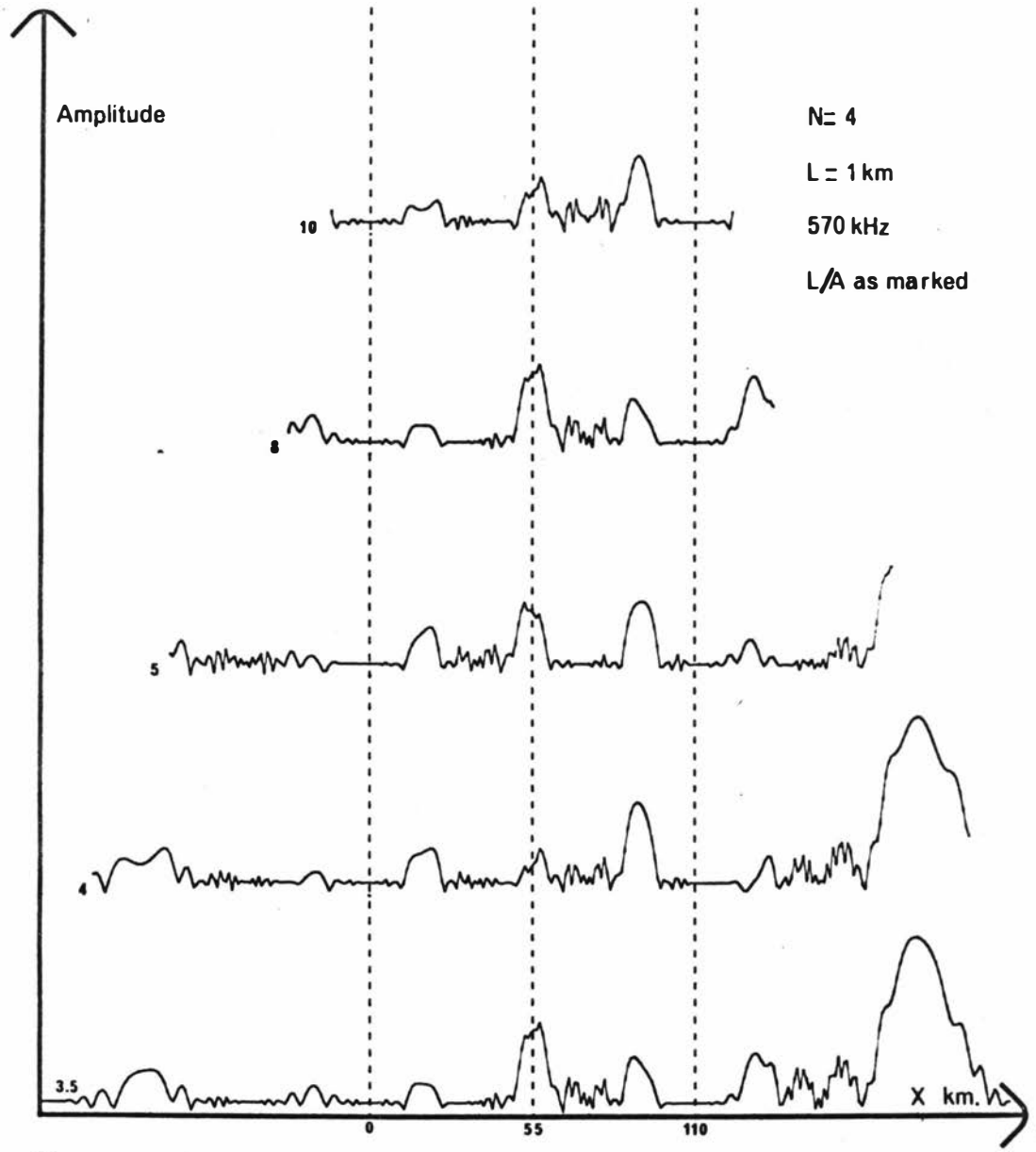


Figure xvi: A set of theoretical curves of sky wave amplitude vs. group position, calculated for D = 110 km, H = 85 km and $N_e = 10^8 m^{-3}$. For clarity the zero on each curve has been offset.

It is obviously not possible to plot all curves that might be required but one may interpolate between curves provided the effect of varying the parameters L, N and L/A is known. The effect of varying L and N can be predicted from a knowledge of diffraction gratings: increasing L beings the peaks closer together in time and increasing N makes the peaks sharper. The effect of varying L/A is not so obvious; the range over which reflections can occur decreases and also the positions of the 'missing orders' changes.

A section of record is considered to be reasonably described by the theory in this work if it meets four conditions. The first of these is that there must be more than three regularly spaced peaks. This prevents the detection of patterns due to very low values of L/A but is adopted in an effort to minimise the probability of counting random peaks as a significant group. Secondly the rank order of height of all peaks other than the central one must be as predicted by theory. Thirdly the area between the regular peaks must look similar to that predicted theoretically, in particular it must not contain any other peak. Finally the group velocity deduced by relating the practical amplitude vs. time curve with the theoretical amplitude vs. position curve must yield approximately $250\,ms^{-1}$, the velocity of sound at a height of 85 km.

This method of processing the results favours intermediate values of L/A. Events due to low values of L/A are unlikely to be noticed, whilst events with high

values of L/A are unlikely to occur without acquiring at least one extra peak due to a random reflection. Events were only noted for which the predicted wavelength L lay in the range 0.75 km to 1.5 km.

4-3-2: Results obtained from high speed records.

Using the selection criteria given above a total of 82 acceptable events were observed from 60 night's recording. A histogram of the occurence of various values of the wavelength L is shown in figure xvii, where a second axis calibrated in frequency is provided (the speed of sound at 85 km being assumed to be 250 ms^{-1}). Since the selection procedure is highly biased this histogram should not be taken as necessarily giving any information about the relative frequency of occurrence of the different wavelengths but rather as a record of the events observed.

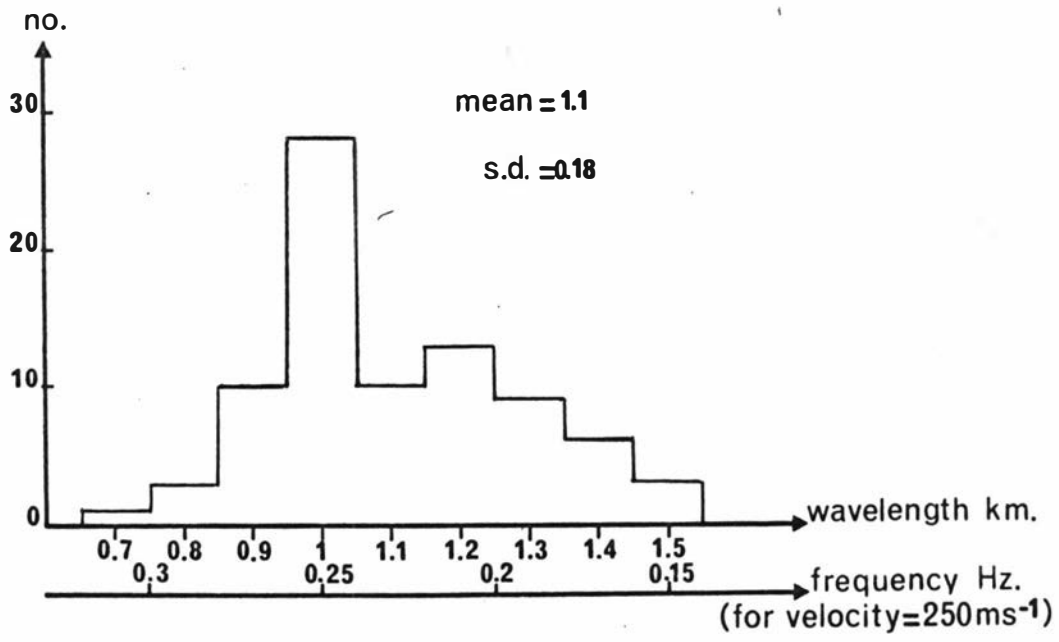


Figure xvii: A histogram of the occurrence of detected values of L in a total population of 82.

An example of the good agreement between theory and experiment obtained by using successive approximations to get a best fit is shown in figure xviii. The fast beating between the specular and non-specular components in the predicted curve would, not if it occurred in practice, be seen on the experimental curve because of the long time constant associated with the recorder circuits. In this section of the record on the

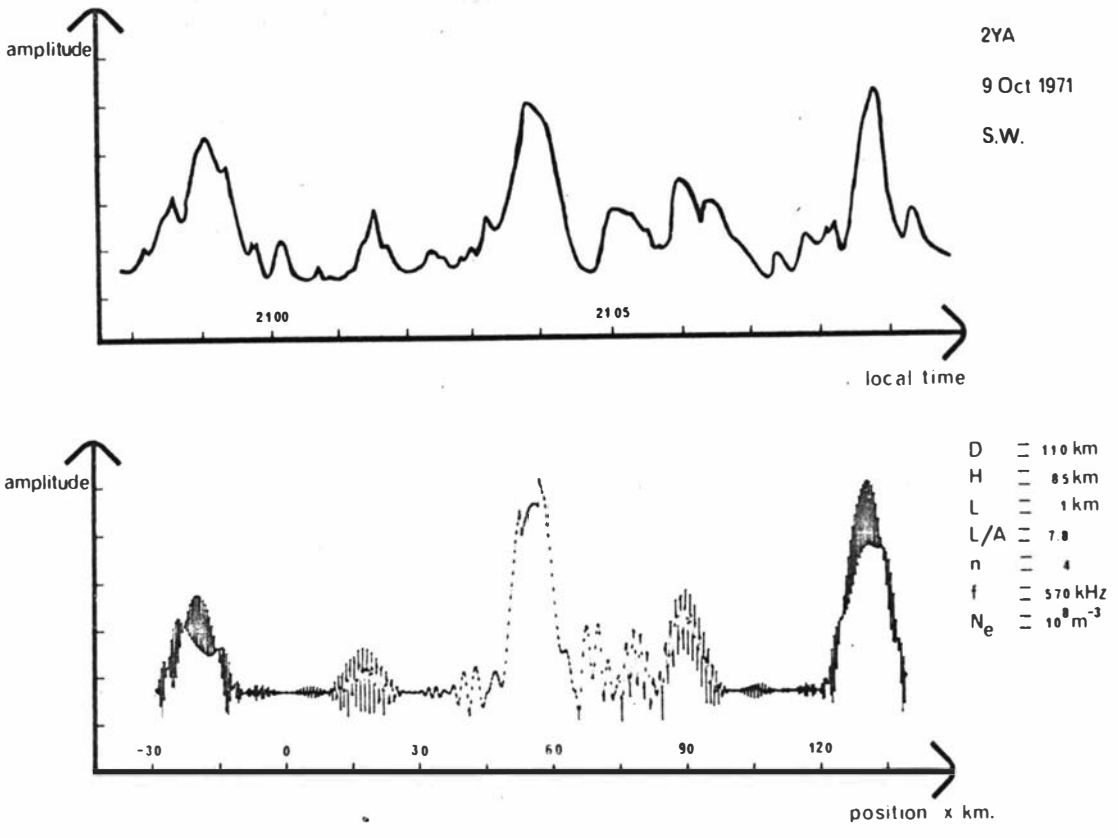


Figure xviii: An example of the good agreement between theory and experiment obtained by using successive approximation to get a best fit.

9th October 1971, the maxima occur at two and a half minute intervals. Examining this record statistically using the Wald-Wolfowitz Runs Test (see, for example, SIEGEL, 1956) the hypothesis that peaks spaced by $2\frac{1}{2}$ minutes do not come from the same population as all other peaks is significant at about the 0.1% level. That is, if all peaks were totally randomly spaced the chance of those peaks spaced by $2\frac{1}{2}$ minutes occurring in the distribution observed on October 9th, 1971, is about 1 in 800. The limitations of the angle of arrival measuring equipment have already been noted. Figure xix shows one

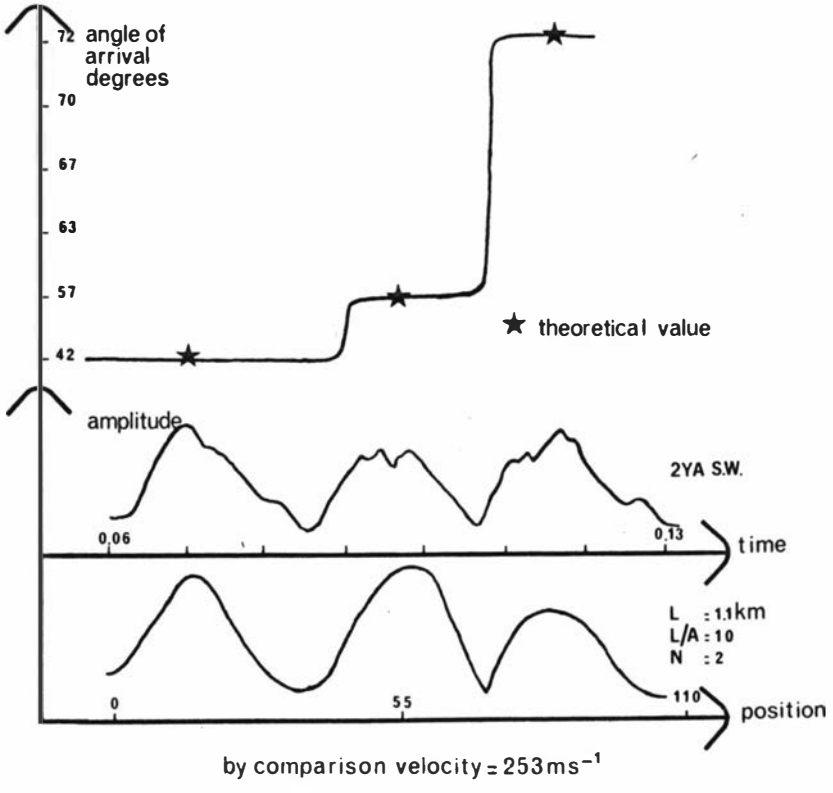


Figure xix: The angle of arrival (top) of the observed sky wave amplitude (middle). The stars give the theoretical values of angle of arrival calculated for the positions of the peaks of the computed curve (bottom).

of the few uncontaminated records of angle of arrival together with the observed sky wave amplitude and a computed curve. The stars indicate the theoretical values of angle of arrival calculated for the positions of the peaks of this computed curve. The dead zone in the comparator leads to an uncertainty in the angle of arrival which is itself a function of the angle of arrival. Typical values for this uncertainty would be ten percent for an angle of arrival of forty degrees decreasing to one percent for an angle of arrival of seventy degrees. As the skyward null in the aerial response pattern always approaches a given angle from below, the experimental value for the angle of arrival will always be low.

Attempts have been made to observe the sky wave from either 2YC or 2YZ at the same time as the sky wave from 2YA. The transmitters for 2YA (570 kHz) and 2YC (660 kHz) are at the same site and therefore sky waves for these frequencies would have very similar paths. The transmitter for 2YZ (630 kHz) lies on an approximately reciprocal bearing to the 2YA transmitter. It was hoped that using the sky waves for 2YA and 2YZ an additional determination of the disturbance velocity could be made, and using 2YA and 2YC two independent measurements could be made of the parameters of the ionospheric disturbance.

A high degree of similarity was noted between sky wave amplitudes observed simultaneously from 2YA and 2YC. Some tentative identifications of pairs of sections of sky wave patterns have been made, the two members of the pair being either simultaneous (2YA and 2YC) or spaced in

time (2YA and 2YZ). These identifications, while giving physically reasonable values for the various parameters, do not agree sufficiently well with theory to be considered more than possible events. The failure, so far, to obtain positive confirmation using a second frequency of the results obtained using one frequency could be due to a number of factors. Both 2YC and 2YZ cease transmission at 11 p.m. thus restricting observations on these frequencies to the relatively less disturbing earlier part of the night. A change in frequency, path length or transmitter radiation pattern each produces changes in the theoretical pattern expected. Figure xx shows a pair of calculated patterns, one for 2YC and one for 2YA, calculated for the same perturbation group.

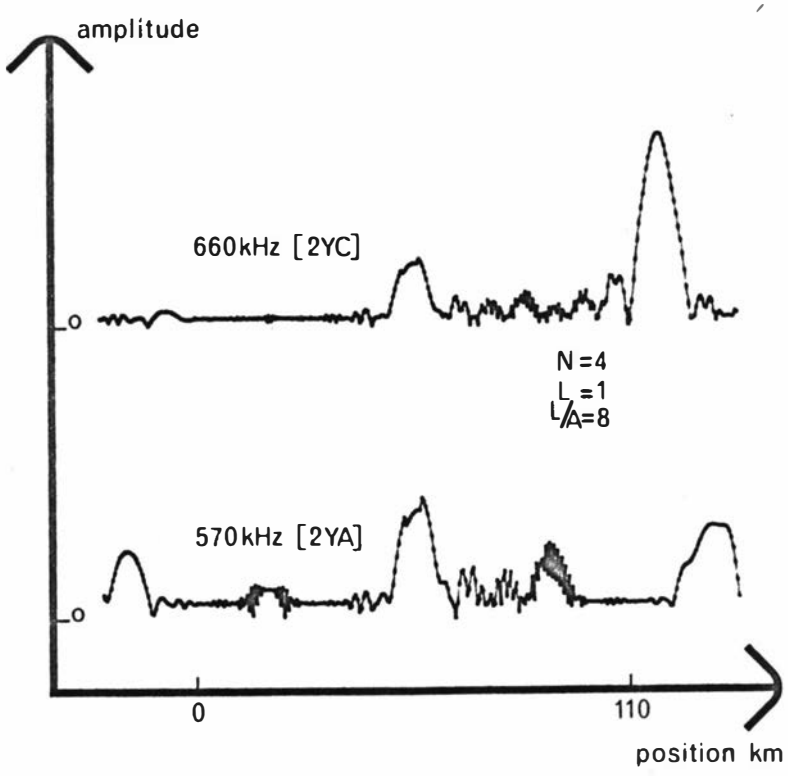


Figure xx: Theoretical sky wave amplitude vs. position curves for both 570 kHz and 660 kHz, all other parameters being the same. For clarity the amplitude zeros have been offset. Comparison of the lower trace with figure xviii shows the effect of the plotting increment on the fine structure.

It can be seen that there is not a great deal of similarity between them, the left half of the 2YC pattern containing no significant amplitude peaks at all. This pattern, even if perfectly reproduced in practice, would not pass the acceptance criteria of section 4-3-1. Thus, even with idealized patterns, severe pattern recognition problems exist. In addition in view of the approximations in the theory, there will not be a high frequency of occurrence of one set of perturbation parameters which would produce theoretical curves that closely corresponded to both of a pair of practically obtained curves. To this must be added the increased probability that one or other of the record pair will be contaminated by stray reflections from areas other than the moving reflection zone.

The predicted patterns for a reflection group moving from south to north would, neglecting the earth's magnetic field, be a mirror image of the patterns for a reflection group moving from north to south. The degree of anisotropy is small, coming mainly from the transmitter and receiver vertical polar diagrams. When events were classified according to apparent direction of travel, it is interesting, and may be significant that only two of the eighty-two events observed were believed to have been travelling in a southerly direction.

4-4: Comparison with other experiments.

As far as the author is aware no other experiment has been described studying similar phenomena at these heights. The range of possible probing frequencies to

conduct an experiment of this type is quite small. As the frequency is increased from those used in this study, the reflection coefficient decreases; as the frequency is decreased the minimum size irregularity which could be observed increases. It is not surprising that phenomena of the scale size considered in this work have not been reported by workers with vertical sounding type experiments as a satisfactory compromise between reflection coefficient and irregularity size resolution cannot be obtained. With the very much shortend total path travelled by a vertically incident radio wave compared to the path travelled by an obliquely reflected wave, the frequency for the vertical sounder needs to be higher for the same Fresnel zone radius. The vertical sounder frequency to give a first Fresnel zone radius of 1 km at a height of 85 km is about 13 MHz. At this frequency the reflection coefficient would be extremely low.

4-5: Conclusion and suggestions for further research.

The purpose of this work is to put forward a hypothesis which seems physically reasonable in the light of current knowledge. The mathematical analysis of and the predicitions made from this hypothesis are only approximate, but nevertheless have been found to correspond quite closely to some observations. It is therefore suggested that some of the variations observed in the received nighttime sky wave signal at about 600 kHz are due to interference between the multiple reflections from a moving distortion of the ionization boundary at a height of about 85 km. Further work will be needed to establish

beyond doubt a definite relationship and the proportion of events caused by this type of process. The apparently more frequent occurrence of ducted microbarums later at night (particularly from 11 p.m. to 3 a.m.) referred to in section one is interesting. It may be a genuine high point for microbarum activity, or it may be an effect of the detection system. As the night goes on the ionization ledge gets sharper (see figure ii), and the height of this ledge changes. It is possible that between 11 p.m. and 3 a. m. the position of the ledge corresponds to the height of maximum ducting.

Brief mention is made below of some of the experimental difficulties associated with the present arrangement and suggestions are made for future research.

The principle problem with the present experimental arrangement is one of data reduction. The manual analysis of charts is a time consuming procedure which depends a great deal on the skill of the operator, although the eye is the best pattern recognising instrument known to man. A considerable improvement in convenience would be for a preliminary selection to be made automatically of some sections of the record which could be then studied in more detail. The generation and storage of a complete set of all possible patterns and the comparison of these against the received signal is obviously not very practical. Even if it were accomplished the selection made would be influenced by the details of the model used, which is hardly desirable.

An alternative approach is to use the expected

correlation between the time shifted signals received over two reciprocal paths as the basis of the selection procedure. This would require conditions along the paths to be very similar and for the disturbing wave group to remain substantially unaltered while travelling between the two transmitters if the signals are to be meaningfully compared. A method of achieving this is to use one central transmitter and two equidistant receivers one lying to the geomagnetic north and the other to the geomagnetic south of the transmitter. The two received signals would ideally be time shifted mirror images of each other and therefore suitable for correlation. Further study of sections for which the correlation coefficient exceeded a certain threshold would then have to be made manually.

One feature missing from the model is a source for the microbarums. The apparent predominance of northerly travelling wave groups, the required intensity ($4\ Wm^{-2}$) and the difficulty of propagating energy from other heights into a duct, all tempt speculation that the energy source could be in the auroral zone. It is not easy to obtain information as to the propagation direction of the groups without which this must remain pure speculation. As the acoustic wave manifests its presence by moving electrons up or down the field lines the probing transmitter and receivers must all in theory lie on a geomagnetic north-south line. (In practice a certain amount of scattering occurs. As can be seen from the frontispiece none of the transmitters lay directly magnetic north or south of

the receiver). Thus the wavelengths and velocities measured are in fact only the north-wouth components. An acoustic wave advancing on a broad front in an east-west direction would only cause a uniform rise and fall over the whole of that front which would not produce an observable change in the interference pattern. The apparent seasonal variation in the occurrence of micro-barums may be genuine or may be due to a seasonal variation in the preferred propagation direction. Again directional information is needed before one can tell.

By using two of the transmitter-dual receiver arrangements described above, each orientated geomagnetic north south but displaced from one another by some few kilometers in the east-west direction, the east-west velocity component could be estimated for some waves travelling at an angle to the north-south line. The closer the wave travelled to the east-west direction the less the interference effect produced to be observed, and therefore waves travelling due east-west still could not be detected.

All the above methods are based on the amplitude of the received signal. Provided a sufficiently stable reference is available at each receiver site measurements can also be made of the phase of the received signal. This may assist in the prelimary selection of records due to the characteristic phase variation across an inter-ference pattern. Lack of a suitable reference, as well as more information from the amplitude measurements than could readily be processed, precluded measurements of this type being made during this work.

Appendix One: The prediction of the sky wave signal to be expected from an acoustic wave group of arbitrary length.

In section 2-6 the possibility was mentioned of using the theory of section two to calculate the effect of a continuous wave as it moves onto, through and finally out of the range from x_{min} to x_{max}. As this has not been used in this experiment the details are given here in an appendix rather than in the main body of the text.

Let the leading edge of the wave be at x_{min} at time t=o, and the wave be artificially divided into groups. Let the velocity of each group be V and the group length be nL. Then, considering only reflections from the range from x_{min} to x_{max}, at time $t=nL/V$ only reflections from the first group are important, and the amplitude and phase can be calculated as above. At time $2nL/V$ the second group now occupies the space (space one) previously occupied by the first group which has moved on so that its centre is at $x = x_{min} + 3nL/2$ (space two). As the groups are assumed indistinguishable the resultant amplitude and phase are given by adding the effect of the group in space two to the previous result. Similarly at $t = 3nL/V$ the resultant is given by adding the effect of a group in space three to the previous result, and so on. The situation at the tail of the wave can be treated similarly by removing the contributions of the groups from space one onward, one by one. A continuous curve can be interpolated from these calculated points.

The result expected from a wave whose length is too great to permit analysis as a single group but too small to ever occupy all the distance from x_{min} to

x_{max} simultaneously can be treated in a similar way. The contributions from all groups centred in steps of $x = nL$ are calculated. At a given instant the wave occupies certain of these possible positions and these contributions are summed to give the resultant at this time. At a later time different positions are occupied and these contributions are summed. Again a continuous curve can be interpolated.

Only a wave whose total length is a multiple of nL can be accurately subdivided in this way. However as the wave groups being analysed are idealised this is probably not a severe limitation. As a very long wave train moves into the region of possible reflection the total non-specular reflected sky wave amplitude increases from zero in a series of oscillations which soon die away to leave a fairly consistant non-zero value. This is maintained until the trailing edge of the wave train moves into the region of possible reflection when the trend is reversed. No examples of sky wave amplitude variation of this type were observed during this work.

Appendix Two: Simplified circuit diagrams.

These notes are intended to accompany the circuit diagrams in figure xxi. These diagrams are highly simplified, biasing and all other detail not essential to the circuits operation being omitted. Circuit A illustrates the tuned amplifier which receives the signals sent down the coaxial line from the aerial unit. In this experiment one double tuned and one single tuned stage were used. The only point of note is the A.C. termination of the line whilst the supply for the aerial unit is fed into the line via a radio frequency choke.

Circuit B is a basic phase shifter in which a variable phase output is obtained by mixing different proportions of the two equal amplitude out of phase signals derived from the source and drain of the FET. The output impedance at these two points is not zero and there will therefore be some variation in amplitude of the output signal as its phase is varied. This is not important for the purpose to which this circuit is put, but would be unacceptable for the voltage controlled phase shift shown in circuit C. Here the two signals to be mixed are derived via emitter followers to ensure very low source impedance. For simplicity the capacitance C is varied rather than the resistance R, use being made of a voltage dependent capacitor for this purpose.

The voltage to control the phase shift introduced by circuit C is derived from the relaxation oscillator shown in circuit D. To give a right angled triangular output waveform the timing capacitor is charged from a

constant current source. The voltage across the capacitor is sampled by an insulated gate FET in order to provide a capacitor discharge time constant of hours; the programmable unijunction having been chosen for very low leakage. The sweep rate used was one cycle every few seconds.

Circuit E first amplifies and then detects the output from the sky wave rejection circuit. The gain is set by the D.C. voltage applied to the second gate of the dual insulated gate FET. The rectifying circuit is quite conventional and an extra low pass filter is used to minimise the effect of the signal modulation. This signal is compared in circuit F with a signal from the ground wave reference correction circuit. The variable resistor shown varies the dead band. Provided the voltage from the detector lies between the voltages existing at the ends of the variable resistor the output from the OR-wired comparators is low. When no voltage is applied from the ramp generator of circuit D ($\alpha = \alpha_{min}$), the voltage on gate two of the dual insulated gate FET is arranged so that the output to the comparators is the normal ground wave signal expected. As the voltage from the ramp generator rises on gate one so the voltage supplied to the comparator rises. It should be noticed that rather than providing a variable ground wave reference, a fixed ground wave reference is used as this is easier to provide. The variable gain amplifier in circuit E is used in match the actual day-time ground wave to the ground wave reference provided.

Circuit G is that of a variable time constant filter. The filter is basically a low pass one but in order to get the time constants required the capacitor has

to be produced by capacitance multiplication. The circuit shown produces a capacitance of $(R_1/\Delta R)C$ at the output terminals. Care must be taken to minimize the offset error of the operational amplifier. In this experiment, time constants from ten seconds to one thousand seconds could be produced by this circuit.

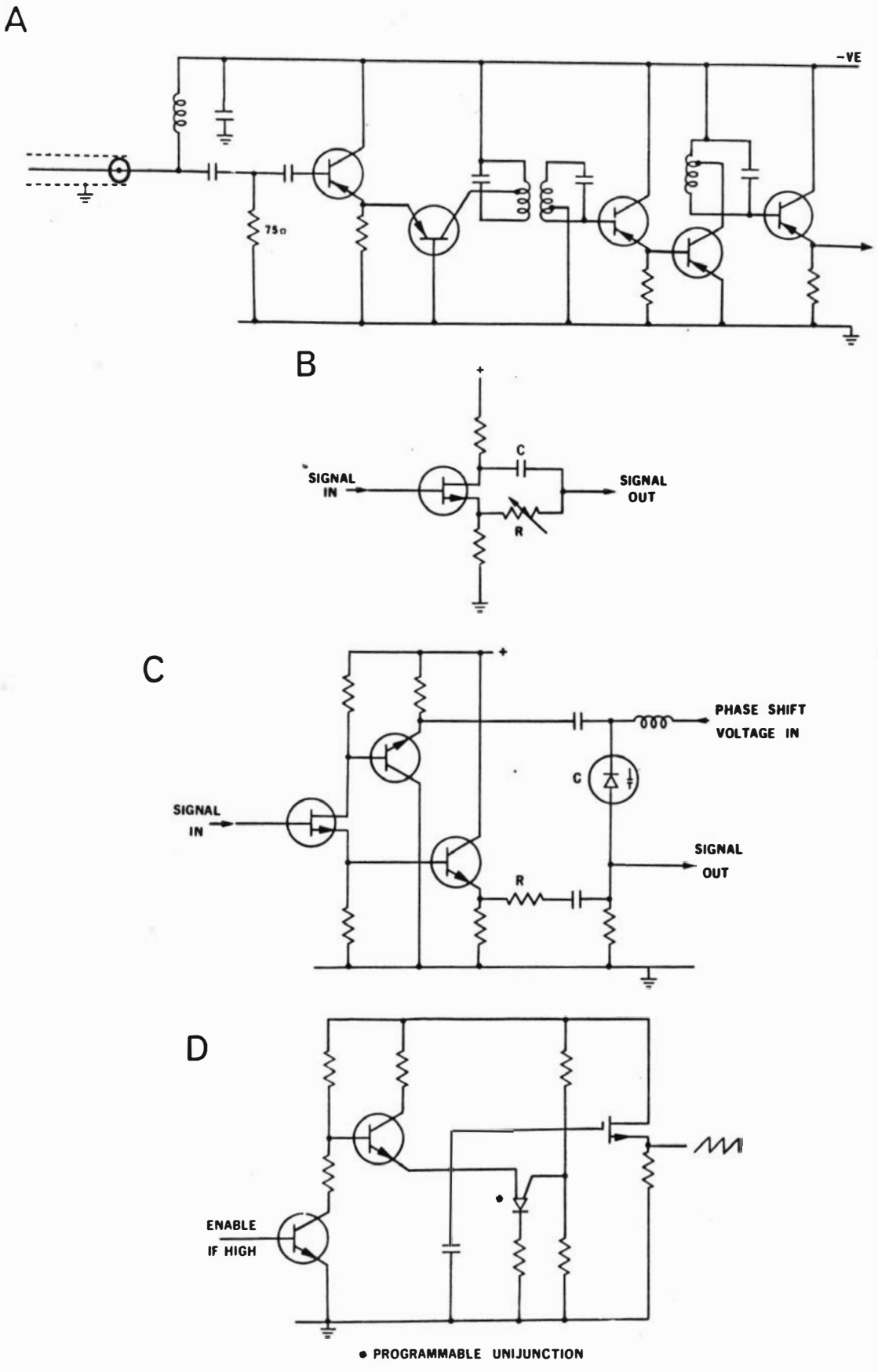
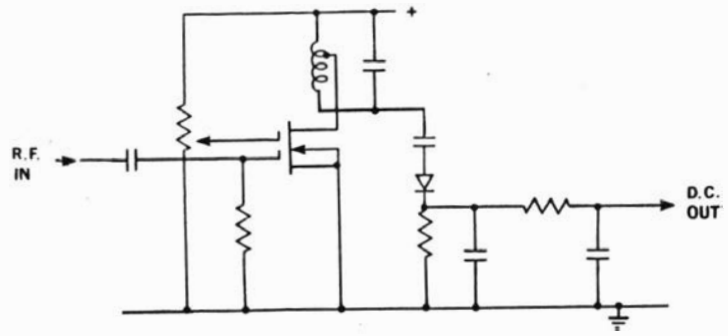
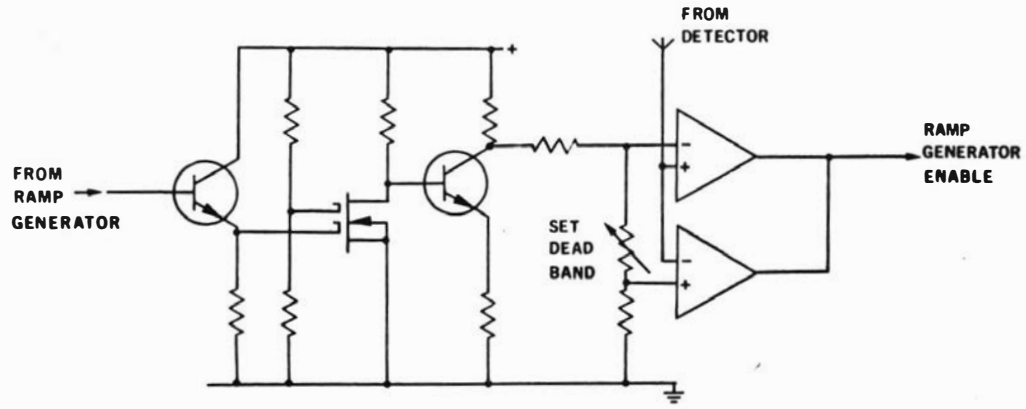


Figure xxi: Simplified circuit diagrams.

E



F



G

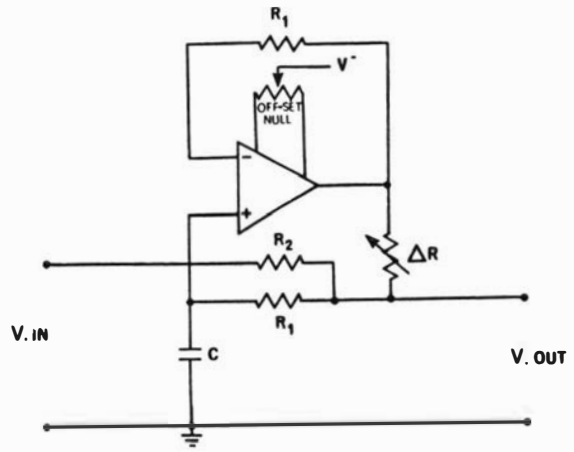
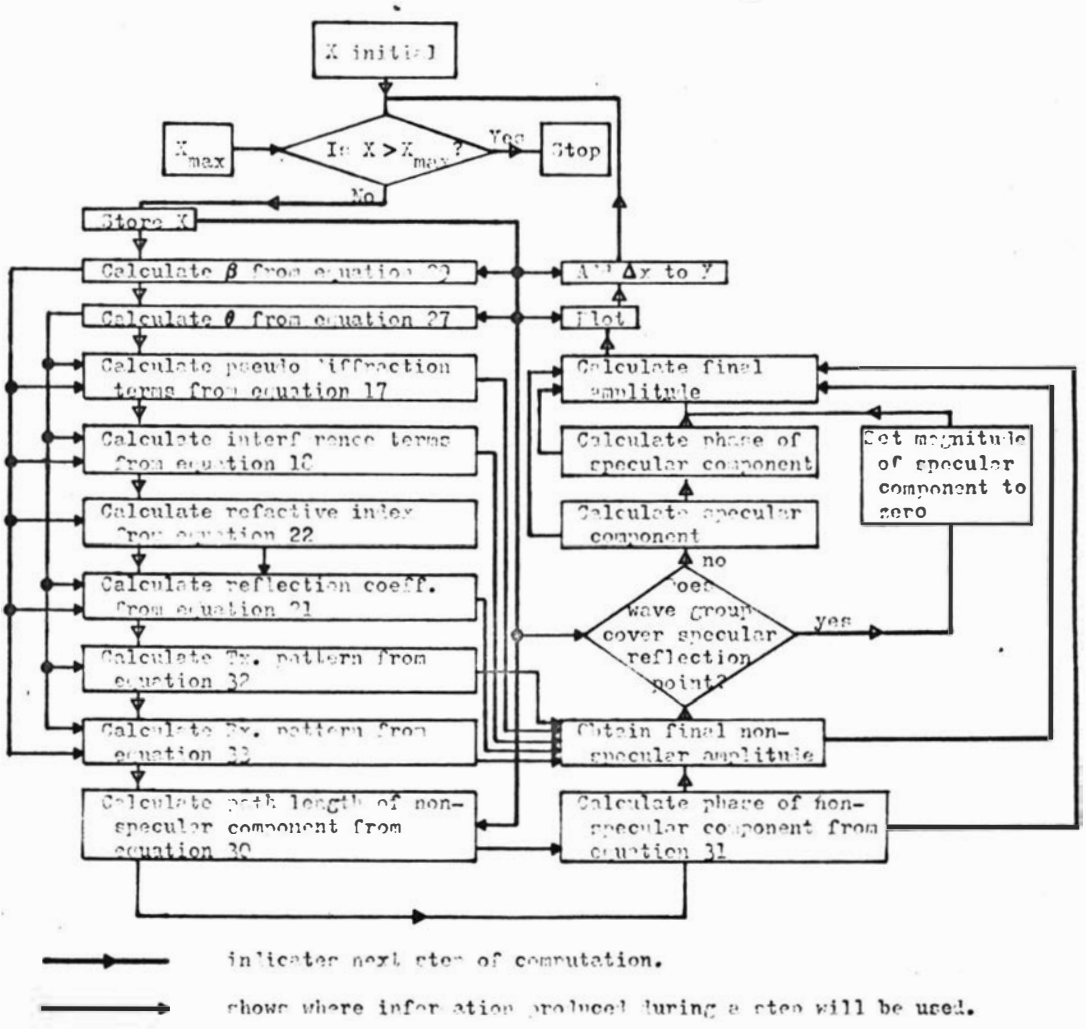


Figure xxi (contd.): Simplified circuit diagrams.



Appendix Three: Flow chart for calculating the expected sky wave amplitude using the theory of section two.

REFERENCES

Bain W.C., Bracewell R.N., Straker T.W.& Westcott C.H.	1952	Proc. Instn. Elect. Engrs. 99, 250.
Bibl K. and Olson K.	1967	Proc. Conf. on Ground-Based Radio Wave Propagation Studies of the Lower Ionosphere, Vol 2, p 538 Ottawa, Canada.
Bracewell R.N., Harwood J. and Straker T.W.	1954	Proc. Instn. Elect. Engrs. 101, 154.
Brooker H.G. and Dyce R.B.	1965	Radio Science, 4. 463.
Budden K.G.	1967	Radio Waves in the Ionosphere, Cambridge University Press.
Cavendish Laboratory	1951	Proc. IEEE, 98, 221.
Coll D.C. and Storey J.R.	1964	J.Res. natn. Bur. Stand. 68D, 1155.
	1965	ibid 69D, 1191.
Deeks D.G.	1966	Proc. R. Soc. (Lond) A291,413.
Doherty R.H.	1970	J.Atmosph. Terr. Phys.32,1519
Down W.L.	1967	Nature, 215, 1469.
Eccles D. and King J.W.	1970	J. Atmosph. Terr. Phys. 32,517.
Elleyett C. and Watts J.M.	1959	J. Res. natn. Bur. Stand.63D, 117.
Felgate D.G. and Golley M.G.	1971	J. Atmosph. Terr. Phys.33,1353
Fenwick R.B. and Barry G.H.	1967	Proc. Conf. on Ground-Based Radio Wave Propagation Studies of the Lower Ionosphere, Vol 2, p 532 Ottawa, Canada.
Fraser G.J.	1965	J. Atmos. Sci. 22,217
	1968	J. Atmosph. Terr. Phys. 30,707.
Fraser G.J. and Vincent R.A.	1970	ibid. 32,1591.
Gardner F.F. and Pawsey J.L.	1953	ibid. 3,321.
Geisler J.E. and Dickinson R.E.	1968	ibid. 30,1505.
Georges T.M.	1967	'Ionospheric Effects of Atmospheric Waves', ESSA Technical Report IER 57-ITSA-54.

REFERENCES contd.

Gossard E.E. and Paulson M.R.	1968	J. Atmosph. Terr. Phys. 30,885
Hale L.C., Hoult D.P. and Baker D.C.	1967	Space Research VIII,320.
Helliwell R.A.	1949	Proc. I.R.E. 37,887
Helliwell R.A., Mallinkrodt A.J.& Kruse F.W.Jr.	1951	J. Geophys. Res. 56,53
Johler and Walters	1960	J. Res. natn. Bur. Stand. 64D, 269
Jones R.M.	1970	Radio Sci. 5, 793
Kantor G. and Pierce A.D.	1968	J. Atmosph. Terr. Phys. 30, 1497.
Lindquist R.	1953	ibid 4, 10.
Mechtly E.A. & Smith L.G.	1968	ibid. 30, 363.
Radicella S.M.	1968	ibid. 30, 1745
Sagalyn R.C., Smiddy M. and Freden S.C.	1967	Space Research VII, 448.
Siegel S.	1956	" Non Parametric statistics for the Behaviourable Sciences" McGraw-Hill
Straker T.W.	1955	Proc. Instn. Elect. Engrs. 102c, 122.
Thomas L. and Harrison M.D.	1970	J.Atmosph. Terr. Phys. 32,1.
Von Biel H.A., Flood W.A. and Camnitz H.G.	1970	J. Geophys. Res. 75,4863.
Watts J.M.	1952	J. Geophys. Res. 57,487.
Weekes K. and Stuart R.D.	1952a	Proc. Instn. Elect. Engrs. 92, 29.
Weekes K. and Stuart R.D.	1952b	ibid 92, 38.