Effects of Ionospheric Scattering on Very-Long-Distance Radio Communication

H. A. WHALE

Director, Radio Research Centre
The University of Auckland
Auckland, New Zealand

PLENUM PRESS • NEW YORK • 1969

Library of Congress Catalog Card Number 76-84765

A Division of the Plenum Publishing Corporation
227 West 17 Street, New York, N. Y. 10011

Printed in the United States of America

PREFACE

In this work I have attempted to present a reasonably coherent account of many aspects of long distance communication by radio that have received little attention in the past. Investigation of these problems was undertaken since New Zealand, by reason of its location and its traditional ties with Europe as a market and a supplier has problems in the field of radio communication which are probably amongst the most difficult in the world. These were particularly troublesome during World War II and it was in an effort to solve these problems that the Seagrove Radio Research Station was set up by the New Zealand Department of Scientific and Industrial Research as a section of the Physics Department of the University of Auckland. This was largely due to the efforts of the New Zealand Post Office and New Zealand Broadcasting Corporation (encouraged by Professor Emeritus P.W. Burbidge of the Physics Department, University of Auckland) who were directly concerned with maintaining long-distance communications. For the last few years this organization has operated as a Department (the Radio Research Centre) of the University of Auckland. Since operations at the station began in 1951 the main interest has been in long-distance propagation and this, because of the nature of the problems which have arisen, has led to an allied interest in ionospheric irregularities.

Even in 1951 it was felt by many that the subject of ionospheric radio propagation had been "worked out". It certainly was not one of the fashionable subjects in physics at that time. It is now evident that the approach to most of the problems which were then recognised was one of producing more accurate solutions to more and more idealized cases. In the following pages it will be seen that, even for short distance radio propagation, there is a limit to the accuracy with which predictions can be made. This does not mean that there is anything basically wrong with our predicting techniques but rather that there are fundamentally unpredictable (or random) factors present. These assume greater importance as the propagation distance increases until, at distances greater than about 15000 kilometres, the essentially random factors become more or less dominant. From this point of view it can be said that the work presented here is an extension of that in the classical text books on radio propagation where essentially short distances are involved but goes a long way towards replacing the classical concepts where the propagation is over long distances.

I have tried to indicate, in this text, that although many questions have been answered, many new problems have been raised. In particular, of course, we have dealt with only one receiving location in most cases. The problem now is whether, with the increasingly widespread use of satellites for communication purposes, this type of investigation should be continued. The answer may be affirmative since, although satellite relay systems are ideal for fixed point communication, they do not as yet provide a solution to two-way communication between isolated people and peoples.

CONTENTS

CHAPTER 1

INTRODUCTION

In the relatively few years since 1904 when Marconi first demonstrated the feasibility of communicating over relatively long distances by means of radio waves, a tremendous volume of work on the subject of radio propagation has been published. Much of it has been concerned with obtaining more accurate agreement between theory and experiment in radio transmission *via* the ionosphere. As the distance over which such experiments are carried out increases, the whole problem becomes more and more complicated since there are always some ionospheric properties which are unknown somewhere along the transmission path. If the path length (i.e. transmission distance) is small, the lack of detailed information on the properties of the ionosphere over some portions of the path may not present a very serious obstacle to obtaining a reasonable estimate of the behaviour of radio waves in traversing it because some idea of the properties of the unknown regions may be obtained by interpolating between those regions where the characteristics of the ionosphere are known. The errors introduced by this procedure tend, however, to be cumulative in their effects on the calculated results so that very large errors may occur if too many interpolations are made. In general it may be said that the problem becomes intractable (even using modern electronic computers) if the path length involved is more than about 10,000 kilometres. It is under these circumstances that it must be recognized that an approach such as that presented here which is essentially statistical in nature and does not attempt to study detailed behaviour may still yield useful results.

The aim in this work has been to attempt to present an account of those general characteristics of long-distance radio propagation which can be deduced from what is essentially a study of the average (and hence statistical) behaviour. In other words, we are forced to accept the fact that we will probably never be able to *specify* the characteristics of the ionosphere along the whole of a particular propagation path even for a short period of time, let alone *predict* these detailed characteristics for future times. It is pointed out in chapter 3 that, although some of the major variations in ionospheric characteristics are predictable (the average diurnal variations), the minor variations are largely unpredictable i.e. they are more or less random in occurrence. While these latter variations may be minor in relative scale they may nevertheless be of paramount importance in determining some of the major propagation effects. We are therefore led to making measurements on the signals as actually received after traversing a long path and to attempting to specify, on the basis of these measurements, the limits of the variations which may occur in the characteristics of these signals.

The major part of the experimental results on which the work is based were obtained in Auckland; since there appeared to be very special problems associated with receiving BBC transmissions in New Zealand, many of the results were obtained using the normal short-wave broadcasts of the BBC as signals. There were considerable difficulties associated with this approach since these transmissions were not generally of the most suitable form for the type of investigation for which they were used. It will be seen that the basic data which is required is, at least, the variation of the signal strength and the direction of arrival of the signal with the time of day, with the frequency and with the location of the transmitting station. The difficulties introduced by using the available standard short-wave broadcasts arise from the fact that, in general, a particular frequency is only transmitted for a limited period each day. Unfortunately, the gaps can not usually be filled in by using other stations since stations in the same area tend to broadcast on similar frequencies at similar times of the day. Some special transmissions (e.g. the pulse transmissions discussed in Chapter 7) have been used in an endeavour to solve some particular problems but the great amount of data normally required for statistical analysis has made this approach impracticable in the general case. It will, in fact, be found that much of the following is based solely upon measurements of the direction of arrival of those signals which were available.

The situation as far as further investigations are concerned is, however, improving since, with the advent of artificial satellites which transmit radio signals, we now have a source that is standardized and that, in time, moves over almost the whole of the earth's surface. It is certain that many of the questions left unanswered here will be solved in the near future as these transmissions become more common. The major problem in utilizing satellites for this type of investigation arises from the fact that, in order that long-distance propagation around the earth may occur, the satellite must usually be located at a height which is below that of the maximum electron density of the ionosphere. This generally means that the satellite must orbit the earth at a height of about 100 miles or less and unfortunately the satellite then has a life-time of only a few weeks so that many consecutive launchings of similar satellites may be required in a long-term investigation.

CHAPTER 2

IONOSPHERIC PROPAGATION IN THE IDEAL CASE

The fundamental phenomena of ionospheric propagation over short distances are well known and are discussed, for example, by MITRA (1947)*. A most comprehensive handbook on the subject of ionospheric radio propagation is that of DAVIES (1965). The main relevant features in a simplified typical case of ionospheric propagation are shown in fig. 2.1. Here we have a transmitter located on the earth's surface at T and we will consider the fate of the four emitted rays T1, T2, T3 and T4, ignoring the fact that, because of the earth's magnetic field, each ray is generally split into two magneto-ionic components (the ordinary and extraordinary rays) which travel along somewhat different paths.

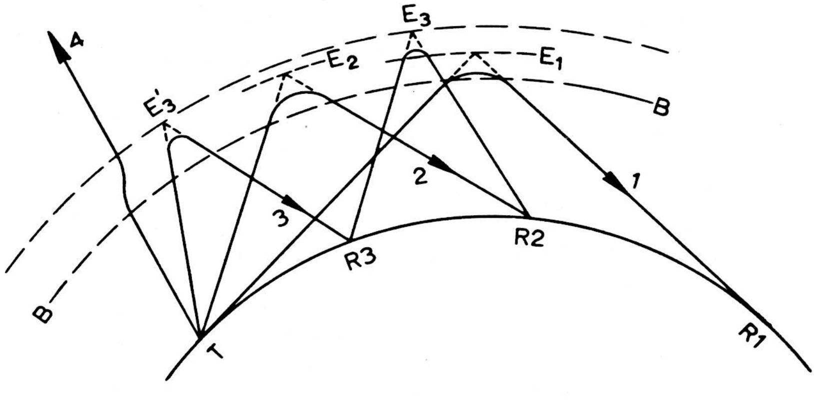

Fig. 2.1. Idealized hop-type ionospheric propagation showing the different effective reflection heights for different angles of incidence of the rays.

---

*All references are given after Chapter 21.

Ray 1 is the ray at the lowest possible angle of elevation i.e. it is tangential to the earth's surface at the transmitter at T. When this ray encounters the ionosphere (whose base is at B) it is refracted so that it returns to the earth, striking it again tangentially at R1. A great simplification is introduced if, instead of considering the actual curved ray path in the ionosphere, we regard the ray as travelling in straight lines and as being returned from a perfectly reflecting layer at E_1. Then the height of E_1 above the surface of the earth is called the "effective" or "virtual" height of the reflecting layer for this particular ray. We note that, in this particular case, even a small tilt of the layer near to E_1 can cause ray 1 either to miss the earth completely or to strike it at an elevation angle which is substantially greater than the tangential angle of zero degrees.

Ray 2 is typical of single hop transmission between T and a receiver placed at R2. The height of E_2 above the earth's surface (i.e. the effective height of the reflection point) is somewhat greater than that of E_1. Similarly, ray 3 leaving the transmitter at a slightly greater elevation angle will be reflected at a slightly greater effective height and will reach the ground at R3, say. It will then be reflected from the ground, will make another encounter with the ionosphere and, if the distance TR3 is exactly half TR2, it will reach the ground again at R2. This is the case of a two-hop transmission mode.

If the ionosphere is sufficiently dense, i.e. if the maximum electron density is sufficiently high, multiple hop modes of very much higher order may be possible.

As a final basic type of propagation, we have the ray T4 which is incident on the ionosphere at such a high elevation angle that it passes right through and will, in general, be lost in space. For this to occur, the transmitted frequency must be greater than the *critical* frequency f_c of the ionosphere. This critical frequency is determined by the maximum electron density of the ionosphere; it is one of the quantities which are measured regularly by the world-wide network of ionospheric sounding (ionosonde) stations and which are the subject of both short and long-term predictions by various government agencies. The best-known of these are the Radio Research Station, Slough, the National Bureau of Standards, Washington and the Ionospheric Prediction Service, Sydney.

In the very simple examples which we have considered above, none of the paths is more complicated than a two-hop mode. However, in order that exact predictions for the characteristics of even these simple paths may be made, it is necessary to have detailed knowledge of the properties of the ionosphere in the general regions near E_1, E_3, E_2 and between E_3' and a point directly above the transmitter. Such complete information is seldom available although some special measurements have been made by various workers under conditions which approached the ideal (i.e. all the relevant properties of the ionosphere were measured at the time of transmission). It is not our purpose to investigate or discuss the use of predicted values of ionospheric parameters in propagation problems since this subject is already very adequately covered in the literature. We will, however, consider (in Chapter 3) measurements made on a relatively simple radio link with the aim of estimating the accuracy and usefulness of such predictions in practice. It can be seen from fig. 2.1 that the propagation mode can be identified if measurements of the elevation angle are made at the receiving site. This is, in itself, a fairly difficult measurement to make but some of the methods which have been used are described in Chapter 4, for example. In practice, measurements of both elevation angle and bearing angle are made and these are generally referred to jointly as the *direction of arrival*. As the ionosphere changes through the day, it is to be expected that the elevation angle, in particular, will show some more or less regular variation. It is, in fact, found that although the averaged diurnal variations of the directions of arrival of the signals show a well-defined behaviour (when the averaging is carried out over about 30 days) there are significant departures from this average on any particular day. This is not to be taken as indicating that the predictions of the characteristics of the propagation path which are themselves based on predicted general ionospheric characteristics are of no value. Rather, it is confirmation of the generally accepted rule that ionospheric changes occurring in a time of less than about one month are essentially random and hence largely unpredictable.

CHAPTER 3

MEASUREMENTS ON A SHORT PATH

From a receiving station at R (in fig. 2.1) it should be possible, as has been pointed out, to determine the propagation mode (whether one-hop or two-hop, for example) by taking measurements on the incoming wave. As an example of the complications which occur in practice on even such a simple communications path we will consider the results of measurements taken at Auckland on transmissions from Brisbane, a distance of 2,250 kilometres. These measurements of the elevation and bearing angles were made on a rotating interferometer located at Seagrove, near Auckland. This instrument is described in the next chapter; at this stage it is sufficient to mention that, being a phase-comparison device, it yields a result which refers to the strongest wave in a group of incoming waves.

Firstly, it is found that, with any ionospherically reflected wave, the received signal *scintillates*. This is a term borrowed from the astronomers who use it to describe the fluctuations in brightness of a star which in their case arise from variations in the refractive index of the atmosphere and are an indication of atmospheric turbulence. They are accompanied by small fluctuations in the apparent position of the star. Similarly, in the case of radio waves, it is found that the received signal varies in its amplitude or "brightness" (i.e. it "fades") with typical fading periods of a few seconds and, at the same time, the apparent incoming direction or "position" also fluctuates with about the same period. This *radio* scintillation arises, not directly from variations in atmospheric density, but from variations in the electron density of the ionosphere. These latter variations do, however, seem to arise largely as a result of variations in the density of the atmosphere in which the ionosphere is "imbedded". In fact, of course, the ionosphere and its associated neutral atmosphere are very closely tied together since any change in electron density must be accompanied by a change in positive ion density, the medium as a whole remaining neutral; these large positive ions will then communicate any velocity changes to the neutral particles (also large compared with electrons) so that the atmosphere as a whole will tend to hang together.

We may note here that the fluctuations in the strength of the radio signal tend to be related to the fluctuations in the incoming direction of the wave, the weak signals tending to occur at times when the deviation of the wave from its mean position is greatest. This effect is discussed in more detail in Chapter 9 although a preliminary discussion of a simple case in which it occurs is given below in this chapter. For measurement purposes we often regard these scintillations as a complicating factor which could, with advantage, be removed. This can be achieved by an averaging process; in the measurements to be described in the section, the averaging time was about half an hour so that the effects of scintillations (which occupy times of the order of a few seconds) were completely removed.

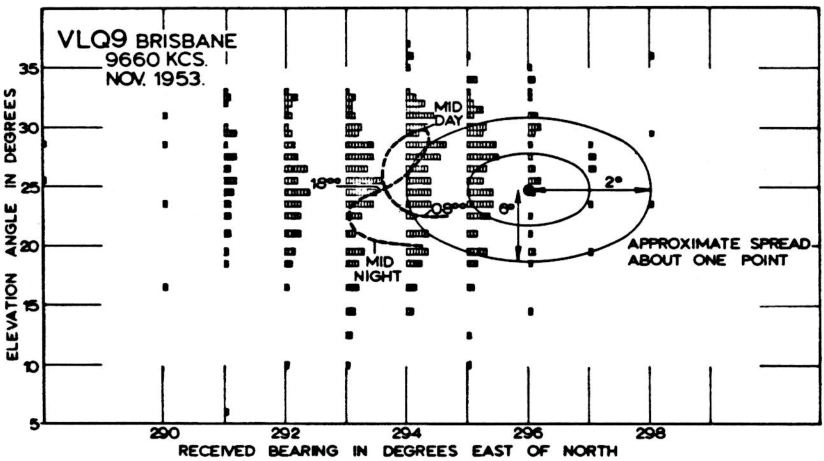

Fig. 3.1 The three types of variation in the received angle of arrival.
Dashed line - the diurnal variation averaged over the whole month.
Small rectangles - the actual measured directions averaged over one hour to remove the scintillation effects.
Ellipses - the approximate spread of incoming directions arising from scintillation effects.