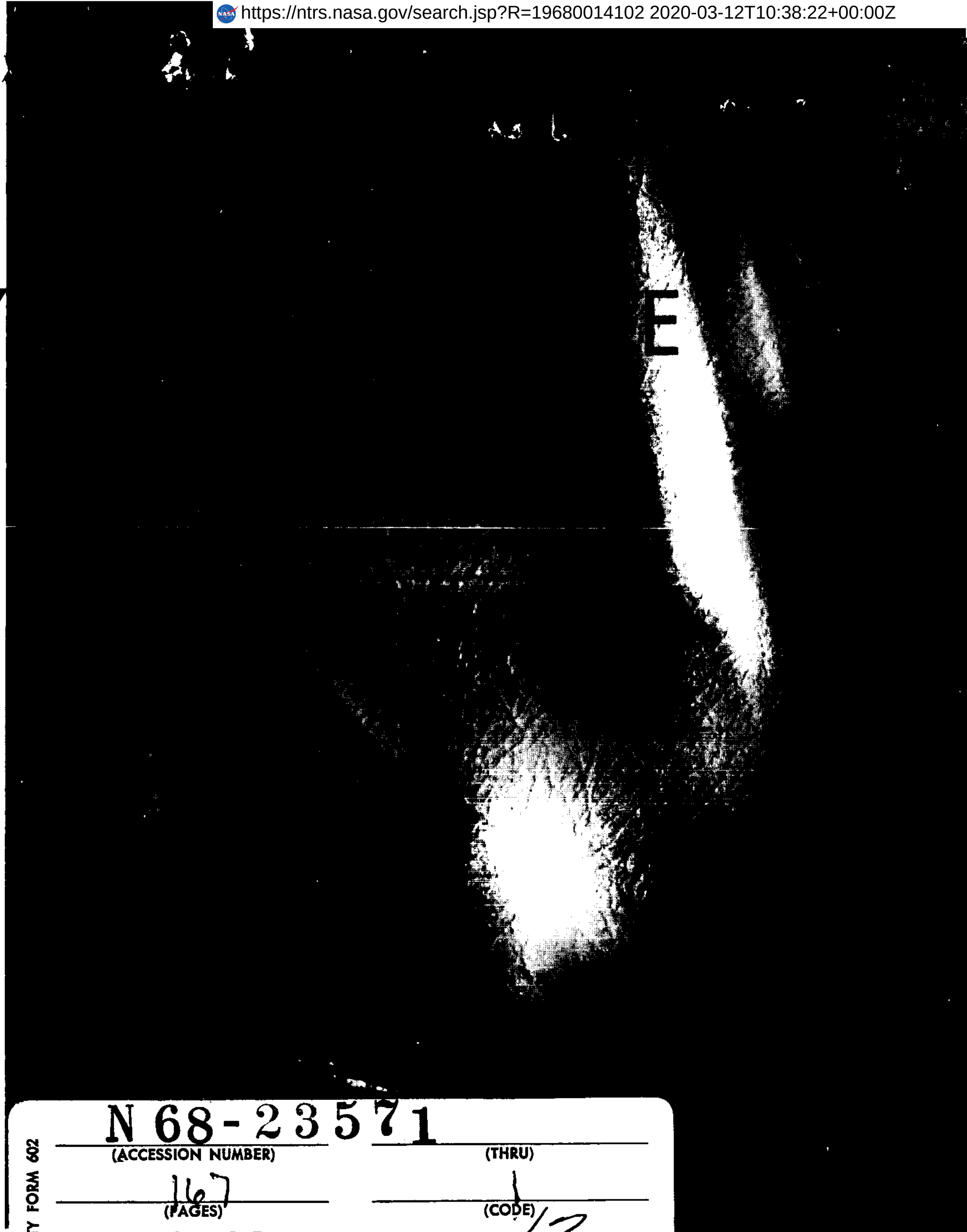
# GEOPHY E



**UNIVERSITY OF ALASKA**

**COLLEGE, ALASKA**

**UAG R-180 Vol. 1**





# RADIOWAVE SCATTERING STRUCTURE IN THE DISTURBED AURORAL IONOSPHERE: SOME MEASURED PROPERTIES

by

**Edward J. Fremouw**

**Scientific Report — Vol. 1**

## THEORETICAL CONSIDERATIONS

**Contract NAS5-3940 of Goddard Space Flight Center, NASA**
**and**
**Grant GP-947 of the National Science Foundation**

**JUNE 1966**

GEOPHYSICAL INSTITUTE

OF THE

UNIVERSITY OF ALASKA

# RADIOWAVE SCATTERING STRUCTURE IN THE DISTURBED AURORAL IONOSPHERE: SOME MEASURED PROPERTIES

by

Edward J. Fremouw

N68-23571

Scientific Report

Research Contract NAS5-3940 of Goddard Space Flight Center
National Aeronautics and Space Administration

and

Grant GP-947 of National Science Foundation

June 1966

Principal Investigator:

Edward J. Fremouw

Report Approved by

K B Mather

Keith B. Mather
Director

## ABSTRACT

A technique for quantitative description of radiowave scattering structure in the disturbed auroral ionosphere is developed in this work. Application is made by means of multi-spacing interferometric observations of a radio star. The work is based on the observed fact that sufficient scattering causes a measurable decrease in correlation of output voltages from neighboring antennas. Such correlation decreases are called visibility fades herein and have been called long-duration fades and radio-star fadeouts by other workers.

Random noise theory is employed, and it is assumed that the angular spectrum of the source, as received at the ground after scattering, is randomly phased. However, the usual assumption of a gaussian autocorrelation function to describe the scattering structure is circumvented, and provision is made for the existence of quasi-periodic structure. Further, the usual assumption of weak (single) or strong (multiple) scatter is avoided. The statistical characteristics of amplitude, phase, and complex signal are developed for the general case of arbitrary degree of scatter, using a numerical method.

The technique is applied to observations with phase-switch and phase-sweep interferometers, yielding two important parameters of the received wavefront, the coherence ratio and the wavefront autocorrelation function. The coherence ratio is defined as the ratio of nonscattered to scattered flux received from the source. The wavefront autocorrelation function is defined as the spatial

autocorrelation function of the scattered portion of the (complex) wavefront.

Two quantities which describe the ionospheric scattering region are obtained from the coherence ratio and wavefront autocorrelation function. First, the optical depth of the region (considered as a purely scattering medium) is determined from the coherence ratio. Second, the ionospheric structural autocorrelation function is established jointly from the wavefront autocorrelation function and the optical depth, yielding a statistical description of the average size and idealized shape of the ion-density irregularities which produced the scattering.

Forty-nine visibility fades observed at College, Alaska, between November of 1964 and February of 1966, inclusive, are analyzed. A majority of the fades revealed optical depths in excess of unity at 68 MHz. Optical depth is numerically equal to mean-square fluctuation in radio-frequency phase across a plane at the base of the scattering region, so the fades were characterized by rms phase deviations in excess of one radian at 68 MHz. An approximately inverse-square dependence of optical depth on frequency was obtained from simultaneous observations at 68, 137, and 223 MHz.

At 68 MHz, tri-spacing observations were carried out on east-west baselines of 110 meters ($25\lambda$), 220 meters ($50\lambda$), and 330 meters ($75\lambda$). The observations seldom were consistent with the demands of a gaussian autocorrelation function, as is commonly assumed. Rather, the disturbed auroral ionosphere displays evidence of quasi-periodic structure in the dimensional range of tens and hundreds of meters. The structure observed is comparable in size to auroral rays.

While most of the observations were consistent with the assumption of a randomly phased angular spectrum, a significant minority was not. Quantitative results could not be obtained in these instances, and they imply the existence of highly developed quasi-periodicity. Theoretical work is needed to bridge the gap between quasi-periodic structure in the sense of random-noise theory and strict periodicity.

Narrow-beam photometers were mounted on one of the interferometer antennas tracking the radio star. Auroral luminosity was recorded along the line of sight during 100% of the visibility fades which occurred at night under clear-sky conditions and during many night-time fades which occurred under cloudy conditions. Thus, VHF radio-star visibility fades in the auroral zone result from scattering by irregularities directly associated with auroral forms, at least at night.

# PREFACE

This report presents, in two volumes, work carried out at the Geophysical Institute under a NASA research contract, NAS5-3940, and an NSF grant, GP-947. It deals with radiowave scattering in the auroral ionosphere under disturbed conditions. Volume 1, containing Chapters I through III, presents some theoretical aspects of ionospheric scattering and its observable result at the ground. Volume 2, containing Chapters IV through VI, describes an experiment for ascertaining certain parameters of ionospheric structure and presents results of the experiment and conclusions to be drawn therefrom.

The fundamental aspects of this report represent research carried out under the Institute's IQSY program, funded by the NSF grant. Reports and papers by other authors describe work on four other problems of high-latitude geophysics carried out under the same program.

The experimental results herein reported provide a basis for more applied research aimed at describing effects of ionospheric scattering on satellite communications in the auroral zone. This aspect of the work was carried out under the NASA research contract. In this report, attention has been focused on certain severe scatter events - called "radio-star visibility fades" - which are of scientific interest in their own right and which represent the most severe ionospheric conditions a satellite communications system may be expected to encounter. Investigations of scattering under less disturbed conditions - as manifested by scintillation of radio-star and satellite signals - will be described in the final report of contract NAS5-3940.

A considerable portion of Volume 1 of the present report is devoted to a detailed review of previous theoretical work by one man, E. N. Bramley. The concepts developed by Bramley are fundamental to the theoretical considerations underlying the experiment. The author found a thorough understanding of Bramley's work essential before progress could be made on interpretation of observations. Accordingly, much of Chapters II and III represent merely foundations for generalization of Bramley's work.

The reader who is familiar with Bramley's work on spaced-aerial reception of scattered waves and with the work by Rice on signal statistics, which underlies Bramley's considerations, probably will find that much of sections IIC, IID1, and IID2 may be omitted. Such a reader also will find the concepts discussed qualitatively in section IIA to be familiar. The reader unfamiliar with these concepts and with Bramley's work probably will find the above-mentioned sections necessary for understanding section IID3, which is a generalization of Bramley's work. It is this generalization which allows experimental determination of quantitative results concerning the disturbed auroral ionosphere.

In Chapter III, the reader who has worked actively on the problem of ionospheric scatter again will find many familiar concepts. In particular, the reader who is thoroughly familiar with Bramley's work on scattering, per se, as opposed to observational considerations, may omit sections D1 and D2. Sections D3 and D4 represent generalizations on this work. Related work is reviewed and generalized very simply in section IIIC. The other sections of Chapter III are essentially reviews of well-known concepts, applied to the experimental problem attacked in later chapters. The generalizations achieved in Chapter III allow

relaxation of assumptions, which were previously necessary, concerning the nature of ionospheric irregularities.

The experimental work reported in Volume 2 involved the efforts of several people. Salient among these were Mr. R. C. Domke and Mr. W. O. Starner. As electronic technicians at the Geophysical Institute, they contributed a great deal of time and ability to development of the experimental apparatus and to gathering of data. Contributions were made by Mrs. Nita Balvin and Mrs. Carolyn Grover in the scaling and reduction of data.

The active cooperation of Mr. J. M. Lansinger is gratefully acknowledged. The reliable and versatile phase-sweep interferometer developed under his direction at Boeing Scientific Research Laboratories in Seattle lay at the heart of the experiment. The author is further indebted to Mr. Lansinger for data reduction of the phase-sweep observations and for making available the computer facilities of the Boeing Company, as part of a joint IQSY effort by BSRL and the Geophysical Institute.

Thanks go also to Dr. Leif Owren, who initiated the research program which led to the work reported herein and who served as a consultant to the project under the NASA research contract.

The cover photograph, showing an auroral drapery silhouetting one of several radio telescopes used at the Geophysical Institute for radio-star observations, was taken by Dr. V. P. Hessler.

## TABLE OF CONTENTS

### Vol. 1 - Theoretical Considerations

TABLE OF CONTENTS (CONT'D)

TABLE OF CONTENTS (CONT'D)

Vol. 2 - Experimental Considerations

TABLE OF CONTENTS (CONT'D)

## LIST OF ILLUSTRATIONS

Page

Page

LIST OF TABLES

# CHAPTER I

## SURVEY OF STUDIES OF IONOSPHERIC IRREGULARITIES

### 1A EARLY STUDIES

The irregular structure of the terrestrial ionosphere is perhaps its most long-standing unexplained major observed feature. The discovery of the ionosphere, whose existence had been inferred by Balfour Stewart as early as 1878[1], is usually dated as 1901, the year of Marconi's first transatlantic wireless transmission. In 1902, Kennelly and Heaviside independently postulated the existence of a high-atmospheric conducting layer to explain Marconi's success. There followed nearly a quarter century of theoretical and experimental work aimed at proving or disproving the existence of such a layer. This early work yielded results, such as development of the magneto-ionic theory, which were later to be of fundamental importance to the study of the ionosphere and to its utilization for long-distance radio communications. Yet, throughout this period, there was doubt as to the ionosphere's very existence. Obviously, descriptive ionospheric research had not yet begun.

It was in 1925 that Appleton and Barnett clearly received a component of BBC transmissions propagating downward after reflection from the ionosphere. Immediately thereafter, a number of workers employed upward directed antennas. Notable among these were Breit and Tuve, who in 1926 employed pulse modulation in the first ionospheric sounder. Now the job of describing the iønosphere could and did begin

[1] Kelso (1964) credits Gauss with a still earlier suggestion, made in 1839.

in earnest. By the time of the second International Polar Year, 1932-33, many of the major features of the ionosphere were known and some were on their way to explanation. The normal E and F layers were recognized and their equivalent heights were established. Some knowledge existed about the seasonal and diurnal variations of the E and F ionization maxima at middle latitudes. Chapman already had developed his theory on the ionizing effect of monochromatic radiation incident on the atmosphere. This theory proved capable of explaining a remarkable number of the observed characteristics of the "regular" ionosphere.

As early as 1930, Appleton observed effects of irregular ionospheric structure. Still today, however, there is no comprehensive theory to explain the existence of localized irregularities of ionization. The abnormal E region ionosonde returns which Appleton had observed have since become known as sporadic E, and the phenomenon received considerable attention during the second Polar Year. Much has been learned about it since, but most present-day ionospheric physicists probably still would add "amen" to the statement of Appleton, Naismith and Ingram (1937) that "we feel that a completely satisfactory explanation of its occurrence on all occasions is still lacking."

While the earliest sporadic E traces reported in the literature were regarded as abnormal, there was no reason to interpret them as being produced by a spatially irregular reflecting region. One of the most direct early observations of ionospheric irregularities, as the term generally is used today, was reported by Eckersley (1937). Eckersley observed "momentary ionospheric echoes from irregular heights, occasionally as low as 60 km, but generally between 100 and 300 km," and commented that "we can scarcely avoid the conclusion that the echo signals are not regularly refracted from a uniform layer in the ionosphere but are scattered from irregularities or clouds." Earlier instances of "lateral deviation" of radio signals reported by Ratcliffe and Pawsey (1933) probably were caused by scattering from similar but longer-lived ionospheric irregularities. It appears that Pawsey actually observed drifts of such irregularities as early as 1935, using spaced receivers (Ratcliffe, 1956).

One of the most commonly observed effects of ion-density irregularities is the spreading of ionosonde F-layer traces, the so-called spread-F phenomenon. To some extent, spread F has been considered a nuisance in the routine scaling of ionosonde records for F-layer parameters such as virtual height and critical frequency. As a consequence, many modern sounders have been adjusted in such a manner as to decrease their utility for studying the irregular structure of the F layer. Still, a considerable amount of effort has gone into study of the spread-F phenomenon itself. The first such effort appears to have been that of Booker and Wells (1938), who interpreted spread-F ionosonde returns in terms of Rayleigh scattering by spatial irregularities in the F layer.

Not surprisingly, little progress in describing ionospheric irregularities was reported in the open literature during the years of World War II. E-layer irregularities were considered by Eckersley (1939) and by Eckersley, Millington and Cox (1944), however, in explaining a number of signal propagation modes. Immediately after the war, exploration of ionospheric structure received renewed attention and the benefit of wartime advances in radio technology. Wells, Watts and George (1946) reported observations of rapidly moving irregularities in the F region during a magnetic storm, using a new ionosonde with increased time resolution. They attributed the irregularities to an influx of extraterrestrial particles and interpreted the motions as being downward from great heights into the F layer. However, their method allowed measurement only of range change and not of the angle of motion. Postwar radio and radar techniques of considerable variety led to ever increasing observations of various types of irregular ionospheric structure. While they did not produce a comprehensive understanding of the structure, the expanded numbers and types of observations did begin to reveal the complexity of the problem.

From advances in the statistical description of electrical signals, was born a series of theoretical studies which allowed more effective use of the new wealth of observational material. One of the first papers in this vein was by Ratcliffe (1948), in which it was suggested that "if the roughness of the ionosphere is supposed to vary in a random manner, then the salient phenomena of the fading of a single wave can be explained." This paper explored the observations to be expected from a vertical-incidence sounder due to "diffractive reflexion" from randomly placed scattering centers, which were assumed to be in random motion. The paper applied to the ionospheric problem the work of Furth and MacDonald (1947) on the statistics of random noise signals.

Of more general application were - and continue to be - the fundamental papers on analysis of random noise by Rice (1944, 1945). An important feature of Rice's work is that it included the case of a sinusoidal signal in the presence of noise. McNicol (1949) recognized the importance of this case to the problem of ionospheric reflection which includes a specular component as well as scattered components. He applied Rice's work to normal-incidence and oblique-incidence receptions at a single point, determining the ratio of specularly reflected to scattered power. By applying also the work of Booker, Ratcliffe and Shinn (1950), then in the process of publication, he was able to calculate an effective velocity of the ionospheric irregularities responsible for the signal fading which he observed. Working with a single receiver, however, McNicol was not able to ascertain whether his effective velocity measured an ordered ionospheric drift, a random motion of irregularities or a combination of the two.

As important as Rice's results to the problem of ionospheric scattering was work by Booker and Clemmow (1950), which generalized the concept of a polar diagram to that of an angular spectrum of plane waves. This generalization allowed familiar Fourier transform techniques to be brought into play. A marriage of the statistical concepts of Rice and the Fourier techniques of Booker and Clemmow brought forth the fundamentally important paper of Booker, Ratcliffe and Shinn (1950).

Booker, Ratcliffe and Shinn, in basing their work on the derivations of Booker and Clemmow, stated their considerations within the context of diffraction processes. To some extent such a context tends tacitly to introduce unnecessary restrictions on application of the results. The term diffraction brings to mind a discontinuous medium, with significant changes in transmission parameters taking place in the propagation direction within distances comparable to the observing wavelength. In point of fact, no such restriction exists in the work. The derivation of the angular spectrum starts with consideration of an irregular electromagnetic wave-field assumed to exist without statement of its originating mechanism. Thus, any form of scattering which can produce an irregular wave-field and an associated angular spectrum can be examined with the mathematical tools provided by this paper. Several workers have used the tools to attack the problem of scattering by an inhomogenous, phase-changing medium - a process which might descriptively be called differential refraction, but which, by and large, has continued to be called diffraction. The situation is similar to use of the term reflection to describe the return of radio signals from the ionosphere. Except for the lowest frequency radio waves, the term refraction more accurately describes the physical process involved, but the two terms are used somewhat interchangeably.

Immediately after publication of the work by Booker, Ratcliffe and Shinn, a number of papers appeared which restated, simplified, and extended its results and underlying concepts and applied them to experimental problems. The experiments performed on this basis generally fell into three classes: single-receiver measurements carried out on ionospherically returned, man-made signals; spaced-receiver measurements of similar signals; and measurements carried out on signals received from radio stars. Representative of the first type of work was that reported by McNicol (1949), mentioned above. The considerable amount of information available from spaced-receiver measurements based on the results of Booker, Ratcliffe and Shinn was described in a rather complete treatment by Briggs, Phillips and Shinn (1950).

## IB INTRODUCTION OF RADIO ASTRONOMICAL TECHNIQUES

The above-mentioned single-receiver and spaced-receiver methods both depend upon return from the ionosphere, by one or several possibly complicated propagation modes, of a signal transmitted from the earth. Discovery during the 1940's of discrete extraterrestrial radio sources, the so-called radio stars, led to an alternative means of investigation in which the basic propagation mode is simple transmission through the ionosphere. Some of the earliest observations of the strong source in Cygnus - Cyg A - revealed irregular fluctuations in the strength of the received signal (Hey, Parsons and Phillips, 1946; Bolton and Stanley, 1948). The fluctuations in the received intensity of Cyg A and of the stronger source discovered by Ryle and Smith (1948) in Cassiopeia - Cas A - became the subject of considerable investigation.

Simultaneous observations at Cambridge and Jodrell Bank by Smith (1950) and Little and Lovell (1950) showed no correlation for most of the variations, demonstrating that they were not inherent in the radio sources. The latter authors carried out spaced receiver measurements on baselines of 100 meters and 3.9 kilometers and concluded that the flucuations originated near the earth. They suggested as a mechanism a process analagous to the twinkling - or scintillation - of optical stars, due, in the radio case, to localized electron clouds in the F region.

The almost simultaneous appearance of the theoretical work on random scattering by Booker, Ratcliffe and Shinn and the experimental conclusions of Little and Lovell led immediately to application of the tools of radio astronomy to ionospheric physics. Perhaps the earliest paper reporting work which might be termed ionospheric radio astronomy was one by Ryle and Hewish (1950). Carrying on from the suggestion of Little and Lovell that radio-star scintillations might be caused by phase variations suffered during passage through F-layer ionization irregularities, Ryle and Hewish established the diurnal variation of scintillation intensity and compared it with that of spread F. They found that, as for mid-latitude spread F, the scintillations which they observed in England represented a nighttime phenomenon. They found a cross-correlation coefficient between indices of the two phenomena of between 42 and 45 percent. From this result they concluded that radio-star scintillations and spread F are related (at middle latitudes) and that the two may be caused by the same F-layer irregularities.

That radio astronomical observations could effectively supplement the older methods of ionospheric measurement was demonstrated in consideration of the diurnal variations of spread F and scintillation. In the case of spread F, it was not known whether the scattering irregularities themselves displayed a diurnal variation of occurrence or rather existed round the clock, being masked during the day by reflection from a normal F layer below them. Since a radio star signal necessarily traverses the entire ionosphere, the observed diurnal variation of scintillation is a direct indication of the diurnal behavior of the scattering irregularities.

The instrumentation developed for radio astronomical observation also supplements that of the earlier ionospheric techniques. In particular, having a need for high angular resolution in observing discrete celestial radio sources, radio astronomers introduced the radio interferometer (Ryle, 1950). From one point of view, the radio interferometer is an extension of the spaced receivers used earlier by ionospheric researchers (although the development was by analogy with an optical instrument, the Michelson steller interferometer). The extension involves retaining phase coherence between the two receivers by using a common local oscillator and adding the intermediate-frequency signals so that the resultant depends upon the phase difference between the two received signals. At the expense of losing discrimination between ordered drift and random change of the amplitude pattern on the ground, the instrument permits observation of the phase as well as the amplitude of the irregular wavefront arriving from the ionosphere.

That the phase distribution of the wavefront contains ionospheric information not available in the amplitude distribution alone was demonstrated by Hewish (1951) in his extension of the work of Booker, Ratcliffe and Shinn (1950) for the case of a structured phase-changing ionospheric layer. From measurements involving both the amplitude and phase of the wavefront, Ryle and Hewish (1950) and Hewish (1951) concluded that ionospheric structure with a horizontal scale of a few kilometers and consisting of less than one percent variation in the phase path thickness of the ionosphere could explain their observed scintillations.

Carrying out a somewhat more extensive program of amplitude observations, Little and Maxwell (1951) also found scales of a few kilometers and a close association with spread F. In addition, using the radio star Cygnus A, they found scintillation amplitude to be very nearly proportional to atmospheric path length when the line of sight traversed the mid-latitude ionosphere but to increase sharply when the source was viewed through the auroral-zone ionosphere.

Meanwhile, Hewish (1952) was extending the observational program based on his earlier theoretical considerations, using both spaced receivers and a greatly improved interferometric instrument, the phase-switch interferometer (Ryle, 1952). He found that many of his observations could be explained by the drift of otherwise fairly stable ionospheric structure, mainly along east-west lines. The scales which he deduced averaged about five kilometers, and the ionospheric drift velocities ranged generally between 100 and 300 meters per second. He found a clear relationship between the velocity and geomagnetic activity. Devising a means of deducing the height of the irregularities from simultaneous measurements of amplitude and phase fluctuations, he estimated the height at about 400 kilometers.

Hewish's considerable success in deducing certain characteristics of the scintillation-producing irregularities was limited by two simplifying assumptions in his theoretical work. First, following Booker, Ratcliffe and Shinn (1950), he worked within the context of a thin defracting screen, which allowed very little consideration of the scattering process actually taking place in the ionosphere. Essentially, Hewish had to postulate a wavefront containing deviations in phase immediately below the scattering layer. He then analyzed what happens to the wave in subsequent propagation to the ground. Second, Hewish emphasized the special case of small phase deviations (less than one radian variation across the screen). While most of Hewish's observations met the requirements for this special case, we shall see later that generalization is required for discussion of certain scatter events, particularly in the auroral zone.

In an ambitious and significant theoretical work, Fejer (1953) seems to have made the first step toward considering a scattering medium of greater physical plausibility, namely a thick slab containing irregular variations in refractive index. In this work, Fejer made use of various methods and concepts previously used in analysis of radio-wave scattering by tropospheric turbulence (Booker and Gordon, 1950), sound-wave scattering by small random variations in refractive index (Ellison, 1952), and scattering of x-rays by particles larger than the wavelength (Dexter and Beeman, 1949). Fejer's analysis included an extension of the work of Booker, Ratcliffe and Shinn by considering the case of a thin screen with irregular structure in two dimensions rather than in one. More fundamentally, he went on to consider also the case of a thick scattering region with three-dimensional irregularities in refractive index.

In considering the thick-slab case, Fejer first made the simplifying assumption that scattering takes place only from the incident wave - i.e., he first considered the case of single scatter. In this case, Fejer's results are essentially the same as those of Booker and Gordon (1950) except for choice of a physically more realistic autocorrelation function to describe the random nature of the irregular dielectric through which the wave passes. The most significant advance made by Fejer's paper was his analysis of multiple scattering by a thick layer having an irregular dielectric constant. For this important case, Fejer worked out the angular spectrum and the autocorrelation function of the scattered wave for a thick medium containing spherically symmetric irregularities. Thus Fejer overcame the first of the two limitations of Hewish's earlier work.

Fejer was able to discuss the physical significance of Hewish's other limitation - namely, the special case of small phase deviations. He showed that if the slab were sufficiently thin and/or its dielectric irregularities sufficiently weak, then the conditions of Hewish's special case would be met. Ouantitatively, these conditions correspond to a scattering layer which allows a fraction of the incident power, equal to or greater than about $e^{-1}$, to emerge unscattered. More generally, Fejer showed that the mean-square phase deviation imposed on the wavefront by the screen is equal to the thick-slab parameter which he called "the effective depth of scattering." This latter parameter was given by Fejer as

$$B_o = \int_o^{z_o} A \, dz$$

where $z_o$ is the thickness of the slab and A is the fraction of flux scattered from an incident beam by a unit thickness of the slab.

Further consideration of the relationship between Fejer's thick-slab model and the thin diffracting screen assumed by Hewish and by Booker, Ratcliffe and Shinn was carried out by Bramley (1954). The primary goal of Bramley's short paper was to show "that the angular spectrum, as derived by considerations of multiple scattering, for transmission through a thick stratum containing normally distributed and statistically isotropic irregularities of dielectric constant with autocorrelation function of the form exp $(-r^2/l^2)$, can equally well be evaluated by considering an equivalent thin phase-changing screen." This he did, and the relationship between a thick scattering layer and an equivalent thin screen, which he elucidated, is useful. There is a danger, however, in thinking in terms of the equivalent screen rather than in terms of the more realistic thick layer. It is often convenient from an observational point of view, but it tends to draw attention away from the ionospheric phenomena which are, after all, the final goals of this whole endeavor. It is to be noted, happily, that Bramley did not fall into his own trap. We shall have occasion in Chapter II to examine more of his work in detail (Bramley, 1951, 1953, 1955).

## IC WORK DURING AND SINCE THE IGY

Much of the work described in the foregoing, as well as the results of other workers, was summarized in an excellent review paper by Ratcliffe (1956). Since Ratcliffe's review, investigations of irregular ionospheric structure have been extended greatly. The International Geophysical Year accelerated experimental investigation in this field as in all branches of geophysics. Regarding ionospheric irregularities, the IGY left primarily a two-fold experimental legacy. First, the standard instruments of the ionospheric physicist - in particular, the ionospheric sounder - were deployed around the world. Second, new techniques were introduced - most notably techniques employing artificial earth satellites.

The very existence of artificial satellites in orbit about the earth and transmitting back to it provides a powerful technique to complement the radio astronomical methods. Just as the signals received from radio stars scintillate due to scattering in the ionosphere, so do signals received from satellites orbiting above the scattering layer. Clearly the relatively fast-moving satellites can be applied in combination with the slower natural radiators for descriptive studies of certain spatial and temporal characteristics of ionospheric irregularities. When rapid advances in miniaturization of electronic circuits allowed placement of ionospheric sounders themselves in orbit, another new and at least equally powerful technique was made available - that of "topside sounding."

Ionospheric scattering layers can now be observed from above as well as from below.

The tremendous number of observational data made available during and after the IGY gave impetus to new theoretical investigations, as is to be expected. On the whole, the theoretical work took a somewhat different turn from the majority of that carried out before Ratcliffe's 1956 review paper. Most of the earlier work necessarily was directed toward interpretation of observations in terms of ionospheric scattering processes. That is, the theories developed strived to describe the scattering processes themselves. This work leaned heavily on the techniques of signal statistics, which arose largely in engineering problems of radar development.

The later theoretical work has taken a more geophysical turn, concentrating on the dynamics of irregularity development and motion. The seeds of such investigation, of course, were sown earlier, and Ratcliffe's review article describes observations and analytical techniques employed in the study of irregularity motion. (Another comprehensive review of this aspect of pre-IGY irregularity studies was given by Briggs and Spencer, 1954.) In the early studies of irregularity motion, the existence of the irregularities usually was accepted and the dynamical aspects of the problem were confined to the drift (possibly coupled with some kind of random motion or growth and decay) of what were taken to be clouds of excess ionization. There was always a question as to the relation between motion of ionization and gross motion of high-atmosphere neutral gas. The later workers have questioned also the relation between ionospheric motions and the very nature of the observed irregularities.

In much of the recent work, attention has been given to the possibility that moving irregularities do not necessarily represent directly the motion of atmospheric ionization - let alone that of the neutral gas. In this case, the irregularities are viewed as propagating waves, which simply perturb the ionization density as they travel. Some of this work has been summarized briefly by Hines (1964). Even the seeds for this line of thought had been sown much earlier. Observations carried out as early as 1937 were interpreted before the IGY in terms of compressional waves or "travelling ionospheric disturbances" (Munro, 1950). This term was originally and usually still is applied only to irregularities with a scale of many tens of kilometers, whose propagation can be traced by ionosondes over distances of many hundreds of kilometers.

The more recent work strives also to interpret smaller-scale irregularities in terms of wave phenomena. For instance, Farley (1963) has investigated the possibility of the production of field-aligned irregularities by means of a two-stream plasma instability in the equatorial electrojet. He concluded that the mechanism ought to produce ion waves having many of the characteristics displayed by one type of equatorial sporadic E.

Thus, not only are planetary-scale dynamic phenomena important to irregularity motions but possibly to irregularity existence as well. Certainly the general interdependence of ionospheric irregularities, however produced, and global circulation effects, especially those of the atmospheric dynamo, will receive much attention in the next few years. The world-wide observations already carried out, however, argue against hope for a simple comprehensive theory to explain irregular ionospheric structure in the way that Chapman's theory explained regular layer formation so early in the development of ionospheric research. We probably are dealing with several originating mechanisms, which may not necessarily be very closely related.

The multiplicity of observed characteristics of ionospheric irregularities has often led to classification according to at least two variables. First, there is often a classification on the basis of height - usually according to particular ionospheric layers or regions. Thus, there exist collected results of E-layer irregularity studies (Smith and Matsushita, 1962) and of F-layer irregularity studies (Newman and Penndorf, 1966). Second, there is often classification according to latitude. The latter classification is sometimes made on the basis of observed characteristics, as for instance the spread-F occurrence minimum near 30 degrees magnetic latitude (Lyon, Skinner and Wright, 1962). On the other hand, the latitude classification may arise only for the obvious reason that individual experimenters find their observations confined to distinct latitude ranges. For instance, Bowles and Cohen (1962) have pointed out similarities between certain types of E-region irregularities observed in equatorial regions and certain aurorally associated irregularities observed at high latitudes. In addition, physical relationships may be expected between irregularities at different heights and have been reported by Thomas (1962) and others.

At least as complicated as the relationship between irregularities at different latitudes and at different heights is the relationship between the effects observed with different experimental techniques. International conferences, in fact, have been devoted to the results of particular observing techniques (Aarons, 1963). Obviously, much work needs to be done on relating observations obtained with different techniques, at different places, and due to irregularities at different heights (when the height is known). In the meantime, there is need for work with specific observing techniques (or judiciously chosen combinations) at specific latitudes. Nowhere is the need for descriptive work more acute than in the auroral and polar regions. Not only are past observations scarcest in high-latitude regions, but also the phenomena are often more complicated there than at lower latitudes.

To a large extent, observations in the polar regions have been restricted to those obtained with ionosondes (ground-based and topside). Calvert and Schmid (1964), for instance, have reported on world-wide spread F as observed from above the F layer. Penndorf (1962, 1964), using bottom-side ionograms, has found permanent maxima of spread F in the northern and southern magnetic polar regions and travelling maxima which circulate around the boreal and austral auroral zones on the midnight meridian.

At auroral-zone latitudes, other observing techniques also have been used. In the present work, VHF observations of radio stars have been employed. Before proceeding to the specifics of this work, let us briefly review the results of possibly related previous work on ionospheric irregularities in the auroral zone. We shall not consider D region irregularities to any appreciable extent, for two reasons. First, far less is known about the auroral D region than about the higher layers of the auroral ionosphere. Second, due to the inherently lower ion density of the D region, its contribution to the scattering of extraterrestrial VHF waves is not likely to be competitive with the contribution of higher regions.

The relatively high collision frequency of the D region suggests the use of cosmic-noise absorption as an indicator of (at least large-scale) irregularities there, which may be related to higher irregularities detectable by radio-star techniques. Ansari (1964) for instance, has reported patches of absorption, using a narrow-beam riometer. In addition, information on smaller-scale irregularities in the auroral D region may be forthcoming in the next few years from application of VHF ionospheric scatter propagation to geophysical problems.

Results at middle latitudes lead first to consideration of the F layer as the likely location for scintillation-producing irregularities. In probably the earliest investigation of radio-star scintillation as a recognized ionospheric phenomenon, Ryle and Hewish (1950) found a definite association with spread F. In another early scintillation study, Little and Maxwell (1951) found a continuous but marked increase in the fluctuation of radio-star amplitude when the line of sight from the source to their mid-latitude observatory traversed the auroral zone. Later, Briggs (1958) performed an extensive study of spread F and scintillation at mid latitude and found his results consistent with the view that they are caused by the same irregularities.

An informative presentation of the characteristics of high-latitude spread F has been given by Penndorf (1962, 1964). For an auroral-zone station such as College, Alaska, which lies in the path of Penndorf's travelling spread-F maximum, the phenomenon shows clear-cut seasonal and diurnal variations. For the solar-maximum period which Penndorf analyzed, spread F at College exhibited a consistent nighttime maximum. In summer, the nightly maximum was simply the culmination of a gradual and deep 24-

hour variation. In winter, it spread for many hours before and after midnight, encroaching on the middle hours, which then showed a short sharp dip in spread-F occurrence. The equinoctial periods appeared as clear-cut transitions between summer and winter conditions. Shimazaki (1962) and other workers have found a decrease in high-latitude spread-F near solar minimum as compared with solar maximum, in contrast to lower-latitude spread F.

Herman (1964) has interpreted Penndorf's results in terms of a competition between photoionization and charged particle ionization in the F layer. In Herman's model, field-aligned irregularities are produced by proton flux. Homogeneous ionization by solar illumination would tend strongly to obscure the irregularities during the day, leaving a travelling maximum of irregular structure on the auroral-zone midnight meridian. Herman's model requires an unspecified magnetospheric mechanism to produce spatially irregular proton streams. This problem has existed for years, however, in regard to auroral electrons.

As pointed out by Shimazaki (1962), spread F as observed on vertical-incidence ground-based ionograms probably does not reveal the whole spread-F picture at high latitudes. For instance, Bates (1959, 1960a, 1960b) has observed oblique HF backscatter from F-layer irregularities. Due to aspect sensitivity, he concluded that the irregularities are field-aligned and he calculated their heights (1960a) to be in the region of 250 to 400 km. The irregularity backscatter was found to be primarily a nighttime phenomenon, with occasional patchy returns observed during the day (1960b). Transitional effects near sunrise and sunset were noted which prompted Bates to write, "It appears that during magnetically quiet periods solar radiation eliminates the random irregularities - - -." This comment is of interest in regard to Herman's proposed model of spread F.

Another oblique-sounder HF return reported by Bates (1960b) and interpreted by him as being due to a (possibly irregular) field-aligned sheet of ionization is of interest in the light of certain topside sounder results. Muldrew (1963) has reported evidence of propagation via field-aligned sheets of ionization as the topside sounding satellite, Alouette, passed over the equatorial F region. Alouette has revealed a variety of apparently field-alinged spread F configurations (Calvert and Schmid, 1964). At least one type displays a strong latitudinal occurrence maximum at high latitudes along with a secondary but definite equatorial peak. Calvert, Knecht and VanZandt (1964) have interpreted certain auroral-zone returns from the fixed-frequency topside sounder Ionsopheric Explorer I (S-48) in terms of sheets of field-aligned irregularities extending from normal F-layer heights at least to the height of the satellite (about 950 km). The interpretation seems quite clear from their published ionograms.

A large collection of evidence for field-aligned irregularities in the auroral E layer has been secured over the years. The phenomenon known as radio aurora and detected by radar returns from aurorally associated ionization has provided a large portion of the evidence. A survey of experimental and theoretical work on the subject was given by Owren (1960), who concluded that auroral radar returns were produced by aspect-sensitive scattering from field-aligned irregularities. Both horizontal and vertical (upward as well as downward) motions of ionization in radio auroral forms have been reported by many workers on the basis of range changes and doppler shift and spread observations (Bowles, 1954; Nichols, 1957).

The obvious question of how closely visual and radio auroras are related (in particular, in regard to their positions) has been argued almost since the discovery of the latter. Early results often were ambiguous or else implied only a loose spatial relationship between radio and visual forms (Harang and Landmark, 1954). Recent work by Kelly (1965) using a narrow-beam (2.2 degrees) antenna in conjunction with a photometer of matching field of view directed along the radio line of sight shows close spatial and temporal relationships between specific radio and visual forms. Strength of radar returns also was found to be directly related to auroral brightness.

The radio aurora is believed to be very closely associated with sporadic E ionization in the auroral zone. In a survey of world-wide sporadic E characteristics, Thomas and Smith (1959) classified auroral-zone sporadic E into four types: so-called "flat, slant, retardation and auroral." They reported little seasonal variation for the overall phenomenon in the auroral zone but found that individual types show varying yearly patterns. All types show geomagnetic correlation. The authors could find no data available on correlation with radio-star scintillations. All types generally show spreading of the ionosonde returns, in common with equatorial sporadic E but in contrast with mid-latitude returns. Rapid vertical and horizontal motions and evidence for field-alingment were reported.

Work carried out both before and after the survey by Thomas and Smith suggested a close relationship between visual aurora and types of sporadic E other than the so-called "auroral" type. Knecht (1956) found "slant" type sporadic E to be related to remote auroras and concluded that there was a close spatial relationship between visual auroral forms and the ionization responsible for his radio observations. He also found the highest frequency present in aurorally associated sporadic E returns to be directly related to auroral brightness.

Bates (1961) later found slant sporadic E to occur only during magnetic disturbance and thought it to be simply the HF manifestation, on vertical-incidence ionosondes, of certain VHF auroral radar returns. Hunsucker and Owren (1962) also pointed out that so-called "auroral" type sporadic E was not the only type related to visual aurora. They found it to be the type usually present on vertical-incidence ionograms obtained in the presence of zenithal auroras but also found "flat" sporadic E returns under such conditions. They found a correlation coefficient of 0.5 between the upper cutoff frequency of sporadic E and a zenithal auroral index. Under conditions of pulsating aurora, they reported almost inevitable strong absorption or complete blackout on HF vertical soundings.

The one type of sporadic E in the auroral zone which does not appear to be related to visual auroral displays is the so-called "retardation" type. An unpublished study by Ansari revealed some similarities in the behavior of retardation sporadic E and spread F in the auroral zone near solar maximum. The similarities did not exist for other types of sporadic E. Ground-based ionosondes, however, cannot be considered reliable indicators of relationships between irregular structure at different ionospheric altitudes. All too often, ionization at low levels obscures the situation at higher levels at times of most interest.

The advent of topside sounders and their coordinated use with ground-based ionosondes is improving our knowledge of trans-ionospheric conditions, although limitations still exist under disturbed conditions. Evidence is starting to come in for the existence of field-aligned structure extending through great altitude ranges. For instance, du Castel and Vila (1964) have reported being able to trace such structure from sporadic E levels up to 1000 km. They raised the conjecture that "such fronts might turn out to be responsible for some oblique-reflection spread F echoes at the polar latitudes - - -." Bates' field-aligned sheets come to mind. It is of interest that Bates (1960b) differentiated between aspect-sensitive backscatter from field-aligned irregularities in the F layer and returns attributed to the field-aligned "sheet." The comment of du Castel and Vila, quoted above, continued "- - - we should not confuse their isolated discontinuous pattern of travelling disturbances with the regular steady periodic structure responsible for spread F." The latter authors, however, concluded that normal spread-F irregularities also are manifestations of very high-latitude structure, stating that "spread F is thus understood as the bottom extension of exospheric sheets of ionization."

According to du Castel and Vila, field-aligned structure in the upper F layer is very widespread. They postulate that special magnetic conditions may be necessary for it to become detectable by ionosondes exploring the bottom of the F layer. They describe the structure as being "ripply" in character, with spatial wavelengths of the order of a kilometer. In one particular case, such ripples with spatial wavelengths estimated at 5 km were observed over England. This is in remarkable agreement with the scales obtained for mid-latitude F-layer structure from early radio-star scintillation measurements (Ryle and Hewish, 1950; Hewish, 1951, 1952; Little and Maxwell, 1951).

In addition to the kilometer-scale structure, du Castel and Villa find a larger-scale "envelope." The envelope delineates the upper boundary of the smaller-scale structure, forming wedge-shaped groups. The sharp apices of the wedges point upward along the magnetic field and are separated by distances of the order of a few hundred kilometers. Lawrence, Jespersen and Lamb (1960) attributed slow angular variations of radio stars to "lens-like ionospheric irregularities having dimensions as large as 200 kilometers."

It would appear that many of the interpretations given to mid-latitude observations of radio-star scintillations are on the verge of corroboration by the powerful techniques of topside sounding. The topside sounders also may be able to fill in descriptive detail of structures which radio-star observations were able to suggest only in idealized fashion. A major point which seems to be emerging is that the F-layer irregularities responsible for mid-latitude radio-star scintillations are intimately related to heretofore little known structure in exospheric ionization. Exospheric irregularities were detected with satellite scintillation techniques in the northern auroral zone as early as 1959 (Basler and DeWitt, 1962), but they seemed to be sporadic. The relative consistency of the topside-sounder observations makes the manner in which such exospheric structure relates to the ionospheric-magnetospheric interface a question of particular interest, especially in the auroral zones. For instance, the theory offered by Axford and Hines (1961) to explain a large number of high-latitude geophysical phenomena relies heavily on the latitudinal dependence of diurnal maximization in high-latitude ionospheric irregularity occurrence.

To some extent, the statistical data used by Axford and Hines could be misleading as regards the morphology of specific events. Akasofu (1964), for instance, has shown that the instantaneous "auroral belt" shows significant departures from the statistical auroral zone. The same situation may be expected for radio auroras. Some radio methods may produce statistical results which in fact are biased against major magnetospheric agitations because of the effects of absorption. Sporadic-E and bottomside spread-F studies are among these.

Observations of radio-star and satellite scintillations at frequencies above about 50 MHz offer unbiased data. The possibility that these techniques could contribute to understanding of high-latitude geophysical phenomena through application at particularly disturbed times has been relatively unexplored. Surveys of world-wide satellite scintillation (e.g. Yeh and Swenson, 1964; Aarons, 1964) imply that such studies might be profitable. The satellite surveys have shown the existence of an often sharply defined zone of enhanced scintillation at auroral and (statistically) subauroral latitudes. This zone seems to coincide closely with that found for certain topside spread-F structure (Calvert and Schmid, 1964). In the case of scintillations, for which there are more observations, the equatorial boundary of the irregular zone is found to be related to geomagnetic K index. These and other observational facts have prompted Aarons (1964) to state, "It is inescapable that the moment to moment variations of irregularity regions are under as strong a control by the state of solar disturbance as are the visible aurora but the physical linking mechanism is nearly completely unknown."

Observations of radio-star scintillation at Saskatoon, Saskatchewan, over a four-year period (1955-58) have been interpreted by Forsyth and Paulson (1961) in terms of a high-latitude region of enhanced irregularities whose southern boundary migrated southward with increasing solar activity. Fremouw (1963), working with periods of particular disturbance (visibility fades observed at College during the IGY) found a dependence on local magnetic K index which is consistent with a southward shift also of the northern boundary of the scattering zone during periods of high magnetic activity. Thus, he found the "importance" (a measure of both duration and severity) of severe scatter events to be directly related to local K index up to an index of 3 but to be inversely related for greater K indices. This is in contrast to lower-latitude observations of similar events (but at different phases of the solar activity cycle), which have shown more direct relationships (Nichols, 1960; Moorcroft and Forsyth, 1963).

In view of the observations mentioned above relating irregularities to high-latitude disturbance phenomena, the presence of high-latitude irregularities also as a normal ionospheric feature is a befuddling - but persistent - fact. Yeh and Swenson (1964) have stated, "scintillation occurs much more frequently than any of these other phenomena, and may provide a more sensitive indication of the (particle) 'dumping' process." This conjecture has not been demonstrated experimentally, although Hook and Owren (1962) have reported one observation of E-layer irregularities beneath a satellite which simultaneously detected an influx of electrons.

Aarons has pointed out that in addition to simultaneous observations of particle flux, topside spread F, and scintillations, "We need more measurements on the strength of the irregularities and their height and size variation with latitude and geophysical disturbance and the lifetime of individual irregularities as well as cloud regions." The primary goal of the present work is to provide some of these measurements at times of enhanced irregular structure in the auroral zone near solar minimum.

Within the present context, "enhanced irregular structure" refers to the relatively strong, small-scale structure responsible for radio-star visibility fades. Visibility fades occur when the flux from a radio star becomes sufficiently scattered to reduce the correlation between signals received at two nearby antennas. They also have been called "long duration fades" (Little et al, 1962) and "radio-star fadeouts" (Flood, 1963; Moorcroft, 1963).

Besides representing an observed phenomenon asking for explanation in its own right, the visibility fade offers itself as a recognizable discrete event against the background of essentially omnipresent auroral-zone scintillations. In addition, it provides an opportunity for quantitative measurement of certain parameters of small-scale ionospheric structure with radio-interferometric techniques, as we shall see in later chapters. The strong direct dependence of auroral-zone ionospheric scattering on the solar activity cycle (Owren, Fremouw and Hunsucker, 1964) suggests solar minimum as the opportune time for an attempt at sorting out "disturbances" from "normal" irregular structure in the auroral zone.

# CHAPTER II

# INTERFEROMETRIC RECEPTION OF RANDOMLY SCATTERED WAVES

## IIA INTRODUCTORY CONCEPTS

### A1 Descriptions of Scattering

When a plane wave passes through a region of irregular refractive index, it emerges as a distorted version of itself - no longer plane but rather containing spatial variations in phase. Suppose, for instance, that a plane wave of light from a distant point source encounters a sheet of glass having irregular thickness. On the source side of the glass, the surfaces of constant phase are parallel planes, and the direction of travel of the wave is easily identified with the perpendicular to these planes. On the opposite side of the glass, the surfaces of constant phase contain ripples as a result of the differential phase retardation caused by the irregular glass.

The manner in which we commonly describe the effect of the irregular glass on the penetrating wave depends on how large a portion of the wavefront we are interested in at a given time. If all of our information comes from a small portion of a single phase ripple or irregularity, the wavefront still appears as approximately a plane. The perpendicular to this quasi-plane is still identified with the direction of travel of light rays, and to the extent that this direction is different from the original propagation direction, we say the wave has been refracted by the glass. Lacking any information from elsewhere in the wavefront, we are apt to conclude that the glass has the not very irregular shape of a simple wedge.

If now we explore a plane near the glass (on the "output" side), successively examining the phase of the wave on small adjacent segments of the same irregularity, we begin to sense that the surfaces of constant phase are not planes. Still identifying the "light rays" with the direction of propagation, we measure a change in direction as we move. After we have explored one phase irregularity in this fashion, the glass seems to us to have the shape of a lens. If we continue to explore the irregular wavefront in this fashion, taking note of the differential refraction imposed by the glass, we think of the glass as a collection of positive and negative lenses.

So long as, at any one time, we receive information from a portion of the wavefront which is small compared with one irregularity, we are content to think of the effect of the glass as refraction. We have no trouble assigning an instantaneous direction of propagation to that portion even if we note that the direction changes with time. If, however, we consider the nature of the wavefront over many irregularities, we no longer can find a unique direction defined by the direction of "rays" which lie perpendicular to the surfaces of constant phase. Indeed, the rays so defined are travelling in a multiplicity of directions, and we refer to the light as having been scattered by the glass rather than refracted.

## A2 The Angular Spectrum

We should like to find a means of describing scattering which will be valid over our whole wavefront region of interest. If our instantaneous interest is confined to a small portion of one irregularity, the concept of rays will suffice, and we can explore a larger portion of the wavefront over a period of time. If, however, our instantaneous interest is over many irregularities, another more useful concept has been made available to us by Booker and Clemmow (1950), the concept of an angular spectrum of plane waves. We shall not review in detail the fundamentals of angular spectra but rather move directly to apply them to our observational problem. The fundamentals have been discussed on several levels.

An elementary discussion has been given by Sokolnikoff and Redheffer (1958), in which it is shown how Fourier transformation of the field existing at the aperture of an antenna gives a plane-wave expansion of the field. The usefulness of this expansion, which constitutes the angular spectrum, arises from the fact that the plane-wave components can be taken as propagating independently away from the antenna - or from a piece of rough glass, or from an irregular ionosphere - and then synthesized by an inverse Fourier transformation to produce the field at any other plane. This was shown rigorously by Booker and Clemmow as well as by Booker, Ratcliffe and Shinn (1950). The correspondence of the field so produced to the Fresnel diffraction pattern also was shown by Booker and Clemmow, who related the concept to diffraction and antenna problems. All these considerations were reviewed with relative brevity and clarity by Ratcliffe (1956).

Booker and Clemmow showed that the angular spectrum is identical to the polar diagram of a spatially limited antenna at sufficiently great distance. It is in this case that the concept of rays becomes useful. The concept of an angular spectrum is valid at any distance from an antenna of any size. It is to be noted that while sequential observation of small portions of a wavefront will allow us eventually to determine the angular spectrum, the sum of rays from all portions is not identical to the angular spectrum. Rays are defined only over very restricted regions of space while components of the angular spectrum are infinite in spatial extent, just as components of a frequency spectrum are infinite in temporal extent. For a number of important special cases, however, the angular spectrum may be thought of as a collection of rays. In particular, at sufficient distance from a random scattering screen, the width of the angular spectrum can be closely approximated by the angular extent of rays emanating from a typical irregularity.

### A3 Amplitude and Phase of Random Signals

For application of the angular-spectrum formalism to a particular problem, such as ionospheric scattering, it is convenient to look for valid simplifications of the general approach. Before we even mentioned the term angular spectrum, in describing the scattering of light by an inhomogeneous glass, we utilized some simplifications of reality. For instance, we talked about a plane wave. A plane wave has a unique direction of propagation and, rather obviously, it has therefore an angular spectrum made up of a single unique component. Beyond this, however, the discussion of wave propagation involves the concept of phase, and we must ask what we really mean by phase and under what conditions it is a meaningful concept.

Our most common encounters with the concept of phase are with time-varying quantities. For instance, suppose we observe the temporal variation of voltage at the terminals of an antenna. If the voltage is varying sinusoidally with time about a mean of zero, we can completely specify the voltage by noting its frequency, its amplitude, and its phase at some reference time. Further, by noting the reference-time phase of the voltage produced by antennas at other positions, we can determine phase as a function of position and ascertain the propagation of an electromagnetic wave with considerable conceptual ease.

If, in the present work, we were dealing with signals from man-made earth satellites, the monochromatic concept of phase alluded to above would be quite adequate. We shall be dealing, however, with signals from natural radiators - radio stars. In this case, the voltage which we observe at the terminals of an antenna has the character of noise. That is, the voltage is a random function of time, where by "random" we simply mean unpredictable. In the monochromatic case, the time-varying real voltage can be specified for all time by a single complex number giving its (real) amplitude and phase. (The complex number may be referred to as the complex amplitude.) In the noise case, no single number can provide us with such a complete description of the voltage. What then can we mean if we talk of "amplitude" and "phase" and how can we describe the propagation of a "wave"?

A straight-forward development of "a complex representation of real polychromatic fields" is given by Born and Wolf (1959, section 10.2). We shall concern ourselves here with one of the resulting concepts of that development. The concept is that of the envelope of a random signal, and we shall demonstrate it empirically rather than analytically. Consider the oscillograms shown in figure 1. At the left is displayed the (amplified) output voltage from a random noise generator. We note the unpredictable nature of the voltage and, referring to its Fourier transform, we are apt to call it "white noise." By this we mean that the

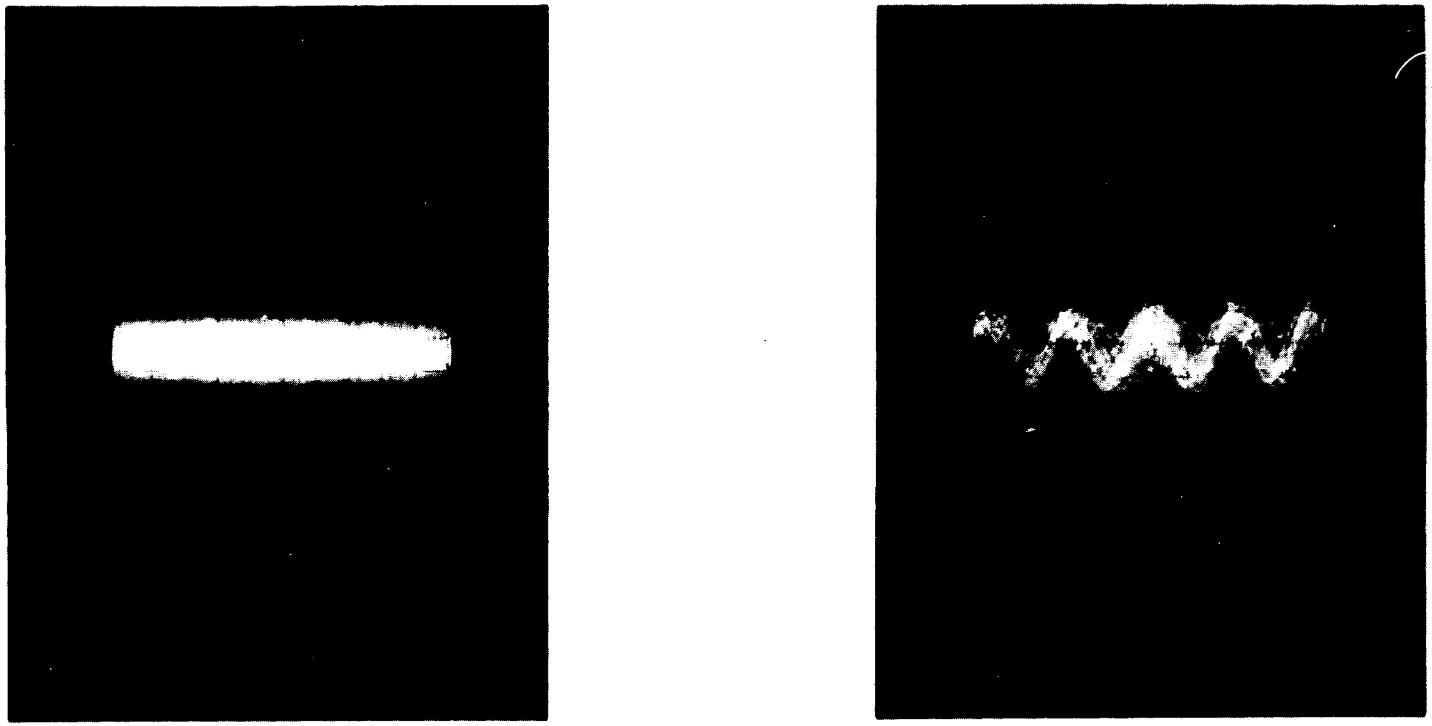


Fig. 1. Left, white noise; right, band-limited noise. Illustrating the meaning of the envelope amplitude and phase of a noise signal.

power spectrum of the signal contains equal contribution from all frequencies. We seldom venture beyond the concept of a power spectrum to consider what might be the relationship between the phases of the Fourier components. This we shall do later, in a different context, but for now we are concerned with the time domain. In looking at the left oscillogram, we can see no quantities which we might term the amplitude or phase of the time-varying signal. The best we can do is to note the fluctuation of instantaneous voltage as a function of time.

In the right-hand oscillogram is shown the output of the same noise generator, after the signal has been passed through a band-limiting filter. In this case, the "filter" was in fact an amplifier of the type employed at intermediate frequency in our radio-star observations. The amplifier passband was centered at 30 MHz and had a half-power bandwidth of about 250 kHz. The oscilloscope sweep rate and the photographic exposure time were equal to those in the left oscillogram. The observed voltage, while still having the general character of noise, is now seen to have some element of predictability, characterized by what we might call its "envelope." The envelope appears as a sort of fuzzy sine wave with a frequency of 30 MHz. To the envelope, we can assign a nebulous amplitude and phase.

If we stop to consider the manner in which the fuzzy sine wave is produced by the oscilloscope, we can give analytic definition to the envelope. The oscillogram is produced by repetitive sweeps of the scope's electron beam, each sweep lasting a few cycles of a 30 MHz oscillation. If the amplitude and phase of the 30 MHz signal vary with time, randomly but slowly compared with the basic 30 MHz oscillation, then successive sweeps of the elctron beam produce the fuzzy sine wave observed.

Thus, the observed signal could be described analytically as a sine wave at the center frequency of our band-limiting filter, with the amplitude and phase themselves being functions of time. Born and Wolf arrived at this result rigorously. Our intent here is to establish some intuitive feeling for what is meant by the amplitude and phase of a random noise signal. Note that these terms and their unifying concept of an "envelope" have meaning only when the bandwidth of the noise signal is limited. We shall make use of this idea of "band-limited white noise" later.

It is important to note that in most observations of natural noise radiators such as radio stars - including the present observations - the signal of the type shown in the oscillogram is detected and averaged. The averaging is over times very long compared with the characteristic times of the signal shown. These characteristic times are the period of the fundamental oscillation (0.033 microseconds in this case) and the longer time of variations in amplitude and phase. This latter characteristic time is determined by the filter bandwidth, being comparable with its reciprocal. It is called the coherence time of the signal and, in the present example, is about four microseconds.

For averaging times very long compared with the coherence time, what we finally deal with are the average amplitude and phase of the signal envelope. These are very well defined quantities, as can be appreciated by considering a smoothed version of the trace shown in the right-hand oscillogram. The ionospheric effects with which we shall be concerned produce deviations from these averages, which are slow compared with all the times we have thus far considered - on the order of many hundreds of milliseconds or a few seconds, in fact. Thus, although we shall be dealing with random noise signals, the concepts of amplitude and phase will continue to be useful and well defined.

From the above rather intuitive introduction to the concepts of angular spectrum and random-noise envelope and the references to rigorous treatments of these subjects, let us proceed to the problem of ionospheric scattering. Along our route, we shall introduce another important concept, that of random phasing in the angular spectrum. This idea will be fundamental to our theoretical considerations and is not in such widespread use as the general concept of an angular spectrum or the envelope parameters of a noise signal. Accordingly, we shall have to explore the meaning and consequences of random phasing rather completely - although only within the context of our particular problem.

## IIB RELATION BETWEEN ANGULAR SPECTRUM AND FREQUENCY SPECTRUM

Let us begin from an observational point of view, considering the reception of an ionospherically scattered signal by a radio interferometer. We are not concerned whether the original signal is deterministic, as from a satellite, or noiselike, as from a radio star, so long as the latter is bandwidth limited in the sense that the signal envelope is well defined. In either case, we shall be concerned with amplitude and phase modulations of the signal imposed by the ionosphere. In accord with what actually is observed for the most part, we shall take the modulation itself to be random. Departures from this condition, which in fact also are observed on occasion, represent interesting exceptions and will be discussed in Chapter V. In much of the present chapter, we shall follow closely the work of Bramley (1951), generalizing on and adding to his development where appropriate to our observational situation.

Let the observing frequency be $\omega/2\pi$ and suppose that the randomly modulated signal is Fourier analyzed. Thus, let the voltage at the terminals of one antenna be described as

$$v_1 = \sum_{n=1}^{N} c_n \cos \; [\omega t + (\omega_n - \omega)t + \phi_n] \qquad 2\text{-}1$$

where N is some large integer. Each Fourier component in the above expression has the nature of a sinusoid at the observing frequency, whose phase is changing linearly with time at the rate $(\omega_n - \omega)$. Let us denote the time-varying phase by $\psi_n(t)$ and investigate how it might be produced in the wavefield impinging on the antenna.

Consider a plane wave at the observing frequency approaching the antenna from a direction making an angle $\alpha_n$ to the vertical. In a coordinate system with origin at the antenna where the z-axis is vertical and the wave's propagation direction is in the xz plane, such a wave could be expressed as

$$e_n = d_n \cos \; [\omega t + \phi_n + (\omega/c)(x \sin \alpha_n + z \cos \alpha_n)] \qquad 2\text{-}2$$

where c is the velocity of light. For any $\alpha_n$, such a wave would produce at the antenna (x = z = 0) a voltage of the form

$$v_{1n} = c_n \cos \; (\omega t + \phi_n) \qquad 2\text{-}3$$

This corresponds to the $n^{\underline{th}}$ Fourier component of equation 2-1, with the time-varying phase $\psi_n(t)$ equal to zero.

If, now in addition to propagation along the wave normal, the wavefront is moving in the x direction with an additional (and relatively very small) velocity u, then $\psi_n(t)$ becomes a non-zero function of $\alpha_n$. This is seen readily by considering that in this case the third term in the argument of equation 2-2 would be

$$(\omega/c)\ [(x - ut) \sin \alpha_n + z \cos \alpha_n] \qquad 2\text{-}4$$

so that at the antenna terminals, we get

$$v_{1n} = c_n \cos [\omega t + \phi_n - (\omega/c)(ut \sin \alpha_n)] \qquad 2\text{-}5$$

Comparison with equation 2-1 reveals at once that

$$\psi_n(t) = -\omega t\ (u/c) \sin \alpha_n = -\ (2\pi/\lambda)\ ut \sin \alpha_n \qquad 2\text{-}6$$

where $\lambda$ = the observing wavelength, and that

$$\omega_n = \omega[1 - (u/c) \sin \alpha_n\ ] = (2\pi/\lambda)(c - u \sin \alpha_n) \qquad 2\text{-}7$$

Thus, the total output $v_1$ of equation 2-1 could be produced by an angular spectrum of plane waves propagating toward the antenna and simultaneously drifing horizontally. Each Fourier frequency component of $v_1$ would arise from a single component of the angular spectrum. This view corresponds to drift past the antenna of an irregular wavefront produced by a drifting but otherwise unchanging irregular ionosphere. It is, of course, not the only manner in which the voltage $v_1$ could be produced.

The frequency components $\omega_n$ could be produced by a purely temporal modulation by a smooth ionosphere. In this case, a neighboring antenna always would produce an output identical to $v_1$, which is inconsistent with a vast collection of observational material from the past, as reviewed in Chapter I, and from the present work. Such a situation would not correspond to ionospheric scattering at all. Combinations of drift and temporal modulation also could occur. In this case each component of the angular spectrum would have a frequency spectrum associated with it. At the observing frequencies used in this work and with the techniques employed, we are not concerned with components which might be introduced by temporal modulation.

Aside from pure drift at a single velocity, there is another kind of wavefront motion to be considered. This is a wavefront which changes its spatial structure as it drifts, which would arise if the various angular components drifted at different velocities. With interferometer techniques alone it is impossible to differentiate between these two kinds of motion, and we shall not concern ourselves with the relationship between angle and velocity. Our experimental results will be unaffected so long as we suitably restrict our conclusions. Attempts to describe wavefront motion in detail would be futile.

## IIC OUTPUT VOLTAGES OF NEIGHBORING ANTENNAS

Equation 2-1 and the ensuing discussion relate to the output voltage of a single antenna, one of a pair of antennas used in a radio interferometer. Now suppose the second antenna is located on the x axis at a distance d from the first. Due to the separation between antennas, a given component of the angular spectrum will arrive with different phases at the two antennas. To account for this effect, we can write the output voltage of the second antenna as in equation 2-1, adding a fourth term in the argument. In order to preserve symmetry, let us also shift the origin of our coordinate system midway between the antennas. Then the two output voltages are

$$v_1 = \sum_{n=1}^{N} c_n \cos(\omega t + \psi_n + \phi_n + \chi_n)$$

2-8

$$v_2 = \sum_{n=1}^{N} c_n \cos(\omega t + \psi_n + \phi_n - \chi_n)$$

where $\psi_n$ is the time varying phase given by equation 2-6 and

$$\chi_n = (\pi d/\lambda) \sin \alpha_n \qquad \text{2-9}$$

The antenna voltages also can be written as

$$v_1 = A_1 \cos(\omega t + \theta_1)$$

2-10

$$v_2 = A_2 \cos(\omega t + \theta_2)$$

The amplitudes A and phases θ will in general be different for the two antennas owing to the fourth terms in equation 2-8. They will be functions of time owing to the time variations of the $\psi_n$. Bramley (1951) assumed the $\psi_n$ to be independent random functions of time and did not consider the explicit relationship between them and the angular spectrum.[1] As discussed above, an irregular wavefront drifting at a constant velocity will produce $\psi_n$'s which are linear functions of time. Bramley's assumption corresponds to time varying drift velocities and was made in order to account for the statistical distributions thought to describe ionospheri-

---

[1]To avoid confusion, it is prudent to point out that Bramley's $\psi_n$ are equivalent to the sum of $\psi_n$ and $\phi_n$ in the present notation, which was chosen to allow explicit separation of time varying and static quantities.

cally scattered radio signals, namely those associated with band-limited white noise. We shall see shortly that the assumption of independence is sufficient to ensure the same distributions.

## IID SIGNAL STATISTICS FOR A RANDOMLY PHASED ANGULAR SPECTRUM

### D1 The Special Case of Completely Scattered Waves

1a Single Antenna: Let us denote the complex amplitudes of the antenna voltages by $V_1$ and $V_2$ and those of their Fourier components by $C_{1n}$ and $C_{2n}$. Then we have, from equations 2-8 and 2-10

$$V_1 = A_1 \exp (i\theta_1) = \sum_{n=1}^{N} C_{1n} = \sum_{n=1}^{N} c_n \exp (i\gamma_{1n})$$

2-11

$$V_2 = A_2 \exp (i\theta_2) = \sum_{n=1}^{N} C_{2n} = \sum_{n=1}^{N} c_n \exp (i\gamma_{2n})$$

where $\gamma_{1n} = (\psi_n + \phi_n + \chi_n)$ and $\gamma_{2n} = (\psi_n + \phi_n - \chi_n)$

Equations 2-11 show that the complex amplitudes $V_1$ and $V_2$ are the phasor sums of the component complex amplitudes $C_{1n}$ and $C_{2n}$, respectively. The summations represent random walk processes if the component phases are "random."

Now the $\psi_n$ and $\chi_n$ are defined by equations 2-6 and 2-9, respectively. We have said nothing about the $\phi_n$ thus far, however. We now assume them to be random variables[1] which are uniformly distributed between zero and $2\pi$ and independent. This is what is meant by a "randomly phased angular spectrum," and for sufficiently large N it is the sole assumption necessary to produce A's and θ's distributed as the envelope amplitudes and

---

[1]Random functions of angle, α, but temporal constants.

phases of band-limited white noise, as we shall now see. The ionospheric requirements for random phasing will be explored in Chapter III.

Note that if the $\phi_n$ are uniformly distributed and independent, the $\gamma_{1n}$ and $\gamma_{2n}$ are also. The summations in 2-11 then are random walks, and the central limit theorem can be evoked to show that the real and imaginary components of $V_1$ and $V_2$ approach normal distributions as N increases without limit. For our purposes, the descriptive treatment of the central limit theorem given by Munroe (1951) and the somewhat more rigorous one by Middleton (1960) are enlightening.

It is pertinent to point out that while independence of the component phasors in 2-11 is neither a necessary nor a sufficient condition for the central limit law to hold, it does allow statement of the central limit theorem in the relatively simple form attributed to Liapounoff and in the less restricted Lindeberg formulation. In this formulation, the most important restriction on the component phasors, from a physical point of view, is that no one of them dominates the aggregate. The detailed density distribution of the components is not specified by the theorem and may take a wide variety of forms. Note that in the present situation, we have not assumed any particular distribution for the $c_n$.

Let us define the real and imaginary components of $V_1$ respectively as

$$A_{1c} = A_1 \cos \theta_1$$

and (2-12)

$$A_{1s} = A_1 \sin \theta_1$$

and similarly for the real and imaginary components of $V_2$. Since $A_{1c}$ and

$A_{1s}$ are normally distributed for a randomly phased angular spectrum, as discussed above, then under this sole assumption we can follow Rice (1945) in his treatment of the envelope of band-limited white noise and establish the statistical characteristics of the envelope amplitude $A_1$ and $\theta_1$. We begin with the joint density distribution of $A_{1c}$ and $A_{1s}$. Since each of these variables is normally distributed, their joint distribution will be the product of two gaussians if they are statistically independent. That they are independent is shown in Appendix 1a.

In view of the result of Appendix 1a, the joint distribution for $A_{1c}$ and $A_{1s}$ can be written as the product of two one-demensional normal distributions. Taking into account the zero means of the A's, we have for the joint probability density distribution function[1]

$$f(A_{1c},A_{1s}) = \frac{p(A_{1c},A_{1s})}{dA_{1c}\, dA_{1s}} = \frac{1}{2\pi\, \sigma_c\sigma_s} \exp\left[-\left(\frac{A_{1c}^2}{2\sigma_c^2} + \frac{A_{1s}^2}{2\sigma_s^2}\right)\right] \qquad 2\text{-}13$$

where $p(A_{1c},A_{1s})$ denotes the probability that $A_{1c}$ and $A_{1s}$ lie in the elementary rectangle $(dA_{1c}\, dA_{1s})$, and the $\sigma$'s are the standard deviations of the A's.

Appendix 1a deals with the cross-products of the real and imaginary components of $A_1$. Straight-forward application of the same development to the squares of the real and imaginary components yields the variances of the components. Thus, it is easily shown that

$$\sigma_c^2 = \sigma_s^2 = \tfrac{1}{2} \sum_{n=1}^{N} c_n^2 \qquad 2\text{-}14$$

[1]We shall work throughout with (differential) probability density distribution functions, denoted by f. For brevity, we shall often use the term "distribution" to mean functions of this sort. In no case will we be referring to the corresponding (integral) probability distribution function.

In view of equation 2-14, let us denote both the standard deviations by $\sigma$. Then the probability in equation 2-13 becomes

$$p(A_{1c},A_{1s}) = \frac{dA_{1c}\,dA_{1s}}{2\pi\sigma^2} \exp\left(- \frac{A_{1c}^2 + A_{1s}^2}{2\sigma^2}\right) \qquad 2\text{-}15$$

Equations 2-12 show that $(A_{1c}^2 + A_{1s}^2) = A_1^2$ and that $(dA_{1c}dA_{1s}) = (A_1 d\theta_1 dA_1)$. Making these changes of variables in equation 2-15 we obtain

$$p(A_1,\theta_1) = \frac{A_1 d\theta_1 dA_1}{2\pi\sigma^2} \exp\ (- A_1^2/2\sigma^2) \qquad 2\text{-}16$$

Equation 2-16 shows that $A_1$ and $\theta_1$ are independent random variables since $p(A_1,\theta_1)$ can be expressed as the product of $p(A_1)$ and $p(\theta_1)$. Integration of $p(A_1,\theta_1)$ with respect to $\theta_1$ over the range zero to $2\pi$ yields

$$p(A_1) = \frac{A_1 dA_1}{\sigma^2} \exp\ (-A_1^2/2\sigma^2) \qquad 2\text{-}17$$

Division of 2-17 into 2-16 then shows that

$$p(\theta_1) = d\theta_1/2\pi \qquad 2\text{-}18$$

Thus the joint distribution function $f(A_1,\theta_1)$ is given by the product of $f(A_1)$ and $f(\theta_1)$, where $f(A_1)$ is the Rayleigh distribution given by

$$f(A_1) = (A_1/\sigma^2) \exp\ (-A_1^2/2\sigma^2) \qquad 2\text{-}19$$

and $f(\theta_1)$ is the uniform distribution between zero and $2\pi$, given by

$$f(\theta_1) = 1/2\pi \qquad 2\text{-}20$$

Equations 2-19 and 2-20 show that the envelope amplitude and phase of the fluctuating voltage produced at the terminals of an antenna by the constant horizontal drift of a randomly phased angular spectrum obey the same statistical laws as do the envelope amplitude and phase of band-limited white noise. In arriving at this result we have relaxed the assumption made by Bramley regarding the angular spectrum. No assumption need be made about the drift velocity of the component waves in the angular spectrum, other than that it be nonzero. The assumption of random phasing is sufficient.

1b <u>Two antennas</u>: Obviously the developemnt of equations 2-19 and 2-20 holds for $A_2$ and $\theta_2$ as well as for $A_1$ and $\theta_1$. We now open the question of the relationship between the voltages at the two antennas. Important quantities are the product of voltages $v_1v_2$, the product of amplitudes $A_1A_2$, and the difference of phases $\theta_2 - \theta_1$. In particular, we shall be concerned with the means of these quantities and with the variance of the phase difference.

The mean $\overline{v_1v_2}$ of the product of voltages is the covariance of voltages and is important to observations with phase-switch interferometers, as we shall see. It is shown in Appendix 1b that this quantity is given by

$$\overline{v_1v_2} = \tfrac{1}{2}\sum_{n=1}^{N} c_n^2 \cos 2\chi_n \qquad \text{2-21}$$

under the assumption of random phasing, where $\chi_n$ is the phase angle defined in equation 2-9.

Since equation 2-21 gives the covariance of the voltages, it provides a measure of the correlation between them. The relations we seek between the A's and θ's also involve the correlation between the signals received at the two antennas. Since there are four variables, the situation may be expected to be more complicated, however. To find the desired relationships, we must investigate the joint probability density distribution of the variables. We begin by considering the joint distribution of the four gaussian random variables $A_{1c}$, $A_{1s}$, $A_{2c}$, and $A_{2s}$.

Since we are dealing with gaussian variables, the joint distribution will be a four-dimensional gaussian. It will not, in general, be simply the product of four one-dimensional gaussians since we have not established independence of all four A's. Fundamentally, the multi-variate normal (or gaussian) distribution if that function resulting from the multi-variate central limit law. It can be derived from consideration of the n-dimensional random walk, as has been done by Middleton (1960), for instance.

A convenient form of the n-dimensional normal distribution function for variables with zero means is given without rigorous development by Bendat (1958). In our present quadri-variate case, the n-dimensional form reduces to

$$f(A_{1c},A_{1s},A_{2c},A_{2s}) = \frac{|M|^{-\frac{1}{2}}}{(2\pi)^2} \exp\left( \frac{-1}{2|M|} \sum_{\substack{i=1\\j=1}}^{4} M_{ij}A_iA_j \right) \qquad 2\text{-}22$$

where the numerical suscripts refer successively to the four double subscripts, 1c, 1s, 2c, and 2s; and $M_{ij}$ denotes the cofactors of the determinate $|M|$, whose matrix is

$$M = \begin{pmatrix} \mu_{11} & \mu_{12} & \mu_{13} & \mu_{14} \\ \mu_{21} & \mu_{22} & \mu_{23} & \mu_{24} \\ \mu_{31} & \mu_{32} & \mu_{33} & \mu_{34} \\ \mu_{41} & \mu_{42} & \mu_{43} & \mu_{44} \end{pmatrix} \qquad 2\text{-}23$$

The elements of the matrix are the second moments of the random variables, and M is known as the moment matrix.

The elements of the moment matrix are discussed in Appendix 1c, and it is shown there that under the condition of random phasing M can be reduced to

$$M = \begin{pmatrix} \sigma^2 & 0 & \mu_c & -\mu_s \\ 0 & \sigma^2 & \mu_s & \mu_c \\ \mu_c & \mu_s & \sigma^2 & 0 \\ -\mu_s & \mu_c & 0 & \sigma^2 \end{pmatrix} \qquad 2\text{-}24$$

where

$$\sigma^2 = \tfrac{1}{2} \sum_{n=1}^{N} c_n^2 \qquad 2\text{-}25$$

$$\mu_c = \tfrac{1}{2} \sum_{n=1}^{N} c_n^2 \cos 2\chi_n \qquad 2\text{-}26$$

$$\mu_s = \tfrac{1}{2} \sum_{n=1}^{N} c_n^2 \sin 2\chi_n \qquad 2\text{-}27$$

All the information we require concerning the antenna voltages, $v_1$ and $v_2$, is contained in equation 2-21 and the elements of the moment matrix M. For a randomly phased angular spectrum, it is more appropriate to consider a spectral continuum than discrete spectral components. Thus, let us now replace equations 2-25 through 2-27 and equation 2-21 with their corresponding integral representations. Noting that $\tfrac{1}{2} c_n^2$ is the power which would be dissipated in a unit resistance by the $n^{th}$ component wave of the angular spectrum, we see that equation 2-25 gives P, the total power (per unit resistance) contained in the spectrum. Let us denote by $p(\alpha)$ the (unitary resistance) angular power density in the spectrum. Then, upon recalling the definition of $\chi_n$ from equation 2-9, we obtain the following:

$$\sigma^2 = \int_{-\pi}^{\pi} p(\alpha)\, d\alpha \qquad \text{2-28}$$

$$\mu_c = \int_{-\pi}^{\pi} p(\alpha) \cos\,[(2\pi d/\lambda) \sin \alpha]\, d\alpha \qquad \text{2-29}$$

$$\mu_s = \int_{-\pi}^{\pi} p(\alpha) \sin\,[(2\pi d/\lambda) \sin \alpha]\, d\alpha \qquad \text{2-30}$$

$$\overline{v_1 v_2} = \int_{-\pi}^{\pi} p(\alpha) \cos\,[(2\pi d/\lambda) \sin \alpha]\, d\alpha = \mu_c \qquad \text{2-31}$$

Now equations 2-22 and 2-24 show that the joint distribution of $A_{1c}, A_{1s}, A_{2c}$, and $A_{2s}$ is given by

$$\frac{1}{4\pi^2[\sigma^4-(\mu_c^2+\mu_s^2)]} \exp\left\{ \frac{-1}{\sigma^4-(\mu_c^2+\mu_s^2)} \left[\sigma^2(A_{1c}^2+A_{1s}^2+A_{2c}^2+A_{2s}^2) -2\mu_c(A_{1c}A_{2c}+A_{1s}A_{2s})-2\mu_s(A_{1s}A_{2s}-A_{1s}A_{2c})\right]\right\} \qquad \text{2-32}$$

To find the joint distribution of $A_1, A_2, \theta_1$, and $\theta_2$, we now transform to these coordinates from the four A's. The result, as is easily shown, is

$$\frac{A_1 A_2}{4\pi^2[\sigma^4 - (\mu_c^2 + \mu_s^2)]} \exp\left\{ \frac{-1}{2[\sigma^4 - (\mu_c^2 + \mu_s^2)]} \left[ \sigma^2 (A_1^2 + A_2^2) - 2\mu_c A_1 A_2 \cos(\theta_2 - \theta_1) - 2\mu_s A_1 A_2 \sin(\theta_2 - \theta_1) \right] \right\} \qquad \text{2-33}$$

It is to be noted that the distribution 2-33 cannot be represented as the product of an amplitude distribution and a phase-difference distribution. Thus, while the amplitude and phase at a single antenna are statistically independent, the amplitude at each of a pair of antennas is statistically dependent upon the amplitude at the other and upon the phase difference between the two antennas.

To find the joint distribution of the amplitudes, it is necessary to integrate over all possible phases. Similarly, to find the distribution of $\theta_2 - \theta_1$, integrations must be carried out over $A_1$ and $A_2$. Rice (1945) gives the joint distribution of $A_1$ and $A_2$ as

$$\frac{A_1 A_2}{\sigma^4 - (\mu_c^2 + \mu_s^2)} \, I_o \left[ \frac{A_1 A_2 (\mu_c^2 + \mu_s^2)^{\frac{1}{2}}}{\sigma^4 - (\mu_4^2 + \mu_s^2)} \right] \exp\left\{ - \frac{\sigma^2 (A_1^2 + A_2^2)}{2[\sigma^4 - (\mu_c^2 + \mu_s^2)]} \right\} \qquad \text{2-34}$$

where $I_o$ represents the zero-order modified Bessel function of the first kind for the argument in brackets. The integration over amplitudes has been carried out by MacDonald (1949), with the following result for the distribution of $\theta_2 - \theta_1$.

$$f(\theta_2-\theta_1) = \frac{\sigma^4-(\mu_c^2+\mu_s^2)}{2\pi\sigma^4} \; \frac{(1-\beta^2)^{\frac{1}{2}} + \beta(\pi - \cos^{-1}\beta)}{(1-\beta^2)^{3/2}} \qquad \text{2-35a}$$

where $\beta = [\mu_c \cos(\theta_2-\theta_1) - \mu_s \sin(\theta_2-\theta_1)]/\sigma^2$ 2-35b

Now, we are interested in certain moments of the distributions 2-34 and 2-35. In the case of amplitudes, we are particularly interested in the mean of their product $\overline{A_1A_2}$. For some applications, for instance a space-diversity receiving system employing square-law detectors and post-detection correlation techniques, the quantity $\overline{A_1^2A_2^2}$ is of interest also. Middleton (1960) outlines the development of these moments, with the following results:

$$\overline{A_1A_2} = \frac{\pi\,\sigma^2}{2} F(-\tfrac{1}{2},-\tfrac{1}{2},1,R^2) = \sigma^2[2E(R) - (1 - R^2)K(R)] \qquad \text{2-36}$$

$$\overline{A_1^2A_1^2} = 4\sigma^2(1 + R^2) \qquad \text{2-37}$$

where F is the gaussian hypergeometric function, E is the complete elliptic integral of the second kind, K is the complete elliptic integral of the first kind, and

$$R = \frac{(\mu_c^2 + \mu_s^2)^{\frac{1}{2}}}{\sigma^2} \qquad \text{2-38}$$

It is pointed out in Apendix 1c that the $\mu$'s are covariances between components of signal at the antennas, while $\sigma^2$ is the variance of the signal voltage at a single antenna. Thus R is a kind of correlation coefficient between the signals at the two antennas. In the special case where the spectrum is an even function (in our case, a symmetrical angular spectrum centered on $\alpha = 0$), $\mu_s$ vanishes (Rice, 1945, p. 78; Middleton, 1960, p. 352). In this case R is identical to the correlation between the voltages $v_1$ and $v_2$, as is evident from equation 2-31 and 2-38. In the more general case with which we must deal, the two are not identical but still are closely related. We shall find that R is a parameter of ionospheric interest which can be determined experimentally.

Now, in general, the correlation between two variables is defined as the ratio of their covariance to the geometric mean of their variances. If the variances are equal, then obviously the correlation is given by the ratio of the covariance to the variance of either. It is easy to show, then, that the correlation between the amplitudes at the two antennas is given by

$$\rho_A = \frac{\overline{A_1 A_2} - \bar{A}_1^2}{\overline{A_1^2} - \bar{A}_1^2} \qquad \text{2-39}$$

and that between the square of the amplitudes by

$$\rho_{A^2} = \frac{\overline{A_1^2 A_2^2} - \bar{A}_1^2}{\overline{A_1^4} - \overline{A_1^2}^2} \qquad \text{2-40}$$

The various terms appearing in 2-39 and 2-40 are given by equations 2-36 and 2-37 and the distributions of $A_1$ and $A_1^2$. The distribution of $A_1$ is the Rayleigh distribution, given by equation 2-19. A simple change in stochastic variable[1] gives the distribution of $A_1^2$ as follows

$$f(A_1^2) = \frac{f(A_1)}{2A_1} = (1/2\sigma^2) \exp(-A_1^2/2\sigma^2) \qquad 2\text{-}41$$

The distributions of $f(A_1)$ and $f(A_1^2)$ yield the following:

$$\overline{A_1}^2 = (\pi/2)\,\sigma^2 \qquad 2\text{-}42$$

$$\overline{A_1^2} = 2\,\sigma^2 \qquad 2\text{-}43$$

$$\overline{A_1^2}^2 = 4\,\sigma^4 \qquad 2\text{-}44$$

$$\overline{A_1^4} = 8\,\sigma^4 \qquad 2\text{-}45$$

Combination of the above with equations 2-36 and 2-37 yields, from equations 2-39 and 2-40, the following:

$$\rho_A = \frac{4[E(R)-\frac{1}{2}(1-R^2)K(R)] - \pi}{4 - \pi} \qquad 2\text{-}46$$

$$\rho_A{}^2 = R^2 \qquad 2\text{-}47$$

Series solutions for the complete elliptic integrals are given in standard tables of integrals (Hodgman, 1955, integrals no. 459 and 460).

---

[1]See Bendat (1958), section 3.4-3, for a discussion of change of variable in density distributions.

Using these, it is easily shown that equation 2-62 reduces to

$$\rho_A = 0.91 \left(R^2 + \frac{1}{16}R^4 + - - - - - - -\right) \qquad 2\text{-}48$$

This gives the approximate (and equation 2-46 the exact) relation between the wavefront correlation function R and the cross-correlation coefficient $\rho_A$ between the amplitudes at the two antennas, for the case of a randomly phased angular spectrum in which no component predominates.

The degree of relation between the amplitudes at the two antennas is given in a meaningful fashion by the correlation coefficients $\rho_A$ and $\rho_A 2$. These quantities are simply related to the wavefront correlation function R as described above. Now, the phase difference between the signals at the two antennas also is related to R. This fact is evident from the distribution function for $(\theta_2 - \theta_1)$ as derived by MacDonald (1949) and given in our equations 2-35 upon recalling the definition of R given in equation 2-38. It is difficult, however, to extract a physically meaningful relationship with the degree of generality retained to this point. Let us, therefore, simplify our considerations and return later to a discussion of the phase characteristics.

Recalling that all the physical information we require about the antenna voltages is contained in the elements of the moment Matrix M, let us return to the integrals which define the elements, equations 2-28 through 2-31. From physical considerations, the integration limits in these equations were set at $-\pi$ and $+\pi$. This allows signal reception from all directions, and we can as well set the limits at $-\infty$ and $+\infty$. The physical information concerning the directional distribution of received signal is contained in the angular power density function $p(\alpha)$, more commonly called the "angular power spectrum."

For an ionospherically scattered satellite or radio-star signal, we expect the flux to be received in a narrow cone centered on the direction of the source.[1] It is reasonable to assume also that the angular power spectrum is an even function about the center direction,[2] without specifying the shape of the spectrum in detail. Thus let the center direction be denoted by $\alpha_o$ and make the following change of variables:

$$\alpha = \alpha_o + \delta \qquad 2\text{-}49$$

$$p(\alpha) = F(\delta) = F(-\delta) \qquad 2\text{-}50$$

A physically reasonable assumption which is persistently verified by observation is that F is appreciable only for values of $\delta$ which are sufficiently small that $\cos\delta$ may be approximated by unity and $\sin\delta$ by $\delta$. The above set of assumptions means that we are assuming a narrow, symmetrical angular spectrum.

It is shown in Appendix 1d that a narrow, symmetrical angular spectrum leads to the following expressions for the signal covariances $\mu_c$ and $\mu_s$ and the wavefront correlation R:

---

[1] The "direction of the source" may be only an apparent direction due to refraction by ionospheric structure of a scale larger than that producing the scattering presently under consideration.

[2] It has been pointed out by K. W. Philip (private communication) that one might expect some skewing of the angular spectrum to result from the non-isotropic nature of a magneto-ionic medium. The possible effect of the geomagnetic field on the shape of the angular spectrum produced by ionospheric scattering has not been considered here, except insofar as it might influence the shape of ion-density irregularities.

$$\mu_c = \cos 2\chi_o \int_{-\infty}^{\infty} F(\delta) \cos \left(\frac{2\pi d}{\lambda} \delta \cos \alpha_o\right) d\delta \qquad 2\text{-}51$$

$$\mu_s = \sin 2\chi_o \int_{-\infty}^{\infty} F(\delta) \cos \left(\frac{2\pi d}{\lambda} \delta \cos \alpha_o\right) d\delta \qquad 2\text{-}52$$

$$R = \frac{\int_{-\infty}^{\infty} F(\delta) \cos [(2\pi d/\lambda) \delta \cos \alpha_o] d\delta}{\sigma^2}$$

$$= \frac{\int_{\infty}^{\infty} F(\delta) \cos [(2\pi d/\lambda) \delta \cos \alpha_o] d\delta}{\int_{-\infty}^{\infty} F(\delta) d\delta} \qquad 2\text{-}53$$

Now, equations 2-21 and 2-26 show that the covariance of the voltages at the two antennas is equal to $\mu_c$. Therefore, under the assumption of a narrow symmetrical angular spectrum, we have

$$\overline{v_1 v_2} = \cos 2\chi_o \int_{\infty}^{\infty} F(\delta) \cos [(2\pi d/\lambda) \delta \cos \alpha_o] d\delta \qquad 2\text{-}54$$

so

$$\overline{v_1 v_2} = \sigma^2 (R \cos 2\chi_o) \qquad 2\text{-}55$$

This relates the covariance of voltages to the wavefront correlation function R and shows that the correlation of voltages is simply $R \cos 2\chi_o$.

The correlation of amplitudes at two antennas receiving a narrow symmetrical angular spectrum is given by equation 2-48 with R evaluated according to equation 2-53. For the square of amplitudes, the correlation is given by $R^2$.

When the angular spectrum is narrow and symmetrical, it becomes possible to obtain a meaningful description of the phase difference between the two antennas. In this case, MacDonald's general expression for the distribution of $\theta_2 - \theta_1$, given by equations 2-35, can be reduced as follows. With $\mu_c$, $\mu_s$, and R given respectively by equations 2-51, 2-52 and 2-53, MacDonald's parameter $\beta$ from our 2-35b becomes

$$\beta = R\,[\cos 2\chi_o \cos(\theta_2 - \theta_1) - \sin 2\chi_o \sin(\theta_2 - \theta_1)]$$

$$= R\,[\cos(\theta_2 - \theta_1 + 2\chi_o)] \qquad \text{2-56}$$

Substitution of this expression for $\beta$ into 2-35a gives the following distribution for $\theta_2 - \theta_1$:

$$\frac{1-R^2}{2\pi}\left(\frac{1}{1-R^2\cos(\theta_2-\theta_1+2\chi_o)} + \frac{R\cos(\theta_2-\theta_1+2\chi_o)}{[1-R^2\cos(\theta_2-\theta_1+2\chi_o)]^{3/2}}\left\{\pi - \cos^{-1}[R\cos(\theta_2-\theta_1+2\chi_o)]\right\}\right) \qquad \text{2-57}$$

The distribution 2-57 is still rather complicated, but two facts are evident upon direct inspection. First the distribution of the phase difference $\theta_2 - \theta_1$ is highly dependent upon the wavefront correlation function R. Second, the distribution is even about $2\chi_o$. It can be seen also that the limiting values of zero and unity for the correlation R yield a uniform distirbution over a range of $2\pi$ and a Dirac delta fuction at $2\chi_o$, respectively. For any R, the average value of $\theta_2 - \theta_1$ is $2\chi_o$.

As pointed out by Bramley (1951), the variance and higher moments of the phase-difference distribution would be difficult to calculate. Bramley did calculate the mean absolute phase difference, obtaining for this observational parameter

$$\overline{|\theta_2 - \theta_1|} = \cos^{-1} R \qquad \text{2-58}$$

However, the entire development to this point contains a physically severe restriction demanded by the Lindeberg formulation of the central limit theorem. Before discussing observational parameters, as such, let us remove this restriction.

D2 The Special Case of Weakly Scattered Waves

The Lindeberg formulation of the central limit theorem requires that no component in our angular spectrum dominates the spectrum. If we were discussing backscatter, for instance, this condition would preclude consideration of a spectrum containing a strong component due to specular reflection. In the forward scatter case we might expect an analagous "nondeviated" or nonscattered component. In Chapter III we shall discuss such a component within the context of ionospheric structure. For now, let us assume its existence and ask how we can handle it in our present observational discussion. Again we shall follow Bramley's application of Rice's work. In order to describe conditions observed in the auroral zone, however, we shall have to extend Bramley's analysis and its mathematical foundation due to Rice. Unfortunately, we shall have to resort to numerical techniques.

In the foregoing sections of this chapter we have been dealing with a signal received on the ground after scattering in the ionosphere. We began in equation 2-1 with a Fourier sum of scatter components, with the general component of angular frequency $\omega_n$ presumed to be the result of ionospheric modulation. Suppose now that in addition to the scatter, or modulation, components there is a unique component at the angular frequency $\omega_o = \omega$. The relationship between the unique component and the remainder of the spectrum is that of a carrier and modulation-produced sidebands.

In the absence of ionospheric modulation by scattering, there would be no sidebands and the entire flux from the source would be contained in the unique component. In the foregoing, we were forced by the central limit theorem to consider the opposite case of complete scattering, where all power is contained in the scatter spectrum. We shall now consider the case of weak scatter in which both the nonscattered and scattered components exist but where the former dominates the overall angular spectrum.

For the case of weak scatter, the expression corresponding to equation 2-1 is identical to 2-1 with the addition of a term outside the summation, giving the contribution of the nonscattered component. Since the phases, $\phi_n$, in 2-1 are assumed to be random, we are free to reference them in such a way that $\phi_o = 0$. In addition, for $\omega_o = \omega$, the second term in the argument of 2-1 vanishes, and the contribution of the nondeviated component is simply $c_o \cos \omega t$. That is, the time varying phase $\psi_o$ is zero. It is only in this case that the concept of "carrier" and "sidebands" can be retained strictly.

It is to be noted that one could take account of large-scale refractive effects by allowing the "nondeviated" component to have a time-varying phase $\psi_o(t)$. In this case, small-scale scatter effects would be described in the frequency domain and large-scale refractive effects in the time domain of the same Fourier pair. If this were done, we would find that we would have to have one foot in each domain of the spatial-angular Fourier pair also. This approach has potential for describing multiple-scaled ionospheric structure, but it becomes somewhat complicated analytically. We shall concern ourselves only with effects which, in the present context, can be assigned to "small-scale scattering."

Whatever the nature of the angular spectrum, the voltages $v_1$ and $v_2$ at the two antennas can be represented by (time-varying) amplitudes $A_1$ and $A_2$ and phases $\theta_1$ and $\theta_2$, as in equations 2-10. We are interested in relations between these quantities for the case of a nonscattered component, which we shall denote by $S \cos \omega t$, accompanied by a scatter spectrum. The situation is mathematically equivalent to a sinusoidal signal in the presence of a noise spectrum. It is the scatter or "noise" component which was described in section D1 of this chapter.

All of the foregoing results can be taken over with a slight change in notation. We can no longer equate the antenna amplitudes and phases with those of the scatter-component resultant alone. We shall retain A's and $\theta$'s for the former and use B's and $\phi$'s for the latter. With this notation in mind and taking into account the separation between antennas, the antenna voltages can be written as

$$v_1 = A_1 \cos (\omega t + \theta_1) = S \cos (\omega t + \chi_o) + B_{1c} \cos \omega t - B_{1s} \sin \omega t \qquad 2\text{-}59$$

$$v_2 = A_2 \cos (\omega t + \theta_2) = S \cos (\omega t - \chi_o) + B_{2c} \cos \omega t - B_{2s} \sin \omega t \qquad 2\text{-}60$$

where the four B's represent components of random-walk resultants as discussed in section D1a of this chaper.

The relations between the various quantities above are easily seen in the phasor diagrams of figure 2, where it is to be remembered that the B's and $\phi$'s are random variables.

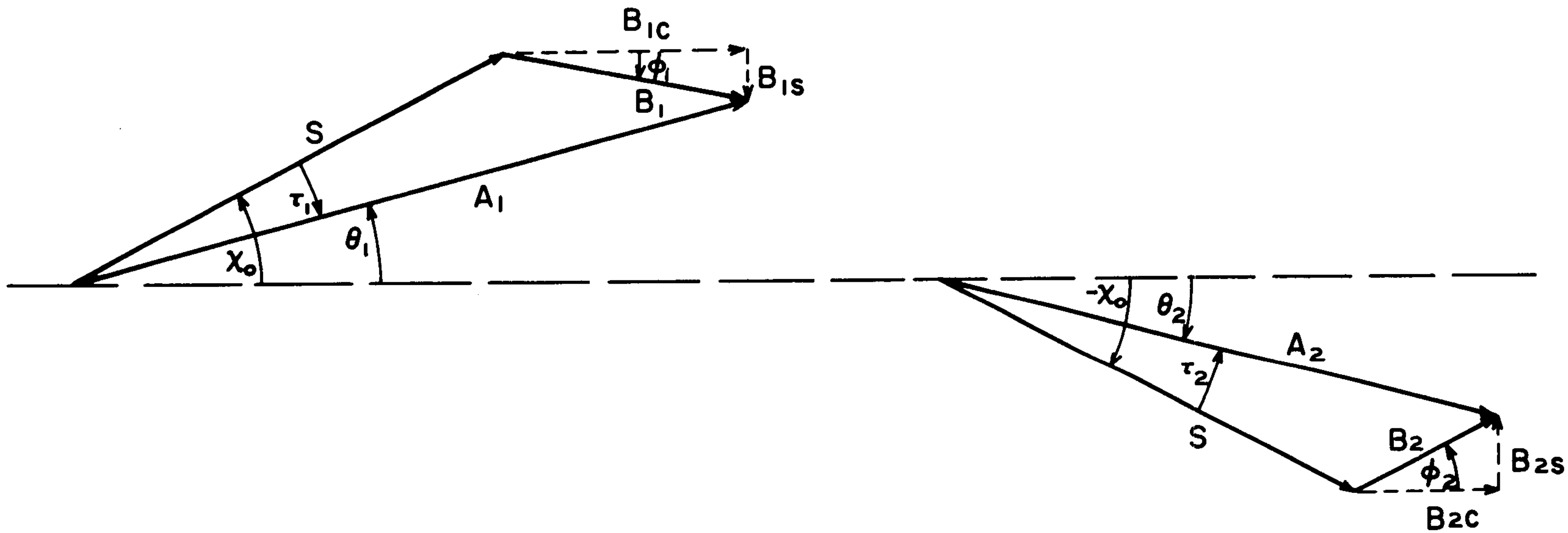



Fig. 2. Phasor relations of scattered and nonscattered signal components at two antennas of an interferometer.

From the phasor diagrams or from equations 2-59 and 2-60, it is obvious that

$$A_1^2 = S^2 + 2S(B_{1c} \cos \chi_o + B_{1s} \sin \chi_o) + B_{1c}^2 + B_{1s}^2 \qquad \text{2-61}$$

$$A_2^2 = S^2 + 2S(B_{2c} \cos \chi_o - B_{2s} \sin \chi_o) + B_{2c}^2 + B_{2s}^2 \qquad \text{2-62}$$

The last two terms in each of the above, when averaged, each give the variance $\sigma^2$ of the scatter components. Further, since $B_{1c}$, $B_{1s}$, $B_{2c}$ and $B_{2s}$ have zero means in accord with the results of section D1a, we get the following upon averaging equations 2-61 and 2-62:

$$\overline{A_1^2} = \overline{A_2^2} = S^2 + 2\sigma^2 = 2P \qquad \text{2-63}$$

where P is the (average) total power (per unit resistance) received from the source. At the frequencies with which we are concerned, where ionospheric absorption may be ignored, the total power is constant.

From equations 2-59 and 2-60 we can find also the covariance of the voltages $v_1$ and $v_2$. It is shown in Appendix 1e that the result is

$$\overline{v_1 v_2} = \tfrac{1}{2}S^2 \cos 2\chi_o + \mu_c \qquad \text{2-64}$$

For a narrow scatter spectrum which is symmetrical about the direction of the nonscattered component, $\mu_c$ is found from equations 2-51 and 2-53 to be

$$\mu_c = R\,\sigma^2 \cos 2\chi_o \qquad \text{2-65}$$

Thus equation 2-64 becomes

$$\overline{v_1 v_2} = (\tfrac{1}{2}S^2 + R\sigma^2) \cos 2\chi_o \qquad \text{2-66}$$

Since the correlation $\rho$ between the voltages is given by the ratio of the covariance to the power P, we have from 2-66 and 2-63

$$\rho = \frac{S^2 + 2R\sigma^2}{S^2 + 2\sigma^2} \cos 2\chi_o \qquad 2\text{-}67$$

It can be seen that if S goes to zero, the variance reduces to that given in equation 2-55 for the complete scattering case and $\rho$ to the corresponding value of correlation, R cos $2\chi_o$.

Equations 2-66 and 2-67 hold for all values of the variables and so are quite general. To discuss the correlation of amplitudes and the phase-difference characteristics of the signals in this general case, however, is difficult. We shall restrict our considerations here to the case of weak scatter, following the example of Bramley (1951). In section D3 we shall extend Bramley's considerations to the more general case, using numerical techniques.

Weak scatter is characterized by the condition that $\sigma$ is small compared with $S/\sqrt{2}$. In this case, according to Bramley, we have approximately

$$A_1 = S + B_{1c} \cos \chi_o + B_{1s} \sin \chi_o + \frac{1}{2S} (B_{1c} \sin \chi_o - B_{1s} \cos \chi_o)^2 \qquad 2\text{-}68$$

$$A_2 = S + B_{2c} \cos \chi_o - B_{2s} \sin \chi_o + \frac{1}{2S} (B_{2c} \sin \chi_o + B_{2s} \cos \chi_o)^2 \qquad 2\text{-}69$$

These approximations can be verified by squaring and comparing with 2-61 and 2-62. If this process is carried out, it is found that the approximation involves ignoring the effect of terms containing factors of order $\sigma/S$ and $\sigma^2/S^2$. From 2-68 and 2-69 we find that

$$\overline{A_1} = \overline{A_2} = S + \frac{\sigma^2}{2S} \qquad 2\text{-}70$$

and

$$\overline{A_1}^2 = \overline{A_2}^2 \approx S^2 + \sigma^2 \qquad \text{2-71}$$

As in the complete scatter case, we shall find that important quantities are the mean product of amplitudes $\overline{A_1A_2}$ and the corresponding quantity for the squares of the amplitudes $\overline{A_1^2A_2^2}$. If, again, we ignore terms of high order in $\sigma/S$, we obtain approximately from 2-68 and 2-69

$$\overline{A_1A_2} = S^2 + \sigma^2 + \mu_c \cos 2\chi_o + \mu_s \sin 2\chi_o \qquad \text{2-72}$$

and from 2-61 and 2-62

$$\overline{A_1^2A_2^2} = S^4 + 4\sigma^2 S^2 + 4S^2 (\mu_c \cos 2\chi_o + \mu_s \sin 2\chi_o) \qquad \text{2-73}$$

The details are given in Appendix 1f.

Now equation 2-39 shows that we can obtain the correlation coefficient for amplitudes from 2-70, 2-71 and 2-72. Thus

$$\rho_A = \frac{\mu_c \cos 2\chi_o + \mu_s \sin 2\chi_o}{\sigma^2} \qquad \text{2-74}$$

But comparison with equations 2-51 and 2-52 and 2-53 shows that, for a narrow symmetrical spectrum, this gives simply

$$\rho_A = R \qquad \text{2-75}$$

In order to find the correlation coefficient for the squares of the amplitudes from equation 2-40, we still need $\overline{A^2}^2$ and $\overline{A_1^4}$. The former we get simply from 2-63. Thus we have approximately

$$\overline{A_1^2}^2 = \overline{A_2^2}^2 = S^4 + 4S^2 \sigma^2 \qquad \text{2-76}$$

The latter is obtained from 2-61 and 2-62 with the approximate result

$$\overline{A_1^4} = \overline{A_2^4} = S^4 + 8S^2\sigma^2 \qquad \text{2-77}$$

Combining these along with 2-73, we get from 2-40

$$\rho_{A^2} = \frac{4S^2\,(\mu_c \cos 2\chi_o + \mu_s \sin 2\chi_o)}{4S^2\,\sigma^2} \qquad \text{2-78}$$

Again this reduces, for the case of a narrow symmetrical scatter spectrum, to

$$\rho_{A^2} = R \qquad \text{2-79}$$

Thus in the weak scatter case the correlation coefficient for amplitudes and that for the square of amplitudes are equal and both are given directly by the wavefront correlation function.

From the phasor diagrams of figure 2, we obtain the following:

$$\tan\theta_1 = \frac{S \sin\chi_o + B_{1s}}{S \cos\chi_o + B_{1c}} \qquad \tan\theta_2 = \frac{-S \sin\chi_o + B_{2s}}{S \cos\chi_o + B_{2c}} \qquad \text{2-80}$$

Now, we should like the distribution of the random variable $(\theta_2 - \theta_1)$. Accordingly, we need the distributions of $\theta_1$ and $\theta_2$. Since $\chi_o$ is constant, the desired distributions will be identical, except for a shift in mean, to the distributions of $\tau_1$ and $\tau_2$, where the latter two random variables are defined as

$$\tau_1 = \theta_1 - \chi_o \quad \text{and} \quad \tau_2 = \theta_2 + \chi_o \qquad \text{2-81}$$

Combining 2-80 and 2-81, we find

$$\tan \tau_1 = \frac{\dfrac{S \sin \chi_o + B_{1s}}{S \cos \chi_o + B_{1c}} - \tan \chi_o}{1 + \dfrac{S \sin \chi_o + B_{1s}}{S \cos \chi_o + B_{1c}} \tan \chi_o} \; ; \; \tan \tau_2 = \frac{-\dfrac{S \sin \chi_o - B_{2s}}{S \cos \chi_o + B_{2c}} + \tan \chi_o}{1 + \dfrac{S \sin \chi_o - B_{2s}}{S \cos \chi_o + B_{2c}} \tan \chi_o}$$

2-82

It is easily shown that 2-82 reduces to

$$\tan \tau_1 = \frac{B_{1s} \cos \chi_o - B_{1c} \sin \chi_o}{S + B_{1c} \cos \chi_o + B_{1s} \sin\chi_o} ; \tan \tau_2 = \frac{B_{2s} \cos \chi_o + B_{2c} \sin \chi_o}{S + B_{2c} \cos \chi_o + B_{2s} \sin\chi_o}$$

2-83

In the weak scatter case where S dominates the angular spectrum, $\tau_1$ and $\tau_2$ will be small and the above can be approximated by

$$\tau_1 = S^{-1}(B_{1s} \cos \chi_o - B_{1c} \sin\chi_o) \; ; \; \tau_2 = S^{-1}(B_{2s}\cos \chi_o + B_{2c} \sin \chi_o)$$

2-84

Equations 2-84 show that, in the weak scatter case, $\tau_1$ and $\tau_2$ are the sums of two independent gaussian random variables. Accordingly, they themselves have gaussian distributions. Their means and variances will be given by the sums of the means and variances of the corresponding B's. Thus both $\tau_1$ and $\tau_2$ have means of zero and, taking into account the constants $S^{-1} \cos \chi_o$ and $S^{-1} \sin \chi_o$, variances of $\sigma^2 S^{-2}$.

From equations 2-84, we obtain

$$\overline{\tau_1\tau_2} = S^{-2} (\overline{B_{1s}B_{2s}} \cos^2 \chi_o - \overline{B_{1c}B_{2c}} \sin^2 \chi_o + \overline{B_{1s}B_{2c}} \sin \chi_o \cos \chi_o$$

$$- \overline{B_{2s}B_{1c}} \sin \chi_o \cos \chi_o)$$

$$= S^{-2}[\mu_c (\cos^2 \chi_o - \sin^2 \chi_o) + 2\mu_s (\sin \chi_o \cos \chi_o)]$$

$$= S^{-2} (\mu_c \cos 2 \chi_o + \mu_s \sin 2 \chi_o) \qquad 2\text{-}85$$

and

$$\overline{\tau_1^2} = \overline{\tau_2^2} = S^{-2} \sigma^2 \qquad 2\text{-}86$$

From equations 2-85 and 2-86 we see that the correlation coefficient between $\tau_1$ and $\tau_2$, and therefore between $\theta_1$ and $\theta_2$, is given by

$$\rho_\tau = \frac{\mu_c \cos 2\chi_o + \mu_s \sin 2\chi_o}{\sigma^2} \qquad 2\text{-}87$$

Again, for the case of a narrow symmetrical scatter spectrum, this reduces to

$$\rho_\tau = R \qquad 2\text{-}88$$

Thus in the case of weak scatter and a narrow symmetrical randomly phased scatter spectrum, the correlation coefficients for all pertinent envelope parameters are given directly by the wavefront correlation R.

The characteristics of $\tau_1$ and $\tau_2$ established above may be summarized by a two-dimensional normal distribution with zero means, variances of $\sigma^2 S^{-2}$, and correlation R. Thus (Bendat, 1958, section 3.4.4)

$$f(\tau_1,\tau_2) = \frac{S^2}{2\pi\,\sigma^2\,\sqrt{1-R^2}}\ \exp\left[-\frac{S^2(\tau_1{}^2 - 2R\tau_1\tau_2 + \tau_2{}^2)}{2\sigma^2\,(1-R^2)}\right] \qquad 2\text{-}89$$

More important, the characteristics of $\tau_1$ and $\tau_2$ allow us to establish the distribution of $\tau_2 - \tau_1$. Middleton (1960, section 7.5-1) shows that the distribution of the sum of two gaussian random variables is always gaussian, even if the two variables are not independent. He also shows that the mean of the sum is equal to the sum of the means and gives a very general expression for the variance of the resulting distribution. In our particular case, the resulting mean is zero and the resulting variance is given by $(2\sigma^2/S^2)(1 - R)$. Hence the distribution of $\tau_2 - \tau_1$ is

$$f(\tau_2 - \tau_1) = \frac{S}{2\sigma[\pi(1-R)]^{\frac{1}{2}}}\ \exp\left[\frac{-S^2\,(\tau_2 - \tau_1)^2}{4\sigma^2\,(1-R)}\right] \qquad 2\text{-}90$$

Now equations 2-81 show that

$$\tau_2 - \tau_1 = \theta_2 - \theta_1 + 2\,\chi_o \qquad 2\text{-}91$$

Thus equation 2-90 gives the distribution of the phase difference $\theta_2 - \theta_1$ with the mean shifted by $2\chi_o$. Therefore the moments of the distribution 2-90 give directly the central moments of $\theta_2 - \theta_1$. In addition, if we define

$$\eta = \tau_2 - \tau_1 \qquad 2\text{-}92$$

we also see that $\overline{|\eta|}$ gives the mean deviation of $\theta_2 - \theta_1$ from its mean value of $2\chi_o$. This observational quantity is given, from the distribution 2-90, by

$$\overline{|\eta|} = \frac{S}{\sigma[\pi(1 - R)]^{\frac{1}{2}}} \int_0^\infty \eta \exp \left[ \frac{-S^2 \eta^2}{4\sigma^2(1-R)} \right] d\eta \qquad 2\text{-}93$$

The integral is of the form $\int_0^\infty \eta \, e^{-a^2\eta^2} \, d\eta$, which by a change of variable to $x = a\,\eta$, is found to be $1/2a^2$ or $2\,\sigma^2(1-R)/S^2$. Thus we obtain

$$\overline{|\eta|} = \frac{\sigma}{S} \, 2 \left(\frac{1 - R}{\pi}\right)^{\frac{1}{2}} \qquad 2\text{-}94$$

which is the weak scatter expression corresponding to equation 2-58 in the complete scatter case.

D3 The General Case of Arbitrary Degree of Scatter

3a <u>The coherence ratio</u>: In sections D1 and D2 of this chapter we have discussed the statistical characteristics of two types of signals. The first was assumed to be totally scattered during transmission through the ionosphere. In the second case it was assumed that only a very small portion of the total flux was scattered, the major portion arriving at the ground without any ionospherically introduced characteristics. This condition of weak scatter corresponds to reality at middle latitudes. The condition of complete scatter was treated by Bramley (1951) largely to provide an analytical foundation for the weak-scatter case. In the auroral zone, ionospheric conditions are such that the complete scatter case also is met occasionally in observations. More often, conditions are intermediate between the two extremes of weak and complete scatter. Thus it is necessary to treat the general case.

In section D2 we introduced an ionospherically nondeviated component of the angular spectrum whose amplitude is S and whose average power therefore is $S^2/2$. The nondeviated component was added to a noiselike spectrum of scatter components of average power $\sigma^2$. Obviously the ratio of the two contributions to total received power can be used as a measure of the degree of scattering. Now, when there is no scattering the received wavefront is fully coherent, and when there is complete scattering the wavefront has the noncoherent or random character of noise. As a measure of the degree of scattering, therefore, let us define the "coherence ratio" b as

$$b = S^2/2\sigma^2 \qquad 2\text{-}95$$

It was pointed out in section D2 that the nondeviated component is mathematically equivalent to a sinusoidal signal buried in noise, the latter arising from the scatter spectrum. Thus the coherence ratio is mathematically equivalent to a (power) signal-to-noise ratio. It is immediately evident that the complete scatter case of section D1 corresponds to a coherence ratio of zero while the weak scatter case of section D2 corresponds to large coherence ratio. Full coherence is achieved as the flux in the scatter spectrum vanishes, forcing the coherence ratio to increase without limit.

The importance of the coherence ratio in describing the effect of ionospheric scattering is obvious. It is a direct measure of the degree to which scattering is taking place during a given observation. This has been recognized for some time and put to use in observations with radiometers and other simple receivers. In analysis of such observations it was possible to make direct use of the work of Rice (1945) on the mathematical analysis of a sinusoidal signal buried in noise (McNicol, 1949). Rice's work, however, stopped short of providing a mathematical basis for analyzing interferometric observations. When two antennas are used and phase coherence between them is retained, one must deal with the correlation which exists between the signals received at the two antennas. The mathematical analogies dealing with the temporal autocorrelation function of a signal were developed by Rice for the case of pure noise but not for the case of a sinusoid buried in noise.

Now Bramley (1951) was able to adapt Rice's work to the case of weak scatter as observed with an interferometer. He did not treat the general case however. His work was adequate for observations at middle latitudes. At equatorial (Koster, 1958) and auroral (Little et al, 1962) latitudes, however, significant reductions are observed in the correlation of voltage at the two antennas of an interferometer. These reductions, which we shall call "visibility fades," can occur only if the coherence ratio decreases to finite values. In the extreme, the coherence ratio may decrease to zero so that complete scatter takes place, but any finite value is possible and indeed a wide range is observed.

Studies of visibility fades to date either have effectively ignored the quantitative ionospheric information available in them or else have treated the phenomenon approximately. The approximate approach precludes direct experimental determination of the coherence ratio. In some of the work which has appeared in the literature, the significance of the coherence ratio and of Bramley's approach to the problem does not appear to have been fully appreciated. It is our intent in this section to extend Bramley's approach to include the general case of arbitrary coherence ratio.

3b Covariance of voltages and fringe visibility: It is widely known that the output of a phase-switch radio interferometer gives directly the covariance of its antenna voltages (Bracewell, 1958; Fremouw, 1963). We have already established (equation 2-66) that this quantity is given for arbitrary b, by

$$\overline{v_1 v_2} = (\tfrac{1}{2}S^2 + R\sigma^2) \cos 2\chi_o \qquad 2\text{-}96$$

Now, in the absence of ionospheric scatter, $\sigma$ vanishes and all power is contained in the nondeviated component of the angular spectrum. Thus, under disturbed ionospheric conditions the covariance of antenna voltages, and therefore the averaged output of a phase-switch interferometer, becomes $P \cos 2\chi_o$. But P, the total flux received from the source, is independent of scatter conditions and in general is given, according to equation 2-63, by $\tfrac{1}{2}S^2 + \sigma^2$. Therefore the ratio of the averaged output of a phase-switch interferometer under conditions of scatter to that under undisturbed conditions is given by

$$\frac{\tfrac{1}{2}S^2 + R\sigma^2}{\tfrac{1}{2}S^2 + \sigma^2} = \frac{b + R}{b + 1} \qquad 2\text{-}97$$

Comparison of 2-97 with equation 2-67 shows that

$$\frac{b + R}{b + 1} \cos 2\,\chi_o = \rho \qquad 2\text{-}98$$

where $\rho$ is the correlation coefficient for voltages.

Now, the covariance is inherently an average quantity, so that identification of it with the output of any instrument implies some sort of averaging of that output. In our case, where we are dealing with a radio interferometer observing a radio star whose signal has been scattered by the ionosphere, our averaging period must be long compared with the scintillations produced by motion of the scattering region. Typically this requires averaging over at least many tens of seconds and preferably over a few minutes.

Statistical stationarity requires that b and R remain constant over the averaging period. In the case of a radio star, however, its direction $\alpha_o$ and the phase angle $\chi_o$ derived therefrom vary with time due to the earth's rotation. We shall see later that a convenient and meaningful averaging period, which turns out in practice to be on the order of minutes, is that during which $2\chi_o$ varies from $-\pi/2$ to $+\pi/2$. Let us denote by $\bar{\rho}$ the average value of $\rho$ during this period. Now $\chi_o$ is approximately a linear function of time so, during the same period, we have $\overline{\cos 2\chi_o} = 2/\pi$. Thus, over such an averaging period, we have

$$\frac{b + R}{b + 1} = \frac{\pi\bar{\rho}}{2} \qquad 2\text{-}99$$

Equation 2-99 relates the coherence ratio b and the wavefront correlation R to the average mathematical correlation coefficient between the voltages at the antenna outputs. The right side of the equation has little physical significance and in fact depends strongly upon our choice of averaging period. The physical parameters we desire are b and R. For simplicity, let us replace $\pi\bar{\rho}/2$ with the symbol r, where

r = the ratio of the average output of a phase-switch interferometer under conditions of scatter to that under undisturbed ionospheric conditions, with the period of averaging in each case corresponding to one-half cycle on the instrument's output pattern.

With this definition, the disturbed coherence ratio b and wavefront correlation R are related observationally by

$$r = \frac{b + R}{b + 1}. \qquad 2\text{-}100$$

The parameter r has strict mathematical meaning through its relation to $\bar{\rho}$, but its physical meaning is obscure from this point of view.[1] It will be given more obvious meaning later when we discuss observational scaling procedures. For now let it suffice to say that r is a quantity which is readily determined from phase-switch interferometer observations. We shall call it "visibility." Its importance lies in its relationship to the physically meaningful quantities b and R. The relation of b and R to ionospheric parameters will be developed in Chapter III.

3c <u>The analytical approach to amplitude and phase characteristics and its limitations</u>: Equation 2-96 gives the covariance of antenna voltages for arbitrary coherence ratio. In order to obtain equally general relations for the amplitude and phase characteristics analytically, we must obtain the appropriate joint probability density functions. Several successors of Rice in the field of signal statistics and communication theory have dealt with the problem of a signal buried in noise. Notable among these is Middleton, who developed the quadri-variate joint distribution for amplitudes and phases and carried out the integrations necessary to obtain the joint distributions for amplitudes and for phases independently (<u>Middleton</u>, 1947).

---

[1] For the reader familiar with radio-astronomy terminology, consideration of equations 2-98 and 2-100 will show that r is the fringe visibility of the interferometer record.

The distributions which Middleton developed are directly applicable to our situation if we include the reasonable assumption of a narrow symmetrical angular spectrum centered on the nondeviated component, with one additional assumption necessary, namely that the slowly varying phase $\chi_o$ is kept at zero. This "phase compensation" is produced and maintained electronically in the pertinent instrument of our experiment. Hence we are free to use Middleton's distributions directly.

Middleton's work is quite general and it includes provision for dealing with modulated signals. This provision could be used in our problem in order to account for refractive effects by large-scale ionospheric structure. Thus changes in apparent direction of the source would produce phase modulation of our nondeviated component and focusing or defocusing by ionospheric lenses would produce amplitude modulation. The same modulations, however, would be imposed on the scatter components, resulting in nonstationarity of the noiselike part of our total signal. Again we shall concentrate on small-scale scattering, having pointed out the potential of Middleton's formulation for dealing with multi-scaled ionospheric scattering.

For the case of small-scale scattering only, equation 5.18 of Middleton (1947) gives directly the joint distribution of amplitudes appropriate to our observations.[1] In our notation the distribution is

$$\frac{A_1 A_2}{\sigma^4(1-R^2)} \exp\left[\frac{-2b}{(1+R)} - \frac{A_1^2+A_2^2}{2\sigma^2(1-R^2)}\right] \sum_{m=0}^{\infty} \varepsilon_m I_m\left[\frac{RA_1A_2}{\sigma^2(1-R^2)}\right] I_m\left[\frac{SA_1}{\sigma^2(1+R)}\right] I_m\left[\frac{SA_2}{\sigma^2(1+R)}\right] \quad \text{2-101}$$

[1]The observations referred to here were obtained with a phase-sweep interferometer employing noncoherent detection and phase compensation. (The phase compensation feature has little practical significance in consideration of amplitudes but considerable importance for phase measurements).

where $\varepsilon_o = 1$, $\varepsilon_m = 2$ for $m \geq 1$, and $I_m$ = the modified Bessel function of the first kind and order m.

While the distribution above is quite complicated, it is fairly easy to see that it reduces to the distribution given by Rice for pure noise, in the event that S and b vanish. In this case the second and third Bessel function arguments above vanish. The corresponding Bessel functions themselves then vanish except for order zero, in which order they have value unity (Watson, 1948, equation 2 in section 3.7 and Table II in the appendix). The distribution 2-101 then reduces to

$$f(A_1,A_2) = \frac{A_1A_2}{\sigma^4(1-R^2)} \exp\left[-\frac{A_1^2 + A_2^2}{2\sigma^2(1-R^2)}\right] I_o\left[\frac{R A_1A_2}{\sigma^2(1-R^2)}\right] \qquad 2\text{-}102$$

Upon recalling the definition of R from equation 2-38, it is readily seen that equation 2-102 is identical with equation 2-34, which is the distribution given by Rice for pure noise (our complete scatter case). In the opposite extreme of full coherence, where $\sigma^2$ vanishes and b increases without limit, the distribution becomes a Dirac delta function at $A_1 = A_2 = S$.

Now Middleton (1947) has developed expressions for second-order moments such as $\overline{A_1A_2}$ and $\overline{A_1^2A_2^2}$ from the distribution 2-101, from which we could derive the correlation coefficients for amplitudes and for the squares of amplitudes. The expressions are so complicated, however, as to preclude their usefulness for our observational problem.

Middleton's joint distribution for the phases $\theta_1$ and $\theta_2$ (equation 5.20 in his 1947 paper) is even more complicated than that quoted in our equation 2-101 for the amplitudes $A_1$ and $A_2$ under the same conditions. In his textbook on statistical communication theory (Middleton, 1960), the author himself points out the limited applicability (due to complexity) of expressions for moments of the joint phase distribution. Further, it is not the joint phase distribution, per se, whose moments we desire for our observational problem, but rather observational parameters derived from the single-variate distribution of $(\theta_2-\theta_1)$. There appears to be no simple means of obtaining our desired distribution and observational parameters derived therefrom, even if we were to use Middleton's joint distribution for $\theta_1$ and $\theta_2$.

3d <u>The numerical approach</u>: In view of the above it appears necessary to abandon hope for an analytical treatment of the general scattering problem. Instead we shall develop a numerical technique which will allow us to compute observational parameters for arbitrary coherence ratio. The technique must produce results which agree with the analytical results of Rice and Bramley in the special cases where analytical solutions exist. In Chapters IV and V we shall use the results of the numerical analysis to reach observational conclusions.

Let us recall that the signals with which we must work in the general case consist of a nonscattered component whose amplitude is S and a spectrum of scatter components whose resultant amplitude is $B_1$ at one antenna and $B_2$ at the other. The phase angles of the nonscattered component are plus and minus $\chi_o$, respectively, at the two antennas. The phase angles of the scatter components are independent of one another and uniformly distributed over a range of $2\pi$.

Now $\chi_o$ is established purely by the position of the source and, for a radio star, is slowly varying and completely predictable. It contains no information about the ionosphere or its scattering process (for the small-scale scattering under consideration here.) It can be (and is in our phase-sweep interferometer) maintained at the constant value of zero by instrumental introduction of time-varying phase compensation. The same phase retardation is introduced into each of the scatter components with the result that their statistical characteristics are unchanged. All the statistical results of section D1 of this chapter are applicable to the scatter spectrum and its resultant, with $\chi_o$ to be taken as zero throughout.

The relationships between the quantities of interest are those shown in the phasor diagrams of figure 2, except that $\chi_o$ is to be taken as zero. Thus we have the relations given in figure 3. It is to be recalled that the B's are random variables resulting from random-walk addition of a randomly phased angular scatter spectrum. Their statistical characteristics, along with the relative magnitude of S, determine the statistical characteristics of the A's and θ's. It is certain average characteristics of these latter random variables which we seek.

The independent variables which we wish to control as input parameters in our numerical technique are the relative magnitude of the non-scattered and scattered components and the correlation between the scatter resultants at the two antennas. That is, we wish to control the coherence ratio b (while maintaining the total power constant) and the wavefront correlation R. It will be recalled that the latter is identical to the correlation of voltage in the scatter spectrum when $\chi_o$ is zero.

Now the total power is given by $P = S^2 + \sigma^2$ and the coherence ratio by $b = S^2/2\sigma^2$. Combination of these two equations shows that

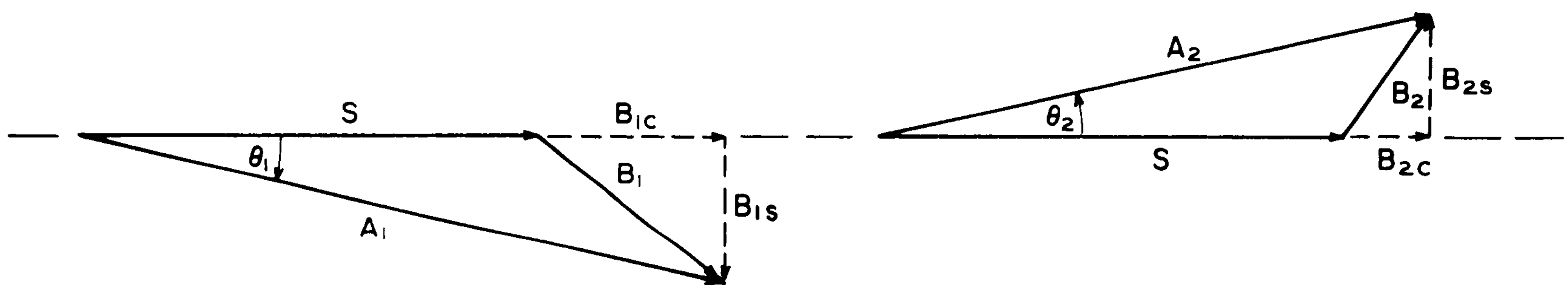



Fig. 3. Effective phasor relations of scattered and nonscattered signal components at two antennas of a phase-compensated interferometer.

$$S = [2bP/(b+1)]^{1/2} \text{ and } \sigma = [P/(b+1)]^{1/2} \qquad 2\text{-}103$$

For constant P, the magnitude of the coherent and nocoherent components are controlled by choice of b through equations 2-103.

The control of correlation is not so obvious but it turns out to be rather simple with digital computer techniques. Modern program libraries contain routines for generating fields of random numbers whose distributions are gaussian. We shall now show how four independent gaussian fields can be combined to produce two correlated gaussian fields in which the correlation coefficient is controllable as an input parameter. We shall find that the two correlated fields together with two of the original four independent fields possess the joint distribution which describes the scatter component of our angular spectrum.

First let us recall the distribution which must be satisfied. It is given in equation 2-32 for the case of general $\chi_o$. Equations 2-51, 2-52 and 2-53 show that, for $\chi_o = 0$, $\mu_s$ vanishes and $\mu_c$ reduces to $R\,\sigma^2$. In this case, 2-32 becomes

$$f(B_{1c},B_{1s},B_{2c},B_{2s}) = \frac{1}{4\pi^2\,\sigma^4(1-R^2)} \exp\left\{\frac{-1}{\sigma^2(1-R^2)}\left[(B_{1c}^2+B_{1s}^2+B_{2c}^2+B_{2s}^2) - 2R(B_{1c}B_{2c}-B_{1s}B_{2s})\right]\right\} \qquad 2\text{-}104$$

This is the quadri-variate gaussian distribution appropriate to our case, namely a randomly phased, narrow and zero-center symmetrical angular spectrum of scatter components. The first and second moments are as follows:

$$\overline{B_{1c}} = \overline{B_{1s}} = \overline{B_{2c}} = \overline{B_{2s}} = 0$$

$$\overline{B_{1c}^2} = \overline{B_{1s}^2} = \overline{B_{2c}^2} = \overline{B_{2s}^2} = \sigma^2 \qquad 2\text{-}105$$

$$\overline{B_{1c}B_{2c}} = \overline{B_{1s}B_{2s}} = R\,\sigma^2$$

$$\overline{B_{1c}B_{1s}} = \overline{B_{2c}B_{2s}} = \overline{B_{1c}B_{2s}} = \overline{B_{2c}B_{1s}} = 0$$

Now suppose we have four independent random variables with gaussian distributions, namely $B_{1c}$, $B_{1s}$, $x$, and $y$, with the following first and second moments:

$$\overline{B_{1c}} = \overline{B_{1s}} = \overline{x} = \overline{y} = 0$$

$$\overline{B_{1c}^2} = \overline{B_{1s}^2} = \overline{x^2} = \overline{y^2} = \sigma^2 \qquad 2\text{-}106$$

Since we specified independence, the crossmoments are zero as follows:

$$\overline{B_{1c}B_{1s}} = \overline{xy} = \overline{B_{1c}x} = \overline{B_{1c}y} = \overline{B_{1s}x} = \overline{B_{1s}y} = 0 \qquad 2\text{-}107$$

Let us form the following linear combinations:

$$B_{2c} = RB_{1c} + (1-R^2)^{\frac{1}{2}}\,x \quad \text{and} \quad B_{2s} = RB_{1s} + (1-R^2)^{\frac{1}{2}}\,y \qquad 2\text{-}108$$

The means of $B_{2c}$ and $B_{2s}$ are simply the sums of the means of the two terms in their respective generating transformations, 2-108. In view of the first equation in 2-106, then, we have

$$\overline{B_{2c}} = \overline{B_{2s}} = 0 \qquad 2\text{-}109$$

Further, since $B_{1c}$ and $x$ are statistically independent and $B_{2s}$ and $y$ are also, the variances of $B_{2c}$ and $B_{2s}$ are the sums of the variances of the terms in the generating transformations. In view of the second equation in 2-106, then, we have

$$\overline{B_{2c}^2} = \overline{B_{2s}^2} = R^2\,\sigma^2 + (1-R^2)\,\sigma^2 = \sigma^2 \qquad \text{2-110}$$

Forming other pertinent combinations from 2-106, 2-107 and 2-108 we get the following crossmoments:

$$\overline{B_{2c}B_{2s}} = R^2\overline{B_{1c}B_{1s}} + (1-R^2)\,\overline{xy} + R(1-R^2)^{\frac{1}{2}}\,\overline{(B_{1c}y + B_{1s}x)} = 0$$

$$\begin{aligned}
\overline{B_{2c}B_{1s}} &= R\overline{B_{1c}B_{1s}} + (1-R^2)^{\frac{1}{2}}\,\overline{xB_{1s}} = 0 \\
\overline{B_{2s}B_{1s}} &= R\overline{B_{1s}^2} + (1-R^2)^{\frac{1}{2}}\,\overline{yB_{1s}} = R\sigma^2 \qquad \text{2-111} \\
\overline{B_{2c}B_{1c}} &= R\overline{B_{1c}^2} + (1-R^2)^{\frac{1}{2}}\,\overline{xB_{1c}} = R\sigma^2 \\
\overline{B_{2s}B_{1c}} &= R\overline{B_{1s}B_{1c}} + (1-R^2)^{\frac{1}{2}}\,\overline{yB_{1c}} = 0
\end{aligned}$$

Summarizing the first and second-order moments resulting from the four independent fields defined by 2-106 and the linear transformations 2-108, we get the following:

$$\begin{aligned}
&\overline{B_{1c}} = \overline{B_{1s}} = \overline{B_{2c}} = \overline{B_{2s}} = 0 \\
&\overline{B_{1c}^2} = \overline{B_{1s}^2} = \overline{B_{2c}^2} = \overline{B_{2s}^2} = 0 \\
&\overline{B_{2c}B_{1c}} = \overline{B_{2s}B_{1s}} = R\sigma^2 \qquad \text{2-112} \\
&\overline{B_{1c}B_{1s}} = \overline{B_{2c}B_{2s}} = \overline{B_{2c}B_{1s}} = \overline{B_{2s}B_{1c}} = 0
\end{aligned}$$

But the moments given in 2-112 are identical with those given in 2-105. Further, a gaussian distribution is defined uniquely by its first and second-order moments. Thus, starting with four independent gaussian random variables, we can transform two of them in such a way that the resulting joint distribution defines voltages produced at two antennas by a randomly phased scatter spectrum with controllable correlation. The linear transformations involved are given in 2-108.

3e <u>Results of the numerical approach</u>: The preceding section describes a means of generating and relating the coherent and noncoherent component phasors shown in figure 3. With this accomplished, the trigonometry of the phasor diagrams allows straight-forward calculation of statistical characteristics of the resultant phasors as functions of b and R. For an extension of the work by Bramley, we should like to calculate the amplitude and amplitude-square correlation coefficients, $\rho_A$ and $\rho_A 2$, and the mean absolute phase difference $\overline{|\eta|}$.

We saw in section D2 that $\rho_A$ and $\rho_A 2$ are identical in the case of weak scatter and in section D1b that they are nearly so in the case of complete scatter. Thus, these two quantities are essentially equal throughout the range of b from zero to infinity and we shall concern ourselves only with $\rho_A$. Recalling that $\rho_A$ is given by

$$\rho_A = \frac{\overline{A_1 A_2} - \overline{A_1}^2}{\overline{A_1^2} - \overline{A_1}^2} \qquad \text{2-113}$$

it is easily seen from the phasor diagrams how $\rho_A$ can be calculated once the phasors S, $B_{1c}$,$B_{1s}$,$B_{2c}$, and $B_{2s}$ have been generated for a given combination of b and R, as described in section D3d.

Recalling also that $\overline{|\eta|}$ is given by

$$\overline{|\eta|} = \overline{|\theta_2 - \theta_1|} \qquad \text{2-114}$$

when $\chi_o$ is zero, this quantity[1] too is easily found from the trigonometry. In performing this calculation, the phase difference is kept within the range of $-\pi$ to $+\pi$, corresponding to the observational fact that ambiguities are not resolved. This corresponds also to the fact that phase distributions in the work of Rice, MacDonald, and Middleton refer to a $2\pi$ range.

In figure 4 are given the results of numerical calculation of $\rho_A$ and $|\eta|$ as functions of R for some finite values of b, along with the same quantities as obtained from Bramley's expressions for the limiting cases of large b and b = 0. These results are included to show how our numerical results relate to the analytical ones of Bramley. Certain other quantities are of more observational interest than $\rho_A$ and $|\eta|$. Let us first explore the statistical characteristics of amplitude-related quantities and then examine the characteristics of phase.

Many scintillation studies have employed "scintillation indices" based on the fractional fluctuation of amplitude or power. These indices take no account of reductions in correlation between the signals received at two antennas of an interferometer. The most direct interferometric method of measuring amplitude or power fluctuations is with a phase-sweep (or "lobe-sweep") interferometer which employs simple rectification of the audio signal rather than coherent (or "phase-sensitive") detection at the sweep frequency. Aside from a constant of proportionality, the output of such a device is given in our notation by $A_1A_2$. This quantity,

[1]Sometimes called the "difference correlation coefficient" for phase (Ratcliffe, 1956, section 8.4).

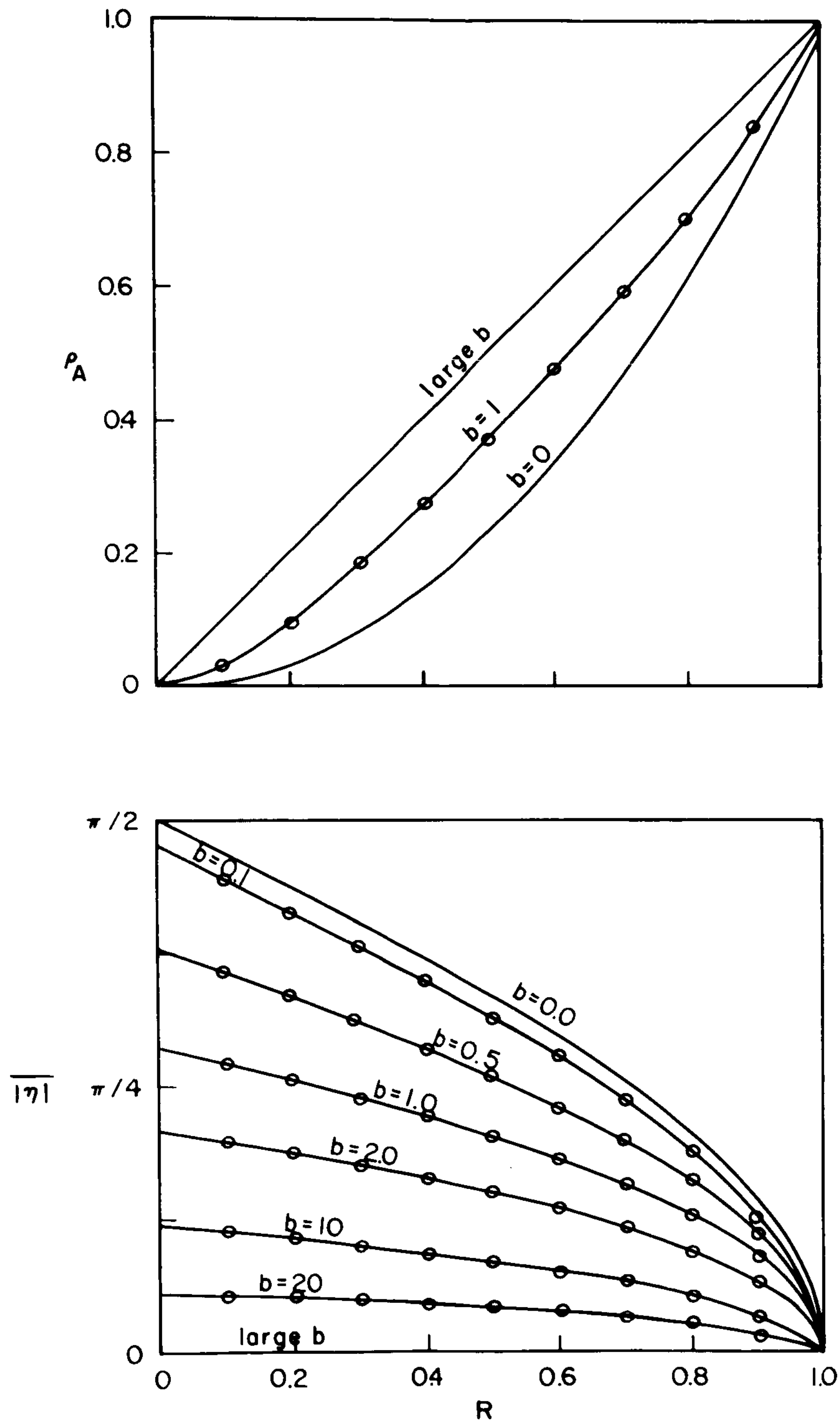



Fig. 4. Top, amplitude correlation coefficient computed numerically for b = 1 and analytically for limiting values of b; bottom, mean absolute phase difference computed numerically for finite values of b and analytically for limiting values of b.

or its square root, is of fundamental importance in scintillation indices as well as in the amplitude correlation coefficient $\rho_A$. Of interest, then, are the statistical distributions of $A_1A_2$ and $\sqrt{A_1A_2}$. The distribution of either of these quantities will suffice for our purpose, the other being easily obtained therefrom.

We shall concern ourselves with the distribution of $\sqrt{A_1A_2}$ because it can be compared directly to analytical results in a limiting case. This distribution can be obtained as a function of b and R by our numerical technique. Figure 5 gives the resulting histograms for twelve combinations of b and R. The histograms show that the average value of $\overline{\sqrt{A_1A_2}}$, and therefore also the average value of $A_1A_2$, is not very strongly dependent on either b or R. Thus, these average quantities cannot be expected to yield much information about small-scale ionospheric scattering. Some workers have relied on $\overline{A_1A_2}$ in attempts to explain the nature of visibility fades. Others, notably Flood (1963) and Moorcroft (1963), have concentrated on phase characteristics during visibility fades.

While the average value of $\overline{\sqrt{A_1A_2}}$ is not strongly dependent on b or R, figure 5 shows that the spread in the distribution is strongly controlled by b and to a lesser extent by R. The histograms clearly display the approach of the distribution toward a Dirac delta function as b is increased, which is to be expected from Bramley's analytical work. Less obvious, but nevertheless consistently displayed, is a sharper peaking of the distribution for a given b as R is decreased. This latter fact means that an interferometer record will show somewhat less fluctuation, for a given degree of ionospheric scatter, as amplitude scintillations at the two antennas become less correlated.

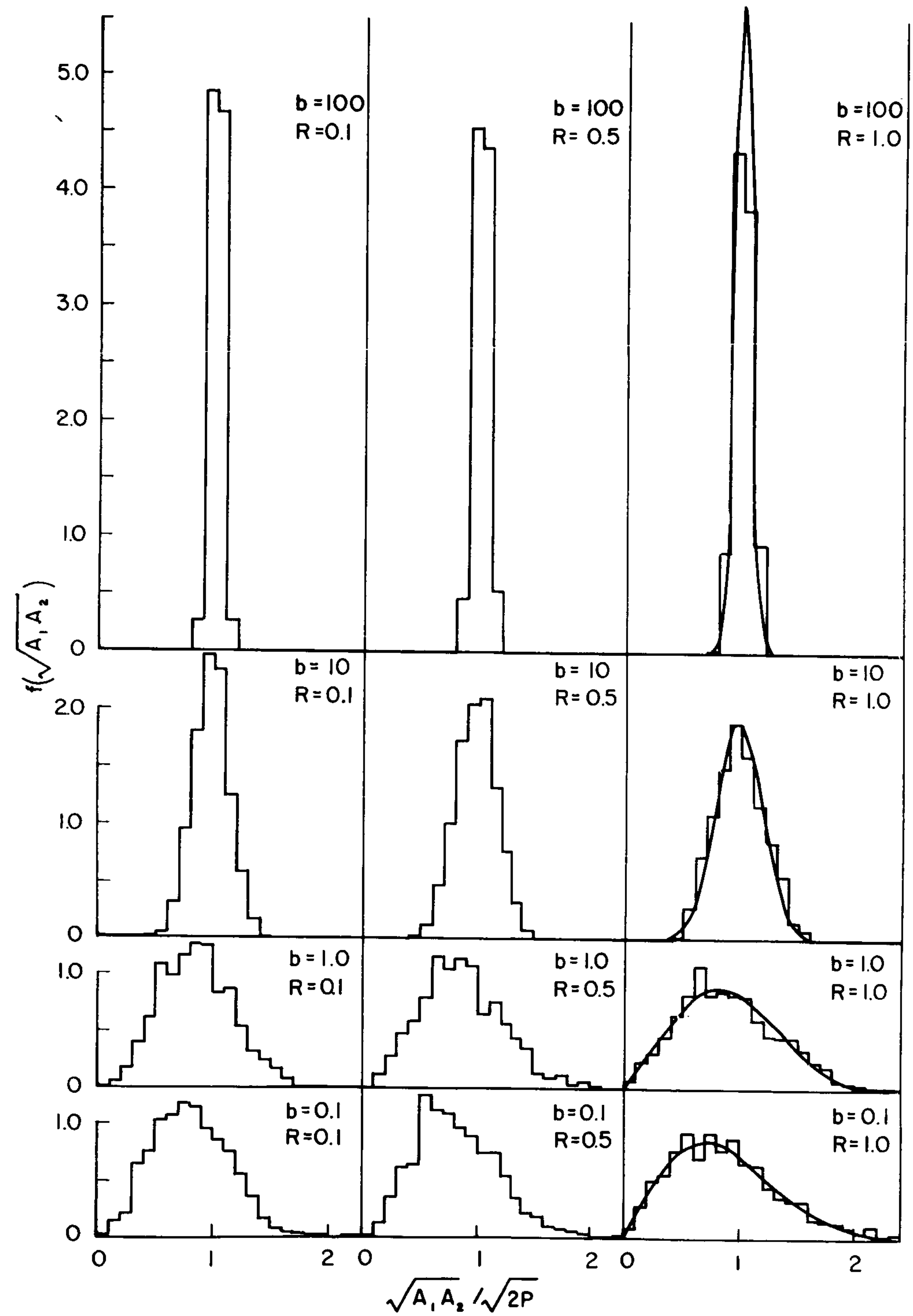



Fig. 5. Effective amplitude distribution for 12 combinations of coherence ratio, b, and wavefront correlation, R. Histograms computed numerically. Smooth curves computed from Rice distribution.

In section D2 it was pointed out that Rice's work on mathematical analysis of a sinusoidal signal buried in noise is directly applicable to observations of ionospheric scatter by single receivers. For the case of unity correlation, interferometric observations become identical with radiometric observations insofar as amplitude characteristics are concerned. Therefore, our distributions for unity correlation should agree with analytical results taken from Rice (1945). Under conditions of unity correlation, the output of a noncoherently detecting phase-sweep interferometer is simply proportional to received power and its square root is proportional to amplitude. The histograms of figure 5 under this condition must be consistent with the Rice distribution for amplitude, given in his equation 3.10-11. In our notation the Rice distribution is

$$f(A_1) = \frac{A_1}{\sigma^2} \exp\left[- \frac{A_1^2 + S^2}{2\,\sigma^2}\right] I_o\left(\frac{A_1 S}{\sigma^2}\right) \qquad 2\text{-}115$$

or

$$f(A_1) = \frac{(b+1)A_1}{P} \exp\left[ - b - \frac{(b+1){A_1}^2}{2P}\right] I_o\left(\frac{\sqrt{2b(b+1)}A_1}{\sqrt{P}}\right) \qquad 2\text{-}116$$

In the case $b = 0$, for which $P = \sigma^2$, the Rice distribution reduces to the Rayleigh distribution, given in equation 2-19.

The smooth curves given in the unity correlation column of figure 5 were calculated from equation 2-116. They show the degree of agreement between our numerically derived histograms and analytically derived distributions, where an analytical method is available. The fluctuations evident in the histograms resulted from the finite number (1000) of input data used in their generation. In computing observational quantities, up to 10,000 input values were used for each of the four random variables $B_{1c}$, $B_{1s}$, x, and y. In addition, some graphical smoothing was employed. Thus, the results to be discussed shortly are relatively free of statistical fluctuations. We shall find later that most of our experimental data contain considerably fewer independent input values. In order to estimate errors due to statistical fluctuations, it was considered necessary to reduce such fluctuations to negligible levels in the theoretical computations.

In section D2 we defined the observational quantity r, called visibility. This quantity is obtained from phase-switch (or coherently detecting phase-sweep) observations by taking the ratio of the average output under scatter conditions to that under undisturbed conditions. Under conditions of severe scatter - i.e., during a visibility fade - the quantity r is observed to decrease below its undisturbed value of unity. Equation 2-100 shows that a sufficient condition for zero visibility is $b = R = 0$. The condition $b = 0$, it will be recalled, represents the condition of complete ionospheric scatter.

It was mentioned earlier in the present section that some attempts have been made to interpret visibility fades by their effect on the average output of a noncoherently detecting phase-sweep interferometer. Where this has been attempted, the observed effect has been less than obvious (fig. 14, Little et al, 1962). Figure 6 indicates quantitatively that this is to be expected, as was inferred qualitatively from the distribution of figure 5. The curves of figure 6 give the predicted ratio, $\overline{A_1 A_2}/2P$, of the average output of a noncoherently detecting phase-sweep interferometer during a visibility fade to that under undisturbed

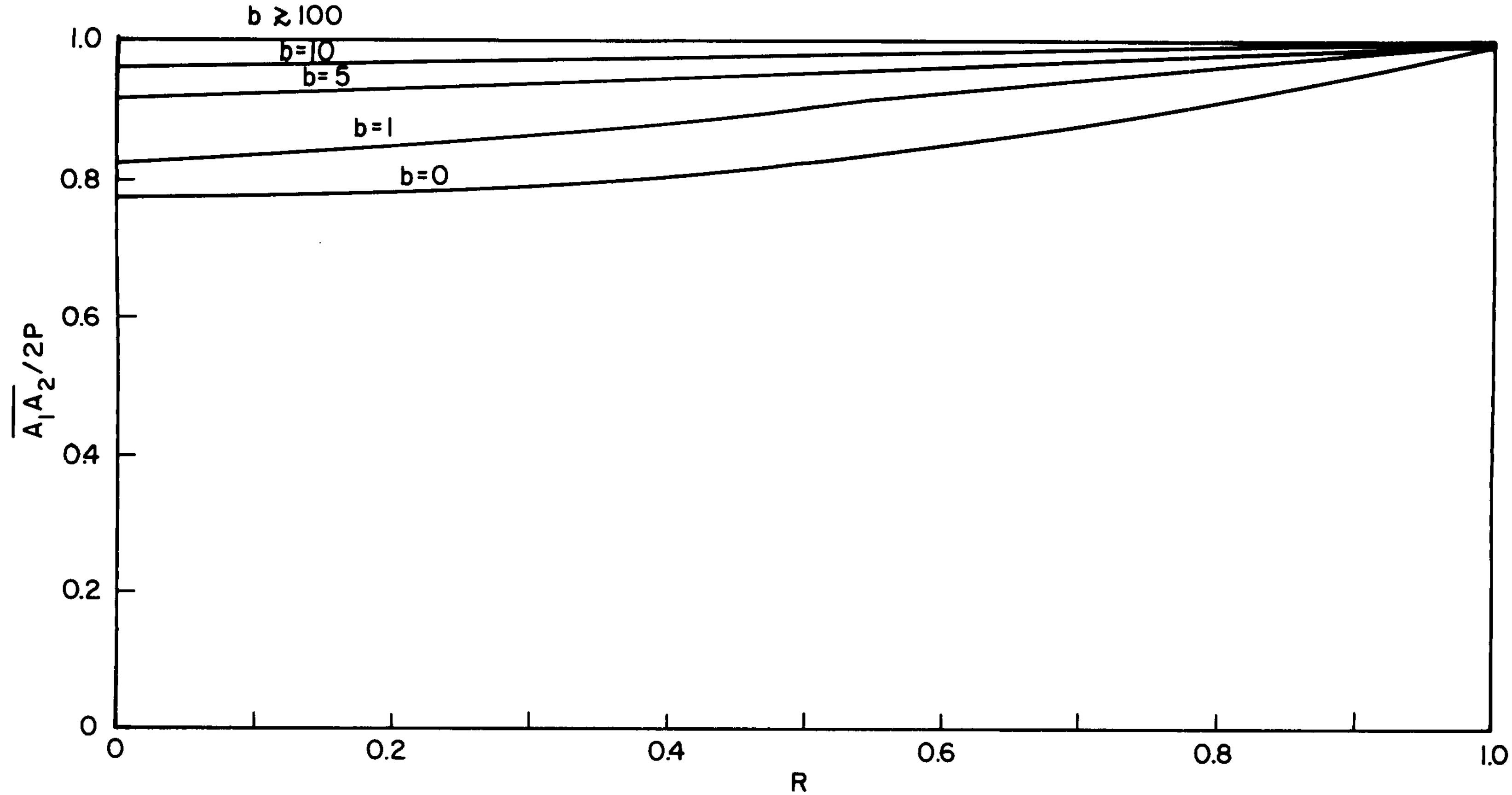



Fig. 6. Relative average output of a noncoherently detecting phase-sweep interferometer as a function of b and R.

ionospheric conditions. The ratio is given as a function of R and b. It is seen that a very severe visibility fade, one which produces the condition b = R = 0, results in only about a 22% decrease in the average output of a noncoherently detecting phase-sweep interferometer.

Whereas figure 6 displays the dependence of the average value of $A_1A_2$ on b and R, figures 7 and 8 display the dependence of the fluctuation of amplitudes on these two parameters. The ordinate in figure 7 is the fractional mean-square fluctuation of $\sqrt{A_1A_2}$, given by

$$\Delta_A = \frac{\overline{(\sqrt{A_1A_2} - \overline{\sqrt{A_1A_2}})^2}}{(\overline{\sqrt{A_1A_2}})^2} \qquad 2\text{-}117$$

When R is near unity, the amplitudes are essentially identical, and the above reduces to the fractional mean-square amplitude fluctuation given by

$$\Delta_A = \frac{\overline{(A_1-\overline{A}_1)^2}}{\overline{A}_1^2} = \frac{\overline{A_1^2} - \overline{A}_1^2}{\overline{A}_1^2} \qquad 2\text{-}118$$

Even at auroral latitudes, most scintillation observations are taken under conditions of near-unity wavefront correlation. The quantity $\Delta_A$, or an amplitude "scintillation index" based on it, is then a direct measure of the coherence ratio b, corresponding to the relationship between $\Delta_A$ and b existing at the right-hand edge of figure 7. Under visibility-fade conditions, however, when R is reduced, the relationship between $\Delta_A$ and b is altered. Under such conditions, the apparent scintillation index is depressed. For instance, for b = 1.0 and R = 0, figure 7 shows that the interferometrically observed value of $\Delta_A$, defined by equation 2-117, is 43% below the fractional mean-square amplitude fluctuation, defined by equation 2-118, which would be observed under the same conditions by a radiometer.

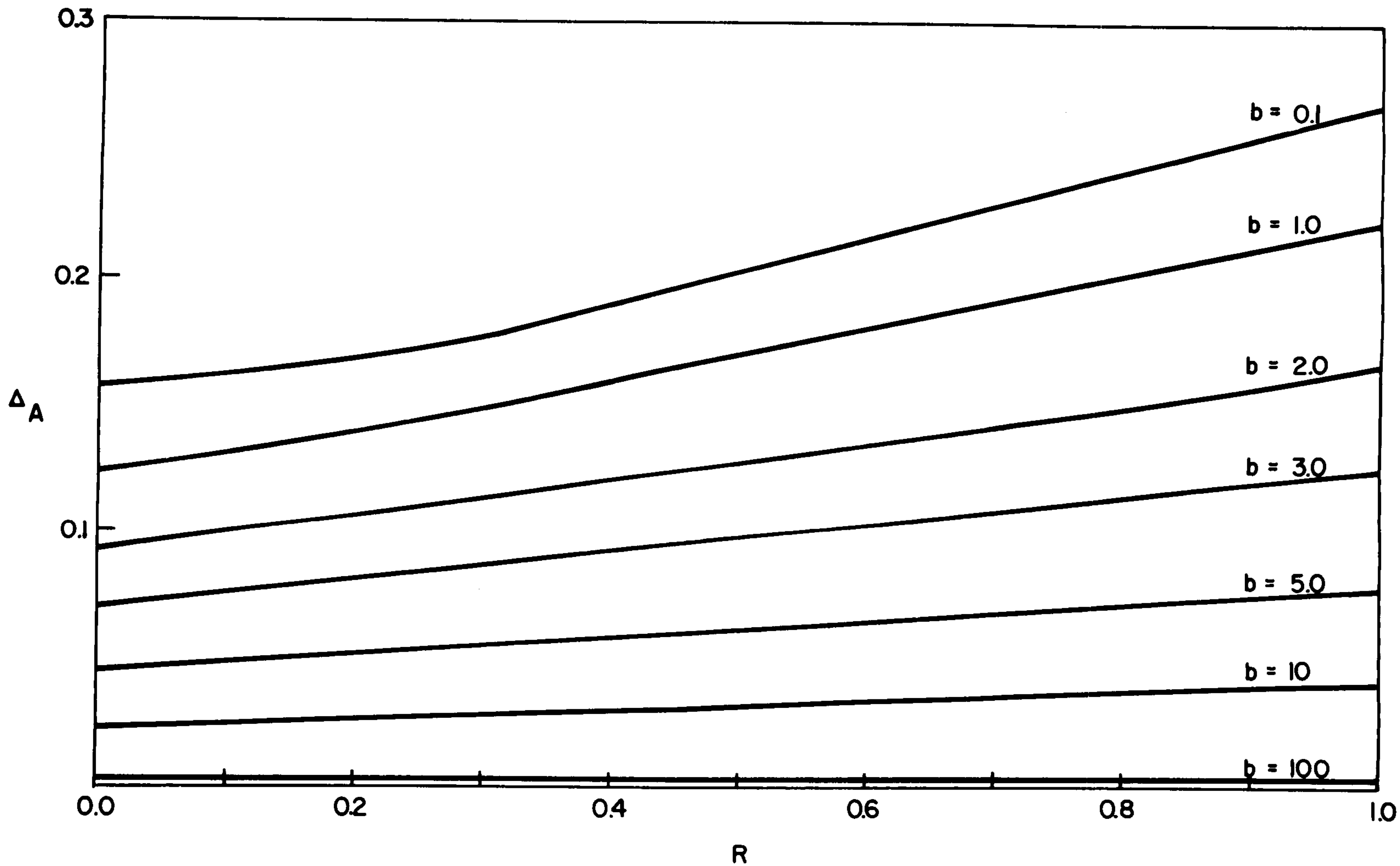



Fig. 7. Effective fractional mean-square amplitude fluctuation as a function of b and R.

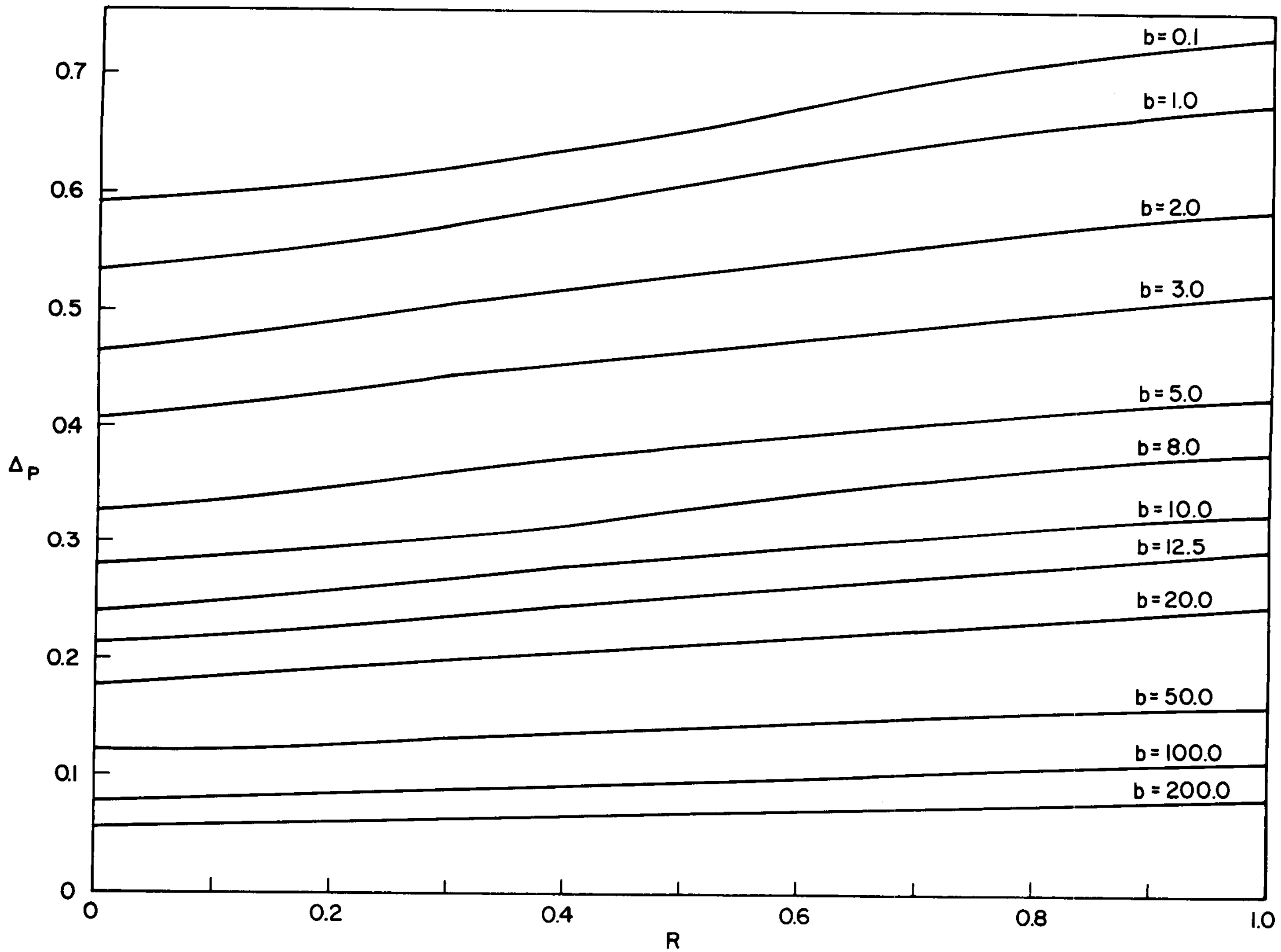



Fig. 8. Effective mean fractional power fluctuation as a function of b and R.

Another measure of amplitude scintillation sometimes used (Little, et al, 1962) is based on the mean fractional fluctuation in power received from the source. Under conditions of near-unity correlation, this latter quantity is given by

$$\Delta_P = \frac{\overline{|A_1^2 - \overline{A_1^2}|}}{\overline{A_1^2}} \qquad 2\text{-}119$$

Under visibilitv fade conditions, this generalizes to

$$\Delta_P = \frac{\overline{|A_1A_2 - \overline{A_1A_2}|}}{\overline{A_1A_2}} \qquad 2\text{-}120$$

The quantity given in 2-120 is plotted as a function of b and R in figure 8, reducing to that given in 2-119 at the right-hand edge. Again the effect of reduced correlation on the apparent fluctuation is evident. For b = 1.0, zero correlation produces a 21% reduction as compared with the unity-correlation value.

Let us turn now to the characteristics of phase as calculated by our numerical technique. The quantity of observational interest is the phase difference $\eta$. The most complete description of $\eta$, of course, is given by its distribution function $f(\eta)$. In figure 9, $f(\eta)$ is plotted for twelve representative combinations of b and R. It is obvious that the technique yields the results predicted by the work of Rice (1944, 1945), MacDonald (1949), and Bramley (1951) in the limiting cases. In particular, for low b and R, $f(\eta)$ approaches a uniform distribution at the value $1/2\pi$ and is approximated by a Gaussian function for large b. The technique also yielded a Dirac delta function at zero for unity correlation and any value of b, although this fact is not shown in the figure.

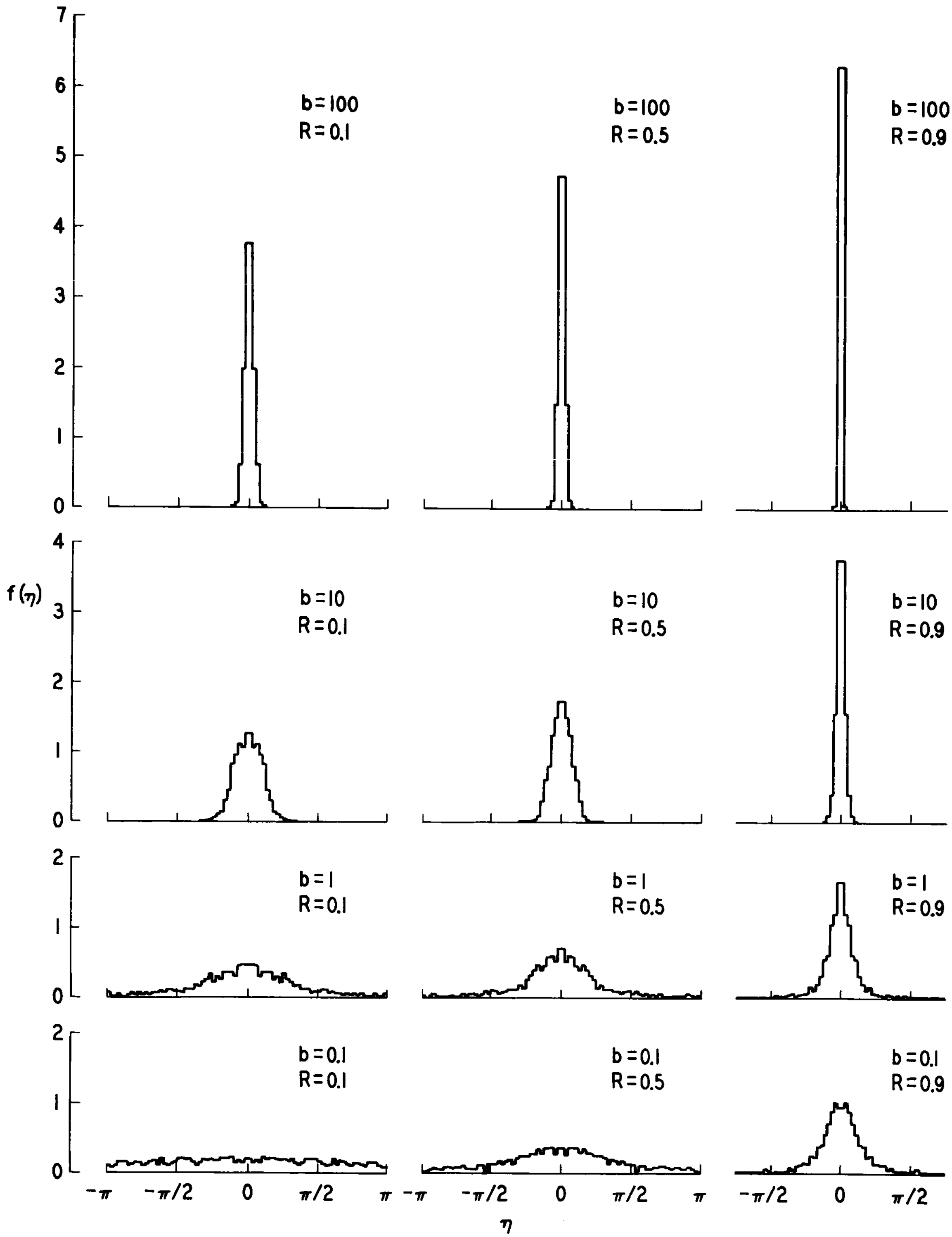



Fig. 9. Phase-difference distribution for 12 combinations of coherence ratio, b, and wavefront correlation, R.

It will be noted that for small b, the distribution departs from a delta function very quickly as the correlation R is decreased.

For actual use in observational analysis, what one desires rather than the full distribution of $\eta$ is some average quantity derived therefrom. One possible choice is the variance $\overline{\eta^2}$, which is shown plotted as a function of b and R in figure 10. It will be found that the curves of figure 10 approximate $\overline{\eta^2} = \frac{(1-R)}{b}$ for sufficiently large b, in agreement with Bramley's results for the limiting case (Bramley, 1951, section 3). For b smaller than about 5, significant departures from the approximation are found.

## IIE SUMMARY

The primary purpose of this chapter has been to provide a description of the observational results to be expected from interferometric observations of randomly scattered waves. The necessary concepts of the angular spectrum and the amplitude and phase of random signals were discussed briefly in section A. In section B, the relationship between the angular spectrum and the frequency spectrum of a signal received by a single antenna was explored and presented analytically for the case of a drifting but otherwise unchanging scattering layer. The discussion was extended to include a second antenna, thus introducing interferometric considerations, in section C.

The signal characteristics to be expected were developed and described in Section D, which was based on the assumption of random phasing in the angular spectrum. The meaning of this assumption as regards the received signal was explored in some detail in section D1a, which dealt with the signal statistics of single-antenna observations. The

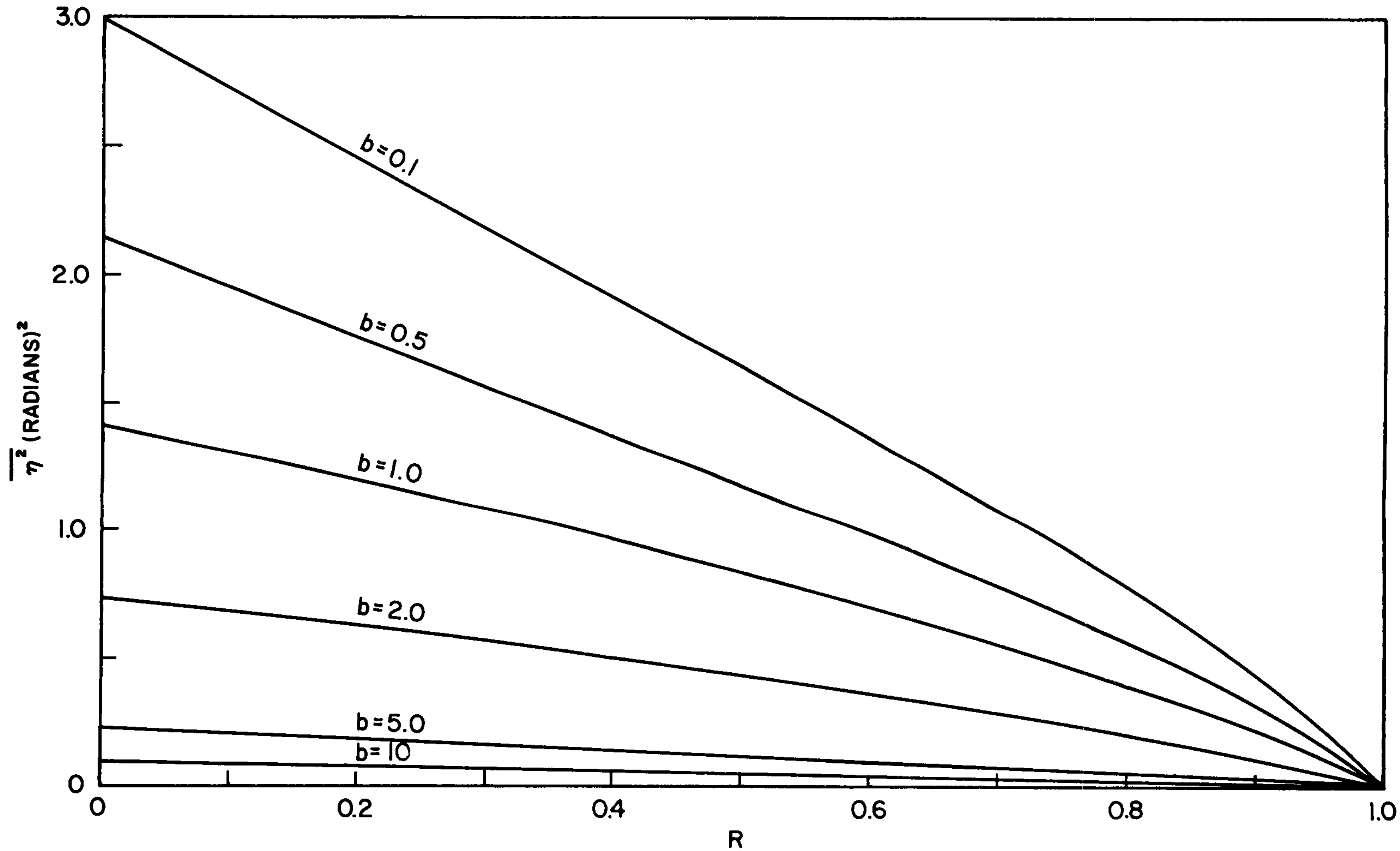



Fig. 10. Variance of the phase difference as a function of b and R.

discussion was extended to interferometric observations in section D1b. Sections D1 and D2 discussed the special cases of completely scattered and weakly scattered waves, respectively. The development followed closely the work of Bramley (1951), who applied the results of Rice (1944, 1945) to the problem of scattered-wave reception. The purpose of sections D1 and D2 was two-fold: to explore and relax slightly the assumption underlying Bramley's work and to lay the analytical framework for section D3.

The purpose of section D3 was to develop and present the signal statistics to be expected from random scattering for the general case of arbitrary degree of scattering. In section D3a, the coherence ratio b - being the ratio of nonscattered to scattered flux - was introduced as a measure of the degree of scattering. In section D3b, the covariance of voltages at two antennas of an interferometer was related to the coherence ratio and the wavefront correlation R, which was defined in section D1b. The observational parameter r, called the visibility, was related to the covariance of voltages and to b and R in section D3b. The resulting relationship, which is important to our experimental problem, was given in equation 2-100 as

$$r = \frac{b + R}{b + 1} \qquad \text{2-100}$$

The covariance of voltages and the visibility are of importance because they are directly observable by means of a phase-switch or coherently detecting phase-sweep interferometer. When noncoherent detection is employed with a phase-sweep interferometer, pure amplitude information can be obtained. Such an interferometer also offers a means of obtaining pure phase information. We shall see later that a combination of pure amplitude information with the complex information inherent in equation 2-100 allows measurement of certain ionospheric parameters related to b and R.

The graphical results of section D3e represent relations between b, R, and various statistical characteristics of the amplitude and phase of randomly scattered signals. The most complete description of the statistical characteristics is given in the density distribution histograms of figures 5 and 9. In our experimental problem, we shall use the information contained in the curves of figure 7, which relates to the fluctuation in the output of a noncoherently detecting phase-sweep interferometer as a function of b and R.

# CHAPTER III

# IONOSPHERIC PRODUCTION OF RANDOMLY SCATTERED WAVES

## IIIA REQUISITES OF A RANDOMLY PHASED ANGULAR SPECTRUM

In Chapter II we discussed the statistical characteristics of an ionospherically scattered signal as received by an interferometer at the ground. Our point of view was observational with little consideration given to the manner in which the observed characteristics are produced. In this chapter we shall turn our attention from conditions at the ground to conditions at the ionospheric scattering layer and relate the two. It will be recalled that the fundamental assumption upon which the work of Chapter II was founded was that of random phasing in the angular spectrum. Our first job in the present chapter is to explore the feasibility and consequences of this assumption.

The scattering effect we are considering arises from the differential phase shift imposed on a plane wavefront as it passes through an irregular ionospheric layer. At the base of the layer the effect can be described in terms of the distribution of phase across a plane. For the scattering we are considering - as opposed to refraction in individual ionospheric lenses - we take the distribution to be continuous outside our range of interest. It does not die off in the manner of a wave packet but continues in the manner either of a periodic function or of a random function.

If the phase distribution is periodic, it can be represented as a Fourier series. If it is random, it can be represented as a Fourier integral. In either case, the simplest configuration is the limiting one of a single Fourier component. More complicated periodic configurations then can be built up by adding harmonic components. Random configurations can be built by expanding the "bandwidth" limits on the Fourier integral. In the latter procedure, the resulting configuration is quasi-periodic for narrow fractional bandwidths in the manner of band-limited white noise. As the fractional bandwidth is increased, less and less ordered configurations result.

Let us begin by considering a simple cosinusoidal configuration of phase at the base of the scattering layer. That is, let the phase as a function of distance along one direction of a plane be

$$\theta(x) = \theta_o + \Delta\theta \cos(2\pi x/D) \qquad 3\text{-}1$$

In the above, $\theta_o$ represents the phase which the radio wave would have at the base of the layer if the layer had no irregular structure. $\Delta\theta$ represents the maximum phase deviation across the plane. The spatial period of the phase structure is D. $\theta_o$ is of little consequence, and we can reference our phase so that $\theta_o = 0$. In addition, we are not concerned with the amplitude of the radio wave at the base of the scattering layer. It is constant across the plane, and we shall take it as unity. With these simplifications, the "aperture distribution" describing the complex amplitude of the radio wavefront across the base plane is

$$V(x) = \exp[i\Delta\theta \cos(2\pi x/D)] \qquad 3\text{-}2$$

For the moment, let us consider the special case where $\Delta\theta$ is very small. In this case, equation 3-2 reduces approximately to

$$V(x) = 1 + i\Delta\theta \cos(2\pi x/D) \quad 3\text{-}3$$

Now the angular spectrum associated with V(x), expressed as a function of the sine s of the propagation angle, is given by the Fourier transform of V(x) if distances are measured in units of the radio wavelength $\lambda$ (Booker and Clemmow, 1950; Ratcliffe, 1956). The Fourier transforms of unity and cos $(2\pi x/D)$ are expressible in terms of the Dirac delta function. The former is a delta function at $s = 0$, and the latter is a pair of delta functions at $s = \pm 1/D$, or $s = \pm \lambda/D$ in unit measure rather than wavelength measure. Thus the angular spectrum of V(x) is made up of a nondeviated component wave of very nearly unit amplitude and two sidewaves of complex amplitude $i\Delta\theta/2$.

Ratcliffe (1956) has discussed the propagation of the angular spectrum described above. As the three waves propagate downward from the ionosphere, their relative phases change. The change in relative phase produces field distributions across lower planes which are different from that existing at the base of the scattering layer. The distributions may be thought of as Fresnel diffraction patterns. While the diffraction patterns differ at different levels, they are quite predictable owing to the small number of components in the angular spectrum and the simple phase relationship between them at all levels. The angular spectrum is far from being randomly phased. Under these simple conditions it is found that the diffraction patterns repeat themselves at periodic distances from the base of the scattering layer. Near the screen the field distributions are dominated by phase fluctuations, farther from it by amplitude fluctuations, still farther by phase fluctuations, and so forth.

If now the maximum phase deviation $\Delta\theta$ is allowed to be large, a more complicated angular spectrum is produced. The Fourier transform for the resulting waveform is well known in the theory of phase modulation. As pointed out by Ratcliffe (1956), the angular spectrum in this case is made up of discrete components at $s = \pm n\lambda/D$ where $n = 0, 1, 2, 3, - - -$. The amplitudes of the components are distributed as a series of Bessel functions of the first kind, with the amplitude of the $n^{th}$ component being given by $J_n(\Delta\theta)$. The number of components with appreciable amplitude is approximately $\Delta\theta$. The phases of components at the base plane are simply related, with the phase of the $n^{th}$ component being $n\pi/2$.

As the angular components propagate downward from the base of the scattering layer, again their relative phases change. If the normal to the base plane is taken as the z axis of the coordinate system and $\sin^{-1} s = \alpha$, then the propagation geometry shows that the phase of the $n^{th}$ component varies along the z axis as $\frac{2\pi z}{\lambda} \cos \alpha_n$. It can be shown that nowhere beneath the base plane do the phases return to their initial interrelationship. Thus, at no plane is the Fresnel diffraction pattern identical to the aperture distribution.

Hewish (1951) has calculated the amplitude and phase distributions produced at various distances from the scattering layer by phase modulation having various values of maximum phase deviation $\Delta\theta$. He found that phase deviations predominate near the layer, as in the case of very small $\Delta\theta$. At greater distances from the screen, amplitude deviations are built up. At still greater distances the strengths of amplitude and phase deviations remain roughly constant. The amplitude and phase deviations do not build and wane alternately as the distance is increased, as in the case of very small $\Delta\theta$.

The fact that the diffraction patterns do not show alternate phase and amplitude domination as a function of distance from the scattering layer is a consequence of the fact that the phases in the angular spectrum do not cyclicly repeat their interrelationship. This does not mean however, that, at any individual plane, the angular spectrum is randomly phased. At the base plane, all the phases are multiples of $2\pi$. Obviously random phasing does not exist there. At lower planes, the phase of the $n^{th}$ component is given by

$$\theta_n = \frac{n\pi}{2} + \frac{2\pi z}{\lambda} \cos \alpha_n \qquad 3\text{-}4$$

Due to the continuous variation of z, the second term above gradually distributes the component phases more uniformly in the range $\theta$ to $2\pi$. However, this is not enough to produce random phasing. The definition of random phasing requires also that the component phases be independent of one another. Equation 3-4 shows that this is not the case for the angular spectrum under consideration. The $n^{th}$ component phase bears a definite functional relationship to the $(n + 1)^{th}$ component phase, etc.

Suppose now that there is not just one sinusoidal component of phase structure in the wavefront at the base of the scattering layer, but many. Let the maximum radio phase deviation of the $m^{th}$ component be $\Delta\theta_m$ and its spatial phase be $\zeta_m$. Then, instead of equation 3-2, we have the following aperture distribution:

$$V(x) = \prod_{m=1}^{M} \exp \left\{ i\Delta\theta_m \cos [(2m\pi x/D) + \zeta_m] \right\} \qquad 3\text{-}5$$

where M is the number of spatial components in the phase structure of the wavefront and the symbol $\prod_{m=1}^{M}$ represents an M-fold product.

The Fourier components of such a distribution can be obtained by carrying out the multiplication on the n components arising from the simpler distribution of equation 3-2, after accounting for the spatial phases $\zeta_m$. This procedure produces the well-known components at $\pm nm\lambda/D$ and at sum and differences thereof. The amplitudes of the Fourier components are M-fold products of Bessel functions. If the $\zeta_m$ are taken to be zero, the phases of the Fourier components turn out to be integral multiples of $\pi/2$, as in the simpler case. The phases in the simple case arise as coefficients $(i)^n$ of the Bessel functions $J_n(\Delta\theta)$ in the Fourier series (Ratcliffe, 1956). For all the $\zeta_m$ equal to zero, the multiplication required by equation 3-5 simply adds integral exponents to these coefficients.

There is no reason that the $\zeta_m$ must be zero, however. Ratcliffe's equation 19, which gives the Fourier series representation of the simple distribution given in our equation 3-2, can be obtained by a method given by Starr (1953, Appendix 7). A generalization of Starr's procedure yields the following expansion for the distribution of equation 3-5:

$$V(x) = \prod_{m=1}^{M} \sum_{n_m=-\infty}^{\infty} e^{in_m (\frac{\pi}{2} + \zeta_m)} J_{n_m}(\Delta\theta_m) \exp \left(\frac{2\pi i m n_m x}{D}\right) \qquad 3\text{-}6$$

The individual Fourier components, which represent member waves of our angular spectrum with their phase measured at the point x = 0 in the base plane, are identified by a set of integers $n_m$. Thus

$$s = \sum_{m=1}^{M} mn_m\lambda/D \qquad n_m = 0, \pm 1, \pm 2, \text{ - - -} \tag{3-7}$$

The phases of the components are given by

$$\phi(s) = \sum_{m=1}^{M} n_m \left(\frac{\pi}{2} + \zeta_m\right) \tag{3-8}$$

It can be seen from equation 3-8 that any ordered relationship in the spatial phases $\zeta_m$ will result in an ordered - although possibly very complicated - relationship in the phases of the angular spectrum components. In this case, the angular spectrum is not randomly phased at the base plane. We have seen in the discussion of equation 3-4, that propagation downward from the ionosphere redistributes the phases in the angular spectrum but does not destroy correlation between them. Thus an angular spectrum corresponding to structure with ordered spatial phases will not be randomly phased at any plane. If, however, the spatial phases are independent, then the angular spectrum will be randomly phased at some distance beneath the base plane and beyond.

In order to achieve random phasing in the angular spectrum, then, we must have a number of spatial components in the radio phase distribution at the base plane, and the spatial phases must be independent of one another. They do not need to be uniformly distributed, however, since the phases of the angular spectrum are redistributed during subsequent propagation in such a manner as to spread them between zero and $2\pi$. It is to be noted also that the number of components in the angular spectrum is many times as great as the number of spatial components, except in the special case where the $\Delta\theta_m$ are very small. This is another "randomizing influence" on the angular spectrum.

In the above argument, we tacitly assumed that the components of spatial structure are harmonically related. This is of no consequence since we did nothing to restrict the fundamental spatial period. We are free to let it increase without limit so that we are dealing with Fourier integrals instead of Fourier series. In fact, to satisfy the random-phasing requirement that we have a large number of angular components, the basic period D must be large if s is to be restricted to physically meaningful values. Only angular components corresponding to s less than unity are capable of carrying radio energy away from the scattering layer. Greater values of s correspond to evanescent waves (Booker and Clemmow, 1950). A more stringent restriction on s was made in Chapter II, a restriction which is persistently verified observationally. We require s to be sufficiently small that it can be interchanged with its arcsine, $\alpha$.

Let us summarize the implications of assuming random phasing in the angular spectrum. First, the requirement of a large number of independent angular components precludes strict periodicity or other steady-state structure within our region of interest. That is, there must be a degree of randomness in the wavefront phase at the base of the scattering layer. In addition, the requirement of a uniform distribution of phase in the angular spectrum sets a minimum propagation distance between the base plane and our plane of observation for any given base-plane phase structure. These are the implications of the basic assumption of Chapter II. They are plausible but do not in themselves justify the assumption. The justification is observational, and we shall return to it in a later chapter.

It is to be noted that we have not specified the amplitudes of the angular spectrum components or the strengths $\Delta\theta_m$ of the spatial components of phase structure. The assumption of random phasing therefore is not seriously restricting from a physical point of view. It is true that we cannot deal with mathematically pure periodicity in ionospheric structure: a square-wave ionosphere or a sawtooth ionosphere, for instance. We can deal with quasi-periodic structure, however, in addition to completely random structure.

For a given degree of randomness in the structure, the condition of random phasing in the angular spectrum is more closely approximated the stronger the scatter, that is the larger the $\Delta\theta_m$. The distribution of $\Delta\theta$ within the spatial-frequency spectrum and the distribution of radio flux within the angular spectrum, however, are not specified. Thus the power spectrum of spatial structure in the wavefront at the base plane and the angular power spectrum still are quite arbitrary.

## IIIB STATISTICAL DESCRIPTION OF THE POST-SCATTERING WAVEFRONT AND THE MEANING OF VISIBILITY

Any assumption we might have made about either the amplitude or the phase of the angular spectrum would have reduced the amount of ionospheric information we can hope to glean. It would take an uncompromised measurement of both to permit a full description of the ionospheric structure which produced the spectrum. Having made the assumption of random phasing, therefore, we cannot hope to reconstruct fully the phase structure at the base of the scattering layer. A full reconstruction would be practically impossible anyway because the phases of the angular spectrum change during subsequent propagation. The amplitudes of the spectral components do not change, however.

The angular power spectrum contains all of the amplitude and none of the phase information of the (complex) angular spectrum. Since only the phases change, it follows that the angular power spectrum remains constant during propagation from the base of the ionosphere. Booker, Ratcliffe and Shinn (1950) defined a generalized autocorrelation function for complex amplitude as follows:

$$\rho(\xi) = \frac{\overline{V^*(x)\ V(x+\xi)}}{\overline{V^*(x)\ V(x)}} \qquad 3\text{-}9$$

They showed that if V(x) denotes complex amplitude across the wavefront, $\rho(\xi)$ is proportional to the Fourier transform of the associated angular power spectrum and is therefore invariant during post-scattering propagation. In the case of ionospheric scattering, then, the autocorrelation function given by 3-9 is the same at the ground as at the base of the scattering layer.

Now if $V(x) = A(x)\, e^{j\theta(x)}$ represents the complex amplitude of the wavefront across the ground (i.e., of the Fresnel diffraction pattern on the ground), then, aside from a constant of proportionality, the voltage at the output of an antenna located at x is

$$v_1 = \left\{ \text{Re} \quad V(x)\, e^{j\omega t} \right\} = A(x) \cos [\omega t + \theta(x)] \qquad \text{3-10}$$

Similarly, the voltage at the output of an antenna located at $x+\xi$ is

$$v_2 = \left\{ \text{Re} \quad V(x+\xi)\, e^{j\omega t} \right\} = A(x+\xi) \cos [\omega t + \theta(x+\xi)] \qquad \text{3-11}$$

Under conditions of ergodicity, we need not separate averages over t and over x, and we have

$$\overline{v_1 v_2} = \overline{A(x)A(x+\xi) \cos [\omega t + \theta(x)] \cos [\omega t + \theta(x+\xi)]}$$

$$= \tfrac{1}{2} \overline{A(x)A(x+\xi) \{\cos [\theta(x+\xi)-\theta(x)] + \cos [2\omega t+\theta(x+\xi)+\theta(x)]\}}$$

$$= \tfrac{1}{2} \overline{A(x)A(x+\xi) \cos [\theta(x+\xi)-\theta(x)]}$$

$$= \tfrac{1}{2} \overline{\text{Re} [V^*(x)V(x+\xi)]} \qquad \text{3-12}$$

But $\overline{v_1 v_2}$ is just the temporal covariance of the antenna voltages, which is obtained directly from the output of a phase-switch interferometer, and $\overline{\text{Re}[V^*(x)V(x+\xi)]}$ is the spatial autovariance of complex amplitude. It was shown in section D2 of Chapter II that the temporal covariance of voltages is equal to $[\tfrac{1}{2}S^2 + R\sigma^2] \cos 2\chi_o$, where $\tfrac{1}{2}S^2$ is the power in the nondeviated component of the angular spectrum and $\sigma^2$ is the power in the scatter spectrum. It is easily shown also, by letting $\xi$ go to zero in the derivation of equation 3-12 and then using Parseval's equality, that $\frac{1}{2}\ \overline{\mathrm{Re}V^*(x)V(x)} = P$ where P is the total power in the angular spectrum, $\frac{1}{2}S^2 + \sigma^2$. Thus, we have

$$\frac{\overline{V^*(x)V(x+\xi)}}{\overline{V^*(x)V(x)}} = \frac{S^2 + 2R\sigma^2}{S^2 + 2\sigma^2} \cos 2\chi_o = \frac{\overline{v_1 v_2}}{P} \qquad \text{3-13}$$

where the notation Re is to be understood in accordance with the usual convention.

The right side of equation 3-13 is simply the (temporal) cross-correlation coefficient for the antenna voltages, which we denoted by $\rho$ in Chapter II. Equation 3-13 shows that $\rho$, expressed as a function of antenna spacing $\xi$, is identical to the generalized spatial autocorrelation function defined by Booker, Ratcliffe, and Shinn. According to the conclusions of Booker, Ratcliffe, and Shinn, $\rho(\xi)$ measured on the ground-level diffraction pattern is identical to the autocorrelation function of the aperture distribution at the base of the ionospheric scattering layer.

As it stands, $\rho(\xi)$ contains the factor $\cos 2\chi_o$ which arises solely from the angular position of the source under observation. The position factor contains no ionospheric information regarding small-scale scatter. Let us, therefore, concern ourselves with the interferometer fringe visibility r, defined in section D3b of Chapter II, instead of with $\rho$. Comparison of equations 2-98 and 2-100 of that section shows that

$r(\xi) = \rho(\xi)/\cos 2\chi_o$. Thus we have, from equation 3-13,

$$r(\xi) = \frac{S^2 + 2R\sigma^2}{S^2+2\sigma^2} \qquad 3\text{-}14$$

Upon recalling the definition of the coherence ratio b, equation 3-14 reduces to equation 2-100, Thus

$$r(\xi) = \frac{b + R}{b + 1} \qquad 3\text{-}14'$$

Now $r(\xi)$ is the same at the ground as at the base of the scattering layer. Since the angular power spectrum is invariant during post-scattering propagation, the coherence ratio also remains constant. Therefore R must be identical at all planes too. Now S and $\sigma$, and therefore b, are not functions of antenna position and therefore are independent of $\xi$. R, on the other hand, is a measure of the correlation existing between the scatter-component resultants at the two antennas. The $\xi$-dependence of $r(\xi)$, therefore, arises through R, which we may write $R(\xi)$.

Let us return to consideration of equation 3-6, whose terms are components of the angular spectrum for a general aperture distribution $V(x)$[1]. The $0^{th}$ component, for which all the $n_m = 0$, is the non-deviated component of the angular spectrum. If all the $\Delta\theta_m$ are zero, corresponding to no phase fluctuations at the base plane, all power is contained in the nondeviated component, since all the Bessel functions except that of zero order have zero magnitude for zero argument. The higher-order waves, then, represent the scatter spectrum. They are responsible for the phase fluctuations at the base plane and for the

---

[1]Note that V(x) is not, in general, identical at the ground and at the base plane.

fluctuations in amplitude and phase at lower planes.

The wavefront consists of a plane wave identical to the non-deviated component, upon which are superposed spatial fluctuations due to the scattered part of the signal. $R(\xi)$ represents the spatial autocorrelation function of the fluctuating part of the wavefront. It was for this reason that we referred to R in Chapter II as the wavefront autocorrelation. The visibility $r(\xi)$ is essentially the spatial autocorrelation function of the composite wave-field.

## IIIC PHYSICAL MEANING OF THE PARAMETERS b AND R

### C1 The Coherence Ratio b

The above discussion shows that the interference fringe visibility depends upon the strength of the phase fluctuations at the base of the ionospheric scattering layer - through b - and on the wavefront autocorrelation function $R(\xi)$. We should like to relate b and R to ionospheric parameters. First let us relate b more explicity to the strength of the phase fluctuations. In order to simplify matters, we shall assume normal incidence of the nondeviated component relative to the ionosphere and to our antenna baseline at the ground. That is, we shall assume $\alpha_o = \chi_o = 0$, which will allow us to interchange $\rho(\xi)$ and $r(\xi) \cos 2\chi_o = r(\xi) \cos (\frac{2\pi d}{\lambda} \sin \alpha_o)$ and note that

$$\xi = d \cos \alpha_o \qquad 3\text{-}15$$

Equation 3-15 takes account of the fact that for off-normal incidence, the observing interferometer has a foreshortened baseline. That is $\xi$

is a measure of distance along the wavefront, which is equal to distance along the ground only for normal incidence.

The assumption of normal incidence here is not physically restricting, being only a convenient simplification of geometry which can be retracted at will by employing the mathematical relation between $\rho(\xi)$ and $r(\xi)$, together with equation 3-15. We shall now make a more basic assumption, akin to and consistent with that of random phasing in the angular spectrum. The assumption of random phasing requires that the phase of the wavefront at the base of the scattering layer be treated as a random variable. We shall need a distribution function with which to describe the wavefront phase. Following Bramley (1955), let us choose the normal distribution, which is reasonable on physical grounds.

The normal distribution in this instance is not demanded, a priori, by the central limit theorem since we know very little about the manner in which the wavefront phase fluctuations are built up. On the other hand, it does not seem unlikely that the fluctuations are built up by a large number of scatterings by independent ion-density irregularities. In this case, a normal distribution for the wavefront phase may be expected, whatever the origin and detailed distribution of the irregularities themselves.

It is to be noted that, while we are assuming a normal distribution for the wavefront phase at the base plane, we are making no restricting assumptions about the moments which describe either the first or the second-order distribution. In particular, we are putting no restrictions on the variance of the phase nor on its spatial autocorrelation function.

We shall, for convenience and without loss of generality, assume that the average, $\theta_o$, of the phase is zero.

Bramley (1955) has developed, under the assumptions mentioned above, the relationship between $r(\xi)$ and the variance and autocorrelation function of the phase distribution. The result is given in his equation 28, which in our notation is

$$r(\xi) = \exp\left[-\overline{\theta^2}\,(1-\rho_\theta)\right] \qquad 3\text{-}16$$

where $\overline{\theta^2}$ is the variance of the phase and $\rho_\theta$ is its autocorrelation function.

Now, according to the results of Booker, Ratcliffe and Shinn (1950), the Fourier transform of $r(\xi)$ is the normalized angular power spectrum, expressed as a function of $s = \sin\alpha$. For small angles we can replace s with $\alpha$, and we have

$$p(\alpha) = \frac{P}{\lambda}\int_{-\infty}^{\infty}\exp\left[-\overline{\theta^2}\,(1-\rho_\theta)\right]\exp\left(-j2\pi\alpha\xi/\lambda\right)d\xi \qquad 3\text{-}17$$

where $p(\alpha)$ represents the flux per unit angle in the spectrum, P is the total flux, and the wavelength $\lambda$ appears explicitly because we choose to use unit measure rather than wavelength measure. Equation 3-17 can be written also as

$$p(\alpha) = \frac{P}{\lambda}e^{-\overline{\theta^2}}\int_{-\infty}^{\infty}e^{\overline{\theta^2}\rho_\theta(\xi)}\exp\left(-j2\pi\alpha\xi/\lambda\right)d\xi \qquad 3\text{-}18$$

Upon expansion of the real exponential in the integrand into its series representation, the above becomes

$$p(\alpha) = \frac{P}{\lambda} e^{-\overline{\theta^2}} \int_{-\infty}^{\infty} \exp(-j2\pi\alpha\xi/\lambda)\, d\xi + \frac{P}{\lambda} e^{-\overline{\theta^2}} \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \frac{[\overline{\theta^2}\rho_\theta(\xi)]^n}{n!} \exp(-j2\pi\alpha\xi/\lambda) d\xi \qquad \text{3-19a}$$

The first integral results in a delta function at $\alpha=0$. Further, for any physically reasonable autocorrelation function, $\rho_\theta(\xi)$, the infinite series in the second term is uniformly convergent and therefore may be integrated term by term. Thus equation 3-19 becomes

$$p(\alpha) = P\, e^{-\overline{\theta^2}}\, \delta(\alpha-0) + \frac{P}{\lambda} e^{-\overline{\theta^2}} \sum_{n=1}^{\infty} \frac{\overline{\theta^2}^n}{n!} \int_{-\infty}^{\infty} [\rho_\theta(\xi)]^n \exp\,(-j2\pi\alpha\xi/\lambda)\, d\xi \qquad \text{3-19b}$$

Now the integral of $p(\alpha)$ over the whole angular spectrum must equal P. From the definition of the Dirac delta function, it is easily seen that the integral of the first term above is $P\, e^{-\overline{\theta^2}}$. The integral of the second term, then, must be $P(1-e^{-\overline{\theta^2}})$. Thus, the first term represents a wave travelling in the original propagation direction and carrying all of the flux in the event of no phase fluctuations ($\overline{\theta^2} = 0$). It is the nondeviated component of the angular spectrum. The second term carries whatever flux is not contained in the nondeviated component, for any degree of phase fluctuation across the base plane (i.e., for arbitrary $\overline{\theta^2}$). It is the scatter spectrum. The ratio of flux in the two terms, which is the coherence ratio, is given by

$$b = e^{-\overline{\theta^2}} \,/\, (1 - e^{-\overline{\theta^2}}) \qquad \text{3-20}$$

The above expression relating the coherence ratio to the variance of phase at the base of the scattering layer was derived by Bramley (1955) for the special case of a Gaussian autocorrelation function. We have shown here that no such restriction is required

C2 The Wavefront Autocorrelation Function $R(\xi)$

Equation 3-20 relates b to conditions at the base of the scattering layer. Let us turn now to the relationship between $R(\xi)$ and conditions at the base plane. It is easily seen from equation 3-20 that

$$b + R = \frac{e^{-\overline{\theta^2}} + R(1 - e^{-\overline{\theta^2}})}{1 - e^{-\overline{\theta^2}}} \qquad 3\text{-}21$$

and

$$b + 1 = 1/(1 - e^{-\overline{\theta^2}}) \qquad 3\text{-}22$$

Combination with equation 3-14 produces

$$r = e^{-\overline{\theta^2}} + R(1 - e^{-\overline{\theta^2}}) \qquad 3\text{-}23$$

Equating 3-16 with 3-23 results in

$$e^{-\overline{\theta^2}} \; e^{-\overline{\theta^2}\rho_\theta} = e^{-\overline{\theta^2}} + R(1 - e^{-\overline{\theta^2}}) \qquad 3\text{-}24$$

which produces, in view of 3-20

$$\overline{\theta^2} \; \rho_\theta = \ln\left(1 + \frac{R}{b}\right) \qquad 3\text{-}25$$

It is easily shown from equation 3-20 that

$$\overline{\theta^2} = \ln\left(1 + \frac{1}{b}\right) \qquad 3\text{-}26$$

Hence the autocorrelation function $\rho_\theta(\xi)$ of the base-plane phase distribution is given in terms of the wavefront autocorrelation function $R(\xi)$ by

$$\rho_\theta(\xi) = \frac{\ln\left[1 + \frac{R(\xi)}{b}\right]}{\ln\left[1 + \frac{1}{b}\right]} \qquad \text{3-27}$$

Equation 3-27 can be expressed as

$$\rho_\theta = \sum_{n=1}^{\infty} n^{-1} \left[\frac{R/b}{1 + (R/b)}\right]^n \Big/ \sum_{n=1}^{\infty} n^{-1} \left[\frac{1/b}{1 + (1/b)}\right]^n \qquad \text{3-28}$$

For large coherence ratio, b, the series can be approximated by their initial terms with the result that

$$\rho_\theta \approx \frac{1 + (1/b)}{1 + (R/b)} R \approx R \qquad \text{3-29}$$

Thus, for weak scattering, the wavefront autocorrelation function is nearly identical to the autocorrelation function of the phase distribution across the base plane. This result was obtained by Bramley (1955) for the special case of a Gaussian autocorrelation function. Again, the result holds for arbitrary autocorrelation function.

The general expression, equation 3-27, shows that for moderate and strong scattering (ie, for moderate and small coherence ratio), the phase-distribution autocorrelation function $\rho_\theta$ can be obtained only from a knowledge of both b and R. In Chapter IV, we shall describe an experiment for determining b and R, based on the development of Chapter II. Now we must turn to the problem of relating b and R, through their dependence on $\overline{\theta^2}$ and $\rho_\theta$, to parameters of ionospheric structure.

Again we shall base our procedure on the work of Bramley(1955), generalizing where possible.

IIID SCATTERING OF VHF WAVES BY ION-DENSITY IRREGULARITIES

D1 The Variance of Phase at the Base of the Scattering Layer

The operating frequencies in our experiment are well above the gyro, collision, and plasma frequencies of the auroral E and F layers. Under these conditions, the refractive index $\mu$ encountered by our observing wave of frequency $\omega/2\pi$ in passing through a region which may contain scattering irregularities can be obtained from

$$\mu^2 = 1 - (Ne^2/m\varepsilon_o\omega^2) \qquad 3\text{-}30$$

where N = the electron density in the region, e = the electronic charge, m = the electronic mass, and $\varepsilon_o$ = the permitivity of free space. In passing through an elemental thickness dz of such a region, the wave undergoes a phase shift of $(2\pi/\lambda)\mu dz$, where $\lambda$ is its free-space wavelength.

Let us suppose now that the region does in fact contain electron-density irregularities and that the electron density in the dz element under consideration is $\bar{N} + \Delta N$, where $\bar{N}$ is the average value of N in the region. The departure of the refractive index in the element from the average value in the region will be given by

$$\Delta\mu = \int_o^{\Delta N} (d\mu/dN)\, dN \qquad 3\text{-}31$$

From equation 3-30, $d\mu/dN$ is seen to be

$$\frac{d\mu}{dN} = \frac{-e^2}{2\mu m\varepsilon_o\omega^2} \qquad 3\text{-}32$$

If the deviation of N in the element is sufficiently small, $\mu$ may be taken as a constant when equations 3-31 and 3-32 are combined. If this is done, then $d\mu/dN$ can be placed outside the integral in 3-31, and obviously $(\Delta\mu/\Delta N) = (d\mu/dN)$.

Then we have, from 3-32

$$\Delta\mu = \frac{-e^2\,\Delta N}{2\mu m\varepsilon_o\omega^2} \qquad 3\text{-}34$$

The approximation involved in obtaining equation 3-34 has been tested for the observing conditions of our experiment. At the lowest frequency used in the majority of the observations (68 MHz), the approximation results in less than a 2% error in $\Delta\mu$ even for a 50% modulation in electron density in the most dense region of the ionosphere. Thus, even under disturbed conditions in the auroral zone, the approximation is considered acceptable. Furthermore, at the observing frequencies used, the ionospheric refractive index may be expected to depart from unity by at most a few percent. Accordingly, $\mu$ in equation 3-34 will be taken as unity.

As a result of the above considerations, it is easily seen that the magnitude of the excess phase shift $(2\pi/\lambda)\Delta\mu dz$ suffered by the wave because of the deviation in electron density in the dz element is

$$d\theta = \frac{\pi e^2\Delta N}{m\varepsilon_o\omega^2\lambda}\,dz \qquad 3\text{-}35$$

For purposes of quantitative calculation, it is convenient to express the constants in equation 3-35 in terms of the plasma frequency and average electron density of the irregular region. If we denote the plasma frequency by $f_o = \omega_o/2\pi$, then $\frac{d\theta}{dz}$ becomes

$$\frac{d\theta}{dz} = \frac{\pi}{\lambda} \frac{\omega_o^2}{\omega^2} \frac{\Delta N}{N} \qquad 3\text{-}36$$

Now $\Delta N$ represents the deviation from the mean of the electron density in an element of depth dz in the irregular region. All the other quantities in 3-36 are constants or relatively very slowly varying functions of z. Therefore, whatever the statistical distribution of electron density in the region, the variance $\overline{(d\theta/dz)^2}$ of $d\theta/dz$ is obtained from the variance $\overline{(\Delta N)^2}$ of the electron density as follows:

$$\overline{\left(\frac{d\theta}{dz}\right)^2} = \frac{\pi^2}{\lambda^2} \left(\frac{\omega_o^2}{\omega^2}\right)^2 \frac{\overline{(\Delta N)^2}}{\overline{N}^2} \qquad 3\text{-}37$$

The last factor in 3-37, of course, is the mean-square fractional fluctuation in electron density in the irregular region. Note that we have not assumed a distribution function for the electron density. It can be Gaussian as might be expected if the irregularities are produced by a large number of independent ionizing particles. On the other hand, it can be quite different as might result from some more ordered process.

As the wave propagates through the irregular region, the excess elemental phase shift $\frac{d\theta}{dz}(z)$ along the z axis (or along any other representative path of constant x and y) varies in proportion to $\Delta N(z)$.

The function $\frac{d\theta}{dz}(z)$ is a particular sample function of the random process represented by the ensemble of similar functions along all possible paths of constant x and y. If the total thickness of the irregular layer is t, then the phase at the intersection of the z axis and the bottom of the irregular layer is

$$\theta = \int_0^t \frac{d\theta}{dz}(z)\, dz \qquad 3\text{-}38$$

Hence $\theta/t$ is just the finite-interval spatial average of $d\theta/dz$.

In sampling theory, $\theta/t$ would be called the sample mean of $d\theta/dz$. The integral in 3-38 taken along some other path of constant x and y would, in general, be different from that taken along the z axis. If we now inspect the sample mean along the x axis, we obtain $\theta(x)/t$, where $\theta(x)$ is the spatially fluctuating wavefront phase. We can now obtain the variance of the wavefront phase from the well-known expression for the variance of the sample mean (Davenport and Root, section 5-3 and section 4-8). Thus

$$\overline{\theta^2} = 2t\ \overline{\left(\frac{d\theta}{dz}\right)^2} \int_0^t \left(1 - \frac{\tau}{t}\right) \rho_z(\tau)\, d\tau \qquad 3\text{-}39$$

where $\rho_z$ is the spatial autocorrelation function[1] of $d\theta/dz$ and therefore of the electron-density variation[2] in the z direction.

---

[1]In the present context, the term "autocorrelation function" refers to what is sometimes (inaccurately) called the "normalized autocorrelation function." The quantity termed "autocorrelation function" by Davenport and Root is equal to the product of our autocorrelation function and the variance and may aptly be called the autovariance function. Accordingly the variance of $d\theta/dz$ appears explicitly in our equation 3-39, whereas the corresponding quantity is contained implicitly in Davenport and Root's equations 4-83 and 5-13. There are several terms used interchangably by various authors to define what we are calling autocorrelation and to define closely related but not identical quantities. Care must be exercised to avoid confusion even in the case of ergodic processes, which we are considering. For nonergodic processes, still more inconsistencies in nomenclature arise.

[2]See page 133 for the mathematical definition of $\rho_z(\tau)$.

In accord with our assumption that the ionosphere's irregular structure is not strictly periodic, $\rho_z(\tau)$ must approach zero for very large values of $\tau$. For a thick layer in which t is very large compared with the largest value of $\tau$ for which $\rho_z(\tau)$ makes a significant contribution to the integral, equation 3-39 can be replaced by

$$\overline{\theta^2} = 2t \, \overline{\left(\frac{d\theta}{dz}\right)^2} \int_0^t \rho_z(\tau) \, d\tau \qquad \text{3-40}$$

D2 Phase Variance for a Scattering Layer with a Gaussian Structural Autocorrelation Function

The approximation involved in obtaining equation 3-40 has a somewhat different meaning depending upon whether the ion-density structure is strictly random or quasi-periodic in the z direction. In the former case, the layer simply must be several irregularities thick. In the latter case, the layer thickness must be great compared with the spatial wavelength of the basic periodicity and compared with the (larger) distance over which the periodicity is maintained with essentially constant phase. In both cases, the layer must be thick compared with what we might call "the correlation depth" of the scattering layer. In the random case, the correlation depth gives directly a measure of the average size of an irregularity. In the quasi-periodic case, the correlation depth may be many times greater than the size of a single irregularity, how much greater depending upon how well defined the periodicity is.

Bramley (1955) treated the random-layer problem in the special case where the autocorrelation function of the ion-density irregularities in the z direction is given by the ("denormalized") Gaussian function.

That is, he treated the case where

$$\rho_z = \exp\,(-\tau^2/\tau_o^{\,2}) \qquad \text{3-41}$$

The assumption of the so-called Gaussian autocorrelation function yields for the integral of equation 3-40

$$I = \frac{\sqrt{\pi}\ \tau_o}{2}\ \mathrm{erf}(t/\tau_o) \qquad \text{3-42}$$

where $\mathrm{erf}(t/\tau_o)$ stands for the error function, which has value zero for zero argument and tends to unity as t increases. Thus, for t much larger than $\tau_o$, the variance of phase fluctuations at the base of the scattering layer is given by

$$\theta^2 = \sqrt{\pi}\ \tau_o\ t\ \overline{\left(\frac{d\theta}{dz}\right)^2} \qquad \text{3-43 (Gaussian)}$$

In the above, $\tau_o$ represents the distance at which the autocorrelation function falls to $e^{-1}$, thus being a measure of the size of the irregularities.

D3 Phase Variance for other Autocorrelation Functions

Another autocorrelation function sometimes assumed in scattering problems is the exponential function, $\exp\,(-\tau/\tau_o)$. For the exponential case, I is given by $\tau_o(1 - e^{-t/\tau_o})$, and the phase variance for large t is

$$\overline{\theta^2} = 2\tau_o\ t\ \overline{\left(\frac{d\theta}{dz}\right)^2} \qquad \text{3-43 (Exponential)}$$

Let us investigate the phase variance for some other simple autocorrelation functions. One of the simplest of which we might conceive is the rectangular function, which has value unity up to $\tau=\tau_o$ and value zero for larger $\tau$. In this case, equation 3-40 produces

$$\overline{\theta^2} = 2\tau_o\, t\, \overline{(\frac{d\theta}{dz})^2}$$ 3-43 (Rectangular)

The rectangular autocorrelation function produces a result identical in form to that obtained from the exponential autocorrelation function. It must be remembered that the correlation depth $\tau_o$ is necessarily defined differently for the two cases, however. In the rectangular case, $\tau_o$ is the value of $\tau$ at which the autocorrelation function drops abruptly from unity to zero. In the exponential case, $\tau_o$ is the value at which the autocorrelation reaches $e^{-1}$.

Another simple autocorrelation function is that which drops linearly as $1 - \tau/\tau_o$ until $\tau = \tau_o$ and remains at zero for larger values of $\tau$. In this case, equation 3-40 produces

$$\overline{\theta^2} = \tau_o\, t\, \overline{(\frac{d\theta}{dz})^2}$$ 3-43 (Linear)

It is easily shown that autocorrelation functions which follow a positive integral power law of the form $1 - (\tau/\tau_o)^n$, so that they drop off more slowly than the linear function, produce

$$\overline{\theta^2} = \frac{2n}{n+1}\, \tau_o\, t\, \overline{(\frac{d\theta}{dz})^2}$$ 3-43 (Integral power

Similarly, positive fractional power law functions of the form $1 - (\tau/\tau_o)^{1/n}$, which drop off faster than the linear autocorrelation function, yield

$$\overline{\theta^2} = \frac{2}{n+1}\, \tau_o\, t\, \overline{(\frac{d\theta}{dz})^2}$$ 3-43 (Fractional power)

All of the autocorrelation functions considered above describe random scattering layers. Their Fourier transforms, which would represent the power spectrum of electron-density structure in the z direction, could be likened to the frequency response characteristics of low-pass filters. As one last example of a random-layer autocorrelation function, let us arbitrarily choose one-quarter of an ellipse. That is, let us choose

$$\rho_z = \tau_o^{-1} \sqrt{\tau_o^2 - \tau^2} \qquad 0 \le \tau \le \tau_o$$

$$\rho_z = 0 \qquad \tau > \tau_o \tag{3-44}$$

In this case, we have

$$I = \int_o^{\tau_o} \tau_o^{-1} \sqrt{\tau_o^2 - \tau^2}\, d\tau = \pi\tau_o/4 \tag{3-45}$$

and

$$\overline{\theta^2} = \frac{\pi}{2} \tau_o\, t\, \overline{\left(\frac{d\theta}{dz}\right)^2} \qquad \text{3-43 (Elliptical)}$$

A glance at the equations 3-43 shows that for all the autocorrelation functions considered, the variance of phase at the base of the scattering layer is given by

$$\overline{\theta^2} = k\, \tau_o\, t\, \overline{\left(\frac{d\theta}{dz}\right)^2} \tag{3-46}$$

where t represents the thickness of the scattering layer, $\tau_o$ the scale of irregularities in the layer, and k is a dimensionless constant determined by the detailed shape of the autocorrelation function. In most cases, k will be smaller the more sharply the autocorrelation function falls off at small values of $\tau$, although it depends also on the precise definition of $\tau_o$.

Let us turn now to an example of a quasi-periodic scattering layer. Such a layer is characterized by an oscillating autocorrelation function, whose oscillations die out rapidly or slowly depending on whether the periodicity is ill-defined or well-defined. The power spectrum of the electron-density structure could be likened to the response characteristic of a bandpass filter, centered on the frequency corresponding to the quasi-period. The envelope which defines the decay of the autocorrelation-function oscillation is the Fourier transform of the band-limiting function. The shape of the envelope is unimportant for our purpose, the feature of interest being the oscillatory nature of the autocorrelation function. Let us examine the case of a Gaussian envelope, so that the autocorrelation function is given by

$$\rho_z = \exp(-\tau^2/\tau_o^2) \cos (\tau/l_z) \tag{3-47}$$

In the above autocorrelation function, $\tau_o$ is again the correlation depth. The size of irregularities, however, is on the order of $l_z$, which is $(2\pi)^{-1}$ times the quasi-period of the structure.

To evaluate the integral in equation 3-40 for the autocorrelation function of equation 3-47, let us make the following changes in variable:

let $x = \tau/\tau_o$ and let $g = \tau_o/l_z$

Then the integral becomes

$$I = \tau_o \int_0^{t/\tau_o} \exp(-x^2) \cos gx \, dx \tag{3-48}$$

Using integral 313-6 of Grobner and Hofreiter (1961), we find that

$$I = \frac{\sqrt{\pi}\,\tau_o}{4e^{g^2/4}} \left\{ \left[\mathrm{erf}\left(\frac{t}{\tau_o} - \frac{ig}{2}\right) + \mathrm{erf}\left(\frac{t}{\tau_o} + \frac{ig}{2}\right)\right] - \left[\mathrm{erf}\left(-\frac{ig}{2}\right) + \mathrm{erf}\left(\frac{ig}{2}\right)\right] \right\} \qquad \text{3-49}$$

The error function is odd, so the second term in curly brackets is zero.

If the quasi-periodicity of the scattering layer in the z direction is too well defined, $\tau_o$ and g are large, and we must deal with complex arguments in the error functions of the first term in curly brackets. Further, if $\tau_o$ is allowed to become too large, only unrealistically large values of t safeguard the assumption we made in deriving equation 3-40. On the other hand, for weakly defined quasi-periodicity, the real parts of the error-function arguments dominate, and equation 3-49 becomes approximately

$$I = \frac{\sqrt{\pi}}{\alpha}\,\tau_o\, e^{-g^2/4}\, \mathrm{erf}(t/\tau_o) \qquad \text{3-50}$$

Thus, for sufficiently large t, we obtain

$$\overline{\theta^2} = \sqrt{\pi}\,\tau_o\, t\, e^{-g^2/4}\, \overline{\left(\frac{d\theta}{dz}\right)^2} \qquad \text{3-51}$$

Equation 3-51 is identical to equation 3-43 (Gaussian) except for addition of the factor $e^{-g^2/4}$. The ratio g of correlation depth to quasi-period is a measure of the degree of periodicity in the scattering layer structure. For well-defined quasi-periodicity g is large and $l_z$ represents the size of the dominant structure. For ill-defined quasi-periodicity g is small. Equation 3-51 approaches identity with equation 3-43 (Gaussian) as the structure approaches strict randomness, in which case $\tau_o$ becomes a statistical measure of the structural scale.

On the basis of the above discussion for Gaussian quasi-periodicity, it seems reasonable to generalize equation 3-46 to the following:

$$\overline{\theta^2} = kG\tau_o t \, \overline{\left(\frac{d\theta}{dz}\right)^2} \qquad 3\text{-}52$$

where G represents a factor which is a measure of the quasi-periodicity in the scattering layer's structure in the z direction. The factor G is unity for a strictly random layer, in which case $\tau_o$ is the statistical scale of irregularities. For weakly defined quasi-periodicity, G departs from unity and also depends upon the scale of the structure. The derivation of equation 3-52 does not hold for well-defined quasi-periodicity.

Substituting from equation 3-37, we have finally for the variance $\overline{\theta^2}$ of the phase at the base of a scattering layer whose autocorrelation function in the z direction may take a variety of forms

$$\overline{\theta^2} = kG \, \tau_o \, t \, \left(\frac{\pi}{\lambda}\right)^2 \left(\frac{\omega_o}{\omega}\right)^4 \frac{\overline{(\Delta N)^2}}{\overline{N}^2} \qquad 3\text{-}53$$

The special case of a Gaussian autocorrelation function, considered by Bramley (1955), is simply a case in which $kG = \sqrt{\pi}$.

D4 Relation Between the Structural Autocorrelation Function and the Phase-Distribution Autocorrelation Function

Equation 3-53 relates $\overline{\theta^2}$ to ionospheric parameters. We still must relate the autocorrelation function $\rho_\theta$ of the phase at the base of the scattering layer to ionospheric parameters. Bramley (1955) showed that for the special case of isotropic irregularities having a Gaussian autocorrelation function, $\rho_\theta$ also is Gaussian. Although Bramley's development contains an error[1], we shall show that $\rho_\theta$ is identical to the x-direction autocorrelation function $\rho_x$ of the scattering layer structure under rather general conditions. Bramley's special-case result follows correctly as an application of our general one.

First let us note the definition of $\rho_\theta$ under our choice of phase reference, which sets $\theta_o$ (the mean value of phase at the bottom of the layer) equal to zero. We have

$$\rho_\theta(\xi) = \overline{\theta(x)\theta(x+\xi)} / \overline{\theta^2} \qquad 3\text{-}54$$

where the bars denote averages over all values of x. From equation 3-38, it is seen that 3-54 can be written as

$$\rho_\theta(\xi) = \frac{1}{\overline{\theta^2}} \overline{\int_o^t \int_o^t \frac{d\theta}{dz_1}(x,z_1) \frac{d\theta}{dz_2}(x+\xi,z_2)\, dz_1\, dz_2} \qquad 3\text{-}55$$

Substituting from equation 3-36, we obtain

$$\rho_\theta(\xi) = \frac{1}{\overline{\theta^2}} \left(\frac{\pi}{\lambda}\right)^2 \left(\frac{\omega_o^2}{\omega^2}\right)^2 \frac{1}{\overline{N}^2} \overline{\int_o^t \int_o^t \Delta N(x,z_1)\, \Delta N(x+\xi,z_2)\, dz_1\, dz_2} \qquad 3\text{-}56$$

Now let us denote the double integral under the x-averaging bar as J. The geometry involved in evaluating J is shown in figure 11. The basic

---

[1] In going from his equation (13) to equation (14), Bramley incorrectly replaced a one-dimensional autovariance function with a two-dimensional autovariance function. His result (under a subsequent approximation) followed only as a result of the special transformation properties of the Gaussian function, and his procedure would not lead to a valid result in more general considerations.

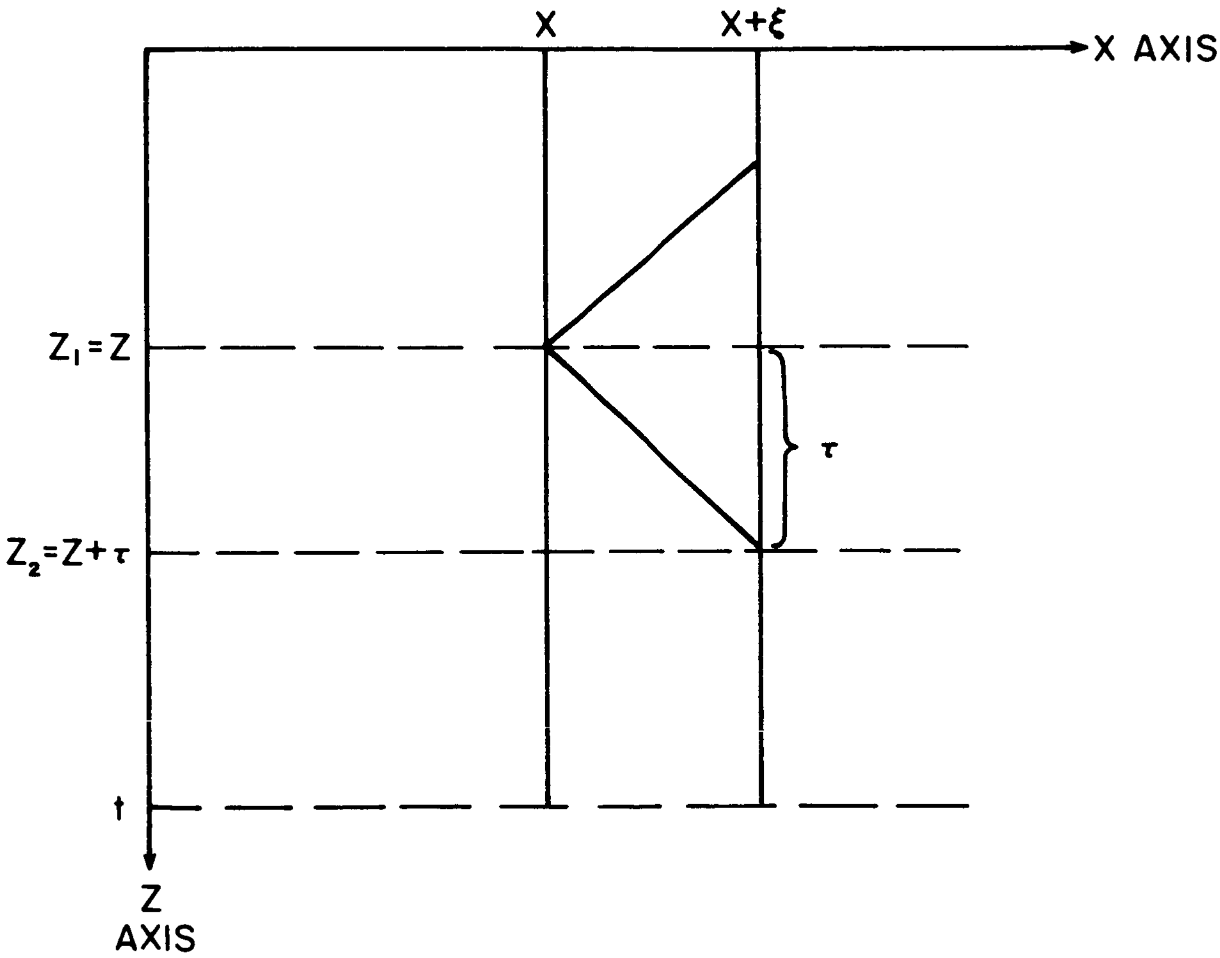



Fig. 11. Integration geometry for evaluating J.

coordinate system involved is indicated by the general x and z axes, with the origin at the upper left corner of the figure. The integrations are taken through the thickness t of the irregular scattering layer, along the z paths located at x and $x+\xi$. The integrand is the product of $\Delta N$ at each pair of points $z_1$ and $z_2$ located on the two paths. Now let us denote $z_1$ by z and $z_2$ by $z+\tau$. Then the double integration can be taken over z and $\tau$ as follows:

$$J = \int_0^t \left[ \int_0^{t-\tau} \Delta N(x,z)\Delta N(x+\xi,z+\tau)dz + \int_\tau^t \Delta N(x,z)\Delta N(x+\xi,z-\tau)dz \right] d\tau \qquad 3\text{-}57$$

Consider the first term in equation 3-57. The integrand is the product formed at the ends of the lower diagonal line in figure 11. The integral over z is obtained as the line moves down through the scattering layer. There is no contribution for z less than zero because $\Delta N(x,z)$ is then zero. There is no contribution for z greater than $t-\tau$ because $\Delta N(x+\xi,z+\tau)$ is then zero. The first term, with the limits as shown, thus contains the total contribution of products between the general point z on the path at x and points at a distance $\tau$ lower on the path at $x+\xi$. Similarly, the second term contains the total contribution of products between the general point z and points a distance $\tau$ above it. There is no contribution for z less than $\tau$ or z greater than t because in each case one end or the other of the product path (the upper diagonal line) is outside the irregular layer. Integrating over $\tau$ from zero to t collects all the non-zero contributions to conclude the evaluation of the double integral J.

A simple origin shift in the second term of equation 3-57 produces

$$J = \int_0^t \left[ \int_0^{t-\tau} \Delta N(x,z)\Delta N(x+\xi,z+\tau)\, dz + \int_0^{t\mp\tau} \Delta N(x,z+\tau)\Delta N(x+\xi,z)\, dz \right] d\tau \qquad 3\text{-}58$$

Now it is $\overline{J}$, the x-average of J, which we desire for use in equation 3-56. We are free to interchange the order of averaging over x and integrating over z and $\tau$ since the limits and integration variable in the averaging process are independent of those in the explicit integrations. We have, therefore,

$$\overline{J} = \int_0^t \left[ \int_0^{t-\tau} \overline{\Delta N(x,z)\Delta N(x+\xi,z+\tau)}\, dz + \int_0^{t-\tau} \overline{\Delta N(x,z+\tau)\Delta N(x+\xi,z)}\, dz \right] d\tau \qquad 3\text{-}59$$

Aside from a factor of t-$\tau$, the inner integrals of equation 3-59 each represent an average over x and z of the product of $\Delta N$ at points in the xz plane separated by a distance $\xi$ in the x direction and $\tau$ in the z direction. If the range of z averaging is sufficiently large, then, each of the inner integrals is the two-dimensional autovariance function of the electron density N in the scattering layer. Recalling that the variance of N in the layer is $\overline{(\Delta N)^2}$ and denoting the two-dimensional spatial autocorrelation function of N by $\rho_N(\xi,\tau)$ we find that $\overline{J}$ becomes

$$\overline{J} = \overline{(\Delta N)^2} \int_0^t 2\,(t-\tau)\, \rho_N(\xi,\tau)\, d\tau = 2t\, \overline{(\Delta N)^2} \int_0^t (1-\frac{\tau}{t})\rho_N(\xi,\tau)\, d\tau \qquad 3\text{-}60$$

Let us assume now that $\rho_N$ can be expressed as the product of two independent but otherwise arbitrary one-dimensional autocorrelation functions $\rho_x(\xi)$ and $\rho_z(\tau)$, where

$$\rho_x(\xi) = \frac{\overline{\Delta N(x,z)\Delta N(x+\xi,z)}}{\overline{[\Delta N(x,z)]^2}}$$

and

$$\rho_z(\tau) = \frac{\overline{\Delta N(x,z)\Delta N(x,z+\tau)}}{\overline{[\Delta N(x,z)]^2}}$$

Then, equation 3-60 can be written as

$$\overline{J} = [2t\ \overline{(\Delta N)^2} \int_0^t (1 - \frac{\tau}{t})\ \rho_z\ (\tau)\ d\tau]\ \rho_x(\xi) \qquad 3\text{-}61$$

But the integral in equation 3-61 is just that which appears in equation 3-39. Hence $\overline{J}$ is given by

$$\overline{J} = [\overline{\theta^2}\ \overline{(\Delta N)^2}\ /\ \overline{(\frac{d\theta}{dz})^2}]\ \rho_x(\xi) \qquad 3\text{-}63$$

Let us now replace the averaged double integral in equation 3-56 with $\overline{J}$ from 3-62 and substitute the right-hand side of equation 3-37 for $\overline{(\frac{d\theta}{dz})^2}$. Upon doing so, we obtain

$$\rho_\theta(\xi) = \rho_x(\xi) \qquad 3\text{-}63$$

Equation 3-63 states that the spatial autocorrelation function of the phase distribution at the base of the scattering layer is identical to the parallel component of the autocorrelation function of structure in the scattering layer. The prime assumption upon which this simple result is based is that the ion-density autocorrelation function can be expressed as the product of mutually independent component functions in the directions of the coordinate system. The assumption does not appear very restrictive. We have considered only two dimensions for simplicity and because the experiment we shall describe in Chapter IV is not dependent upon conditions in the third dimension (and, of course, cannot therefore yield information concerning structure in the third dimension). Generalization to three dimensions is straight-forward.

Along the route to equation 3-63, we made another incidental assumption, that of a "sufficiently large" range of z integration. This assumption means that the layer thickness t is large compared with the largest value of $\tau$ for which there is significant contribution to the integrals. This is the same assumption made earlier in approximating equation 3-39 by equation 3-40. Consequently, we could make the same approximation in equation 3-61. Nothing would be gained however, and our result would be the same.

The assumption that the layer is thick compared with the z-direction correlation distance (which we termed the "correlation depth" earlier) also means that our result does not depend explicitly on the upper limit placed on the $\tau$ integration in evaluating J. The careful reader may have felt uneasy about the choice of t as a limit since there is no a priori reason for this choice contained in the geometry. The justification lies in the thick-layer assumption. Any limit equal to or greater than t would produce the same result, as it would in equation 3-40, since the contribution outside the range 0 to t is negligible under the assumption.

## IIIE SUMMARY AND DISCUSSION

### E1 Assumptions and Results

Let us summarize the important results of this chapter. First, it was shown that our assumption in Chapter II that we are dealing with a randomly phased angular spectrum means that we must exclude certain simple kinds of ionospheric structure from consideration. In particular, we are not here considering refraction in one or a few ionospheric lenses or scattering by strictly periodic ionospheric structure. Although both of these problems can be attacked using the concept of the angular spectrum, the assumption of random phasing excludes them. The simple refraction problem is more conveniently attacked from the ray point of view anyway. Strict periodicity of ionospheric structure is not likely in practice, and quasi-periodicity is not excluded by the assumption of random phasing.

Another assumption, which we made in the present chapter, is that the scattering layer under consideration is thick compared with the correlation depth of its structure. For purely random structure the thick-layer assumption means that the layer is several irregularities deep. For quasi-periodic structure along the radio line of sight, the layer must be deep also compared with the distance over which the quasi-periodicity is sustained with approximately constant phase. Any given layer thickness therefore sets a limit on the degree to which the quasi-periodicity can be defined.

A third assumption, also made in the present chapter, is that the phase at the base of the scattering layer has a normal distribution, such as might result from a large number of independent scatterings. Under the second and third assumptions and two others of less physical importance, the following major results were obtained:

$$\overline{\theta^2} = \ln\left(\frac{1}{b} + 1\right) \qquad \text{(from equation 3-20)} \qquad 3\text{-}64$$

$$\rho_\theta(\xi) = \frac{\ln\,[1 + \frac{R(\xi)}{b}\,]}{\ln\,[1 + \frac{1}{b}]} \qquad \text{(from equation 3-27)} \qquad 3\text{-}65$$

$$\overline{\theta^2} = K\,\tau_o\,t\,\frac{\pi^2}{c^2}\,\frac{f_o^4}{f^2}\,\frac{\overline{(\Delta N)^2}}{\bar{N}^2} \qquad \text{(from equation 3-53)} \qquad 3\text{-}66$$

(where c = velocity of light and K = kG)

$$\rho_\theta(\xi) = \rho_x(\xi) \qquad \text{(from equation 3-63)} \qquad 3\text{-}67$$

The first two expressions above relate the variance $\overline{\theta^2}$ and autocorrelation function $\rho_\theta(\xi)$ of the phase distribution at the base of the ionospheric scattering layer to the observational quantities b and $R(\xi)$. The second pair of equations relates $\overline{\theta^2}$ and $\rho_\theta(\xi)$ to ionospheric parameters. Thus, the set of four equations provides a route to evaluation of certain ionospheric parameters from interferometric observations at the ground.

Each of the above four equations either is identical to a corresponding result given by Bramley (1955) for the special case of isotropic ionospheric structure having a Gaussian autocorrelation function or is readily reduced thereto. We have not, however, assumed isotropy or a Gaussian autocorrelation function in the present work. In addition, there is no restriction on the magnitude of the coherence ratio b or the variance $\overline{\theta^2}$ of the base-plane phase. Therefore, the results hold for strong as well as weak scatter.

## E2 Ionospheric Optical Thickness and Scattering Coefficient

Bramley (1954) pointed out that the analysis contained in his 1955 paper (carried out earlier but delayed relatively in press) afforded a means of relating the supposed thin diffracting screen of Booker, Ratcliffe and Shinn (1950) and other workers to a more realistic thick scattering layer. The terms "strong" and "weak" scatter which we have used from time to time above are directly applicable only to the equivalent thin diffracting screen. It is not to be supposed that a thin layer of the terrestrial ionosphere is capable of strong single scatter of waves in the frequency range we are considering.[1] It would take a more highly ionized or denser medium to do so. A succession of weak scatterings in a thick layer, however, can result in a decreased coherence ratio and produce a resultant field identical to that produced by a single strong scattering in the equivalent thin screen.

Fejer (1953) specifically analyzed the process of multiple weak scatterings - ie, scatterings for each of which the Born approximation[2] holds - in a thick irregular medium. In his analysis, Fejer assumed an isotropic scattering layer with a Gaussian autocorrelation function. Bramley (1954) demonstrated the equivalence of his own and Fejer's results for this special case, insofar as the final resultant field is concerned. Owren (1962) extended Fejer's analysis to a nonisotropic Gaussian scattering layer and showed that the procedure amounts to a technique for solving the equation of radiative transfer, treated more generally by Chandrasekhar (1960).

---

[1] A possible exception is the small number of observations at 26.3 MHz.

[2] Also known as Born's first approximation and stating that the scatter field is weak compared with the incident field.

The radiative transfer approach has the great conceptual advantage of allowing description of the wave-field within the scattering medium itself. On the other hand, when one can observe only the resultant wave after emergence from the medium - as is the case in most ionospheric experiments, including ours - the relative simplicity of Bramley's approach renders it the more useful. Thus, for the most part, we shall use Bramley's results as generalized herein.

Based on the demonstrated equivalence of Bramley's and Fejer's results in the special case of a Gaussian autocorrelation function, however, we shall freely use the descriptive terminology employed by Fejer and by Owren. In particular, for all of our VHF observations (where the Born approximation may safely be assumed to hold), we shall take a measurable decrease in coherence ratio to indicate the occurrence of multiple scatter.

Now Fejer (1953) and Bramley (1954) agreed that the "effective depth of scattering" of the thick layer is equal to the variance $\overline{\theta^2}$ of phase at the base of the layer. It is evident from the work of Owren (1962) that Fejer's "effective depth of scattering" is identical to the optical thickness of a purely scattering layer. Thus $\overline{\theta^2}$ represents the optical thickness of our scattering layer. The plausibility of this result for a general autocorrelation function can be seen by comparing the integral over angles of our equation 3-19b with the formal solution of the equation of radiative transfer, as given for instance by Chandrasekhar (1960) in equation 50 of his Chapter I.[1]

[1] The source function $\tilde{\mathcal{J}}$ (s') which appears in the integral of Chandrasekhar's expression is given, for a scattering atmosphere, in his equation 41. The latter equation involves an integral over all (scatter) angles of the "phase function" of the scattering process, which is essentially our angular power spectrum. An arbitrary phase function or angular power spectrum implies an arbitrary spatial autocorrelation function.

Equation 3-64 shows that determination of the coherence ratio b measured at the ground allows direct calculation of the optical thickness of the ionospheric scattering layer. The combination of equations 3-65 and 3-67 shows that the additional measurement of the wavefront auto-correlation function $R(\xi)$ at the ground allows calculation of the spatial autocorrelation function of the scattering layer in the direction parallel to the interferometer baseline.

Equation 3-66 relates the optical thickness to other physical parameters of the scattering layer. The optical thickness of the layer can be defined as

$$\overline{\theta^2} = \int_0^t \iota \, dz \qquad 3\text{-}68$$

where $\iota$ is the (linear) scattering coefficient of the layer, defined as the flux scattered per unit path length from an incident beam of unit flux density.[1] If the scattering coefficient is constant through the geomagnetic thickness of the layer, then obviously the optical thickness is simply the product of the scattering coefficient and the geometric thickness. Comparison with equation 3-66, then, shows that the scattering coefficient of the layer is given by

$$\iota = K \tau_o \frac{\pi^2}{c^2} \frac{f_o^4}{f^2} \frac{\overline{(\Delta N)^2}}{\overline{N}^2} \qquad 3\text{-}69$$

[1] The equivalent quantity in Chandrasekhar's work is $\kappa\rho$ where $\kappa$ is the mass scattering coefficient and $\rho$ is the mass density of the scattering constituent of the atmosphere. (See equation 51 of his Chapter I.)

The assumption that $\iota$ is constant holds under our tacit assumption that the scattering layer is statistically uniform (ie, displays spatial stationarity) in the z direction, which ensures that all the quantities on the right of equation 3-69 can be treated as constants. Moderate departures from the assumed condition mean that equation 3-69 must be interpreted as giving a weighted average scattering coefficient for the layer.

The scattering coefficient given by equation 3-69 is analagous to the absorption coefficient of the magneto-ionic theory (Ratcliffe, 1959, section 4.4). The absorption coefficient determines the exponential rate of attenuation suffered by a wave in traversing an absorbing region. The scattering coefficient determines the exponential rate of attenuation suffered by the nondeviated component of a wave in travelling through a scattering region. In the case of absorption, electromagnetic energy from the wave-field is transformed into heating of the medium by collisions, decreasing the strength of the wave. In the scattering case, energy is transferred from the nondeviated component into the scatter spectrum, decreasing the coherence ratio while maintaining constant total wave-field energy.

Reinserting the explicit expression for critical frequency into equation 3-69 produces

$$\iota = \left(\frac{e^2}{4\pi c \epsilon_o m}\right)^2 f^{-2} K \tau_o \overline{(\Delta N)^2} \qquad 3\text{-}70$$

Equation 3-70 shows that the scattering coefficient displays a wavelength-squared frequency dependence and depends on the structure of the scattering layer through K, $\tau_o$ and $\overline{(\Delta N)^2}$. It will be recalled that K is a dimensionless constant determined by the line-of-sight spatial autocorrelation function of the layer. It was developed as the product kG, where k depends upon the overall shape of the autocorrelation function (the envelope in the quasi-periodic case) and G depends upon the degree to which quasi-periodicity is developed, reducing to unity for a strictly random layer. The correlation depth of the layer is given by $\tau_o$, which reduces to the statistical scale of irregularities in the line-of-sight direction for the case of a strictly random layer. The final factor, of course, gives the strength of the irregularities as the variance of electron density in the layer. The elements of the first factor have their usual meaning in rationalized MKS units, so that the factor has the numerical value $4.53 \times 10^{17}$ meter$^4$/second$^2$.

APPENDIX I

## APPENDIX 1

## SOME CALCULATIONS PERTAINING TO SIGNAL STATISTICS

### 1a Statistical Independence of Real and Imaginary Components of the Complex Amplitude of Antenna Voltage

Let $v_1$ represent the varying antenna voltage arising from observation of a randomly phased angular spectrum, and let $V_1$ represent its complex amplitude. Then the real and imaginary components of $V_1$, $A_{1c}$ and $A_{1s}$ respectively, are given by

$$v_1 = \text{Re}\left\{V_1 e^{i\omega t}\right\} = A_{1c} \cos \omega t - A_{1s} \sin \omega t \tag{1}$$

Now let $v_1$ be Fourier analyzed. Thus

$$v_1 = \sum_{n=1}^{N} c_n \cos (\omega t + \gamma_{1n}) = \cos \omega t \sum_{n=1}^{N} (c_n \cos \gamma_{1n}) - \sin \omega t \sum_{n=1}^{N} (c_n \sin \gamma_{1n}) \tag{2}$$

For equations (1) and (2) to hold simultaneously for all t, we must have

$$A_{1c} = \sum_{n=1}^{N} c_n \cos \gamma_{1n} \quad \text{and} \quad A_{1s} = \sum_{n=1}^{N} c_n \sin \gamma_{1n} \tag{3}$$

Since, in any observational situation, N is restricted to finite values, we are free to multiply the sums in equations (3) term by term. Thus

$$A_{1c}A_{1s} = \sum_{n=1}^{N} c_m c_n \cos \gamma_{1m} \sin \gamma_{1n} \tag{4}$$

Since the average of any sum of stochastic variables is equal to the sum of the averages (cf, Munroe, 1951, p. 104-5), we can write the following, where the bars denote averages[1]

$$\overline{A_{1c}A_{1s}} = \sum_{\substack{m=1\\n=1}}^{N} c_m c_n \, \overline{\cos \gamma_{1m} \sin \gamma_{1n}} \qquad (5)$$

For a randomly phased angular spectrum, $\gamma_{1m}$ and $\gamma_{1n}$ are independent and uniformly distributed between zero and $2\pi$. Under these conditions the average on the right of equation (5) is zero. Thus we have

$$\overline{A_{1c}A_{1s}} = 0 \qquad (6)$$

For the same reason, $A_{1c}$ and $A_{1s}$ each have zero mean. Therefore equation (6) is the condition of zero covariance for the A's. For gaussian random variables, zero covariance is a necessary and sufficient condition for statistical independence. Hence, $A_{1c}$ and $A_{1s}$ are statistically independent.

1b Covariance of Antenna Voltages

Now let $v_2$ represent the varying output voltage of a neighboring atenna. Then, from equation (2) above, we obtain

$$\overline{v_1 v_2} = \sum_{\substack{m=1\\n=1}}^{N} c_m c_n \, \overline{\cos (\omega t + \gamma_{1m}) \cos (\omega t + \gamma_{2n})} \qquad (7)$$

[1]We make no distinction between ensemble and time averages. This is justified if we are dealing with ergodic processes. For the gaussian random process with which we are concerned, ergodicity is insured by stationarity under very general conditions. Our identification of time averages with ensemble averages here implies only that we are assuming stationary conditions.

which expands into

$$\overline{v_1 v_2} = \sum_{\substack{m=1\\n=1}}^{N} c_m c_n \, [\overline{\cos\gamma_{1m}\cos\gamma_{2n}\cos^2\omega t + \sin\gamma_{1m}\sin\gamma_{2n}\sin^2\omega t}$$

$$\overline{- (\cos\gamma_{1m}\sin\gamma_{2n} + \sin\gamma_{1m}\cos\gamma_{2n})\sin\omega t\cos\omega t}] \tag{8}$$

which reduces to

$$\overline{v_1 v_2} = \tfrac{1}{2}\sum_{\substack{m=1\\n=1}}^{N} c_m c_n \, \overline{(\cos\gamma_{1m}\cos\gamma_{2n} + \sin\gamma_{1m}\sin\gamma_{2n})} \tag{9}$$

Under the assumption of random phasing, equation (9) yields the following simple expression for the covariance of voltages:

$$\overline{v_1 v_2} = \tfrac{1}{2}\sum_{n=1}^{N} c_n^2 \cos(\gamma_{1n} - \gamma_{2n}) = \tfrac{1}{2}\sum_{n=1}^{N} c_n^2 \cos 2\chi_n \tag{10}$$

where $\chi_n = \tfrac{1}{2}(\gamma_{1n} - \gamma_{2n})$.

## 1c <u>The Elements of the Moment Matrix</u>

Let M denote the moment matrix for the four gaussian random variables $A_{1c}$, $A_{1s}$, $A_{2c}$, and $A_{2s}$. The diagonal elements $\mu_{ii}$ of M are the variances, equal for the four variables and denoted by $\sigma^2$. The off-diagonal elements $\mu_{ij}$ are the covariances.

Equation (6) above shows that $\mu_{12} = 0$. Similarly, $\mu_{34} = 0$. Further, the same result would arise upon commutation of the factors on the left of equation (6). Under the condition of random phasing, we have, further that

$$\mu_{13} = \overline{A_{1c}A_{2c}} = \sum_{n=1}^{N} c_n^2 \, \overline{\cos\gamma_{1n}\cos\gamma_{2n}} = \sum_{n=1}^{N} c_n^2 \, \overline{\cos\gamma_{1n}\cos(\gamma_{1n} - 2\chi_n)} \tag{11}$$

$$\text{or} \qquad \mu_{13} = \sum_{n=1}^{N} c_n^2 [\overline{\cos^2 \gamma_{1n} \cos 2\gamma_n} + \overline{\cos \gamma_{1n} \sin \gamma_{1n} \sin 2\chi_n}] \tag{12}$$

$$\text{so} \qquad \mu_{13} = \tfrac{1}{2}\sum_{n=1}^{N} c_n^2 \cos 2\chi_n \tag{13}$$

In similar fashion, it can be shown that

$$\mu_{31} = \mu_{24} = \mu_{42} = \tfrac{1}{2}\sum_{n=1}^{N} c_n^2 \cos 2\chi_n \tag{14}$$

and

$$-\mu_{14} = -\mu_{41} = \mu_{23} = \mu_{32} = \tfrac{1}{2}\sum_{n=1}^{N} c_n^2 \sin 2\chi_n \tag{15}$$

The above can be summarized by writing the moment matrix as follows:

$$M = \begin{pmatrix} \sigma^2 & 0 & \mu_c & -\mu_s \\ 0 & \sigma^2 & \mu_s & \mu_c \\ \mu_c & \mu_s & \sigma^2 & 0 \\ -\mu_s & \mu_c & 0 & \sigma^2 \end{pmatrix} \tag{16}$$

$$\text{where} \qquad \sigma^2 = \tfrac{1}{2}\sum_{n=1}^{N} c_n^2 \tag{17}$$

$$\mu_c = \tfrac{1}{2}\sum_{n=1}^{N} c_n^2 \cos 2\chi_n \tag*{(17)}$$

$$\mu_s = \tfrac{1}{2}\sum_{n=1}^{N} c_n^2 \sin 2\chi_n \tag{19}$$

## 1d The Matrix Elements and the Wavefront Correlation for a Narrow, Symmetrical Angular Spectrum

Let the elements of the moment matrix be given by their integral definitions. Thus,

$$\sigma^2 = \int_{-\infty}^{\infty} p(\alpha)d\alpha \tag{20}$$

$$\mu_c = \int_{-\infty}^{\infty} p(\alpha) \cos 2\chi \, d\alpha \tag{21}$$

$$\mu_s = \int_{-\infty}^{\infty} p(\alpha) \sin 2\chi \, d\chi \tag{22}$$

where $\chi = (2\pi d/\lambda) \sin \alpha$. Now make the following change of variables:

$$\alpha = \alpha_o + \delta \tag{23}$$

$$p(\alpha) = F(\delta) = F(-\delta) \tag{24}$$

Then equation (20) can be replaced by

$$\sigma^2 = \int_{-\infty}^{\infty} F(\delta) \, d\delta \tag{25}$$

By direct substitution, we have also that

$$\mu_c = \int_{-\infty}^{\infty} F(\delta) \cos [(2\pi d/\lambda) \sin (\delta + \alpha_o)] \, d\delta \tag{26}$$

and

$$\mu_s = \int_{-\infty}^{\infty} F(\delta) \sin [2\pi d/\lambda) \sin (\delta + \alpha_o)] \, d\delta \tag{27}$$

Under the condition that F is appreciable only for values of δ which are sufficiently small that cos δ may be approximated by unity and sin δ by δ, we may make the following substitutions:

$$\cos [2\pi d/\lambda) \cos (\delta + \alpha_o)] = \cos [(2\pi d/\lambda) (\delta \cos \alpha_o + \sin \alpha_o)] \quad (28)$$

$$\sin [2\pi d/\lambda) \cos (\delta + \alpha_o)] = \sin [(2\pi d/\lambda) (\delta \cos \alpha_o + \sin \alpha_o)] \quad (29)$$

which yield the following:

$$\mu_c = \int_{-\infty}^{\infty} F(\delta) [\cos(\frac{2\pi d}{\lambda} \delta \cos\alpha_o)\cos(\frac{2\pi d}{\lambda} \sin\alpha_o)$$

$$- \sin(\frac{2\pi d}{\lambda} \delta \cos\alpha_o)\sin(\frac{2\pi d}{\lambda} \sin\alpha_o)] \, d\delta \quad (30)$$

$$\mu_s = \int_{-\infty}^{\infty} F(\delta) [\sin(\frac{2\pi d}{\lambda} \delta \cos\alpha_o)\cos(\frac{2\pi d}{\lambda} \sin\alpha_o)$$

$$+ \cos(\frac{2\pi d}{\lambda} \delta \cos\alpha_o)\sin(\frac{2\pi d}{\lambda} \sin\alpha_o)] \, d\delta \quad (31)$$

If F (δ) is assumed to be an even function, the terms within the brackets above which are odd functions of δ vanish in the integration. Recalling that $\chi = (2\pi d/\lambda) \sin \alpha$, we are left with

$$\sigma^2 = \int_{-\infty}^{\infty} F(\delta) \, d\delta \quad (32)$$

$$\mu_c = \cos 2\chi_o \int_{-\infty}^{\infty} F(\delta) \cos [(2\pi d/\lambda) \delta \cos \alpha_o] \, d\delta \quad (33)$$

$$\mu_s = \sin 2\chi_o \int_{-\infty}^{\infty} F(\delta) \cos [(2\pi d/\lambda) \delta \cos \alpha_o] \, d\delta \quad (34)$$

Now the wavefront correlation R is defined as

$$R = \frac{(\mu_c^2 + \mu_s^2)^{\frac{1}{2}}}{\sigma^2} \tag{35}$$

Substitution of equations (32), (33), and (34) into equation (35) yields

$$R = \frac{\int_{-\infty}^{\infty} F(\delta)\,\cos\,[(2\pi d/\lambda)\,\delta\,\cos\alpha_o]\,d\delta}{\int_{-\infty}^{\infty} F(\delta)\,d\delta} \tag{36}$$

le <u>Covariance of Antenna Voltages in the Presence of a Nondeviated Component in the Angular Spectrum</u>

If a nondeviated component of amplitude S is added to a randomly phased angular scatter spectrum of variance $\sigma^2 = \overline{B_{1c}^2} = \overline{B_{1s}^2}$, then the antenna voltages, $v_1$ and $v_2$, become

$$v_1 = S\cos(\omega t + \chi_o) + B_{1c}\cos\omega t - B_{1s}\sin\omega t \tag{37}$$

$$v_2 = S\cos(\omega t - \chi_o) + B_{2c}\cos\omega t - B_{2s}\sin\omega t \tag{38}$$

The covariance of voltages then is given by

$$\begin{aligned}\overline{V_1V_2} = {} & S^2\,\overline{\cos(\omega t+\chi_o)\cos(\omega t-\chi_o)} + \overline{B_{1c}B_{2c}}\;\overline{\cos^2\omega t} + \overline{B_{1s}B_{2s}}\;\overline{\sin^2\omega t} \\ & + S\,\cancelto{0}{\overline{B_{2c}}}\;\overline{\cos(\omega t+\chi_o)\cos\omega t} - S\,\cancelto{0}{\overline{B_{2s}}}\;\overline{\cos(\omega t+\chi_o)\sin\omega t} \\ & + S\,\cancelto{0}{\overline{B_{1c}}}\;\overline{\cos(\omega t-\chi_o)\cos\omega t} - S\,\cancelto{0}{\overline{B_{1s}}}\;\overline{\cos(\omega t-\chi_o)\sin\omega t} \\ & + \overline{B_{1c}B_{2s}}\;\cancelto{0}{\overline{\cos\omega t\,\sin\omega t}} + \overline{B_{1s}B_{2c}}\;\cancelto{0}{\overline{\cos\omega t\,\sin\omega t}}\end{aligned} \tag{39}$$

so

$$\overline{V_1V_2} = \tfrac{1}{2} S^2 [\cos 2\chi_o + \overset{o}{\overline{\cos 2\omega t}}] + \tfrac{1}{2} \mu_{13} + \tfrac{1}{2} \mu_{24} \tag{40}$$

or

$$\overline{V_1V_2} = \tfrac{1}{2} S^2 \cos 2\chi_o + \mu_c \tag{41}$$

## 1f The mean Amplitude and Power Products in the Case of Dominance by the Nondeviated Component

Now designate the amplitudes and phases of the two atenna voltages by $A_1$, $\theta_1$, $A_2$, and $\theta_2$, so that

$$v_1 = A_1 \cos(\omega t + \theta_1) \quad \text{and} \quad v_2 = A_2 \cos(\omega t + \theta_2) \tag{42}$$

Comparison of equations (42) with equations (37) and (38) yields, for $\frac{S}{\sqrt{2}\sigma} >>$ unity[1],

$$\begin{aligned}\overline{A_1A_2} = {} & S^2 + S(\overset{o}{\overline{B_{2c}}} \cos\chi_o - \overset{o}{\overline{B_{2s}}} \sin\chi_o + \overset{o}{\overline{B_{1c}}} \cos\chi_o + \overset{o}{\overline{B_{1s}}} \sin\chi_o) \\ & + \tfrac{1}{2} (\overset{\sigma^2}{\overline{B^2_{2c}}} \sin^2\chi_o + \overset{\sigma^2}{\overline{B^2_{2s}}} \cos^2\chi_o \\ & + 2\overset{o}{\overline{B_{2c}B_{2s}}} \cos\chi_o \sin\chi_o + \overset{\sigma^2}{\overline{B^2_{1c}}} \sin^2\chi_o + \overset{\sigma^2}{\overline{B^2_{1s}}} \cos^2\chi_o - 2\overset{o}{\overline{B_{1c}B_{1s}}} \cos\chi_o \sin\chi_o) \\ & + \overset{\mu_{13}}{\overline{B_{1c}B_{2c}}} \cos^2\chi_o - \overset{\mu_{14}}{\overline{B_{1c}B_{2s}}} \cos\chi_o \sin\chi_o + \overset{\mu_{32}}{\overline{B_{2c}B_{1s}}} \cos\chi_o \sin\chi_o \\ & - \overset{\mu_{24}}{\overline{B_{1s}B_{2s}}} \sin^2\chi_o\end{aligned} \tag{43}$$

$$\overline{A_1A_2} = S^2 + \sigma^2 + \mu_c(\cos^2\chi_o - \sin^2\chi_o) + 2\mu_s \cos\chi_o \sin\chi_o \tag{44}$$

---

[1] See also equations 2-68 and 2-69 in Chapter II.

so

$$\overline{A_1A_2} = S^2 + \sigma^2 + \mu_c \cos 2\chi_o + \mu_s \sin 2\chi_o \tag{45}$$

We have also[2]

$$\overline{A_1^2A_2^2} = S^4 + 2S^3 [(\overline{B_{1c}} + \overline{B_{2c}}) \cos \chi_o + (\overline{B_{1s}} - \overline{B_{2s}}) \sin \chi_o] + S^2 (\overline{B_{1c}^2} + \overline{B_{1s}^2} + \overline{B_{2c}^2} + \overline{B_{2s}^2}) + 4S^2(\overline{B_{1c}B_{2c}} \cos^2\chi_o - \overline{B_{1s}B_{2s}} \sin^2\chi_o + \overline{B_{1s}B_{2c}} \sin \chi_o \cos \chi_o - \overline{B_{1c}B_{2s}} \sin \chi_o \cos \chi_o) \tag{46}$$

$$\overline{A_1^2A_2^2} = S^4 + 4\sigma^2 S^2 + 4S^2 [\mu_c(\cos^2 \chi_o - \sin^2 \chi_o) + 2\mu_s \sin \chi_o \cos \chi_o] \tag{47}$$

$$\overline{A_1^2A_2^2} = S^4 + 4\sigma^2 S^2 + 4S^2 (\mu_c \cos 2\chi_o + \mu_s \sin 2\chi_o) \tag{48}$$

[2]See equations 2-61 and 2-62 in Chapter II.