621.396.11

Paper No. 922
RADIO SECTION

THE CONCEPT OF AN ANGULAR SPECTRUM OF PLANE WAVES, AND ITS RELATION TO THAT OF POLAR DIAGRAM AND APERTURE DISTRIBUTION

By Prof. H. G. BOOKER, M.A., Ph.D., Associate Member, and P. C. CLEMMOW, M.A.

(*The paper was received* 12*th April*, 1949.)

SUMMARY

A critical examination is made of the somewhat loose and incomplete statement that a polar diagram is the Fourier transform of an aperture distribution. By aperture distribution it is necessary to understand, in the two-dimensional case, distribution across the aperture of the component along the aperture plane of the electromagnetic field in the plane of propagation. Furthermore, the concept of the polar diagram has to be replaced by that of an angular spectrum, except in the common case when the aperture may be considered more or less limited in width, and the field is being evaluated at a point whose distance from the aperture is large compared with the width of the aperture (and the wavelength). For example, it is convenient for some purposes to regard the problem of diffraction of a plane wave by a semi-infinite plane screen, with a straight edge, as a problem about an aperture distribution in the plane of the screen. This is a case for which the concept of a polar diagram is not in general applicable, and has to be replaced by that of an angular spectrum. The field at all points in front of a plane aperture of any distribution may be regarded as arising from an aggregate of plane waves travelling in various directions. The amplitude and phase of the waves, as a function of their direction of travel, constitutes an angular spectrum, and this angular spectrum, appropriately expressed, is, without approximation, the Fourier transform of the aperture distribution. If the aperture distribution is of such a nature that the concept of the polar diagram is applicable at sufficiently great distances, then the polar diagram is equal to the angular spectrum. But the angular spectrum is a concept that is always applicable, whereas the polar diagram is one that is liable to be invalid (for example, in the Sommerfeld theory of propagation over a plane, imperfectly reflecting earth).

(1) INTRODUCTION

During the past ten years it has become quite well known that the distribution of field across the aperture plane of an aerial (such as a paraboloid or a horn) and the polar diagram of the aerial are Fourier transforms of each other. This statement is, however, rather loose. The distribution of field across the aperture plane might refer to the electric field or the magnetic field, and in each of these two cases there are three vector components to be considered. The distribution of field across the aperture plane could, in fact, mean any one of several different functions, and it is only one of these that can be the Fourier transform of the polar diagram.

Again, the concept of the polar diagram assumes that the linear dimensions of the aerial system are limited in extent, so that it is possible to go to a distance large compared with these linear dimensions (and compared with the wavelength) in order to assess the strength of radiation in different directions. The concept of polar diagrams is, in fact, applicable only when the distribution of exciting field over the aperture plane is more or less limited to a finite part of the complete plane. But there are important applications in which it is desirable to be able to consider aperture distributions that are not restricted in this way. Consider, for example, diffraction of a plane wave by a semi-infinite plane screen having a straight edge. The plane of the screen may be considered as an aperture plane, but the aperture distribution across it extends to infinity in directions away from the screen, and does not even tend to zero in these directions. The concept of the polar diagram cannot, in general, be applied to such an aperture distribution. There is, however, a related concept, that of an angular spectrum of plane waves, which can always be applied, even when that of the polar diagram is invalid. The angular spectrum associated with an aperture distribution gives the polar diagram if this concept is applicable, but retains a useful meaning even when it is not possible to talk of a polar diagram.

It is the purpose of the paper to examine critically what lies behind the somewhat loose and incomplete statement that the polar diagram is the Fourier transform of aperture distribution. In doing so, we shall restrict ourselves to two-dimensional problems. In Section 2 the features of Fourier analysis required for the discussion are briefly stated. In Section 3 the application of Fourier analysis to two-dimensional aperture distributions is given, and in Section 4 the distinction between the concepts of angular spectrum and polar diagram is explained. To illustrate this distinction, Section 5 deals with diffraction of a plane wave by a straight edge, regarded as a problem about an aperture distribution. Other applications are mentioned in Section 6, but are not dealt with in detail.

(2) SOME RESULTS CONCERNING TIME-FLUCTUATIONS AND FREQUENCY-SPECTRA

It is well known that, if a series of oscillations of different frequencies with different amplitudes and phases are added together, quite a complicated fluctuation is generally obtained. This result is usually expressed in the form

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} s(\omega) \exp (j\omega t) d\omega \quad . \quad . \quad . \quad . \quad (1)$$

where t is time, and ω the angular frequency: $f(t)$ represents the fluctuation, and $s(\omega)$ its frequency spectrum. Both $f(t)$ and $s(\omega)$ are generally complex. In the Argand diagram, $s(\omega) \exp (j\omega t)$ is a vector rotating with angular velocity ω in the positive direction. Its length is $|s(\omega)|$, and at $t = 0$ the vector is at an angle $\arg s(\omega)$ with the initial line. The component of the vector along the initial line is an oscillation of angular frequency ω, with amplitude $|s(\omega)|$ and phase $\arg s(\omega)$. The result of adding together many such oscillations is the component, along the initial line, of the vector sum of the corresponding vectors, and it is this vector addition that is exhibited in eqn. (1). The vectors summed on the right-hand side of eqn. (1) include those rotating negatively as well as positively, and it should be noticed that contributions to the oscillation of frequency ω arise from components along the initial line of vectors rotating in both directions with angular velocity ω.

Furthermore, it is well known that, no matter what fluctuation may be given, it may be analysed into harmonic oscillations of

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Prof. Booker is at Cornell University, New York, and Mr. Clemmow is at Imperial College, London University.

different frequencies, each with its own amplitude and phase, and that the frequency spectrum $s(\omega)$ of a given fluctuation $f(t)$ is given by

$$s(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-j\omega t)dt \quad . \quad . \quad . \quad . \quad (2)$$

Examples of corresponding fluctuations and spectra that we shall require are the unit impulse-function and the unit step-function (see Figs. 1 and 2).

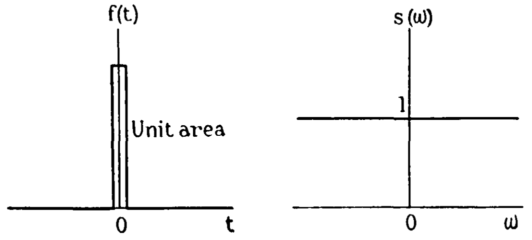

Fig. 1.—The unit impulse-function and its spectrum.

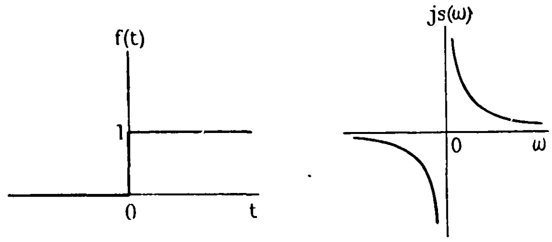

Fig. 2.—The unit step-function and its spectrum.

The unit impulse-function is a fluctuation, usually denoted by $\delta(t)$, which vanishes except near $t = 0$, where it has such a large value that the area under the impulse is unity. If

$$f(t) = \delta(t) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

then the appropriate frequency spectrum is given by

$$s(\omega) = 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

The unit step-function is a fluctuation that vanishes for $t < 0$ and is unity for $t > 0$. We shall denote it by $u(t)$. It is well known that, if

$$f(t) = u(t) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

then

$$s(\omega) = \frac{1}{j\omega} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

We shall also require the shift-rules of Fourier analysis, which are those for delaying a fluctuation and modulating a carrier.

If the fluctuation $f(t)$ has spectrum $s(\omega)$, then the fluctuation $f(t - t_0)$ has spectrum $s(\omega) \exp(-j\omega t_0)$. This means that to delay a fluctuation introduces an additional linear variation of phase with frequency. Applying this shift-rule to eqns. (3) and (4), we see that, if

$$f(t) = \delta(t - t_0) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

then

$$s(\omega) = \exp(-j\omega t_0) \quad . \quad . \quad . \quad . \quad (8)$$

If $s(\omega)$ is the spectrum of the fluctuation $f(t)$, then $s(\omega - \omega_0)$ is the spectrum of the fluctuation $f(t) \exp(j\omega_0 t)$. This means that the spectrum of a modulated carrier of angular frequency ω_0 is the spectrum of the modulation, shifted ω_0 along the frequency axis. Applying this shift rule to eqns. (5) and (6), we see that, if

$$f(t) = u(t) \exp(j\omega_0 t) \quad . \quad . \quad . \quad . \quad . \quad (9)$$

then

$$s(\omega) = \frac{1}{j(\omega - \omega_0)} \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The above features of Fourier analysis and synthesis have been well known for a long time. We now turn our attention to some features which were not so widely appreciated ten years ago, but which are now quite familiar as a result of war-time applications.[1]

(3) THE CONCEPT OF AN ANGULAR SPECTRUM OF PLANE WAVES

Suppose that, in a system of Cartesian co-ordinates, as indicated in Fig. 3, the region $x > 0$ is filled with a homogeneous

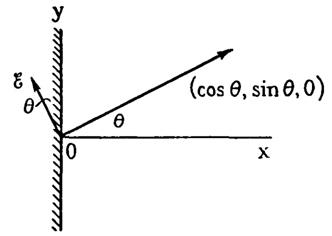

Fig. 3.—Radiation of a plane wave from an aperture plane $x = 0$ into the region $x > 0$.

medium having propagation coefficient k and characteristic admittance η.[2] For a plane wave in the medium, k is the increase of phase difference per unit distance in the direction of propagation, and η is the ratio of magnetic to electric field. Suppose that, in the plane $x = 0$, an electromagnetic field is maintained having components along the plane which vary with time, and with position in the plane, in a prescribed manner. The field maintained in the plane $x = 0$ is propagated into the region $x > 0$, and we may think of the plane $x = 0$ as the aperture plane of an aerial system radiating into the region $x > 0$.

The following simplifications will be adopted:

(*a*) The electromagnetic field will be supposed to vary harmonically at a frequency corresponding to a wavelength λ, and the corresponding complex oscillation-function will be suppressed.

(*b*) The field will be supposed to be two-dimensional, and will be taken as independent of the z-coordinate.

(*c*) The magnetic field will be taken parallel to the z-axis, and the electric field parallel to the xy-plane. (We could equally well take the electric field parallel to the z-axis, and the magnetic field parallel to the xy-plane.)

With these simplifications, what are specified in the aperture plane $x = 0$ are the y-component of electric field and the z-component of magnetic field as functions of y. Only one of these functions is to be thought of as arbitrary, the other being deducible from it. Suppose, for example, that we specify the distribution of tangential magnetic field, and that for $x < 0$ there shall be no magnetic field in the medium. Then there is a discontinuity of tangential magnetic field at the plane $x = 0$, and this is the current per unit width that must flow in the aperture plane to produce the field. From this current we can evaluate the entire electromagnetic field, including the tangential component of electric field at the aperture plane. The aperture distribution is therefore completely specified by the z-component of magnetic field; it is just as completely specified by the y-component of electric field, which is the voltage per unit width that must act along the aperture plane to maintain the field.

Since the z-component of magnetic field is in fact the resultant magnetic field in accordance with simplification (c), it might seem easiest to use it for specifying the aperture distribution in preference to the y-component of electric field. It turns out, however, to be more convenient to describe the aperture distribution by means of the variation with y of the y-component of electric field, in spite of the fact that there is also a component of electric field normal to the aperture plane.

Let us imagine, first of all, that the wave radiated into the region $x > 0$ is a plane wave (Fig. 3) travelling in the direction $(\cos\theta, \sin\theta, 0)$ and consider what aperture distribution would be required to maintain it. By putting

$$C = \cos\theta,\ S = \sin\theta \quad . \quad . \quad . \quad . \quad (11)$$

so that

$$C^2 + S^2 = 1 \quad . \quad . \quad . \quad . \quad . \quad (12)$$

the electric field $\mathscr{E}$ and magnetic field H of the wave at the point (x, y) may be written

$$\begin{cases} \mathscr{E}(x, y) = A(-S, C, 0)\exp[-jk(Cx + Sy)] & . \quad (13) \\ H(x, y) = \eta A(0, 0, 1)\exp[-jk(Cx + Sy)] & . \quad (14) \end{cases}$$

The aperture distribution required to maintain this field is obtained by putting $x = 0$ in eqns. (13) and (14). If we express it, as described above, in terms of the component of electric field tangential to the aperture plane, the aperture distribution is

$$\mathscr{E}_y(0, y) = AC\exp(-jkSy) \quad . \quad . \quad . \quad . \quad (15)$$

We notice that eqn. (15) represents a wave travelling over the aperture plane in the y-direction with propagation coefficient kS. This is the component along the aperture plane of the propagation vector

$$k(C, S, 0) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

of the wave. Any wave such as that given by eqn. (15) travelling over the aperture plane with a propagation coefficient kS produces in the medium (propagation coefficient k) a plane wave in a direction making an angle θ ($\sin\theta = S$) with the normal to the aperture plane.

If we were to consider a wave such as that given by eqn. (15) travelling over the aperture plane with a propagation coefficient kS greater than the propagation coefficient k of the medium, then S would exceed unity, and C, from eqn. (12), would be imaginary. From the wave-function in eqns. (13) and (14), we see that the field produced in the region $x > 0$ would decrease in the x-direction like

$$\exp[-k(S^2 - 1)^{\frac{1}{2}}x] \quad . \quad . \quad . \quad . \quad (17)$$

while travelling in the y-direction like eqn. (15). The field produced by a wave such as that of eqn. (15), travelling over the aperture plane with propagation coefficient exceeding that of the medium, is one that hugs the aperture plane and is not propagated away from it. Such a wave is called an evanescent or reactive wave, and is concerned with the oscillation of electromagnetic energy backwards and forwards across the aperture plane, instead of with the permanent removal of energy from the plane.

We now proceed to consider an aperture distribution which possesses the three simplifications enumerated above, but which depends on the co-ordinate y across the aperture plane in any manner, instead of in the manner of a travelling wave as given by eqn. (15). Such an arbitrary aperture distribution may, in fact, be reduced by Fourier analysis to waves like eqn. (15) travelling over the aperture and having different propagation coefficients along the y-axis. Each wave such as this, travelling across the aperture plane, gives rise to a plane wave in the medium travelling in a particular direction (or to an evanescent wave), as already described. The result is that the field produced in the region $x > 0$ by an arbitrary aperture distribution may be expressed as a combination of waves of the type given by eqns. (13) and (14) with different values of θ. Each wave has its own amplitude and phase, which in general vary with θ, the whole forming what is known as an angular spectrum of plane waves.

In eqns. (13), (14) and (15), A is a complex number whose modulus and argument determine the amplitude and phase of the plane wave at the origin; if we regard A as a function of θ, and integrate these equations with respect to θ, we arrive at an angular spectrum of plane waves. It is, in fact, more convenient to work in terms not of θ but of S, the sine of the angle that a plane wave makes with the normal to the aperture plane. This is because, to incorporate structure in the aperture distribution finer than the wavelength, it is necessary to include evanescent waves, which correspond to a value of θ that is complex, but to a value of S that is real, although numerically greater than unity. The product AC in eqn. (15) is thus to be thought of as a function of S. As this product is the function of S with which we shall describe the angular spectrum of plane waves, in preference to $A(S)$, it is convenient to define

$$P(S) = \lambda CA(S) \quad . \quad . \quad . \quad . \quad . \quad (18)$$

λ being the wavelength, $2\pi/k$, of a plane wave in the medium. Substituting for $A(S)$ from eqn. (18) into eqns. (13), (14) and (15), and integrating with respect to S from $-\infty$ to $+\infty$, we obtain the electromagnetic field

$$\begin{cases} \mathscr{E}(x, y) = \dfrac{1}{\lambda}\displaystyle\int_{-\infty}^{\infty} P(S)(-S, C, 0)\exp[-jk(Cx + Sy)]\frac{dS}{C} & (19) \\ H(x, y) = \dfrac{\eta}{\lambda}\displaystyle\int_{-\infty}^{\infty} P(S)(0, 0, 1)\exp[-jk(Cx + Sy)]\frac{dS}{C} & . \quad (20) \end{cases}$$

which is produced by an aperture distribution

$$\mathscr{E}_y(0, y) = \frac{1}{\lambda}\int_{-\infty}^{\infty} P(S)\exp(-jkSy)dS \quad . \quad . \quad . \quad (21)$$

It may be noted that dS/C in eqns. (19) and (20) is simply $d\theta$ from eqn. (11).

Eqn. (21), giving the aperture distribution $\mathscr{E}_y(0, y)$ that produces the angular spectrum $P(S)$ of plane waves, is to be compared with eqn. (1), giving the time-fluctuation $f(t)$ corresponding to a frequency-spectrum $s(\omega)$. We notice that eqn. (1) becomes eqn. (21) if we replace

$$t,\ f(t),\ \omega \text{ and } s(\omega) \quad . \quad . \quad . \quad . \quad . \quad (22)$$

by

$$-y,\ \mathscr{E}_y(0, y),\ kS \text{ and } P(S) \quad . \quad . \quad . \quad (23)$$

respectively. The role of frequency is thus played by the component of the propagation vector along the aperture plane. Using the above replacements in eqn. (2), we see that the angular spectrum is expressed in terms of aperture distribution by the equation

$$P(S) = \int_{-\infty}^{\infty} \mathscr{E}_y(0, y)\exp(jkSy)dy \quad . \quad . \quad . \quad (24)$$

The following is an example of a simple aperture distribution. Suppose that the aperture plane is filled with a perfectly conducting metal sheet in which is cut a narrow slot along the z-axis (see Fig. 4). Suppose that an alternating voltage of amplitude V is applied across the slot, and that we have to

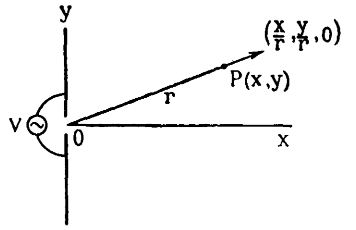

Fig. 4.—Radiation from a slot located along the z-axis.

consider the resulting field in the region $x > 0$. Then, in terms of the impulse-function [eqn. (3)], the aperture distribution may be taken as

$$\mathscr{E}_y(0, y) = V\delta(y) \quad . \quad . \quad . \quad . \quad . \quad (25)$$

The corresponding angular spectrum, from eqn. (4), is

$$P(S) = V \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)$$

The electromagnetic field radiated from the slot is, therefore, from eqns. (19) and (20)

$$\begin{cases} \mathscr{E}(x, y) = \dfrac{V}{\lambda}\displaystyle\int_{-\infty}^{\infty}(-S, C, 0)\exp\left[-jk(Cx + Sy)\right]\dfrac{dS}{C} & (27) \\ H(x, y) = \dfrac{\eta V}{\lambda}\displaystyle\int_{-\infty}^{\infty}(0, 0, 1)\exp\left[-jk(Cx + Sy)\right]\dfrac{dS}{C} & (28) \end{cases}$$

The fact that the angular spectrum [eqn. (26)] is independent of the direction of a plane wave means that all the plane waves have the same amplitude. Moreover, all the plane waves in the angular spectrum are in the same phase on the z-axis, where the slot is located. At a point in the region $x > 0$, however, the various waves have different phases. Consider a point $P(x, y)$ whose distance r from the slot is appreciably greater than a wavelength. A plane wave directed radially from the slot towards P is travelling in the direction

$$\left(\frac{x}{r}, \frac{y}{r}, 0\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

Waves travelling in approximately this direction are nearly in the same phase at P, and therefore interfere constructively, whereas waves travelling in directions remote from (29) are in quite different phases at P, and so interfere destructively. In the neighbourhood of P we have, therefore, a field travelling radially outwards from the slot in the direction (29). It is, in fact, the field due to a "magnetic current" V flowing along the z-axis and radiating into the region $x > 0$ (but not $x < 0$). This is a cylindrical wave, axially symmetrical round the z-axis, and may be written

$$\begin{cases} \mathscr{E}(x, y) = V\left(-\dfrac{y}{r}, \dfrac{x}{r}, 0\right)\exp\left[-j(kr - \tfrac{1}{4}\pi)\right]/(r\lambda)^{\frac{1}{2}} & (30) \\ H(x, y) = \eta V(0, 0, 1)\exp\left[-j(kr - \tfrac{1}{4}\pi)\right]/(r\lambda)^{\frac{1}{2}} \quad . & (31) \end{cases}$$

where r is appreciably greater than λ. Eqns. (30) and (31) are thus the evaluations of eqns. (27) and (28) at distances from the slot appreciably greater than the wavelength, as may be proved by performing the integrations by the method of stationary phase or steepest descent.[3,4]

In this section we have seen that, given any two-dimensional aperture distribution $\mathscr{E}_y(0,y)$, we can find its Fourier transform $P(S)$ by means of eqn. (24), and then substitute this into eqns. (19) and (20), thereby expressing the electromagnetic field at all points in front of the aperture plane as the superposition of an angular spectrum of plane waves.

(4) THE RELATION BETWEEN ANGULAR SPECTRUM AND POLAR DIAGRAM

Instead of referring to $P(S)$ as the angular spectrum, it is quite common to express it as a function of θ and call it the polar diagram. This procedure is appropriate if the aperture distribution is of such a nature that $\mathscr{E}_y(0, y)$ becomes negligible as y tends to $\pm\infty$, so that the aperture is more or less limited to a finite part of the plane $x = 0$, which is normally the case for ordinary aerial systems. At a distance from the aerial appreciably greater than the wavelength, and large compared with the width of aperture over which the aperture distribution is important, radiation is roughly radial from the aerial, and $P(S)$ measures the dependence of field strength on direction from the aerial. Eqn. (24) is in fact the usual formula for the polar diagram of an aerial in terms of the aperture distribution. If we think of the aperture as limited to a part of the plane $x = 0$ in the neighbourhood of the z-axis, points of the aperture having different values of y are joined to a distant point by sensibly parallel lines, as indicated in Fig. 5. But, in adding up the

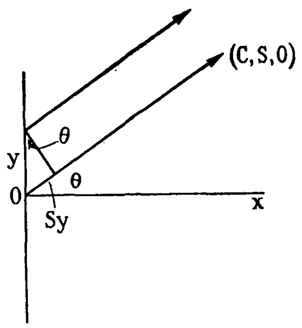

Fig. 5.—Calculation of the polar diagram of an aerial.

contributions from various parts of the aperture, different phase differences must be allowed for. The phase of radiation in the direction $(C, S, 0)$ from the point y of the aperture is ahead of that from the origin by kSy, since Sy is the projection of a length y of the aperture plane on a line making an angle θ ($\sin\theta = S$) with the normal to the plane. The factor $\exp(jkSy)$ in eqn. (24) thus allows for this phase difference in the calculation of the polar diagram $P(S)$ from the aperture distribution $\mathscr{E}_y(0, y)$.

That $P(S)$ is the polar diagram of an aperture more or less limited to a part of the plane $x = 0$ near the z-axis may also be seen from eqns. (19) and (20). For at a large distance from the z-axis the only plane waves in the angular spectrum that interfere constructively are those travelling almost radially outwards from the aerial in the direction (29), corresponding to

$$S = \frac{y}{r} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (32)$$

$P(S)$ hardly varies from one of these waves to another, and may therefore be brought outside the integrals in eqns. (19) and (20) at the value $P(y/r)$. These integrals are then the same as those in eqns. (27) and (28), and so are equivalent to the expressions in eqns. (30) and (31) with $V = 1$. The electromagnetic field given by eqns. (19) and (20) therefore becomes, at sufficiently large distances from the z-axis,

$$\begin{cases} \mathscr{E} = P\left(\dfrac{y}{r}\right)\left(-\dfrac{y}{r}, \dfrac{x}{r}, 0\right)\exp\left[-j(kr - \tfrac{1}{4}\pi)\right]/(r\lambda)^{\frac{1}{2}}. & (33) \\ H = \eta P\left(\dfrac{y}{r}\right)(0, 0, 1)\exp\left[-j(kr - \tfrac{1}{4}\pi)\right]/(r\lambda)^{\frac{1}{2}} \quad . \quad . & (34) \end{cases}$$

Alternatively, if we use $(C, S, 0)$ for the direction cosines (29) of the vector of length r joining the origin to a distant point

$(x, y, 0)$, instead of for the direction cosines of an individual plane wave of the angular spectrum involved in eqns. (19) and (20), then eqns. (33) and (34) may be written

$$\begin{cases} \mathscr{E} = P(S)(-S, C, 0) \exp[-j(kr - \tfrac{1}{4}\pi)]/(r\lambda)^{\frac{1}{2}} & (35) \\ H = \eta P(S)(0, 0, 1) \exp[-j(kr - \tfrac{1}{4}\pi)]/(r\lambda)^{\frac{1}{2}} & (36) \end{cases}$$

Thus, for an aperture distribution more or less limited to part of the aperture plane $x = 0$ near the z-axis, the angular spectrum $P(S)$ is also the polar diagram, and describes the amplitude and phase of radiation from the aperture in the direction making an angle θ with the normal to the aperture plane.

It is important to realize, however, that interpretation of $P(S)$ as a polar diagram is generally applicable only if the aperture may be regarded as limited in width, whereas interpretation of $P(S)$ as an angular spectrum of plane waves applies to any aperture distribution whatever. Moreover, when $P(S)$ is used in eqns. (19) and (20), it gives the electromagnetic field at all points in front of the aperture, however close they may be to the aperture plane.

(5) APPLICATION OF FOURIER ANALYSIS TO DIFFRACTION BY A STRAIGHT EDGE

There are a number of important applications for which the concept of the polar diagram is not wholly appropriate, and $P(S)$ must then be thought of more strictly as defining an angular spectrum of plane waves. There are cases in which the aperture cannot be regarded as limited, so that, at any rate in some directions, no point is sufficiently remote from the aperture to apply the concept of the polar diagram.

Consider an aperture distribution given by the step-function:

$$\mathscr{E}_y(0, y) = Au(-y) \quad . \quad . \quad . \quad . \quad (37)$$

This is an aperture distribution which vanishes for $y > 0$, but for $y < 0$ involves an oscillation whose amplitude and phase do not vary over the aperture plane. Such an aperture distribution could be imagined to arise as a result of a plane wave, travelling in the direction of the positive x-axis in the region $x < 0$, encountering, in the plane $x = 0$, a "black" screen

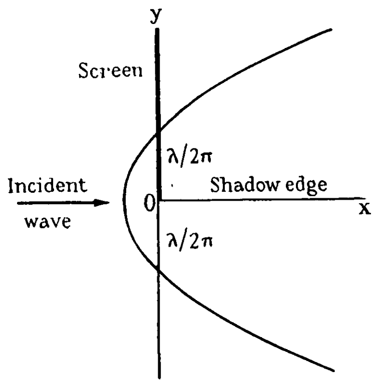

Fig. 6.—Diffraction of plane wave travelling along x-axis by "black" screen, $x = 0, y > 0$.

occupying the portion $y > 0$, as indicated in Fig. 6. The angular spectrum of plane waves corresponding to eqn. (37) is [cf. eqns. (5) and (6)]

$$P(S) = \frac{A}{jkS} \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)$$

This means that the electromagnetic field in the region $x > 0$ is given by eqns. (19) and (20), with the value given by eqn. (38) for $P(S)$.

The general nature of the electromagnetic field arising from diffraction of a plane wave by a black screen with a straight edge is well known.[5, 6, 7] In Fig. 6 the region $x > 0, y > 0$ is referred to as the shadow, and the remainder of space as the illuminated region, while the part of the plane $y = 0$ for which $x > 0$ is the shadow edge. To a first approximation, the wave incident from the left continues, beyond the plane $x = 0$, undisturbed in the region $y < 0$, but is obliterated in the shadow. To a second approximation, some energy is diffracted under the edge of the screen into the shadow, giving, well within the shadow, a wave which appears to radiate from the edge of the screen, known as the edge-wave. Moreover, in the neighbourhood of the shadow edge the field is so modified that no discontinuity exists at the shadow edge.

Let us first consider the edge-wave that exists in the shadow. We are dealing with an aperture distribution in the plane $x = 0$ which cannot be considered as limited in width, so that, in general, eqn. (38) cannot be regarded as the polar diagram of radiation at a large distance from the diffracting edge. If, however, we are sufficiently far within the shadow for the field to take the form of a wave radiated from the diffracting edge located along the z-axis, then eqn. (38) is its polar diagram. Substituting from eqn. (38) into eqns. (35) and (36), we see that the electromagnetic field at a point $r(C, S, 0)$ sufficiently within the shadow is the edge-wave

$$\begin{cases} \mathscr{E} = \dfrac{A}{jkS}(-S, C, 0) \exp[-j(kr - \tfrac{1}{4}\pi)]/(r\lambda)^{\frac{1}{2}} & (39) \\ H = \dfrac{\eta A}{jkS}(0, 0, 1) \exp[-j(kr - \tfrac{1}{4}\pi)]/(r\lambda)^{\frac{1}{2}} & (40) \end{cases}$$

Thus, radiation in the shadow falls off inversely as the sine of the angle from the shadow edge, this result applying not too close to the shadow edge, because the concept of a polar diagram ceases to be applicable there, and not too close to the screen (or at least to the diffracting edge) because of the effect of evanescent waves.

If we now approach the shadow edge and cross into the illuminated region, it is in general impossible to regard eqn. (38) as a polar diagram at any distance from the screen, however large, because the aperture cannot be regarded as having a finite width. However, at points in the region $x > 0, y < 0$ not too near the shadow edge there is an indirect way in which eqn. (38) may be regarded as a polar diagram, for in this region the edge-wave given by eqns. (39) and (40) gives the correction required to the undisturbed incident wave. This is proved simply by applying the result in the previous paragraph after subtracting the undisturbed incident-wave everywhere in the region $x > 0$, thereby converting the aperture-distribution $Au(-y)$ into $-Au(y)$. We then have a plane wave diffracted over a black screen occupying the portion $y < 0$ of the plane $x = 0$ (the complementary screen of Babinet's principle). The position, therefore, is that if we add the edge-wave given by eqns. (39) and (40) to the field that would be expected from geometrical optics, we obtain the required field, except near the shadow edge.

In the region near the shadow edge there are no means whereby eqn. (38) may be regarded as a polar diagram. Any attempt to describe the field near the shadow edge as radiating from the diffracting edge according to a polar diagram merely leads to the conclusion that the diagram changes with distance from the edge, which means that the field is not expressible as a polar diagram at all. There is, however, no difficulty in expressing the field near the shadow edge as an angular spectrum of plane waves. The field at all points in the region $x > 0$, including points near the shadow edge, is given by substituting for $P(S)$ from eqn. (38) into eqns. (19) and (20), thereby obtaining

$$\left\{\begin{aligned} \mathscr{E}(x, y) &= \frac{A}{2\pi j}\int_{-\infty}^{\infty}\frac{1}{S}(-S, C, 0)\exp\left[-jk(Cx+Sy)\right]\frac{dS}{C} \quad (41)\\ H(x, y) &= \frac{\eta A}{2\pi j}\int_{-\infty}^{\infty}\frac{1}{S}(0, 0, 1)\exp\left[-jk(Cx+Sy)\right]\frac{dS}{C} \quad . \quad (42)\end{aligned}\right.$$

Eqns. (39) and (40) constitute an approximate evaluation of the integrals given in eqns. (41) and (42), but this evaluation is not valid near the shadow edge.

Near the shadow edge we may evaluate eqns. (41) and (42) by the method of impulses, known in optics as Huygens's principle. The part of an aperture distribution $\mathscr{E}_y(0, y)$ between $y = \xi$ and $y = \xi + d\xi$ may be regarded as a slot, such as that illustrated in Fig. 4, located along the line $x = 0$, $y = \xi$, and excited with a voltage

$$V = \mathscr{E}_y(0, \xi)d\xi \quad . \quad . \quad . \quad . \quad . \quad (43)$$

We then use this value of V in eqns. (30) and (31), r being the distance from the point $(0, \xi, 0)$ to the point $(x, y, 0)$, and integrate with respect to ξ over the aperture. This integral, for the aperture distribution given by eqn. (37), leads to the complex Fresnel integral.

$$F(u) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}}\int_u^{\infty}\exp(-jv^2)dv \quad . \quad . \quad . \quad (44)$$

which, when plotted in the Argand diagram as a function of the real parameter u, gives the Cornu spiral illustrated in Fig. 7.

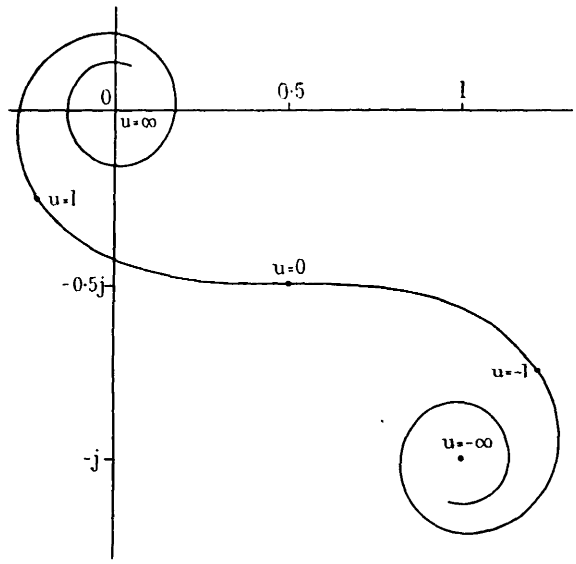

Fig. 7.—The Cornu spiral.

In terms of the Fresnel integral, the expression for the electromagnetic field at the point $r(C, S, 0)$ near the shadow edge is

$$\left\{\begin{aligned} \mathscr{E} &= \frac{1}{\sqrt{2}}AF\left[S\left(\frac{\pi r}{\lambda}\right)^{\frac{1}{2}}\right](-S, C, 0)\exp\left[-j(kx - \tfrac{1}{4}\pi)\right] \quad (45)\\ H &= \frac{1}{\sqrt{2}}AF\left[S\left(\frac{\pi r}{\lambda}\right)^{\frac{1}{2}}\right](0, 0, 1)\exp\left[-j(kx - \tfrac{1}{4}\pi)\right] \quad . \quad (46)\end{aligned}\right.$$

this result applying for values of r appreciably greater than λ. Eqns. (45) and (46) describe the transition from the illuminated region to the shadow, and become identical with eqns. (39) and (40) when the argument $S\sqrt{(\pi r/\lambda)}$ of the Fresnel integral becomes large compared with unity.

The region within which the edge-wave approximation given by eqns. (39) and (40) to the diffracted wave given by eqns. (45) and (46) may be used is illustrated in Fig. 6 by means of a parabola with focus on the diffracting edge, axis along the shadow edge, and semi-latus-rectum $\lambda/2\pi$. Inside this parabola, the argument of the Fresnel integral in eqns. (45) and (46) is less than unity in magnitude, while outside it is greater than unity. Thus, the edge-wave approximation given by eqns. (39) and (40) may be regarded as applying in the region of shadow outside the parabola. In the illuminated region outside the parabola, the field is given by adding the edge-wave to the undisturbed incident wave. Within the parabola, however, the complete eqns. (45) and (46), involving the Fresnel integral [eqn. (44)], must be used, and even these are reliable only at distances from the diffracting edge that are large compared with $\lambda/2\pi$.

It is usual to assume that the diffracted wave is substantially independent of the angle between the normal to the screen and the direction of the incident wave, and depends only on the position of the diffracting edge in relation to the point at which the diffracted field is being evaluated.[5] If we consider only small angles between the normal to the screen and the direction of the incident wave, we may proceed as follows. Take the screen shown in Fig. 6, but let the direction of incidence make an angle θ_0 ($S_0 = \sin\theta_0$, $C_0 = \cos\theta_0$) with the normal to the screen, as shown in Fig. 8. Let the incident wave be

$$\left\{\begin{aligned} \mathscr{E} &= A(S_0, C_0, 0)\exp\left[-jk(C_0x - S_0y)\right] \quad . \quad . \quad (47)\\ H &= \eta A(0, 0, 1)\exp\left[-jk(C_0x - S_0y)\right] \quad . \quad . \quad (48)\end{aligned}\right.$$

so that the aperture distribution in the plane $x = 0$ is

$$\mathscr{E}_y(0, y) = AC_0\exp(jkS_0y)u(-y) \quad . \quad . \quad . \quad (49)$$

From eqns. (37) and (38), the angular spectrum corresponding to an aperture distribution $AC_0u(-y)$ is AC_0/jkS. It follows

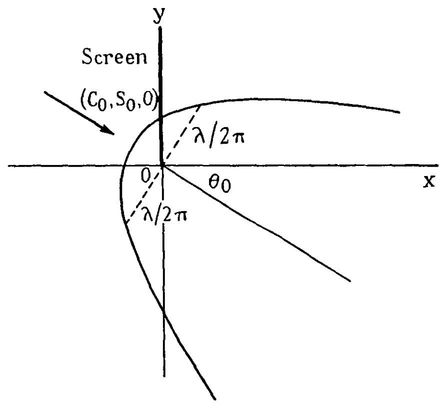

Fig. 8.—Diffraction of plane wave by black screen—oblique incidence.

by the second shift rule in Section 2, using the replacements given by (22) and (23), that the angular spectrum corresponding to the aperture distribution of eqn. (49) is

$$P(S) = \frac{AC_0}{jk(S + S_0)} \quad . \quad . \quad . \quad . \quad . \quad (50)$$

This reduces to eqn. (38) when $S_0 = 0$, $C_0 = 1$. Outside the parabola of semi-latus-rectum $\lambda/2\pi$, shown in Fig. 8, eqn. (50) may be used as a polar diagram giving the edge-wave. Inside the parabola, a formula involving the Fresnel integral [eqn. (44)] must be used, and this is

$$\left\{\begin{aligned} \mathscr{E} &= \frac{1}{\sqrt{2}}AF\left[(S+S_0)\left(\frac{\pi r}{\lambda}\right)^{\frac{1}{2}}\right](-S, C, 0) \\ &\qquad \exp\left\{-j[k(C_0x - S_0y) - \tfrac{1}{4}\pi]\right\} \quad (51) \\ H &= \frac{1}{\sqrt{2}}\eta AF\left[(S+S_0)\left(\frac{\pi r}{\lambda}\right)^{\frac{1}{2}}\right](0, 0, 1) \\ &\qquad \exp\left\{-j[k(C_0x - S_0y) - \tfrac{1}{4}\pi]\right\} \quad (52) \end{aligned}\right.$$

We thus see that, if, in the course of handling angular spectra, we encounter an angular spectrum of the form

$$\frac{1}{S+S_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (53)$$

we shall immediately recognize, by comparison with eqn. (50), that the angular spectrum could be produced by diffraction of a certain plane wave at a certain diffracting edge. We can then write down the corresponding field from eqns. (51) and (52), or from the edge-wave approximation, whichever is more appropriate.

(6) SOME OTHER APPLICATIONS

The angular spectrum (53) has an important application in connection with the Sommerfeld theory[8] of propagation over a plane, imperfectly conducting earth. In that theory, particularly in the version due to Weyl,[9] radiation from the transmitter is expressed as an angular spectrum of plane waves, and, after reflection from the earth, these give rise to an angular spectrum of reflected waves. This reflected angular spectrum involves a singularity of the type (53), and it is this that leads to the complication involved in the theory. It follows from Section 5 that this complication must be expressible as diffraction of a certain plane wave at a certain diffracting edge. This leads to a new approach to the Sommerfeld theory of propagation over a flat, imperfectly conducting earth.[10]

There are also certain problems in connection with ionospheric reflection and scattering, in which the concept of angular spectrum rather than polar diagram is required. These will be dealt with in other papers.

(7) REFERENCES

(1) Ratcliffe, J. A.: *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 28.
(2) Booker, H. G.: *ibid.*, 1947, **94**, Part III, p. 181.
(3) Watson, G. N.: "Theory of Bessel Functions" (Cambridge University Press, 1944), p. 235.
(4) Jeffreys, H. and B. S.: "Methods of Mathematical Physics" (Cambridge University Press, 1946), p. 472.
(5) Baker, B. B., and Copson, E. T.: "Huygens' Principle" (Oxford University Press, 1939), p. 84.
(6) Wood, R. W.: "Physical Optics" (Macmillan, 1934), chap. 7.
(7) Born, M.: "Optik" (Springer, Berlin, 1933), chap. 4.
(8) Sommerfeld, A.: *Annalen der Physik*, 1909, **28**, p. 665.
(9) Weyl, H.: *ibid.*, 1919, **60**, p. 481.
(10) Booker, H. G., and Clemmow, P. C.: See page 18.

CAMBRIDGE RADIO GROUP: CHAIRMAN'S ADDRESS

By C. W. OATLEY, M.A., M.Sc., Associate Member.

"SEMI-CONDUCTORS AND THE RADIO ENGINEER"

(Abstract *of Address delivered at* Cambridge, *11th October*, 1949.)

During the past 50 years the electron theory of metallic conduction has been developed to a stage where it is capable of providing a satisfactory explanation of most of the experimental facts. However, a number of substances, which have been broadly classed as semi-conductors, have properties which can be explained only by an extension of the basic theory. This extension involves a detailed consideration of the permitted energy-levels of electrons in a crystalline solid and of the effect on these levels of impurity atoms in the solid. It is then possible to explain the high resistivities and the large negative temperature coefficients of resistance, which are characteristic of semi-conductors.

The properties of semi-conductors are becoming increasingly important to the radio engineer, and several well-known devices depend upon them. Thermistors are semi-conductors which find application whenever a temperature-dependent resistance is needed. The change of resistance with temperature which they exhibit is many times as great as can be obtained with any pure metal. Rectifiers of the selenium and copper-oxide types depend for their action on the properties of a junction where a semi-conductor is either in contact with a metal or is separated from the metal by a very thin layer of some insulating material. The mechanism of crystal diodes is somewhat similar. Finally, it has recently been found possible to construct crystal triodes, or transistors. These are, at present, in an early stage of development, but they may ultimately replace thermionic valves in many applications.

Mr. Oatley is at the Engineering Laboratory, Cambridge University.

621.396.11:538.566

Paper No. 873
RADIO SECTION

A RELATION BETWEEN THE SOMMERFELD THEORY OF RADIO PROPAGATION OVER A FLAT EARTH AND THE THEORY OF DIFFRACTION AT A STRAIGHT EDGE

By Prof. H. G. BOOKER, M.A., Ph.D., Associate Member, and P. C. CLEMMOW, M.A.

(*The paper was first received* 12*th April, and in revised form* 4*th July*, 1949.)

SUMMARY

A new way of visualizing the Sommerfeld theory of propagation over a flat, imperfectly reflecting earth is presented. The Sommerfeld theory arises because the ray theory of propagation from a source in the presence of a flat, imperfectly reflecting earth is only an approximation. The ray theory involves the assumption that the Fresnel reflection coefficient of the earth does not vary rapidly with angle of incidence, and this assumption is not satisfied for glancing incidence of vertically polarized waves on the earth's surface at broadcasting wavelengths. The main object of the new presentation is to facilitate the solution of problems involving propagation near the surface of the earth partly over land and partly over sea, but these applications are not included in the paper.

It is convenient to think of a two-dimensional problem in which the transmitter is a line source parallel to the earth's surface, having a vertical polar diagram of circular shape. Such a source may be Fourier analysed into plane waves whose directions are distributed in a vertical plane of propagation perpendicular to the line source; the amplitudes of all the plane waves are the same and they are in the same phase at the source. When these waves are reflected from the earth they produce an angular spectrum of reflected waves, the amplitudes and phases of which are determined by the Fresnel reflection coefficient. This angular spectrum could be thought of as arising, in the absence of the earth, from an aperture distribution on the vertical plane through the primary line source. The aperture distribution that produces the angular spectrum of reflected waves in this way is the exact image of the primary source in the imperfectly reflecting earth, and is given by the Fourier transform of the Fresnel reflection coefficient. For a perfectly conducting earth this aperture distribution reduces to a line source identical with the primary source and located at the optical image line. The correction required to this when the earth is not perfectly conducting is mainly the following. An aperture distribution extending indefinitely downwards from the image line must be introduced, and this consists essentially of the aperture distribution produced by diffraction of the Zenneck wave under a screen extending from the image line upwards. The field produced by the primary source in the presence of the imperfectly reflecting earth is thus the field that would be produced with an almost perfectly conducting earth, together with the field arising from diffraction of the Zenneck wave under the image line.

If diffraction of the Zenneck wave under the image line is calculated by the edge-wave approximation we merely arrive at the ray theory of reflection from the earth of radiation from the primary source: the edge wave from the image line, together with the wave from an image in an almost perfectly conducting earth, makes up the wave from the Fresnel image for the imperfectly reflecting earth. But, at broadcasting wavelengths, points close to the earth are often too close to the shadow edge, formed by diffraction of the Zenneck wave under the image line, for application of the edge-wave approximation. We then have to apply the full theory of edge-diffraction based on the Cornu spiral, and this gives the Sommerfeld theory.

(1) INTRODUCTION

The theory of radiation from a radio transmitter over a flat, imperfectly reflecting earth was worked out originally by Sommerfeld,[1] and, since then, other ways of deriving the theory have been given,[2, 3, 4, 5] each bringing out some special feature of the problem. It is the object of this paper to present yet another point of view, which promises to be of practical importance for dealing with situations where the path of propagation lies partly over land and partly over sea.

When first approaching the problem of radiation from a point source T (Fig. 1) in the presence of a flat, imperfectly reflecting

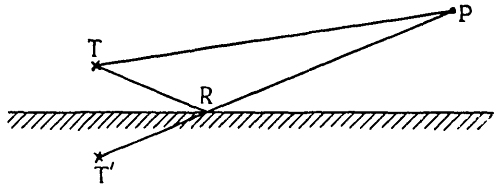

Fig. 1.—Direct and reflected rays.

earth, we are inclined to think that the field strength at a point P may be calculated by combining the results for a direct ray TP and for a ray TRP reflected from the earth, allowing for the Fresnel reflection coefficient of the earth at the appropriate angle of incidence. This is largely true but is misleading in the important practical case where T and P are close to the surface of the earth and the waves are vertically polarized. If T and P are actually on the surface of the earth, the corresponding Fresnel reflection coefficient is -1, which means that the direct and reflected waves neutralize each other and there is no field. This result is true only in the sense that, at a sufficiently large range r, there is no field to the order of $1/r$. But a contribution does exist of the order of $1/r^2$ which is of practical importance. This contribution arises because the Fresnel reflection coefficient referred to above applies only to incident plane waves, and in using it for an incident spherical wave, radiated from T, an approximation is involved. Thus a correction has to be applied to the ray theory, and for communication between points close to the earth's surface the correction is frequently of paramount importance.

The difficulty just described is aggravated in practice by the fact that the (complex) permittivity of the earth is usually large compared with that of the atmosphere. As a result the Fresnel reflection coefficient for vertically polarized waves is practically $+1$ for all directions of incidence on the earth, except nearly glancing directions. As glancing incidence is approached the reflection coefficient swings rapidly from $+1$ to -1. Now the approximation upon which the ray theory is based involves the assumption that the reflection coefficient varies sufficiently slowly with angle of incidence in the neighbourhood of the particular angle of incidence concerned. This assumption, therefore, breaks down for nearly glancing incidence upon the earth of vertically polarized waves, and this is just what is commonly involved in communication between two points close to the earth. As a result, the variation of field strength with range for vertically polarized waves close to the

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Prof. Booker is at Cornell University, New York, and Mr. Clemmow is at Imperial College, London University.

earth's surface often shows no sign of obeying the ray theory for some considerable distance from the transmitter, and even at greater ranges does so only in the sense that it vanishes to the order of $1/r$, leaving correction terms of the order of $1/r^2$ predominant.

The fact that propagation near the surface of an imperfectly reflecting earth involves important corrections to the ray theory implies that the problem is one of diffraction. As a diffraction problem, however, it is of an unusual type. It is therefore a little surprising to find that it is possible to reduce the problem to the well-known one of diffraction at a straight edge. That this is possible and, indeed, convenient is the main point of the paper.

For the most part we shall discuss a two-dimensional instead of a three-dimensional problem, replacing a point source by a line source parallel to the earth's surface. This simplifies the treatment without restricting practical application, because the ratio of field strength in the presence of the earth to that in the absence of the earth is the same for both problems, except within the first wavelength from the source. We shall therefore deal with the two-dimensional problem, express field strength as a ratio to what would be obtained in the absence of the earth, and then use the result in three-dimensional applications.

If T in Fig. 1 now represents a line source parallel to the earth's surface, T′ is the image line. Then T′ in fact forms the diffracting edge involved in the theory. Imagine that the earth is removed and that a plane vertical screen is introduced extending from T′ upwards. Then the theory of radiation from T in the presence of the earth is intimately bound up with diffraction under the edge T′ of a certain plane wave.

The plane wave to be diffracted under T′ is in fact the Zenneck wave. It will be remembered that, prior to Sommerfeld's original paper, Zenneck[6] had published a theory of propagation over the earth's surface. This theory dealt essentially with a vertically polarized plane wave incident upon the earth at the Brewster angle. For this angle there is no reflected wave, but only a wave transmitted into the earth. For the large (complex) dielectric constant usually possessed by the earth, the wave incident in the atmosphere at the Brewster angle is travelling almost horizontally, while the wave refracted into the earth is travelling almost vertically downwards. The Zenneck theory thus presents a wave travelling over the surface of the earth, with energy being abstracted sideways from it by the earth, in the same way as series resistance abstracts energy from a wave travelling along a transmission line. We shall refer to the wave above the earth in the Zenneck theory as the Zenneck wave, which is thus simply a wave incident upon the earth at the Brewster angle. It was no doubt thought at the time of Zenneck's paper that the Zenneck wave represented, to a useful extent, propagation over the earth's surface at some distance from a transmitter. That this is hardly true was one of the points of Sommerfeld's 1909 paper. The Zenneck wave represents radiation from a transmitter so different from those used in practice that simple use of the wave is seriously misleading.

It is a feature of the presentation of the Sommerfeld theory given in this paper that it focuses attention on the extent to which the Zenneck wave is actually involved in propagation over a flat earth. In calculating the field to the right of TT′, in Fig. 2, we remove the earth and insert the vertical screen with its lower edge at T′ as previously described, and then allow the Zenneck wave, with appropriate amplitude, to be incident on the screen from the left. By diffraction under T′ the Zenneck wave makes an important contribution to the required field strength, the remaining contribution being essentially that for a perfectly conducting earth. For an imperfectly reflecting earth the Zenneck wave thus plays an important role. But the wave

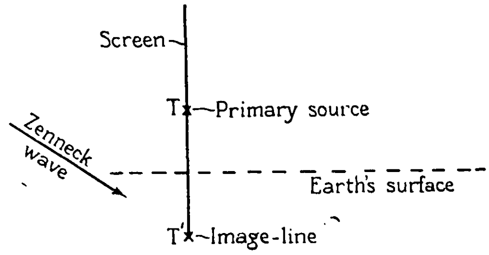

Fig. 2.—Diffraction of Zenneck wave under image line.

is never "seen" direct, and it is only by diffraction under the image line that it makes its contribution.

The Zenneck wave is diffracted at a simple straight edge in Fig. 2, because a two-dimensional problem is being considered. To describe in the same way the three-dimensional problem, the screen in Fig. 3(*a*) (now viewed normally) would be replaced by that in Fig. 3(*b*). The Zenneck wave on the far side of the

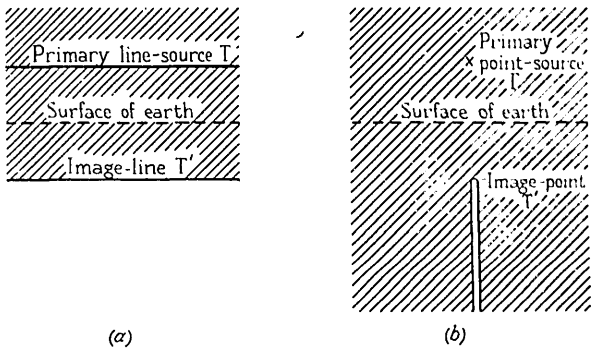

Fig. 3.—Diffracting screens for Zenneck wave.
(*a*) For two-dimensional problem.
(*b*) For three-dimensional problem.

screen would be seen by diffraction through the slot in Fig. 3(*b*), extending from the image point downwards.

If it is true that the field strength due to a line source parallel to an imperfectly reflecting earth may be calculated by adding to the field strength for an almost perfectly conducting earth that due to the Zenneck wave diffracted under the image line, as in Fig. 2, it may be wondered how, in certain circumstances, this is equivalent to the ray theory depicted in Fig. 1. To understand this point, it should be noted that the region in which we have to calculate the field strength in Fig. 2 is to the right of TT′ and above the earth. This is wholly within the shadow of the Zenneck wave cast by the screen extending upwards from T′. Well within the shadow cast by a screen, the diffracted wave may be represented as a wave radiating from the diffracting edge—the so-called edge wave. When the edge wave from T′, due to diffraction of the Zenneck wave, is added to the wave from T′ produced by the image of T in an almost perfectly conducting earth, the combined wave from T′ is simply the ordinary Fresnel image of T in the imperfectly reflecting earth. Thus the ray theory of Fig. 1 is reproduced when diffraction of the Zenneck wave under T′ in Fig. 2 is represented by the edge-wave approximation. However, when the Zenneck wave is travelling almost horizontally, as is usually the case, the edge of the shadow (shadow edge) cast by the screen in Fig. 2 is not far below the position of the earth's surface. Consequently, when evaluating the field near the earth's surface, we are often too close to the shadow edge to use the edge-wave approximation, and diffraction of the Zenneck wave under the image line must then be evaluated by means of the Fresnel integral (Cornu spiral), and this gives precisely the usual Sommerfeld formula for propagation over a flat earth.

The approach to the Sommerfeld theory outlined above is made merely by piecing together some quite well-known results in Fourier analysis. Use of Fourier analysis in the problem was, of course, emphasized originally by Weyl.[2] But in recent times facility with Fourier analysis has become more widespread, particularly in connection with directional aerials, for which the polar diagram and the aperture distribution are Fourier transforms. As a result, the derivation of the Sommerfeld theory given in this paper is now much more obvious than it would have been ten years ago.

In Section 2 the features of Fourier analysis required for the discussion are briefly stated. From these results the exact image of a source, in an imperfectly reflecting earth, is derived in Section 3, and is related to the ray theory of reflection from the earth in Section 4. In Section 5 the special case when the transmitter and receiver are on the earth's surface is studied, and the usual Sommerfeld formula is derived. The advantages of regarding the Sommerfeld extension of the theory for a perfectly conducting earth as arising from diffraction of the Zenneck wave under the image line are illustrated in Section 8, by the difficult problem of propagation along a path partly over land and partly over sea.

(2) SOME RESULTS CONCERNING APERTURE DISTRIBUTION AND ANGULAR SPECTRUM

Suppose that, in a system of Cartesian co-ordinates, the region $x > 0$ is filled with a homogeneous medium having propagation constant k and characteristic admittance η.[7] For a plane wave in the medium, k is the increase of phase lag per unit distance in the direction of propagation, and η is the ratio of magnetic to electric intensity. Suppose that, in the plane $x = 0$ an electromagnetic field is maintained having components along the plane which vary with time, and with position in the plane, in a prescribed manner. The field maintained in the plane $x = 0$ is propagated into the region $x > 0$, and we may think of the plane $x = 0$ as the aperture plane of an aerial system radiating into the region $x > 0$.

The following simplifications will be adopted:

(i) The electromagnetic field will be supposed to vary harmonically in time with a prescribed frequency corresponding to a wavelength λ, and the corresponding complex oscillation-function will be suppressed.

(ii) The field will be supposed two-dimensional and will be taken as independent of the z-co-ordinate.

(iii) The magnetic field will be taken parallel to the z-axis and the electric field parallel to the xy-plane.

As explained in another paper,[8] the field at any point (x, y) in front of the aperture plane $x = 0$ may be expressed as an angular spectrum of plane waves:

$$\left\{\begin{aligned} \mathscr{E}(x, y) &= (1/\lambda)\int_{-\infty}^{\infty} P(S)(-S, C, 0)\exp[-jk(Cx + Sy)]dS/C \quad (1)\\ H(x, y) &= (\eta/\lambda)\int_{-\infty}^{\infty} P(S)(0, 0, 1)\exp[-jk(Cx + Sy)]dS/C \quad (2)\end{aligned}\right.$$

where $k(C, S, 0)$ is the propagation vector of an individual wave of the spectrum.

The aperture distribution is

$$\mathscr{E}_y(0, y) = (1/\lambda)\int_{-\infty}^{\infty} P(S)\exp(-jkSy)dS \quad . \quad . \quad (3)$$

and in terms of this the angular spectrum is given by the Fourier transform:

$$P(S) = \int_{-\infty}^{\infty} \mathscr{E}_y(0, y)\exp(jkSy)dy \quad . \quad . \quad . \quad (4)$$

An example of an aperture distribution in which we shall be interested is an infinite straight slot cut in an infinite, perfectly conducting, metal sheet, with an alternating voltage of amplitude V applied across the slot. If the sheet is in the plane $x = 0$ and the slot along the z-axis, we have

$$\mathscr{E}_y(0, y) = V\delta(y) \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where $\delta(y)$ is the unit impulse-function. The corresponding angular spectrum is

$$P(S) = V \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

At a distance r from the slot, appreciably greater than λ, the wave radiated by the slot is

$$\left\{\begin{aligned} \mathscr{E}(x, y) &= V[-(y/r), (x/r), 0]\exp[-j(kr - \pi/4)]/\sqrt{(r\lambda)} \quad . \quad (7)\\ H(x, y) &= \eta V(0, 0, 1)\exp[-j(kr - \pi/4)]/\sqrt{(r\lambda)} \quad . \quad . \quad (8)\end{aligned}\right.$$

Another example of an aperture distribution in which we shall be interested arises in connection with diffraction of a plane wave by a semi-infinite, "black" screen with a straight edge. Suppose that the screen occupies the portion of the plane $x = 0$ for which $y > 0$, and that the incident wave is given by

$$\left\{\begin{aligned} \mathscr{E} &= A(S_0, C_0, 0)\exp[-jk(C_0x - S_0y)] \quad . \quad . \quad (9)\\ H &= \eta A(0, 0, 1)\exp[-jk(C_0x - S_0y)] \quad . \quad . \quad (10)\end{aligned}\right.$$

so that the aperture distribution in the plane $x = 0$ may be taken as

$$\mathscr{E}_y(0, y) = AC_0\exp(jkS_0y)u(-y) \quad . \quad . \quad . \quad (11)$$

where $u(y)$ is the unit step-function. The corresponding angular spectrum is

$$P(S) = AC_0/jk(S + S_0) \quad . \quad . \quad . \quad . \quad (12)$$

this result being merely a re-interpretation of the well-known frequency spectrum of a carrier wave modulated by a unit step-function. Outside a parabola with focus on the diffracting edge, axis along the shadow edge, and semi-latus-rectum $\lambda/2\pi$, (12) may be used as a polar diagram giving the edge wave. Inside the parabola, however, the concept of polar diagram is inapplicable and (12) has to be regarded strictly as an angular spectrum and substituted into (1) and (2). The resulting field may be evaluated in terms of the Fresnel integral

$$F(u) = \sqrt{(2/\pi)}\int_u^{\infty}\exp(-jv^2)dv \quad . \quad . \quad . \quad (13)$$

the value being

$$\begin{aligned}\mathscr{E} = (1/\sqrt{2})AF[(S + S_0)\sqrt{(\pi r/\lambda)}]&\\ (-S, C, 0)\exp\{-j[k(C_0x - S_0y) - (\pi/4)]\}& \quad . \quad (14)\end{aligned}$$

$$\begin{aligned}H = (\eta/\sqrt{2})AF[(S + S_0)\sqrt{(\pi r/\lambda)}]&\\ (0, 0, 1)\exp\{-j[k(C_0x - S_0y) - (\pi/4)]\}& \quad . \quad (15)\end{aligned}$$

where $r(C, S, 0)$ now denotes the point of observation.

It is an important implication of (12) that if, in the course of handling angular spectra, we come across an angular spectrum of the form

$$1/(S + S_0) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

we shall immediately recognize that the angular spectrum could be produced by diffraction of a certain plane wave at a certain diffracting edge. We can then write down the corresponding field from (14) and (15), or from the edge-wave approximation, whichever is more appropriate.

The results quoted in this Section are explained in detail in Reference (8).

(3) THE EXACT IMAGE OF A SOURCE IN AN IMPERFECTLY REFLECTING PLANE EARTH

Suppose that the atmosphere and the earth may each be regarded as homogeneous media separated by a plane interface $y = 0$, as indicated in Fig. 4. Let k, η be the propagation

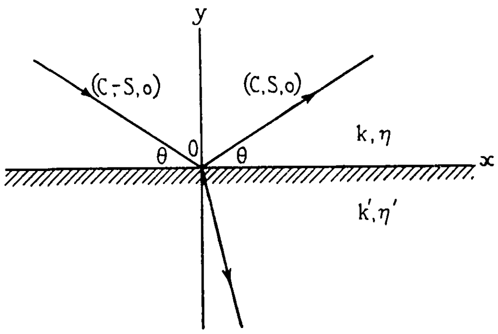

Fig. 4.—Reflection of a vertically polarized plane wave by a plane earth.

constant and characteristic admittance of the atmosphere and k', η' those of the earth. A vertically-polarized plane wave incident upon the earth at an angle of elevation θ ($S = \sin\theta$, $C = \cos\theta$) is

$$\begin{cases} \mathscr{E} = (S, C, 0) \exp[-jk(Cx - Sy)] & (17) \\ H = \eta(0, 0, 1) \exp[-jk(Cx - Sy)] & (18) \end{cases}$$

This gives rise to a reflected wave which may be written

$$\begin{cases} \mathscr{E} = \rho(S)(-S, C, 0) \exp[-jk(Cx + Sy)] & (19) \\ H = \eta\rho(S)(0, 0, 1) \exp[-jk(Cx + Sy)] & (20) \end{cases}$$

where $\rho(S)$ is the Fresnel reflection coefficient of the earth and depends on the angle of elevation of the incident wave. There is also a wave transmitted into the earth, which in practice may be regarded as travelling almost vertically downwards because the complex dielectric constant of the earth is usually quite large. The admittance looking downwards across the earth's surface is therefore η'. That looking downwards in the incident wave is obtained by dividing the horizontal magnetic field strength (18) by the horizontal electric field strength (17) and is η/S. Hence the Fresnel reflection coefficient of the earth is[7]

$$\rho(S) = \frac{\eta' - (\eta/S)}{\eta' + (\eta/S)} \quad . \quad . \quad . \quad . \quad . \quad (21)$$

provided that the transmitted wave is travelling almost vertically downwards. It will be convenient to write (21) as

$$\rho(S) = \frac{S - S_0}{S + S_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

where

$$S_0 = \eta/\eta' \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

The reflection coefficient (22) vanishes when $S = S_0$, so that θ_0 ($\sin\theta_0 = S_0$) corresponds to what is known as the Brewster angle. With η small compared with η', as it usually is in practice, the Brewster angle of elevation is small. Strictly, the Brewster angle of elevation is the angle whose tangent [not *sine* as given by (23)] is η/η'; but for small angles the difference is unimportant. We shall begin by considering an earth whose reflection coefficient is any function of the sine S of angle of elevation, and subsequently pay special attention to the reflection coefficient (22), which covers most practical requirements in connection with the earth.

As described in the Introduction, it will be convenient and adequate to consider two-dimensional radiation from a line source parallel to the earth's surface. For a vertically polarized horizontal line-source we may use an infinitely long horizontal slot cut in an infinite vertical perfectly-conducting sheet, with an alternating voltage of amplitude V across the slot. This situation is depicted in Fig. 5, where the surface of the earth is

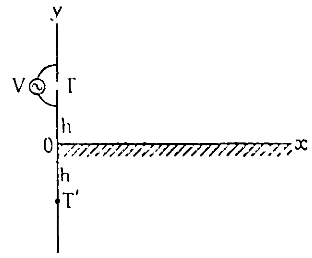

Fig. 5.—Two-dimensional radiation from a horizontal slot parallel to a plane earth.

the plane $y = 0$, and the slot T lies along the line $x = 0$, $y = h$. The image line of the slot in the earth is T′, its location being $x = 0$, $y = -h$. We need consider radiation from the slot only in the region $x > 0$.

If the slot were along the z-axis, the aperture distribution in the plane $x = 0$ would be (5) and the corresponding angular spectrum (6). In fact the slot is shifted along the aperture plane to $x = 0$, $y = h$, so that the aperture distribution is

$$\mathscr{E}_y(0, y) = V\delta(y - h) \quad . \quad . \quad . \quad . \quad (24)$$

By a well-known shift rule,[8] it follows from (6) that the angular spectrum corresponding to the aperture distribution (24) is

$$P(S) = V \exp(jkSh) \quad . \quad . \quad . \quad . \quad . \quad (25)$$

For the angular spectrum into which the given primary source may be analysed, (25) gives the amplitude and phase at the origin of the plane wave whose direction of propagation is $(C, S, 0)$. It follows that the amplitude and phase of the plane wave whose direction of propagation is $(C, -S, 0)$ are given by

$$P(-S) = V \exp(-jkSh) \quad . \quad . \quad . \quad . \quad (26)$$

The corresponding wave reflected from the earth's surface in the direction $(C, S, 0)$ has amplitude and phase at the origin given by

$$P(-S)\rho(S) = V\rho(S) \exp(-jkSh) \quad . \quad . \quad . \quad (27)$$

where $\rho(S)$ is the Fresnel reflection coefficient of the earth. It follows that the angular spectrum of plane waves reflected from the earth as a result of the primary source at $x = 0$, $y = h$ is (27). In order to discover what the earth re-radiates in the presence of the primary source, the field corresponding to the angular spectrum (27) of reflected waves has to be evaluated.

To discuss the angular spectrum (27) of reflected waves it is convenient to shift the origin in Fig. 5 down through a distance h, so that it is on the image line T′, as shown in Fig. 6. This we do by means of the previously mentioned shift rule. The effect is to remove the factor $\exp(-jkSh)$ in (27), leaving the angular spectrum of reflected waves as

$$V\rho(S) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)$$

This means that, in the absence of the earth, a two-dimensional angular spectrum of plane waves, in which the wave travelling at an angle of elevation θ ($\sin\theta = S$) has amplitude and phase given by (28) at the image line T′, reproduces the re-radiation of the earth in the presence of a horizontal slot of voltage V above the earth.

As a simple example of (28), we may consider a perfectly conducting earth, for which $\rho(S)$ is unity. The angular spectrum (28) of reflected waves is then the same as the angular spectrum (6), so that the image consists of the aperture distribution (5) with the origin on the image line. We thus have the well-known result that the image of a slot above a perfectly conducting earth is an identical slot along the image line.

Now, for an imperfectly reflecting earth, the use of images normally leads to the ray theory of reflection from the earth's surface, illustrated in Fig. 1. As explained in the Introduction, this image technique involves an approximation which is liable to break down when transmitter and receiver are sufficiently close to the earth's surface, and this is just the situation that frequently occurs in practice. From what has been said in this Section, however, it is easy to arrive at an exact method whereby the effect of an imperfectly reflecting earth may be represented by means of an image source. We shall now state what is the exact image for a line source parallel to an imperfectly reflecting earth, and then go on, in the next Section, to consider how this exact image is related to the approximate image involved in the ray theory, and when the approximate image fails adequately to represent the exact image.

The exact image of a horizontal primary line-source in a flat imperfectly reflecting earth involves an aperture distribution over the vertical plane through the image line and the primary source. This aperture distribution is calculated from the angular spectrum of reflected waves by substituting for $P(S)$ in (3) the expression (28). The aperture distribution forming the image source is thus the Fourier transform of the angular spectrum (28) of reflected waves. An image consisting of this aperture distribution produces the angular spectrum (28) and therefore represents accurately the re-radiation of the earth in the presence of the primary source. Thus a horizontal slot of unit voltage placed above a plane imperfectly reflecting earth has an exact image in the form of a vertical aperture-distribution, given by the Fourier transform of the earth's Fresnel reflection coefficient.

A more general statement may be made as follows: If any vertical aperture-distribution corresponding to a polar diagram $P(S)$, where S is the sine of the angle of elevation, is placed above a plane earth whose Fresnel reflection coefficient is $\rho(S)$, then re-radiation by the earth is exactly represented by an image in the form of a vertical aperture-distribution, given by the Fourier transform of

$$P(-S)\rho(S) \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

This more general result follows from consideration of the left-hand side of equation (27).

(4) EVALUATION OF THE FIELD REFLECTED BY THE EARTH

Considering a transmitter T in the form of a horizontal slot of voltage V above a flat imperfectly reflecting earth, and taking the z-axis along the image line as indicated in Fig. 6, let us examine the field associated with the angular spectrum (28) of reflected waves. First we will consider the approximate evaluation of this field involved in the ray theory of reflection from an imperfectly reflecting earth. The approximation consists simply in interpreting the angular spectrum (28) of reflected waves as a polar diagram.[8] This means that we regard the wave re-radiated by the earth as a cylindrical wave emanating from the image line and having a vertical polar diagram given by the earth's Fresnel reflection coefficient. Combination of this cylindrical wave from the image line T′ with the cylindrical wave from the primary source T is equivalent to combination of a direct ray from the primary source with a ray reflected from the earth, allowing for the Fresnel reflection coefficient of the earth at the appropriate angle of elevation, as in Fig. 1.

The assumption that the angular spectrum of reflected waves may be regarded simply as a polar diagram is not, however, satisfactory in all practical applications.[8] Let us examine this important point in detail by studying the nature of the angular spectrum of reflected waves, using the approximate expression (22) for the Fresnel reflection coefficient of the earth. Substituting from (22) into (28) we see that the angular spectrum of reflected waves becomes

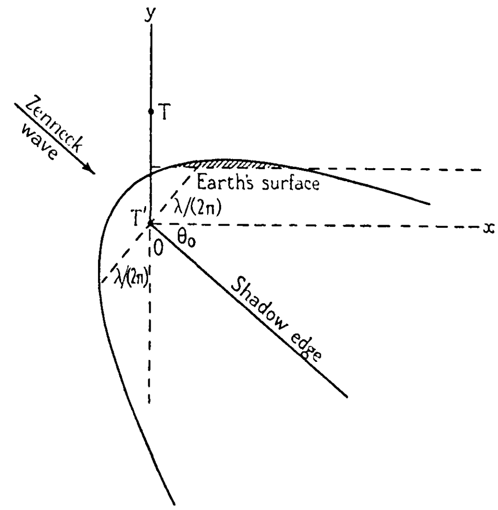

Fig. 6.—Diagram to illustrate the region (shaded) in which the ray theory is inapplicable, for an elevated transmitter.

$$V(S - S_0)/(S + S_0) \quad . \quad . \quad . \quad . \quad . \quad . \quad (30)$$

and this we may split into two angular spectra thus:

$$V - V2S_0/(S + S_0) \quad . \quad . \quad . \quad . \quad (31)$$

Now, according to (5) and (6), the first of these angular spectra corresponds to a slot of voltage V along the z-axis in Fig. 6, i.e. to the image of the primary source in a perfectly conducting earth. The second term in (31) is therefore an angular spectrum which gives the correction to the field for a perfectly conducting earth, required for an imperfectly reflecting earth whose Fresnel reflection coefficient may be taken as (22). If we now concentrate on the correction required to the field for a perfectly conducting earth, the angular spectrum that we have to study is

$$-2VS_0/(S + S_0) \quad . \quad . \quad . \quad . \quad . \quad (32)$$

This expression gives the amplitude and phase at the image line of a plane wave whose angle of elevation is θ ($S = \sin\theta$), θ_0 ($S_0 = \sin\theta_0$) being the earth's Brewster angle of elevation.

We immediately recognize that the angular spectrum (32) is of the same type as (12). It follows that the angular spectrum (32) corresponds to diffraction under the image line T′, in Fig. 6, of a plane wave of the type represented by (9) and (10). According to (23) this wave is travelling in a direction corresponding to the Brewster angle of elevation for the earth, and is therefore the Zenneck wave mentioned in the Introduction. The amplitude and phase of the wave at the image line, obtained by comparing (32) with (12), are represented by the modulus and argument of the complex amplitude

$$A = -2jkVS_0/C_0 \quad . \quad . \quad . \quad . \quad . \quad (33)$$

Substituting this value of A into (9) and (10), we see that the field of the Zenneck wave before diffraction is

$$\begin{cases} \mathscr{E} = -2jkV(S_0/C_0)(S_0, C_0, 0)\exp[-jk(C_0x - S_0y)] & (34) \\ H = -2jk\eta V(S_0/C_0)(0, 0, 1)\exp[-jk(C_0x - S_0y)] \quad . & (35) \end{cases}$$

the z-axis being along the image line. Thus the correction required to the field for a perfectly conducting earth is the field produced by diffraction under the image line of the Zenneck wave (34) and (35).

We are now in a position to see clearly why the ray theory of reflection from the earth's surface has its limitations, and what must be done to remove them. The complete image of a horizontal line-source in a flat imperfectly reflecting earth consists of the image appropriate to a perfectly conducting earth, together with an aperture distribution extending from the image line vertically downwards and given by (11), with the value (33) for A. The exact image of a source in an imperfectly reflecting earth is thus infinitely wide, vertically. Great care is therefore necessary in interpreting the angular spectrum as a polar diagram,[8] and it is precisely this approximation that is made in the ray theory. No difficulty occurs for a perfectly conducting earth, because then the characteristic admittance of the earth is infinite, making the sine of the Brewster angle, S_0, zero by (23), and this makes the amplitude of the Zenneck wave (34) and (35) zero. For an imperfectly reflecting earth, however, the amplitude of the Zenneck wave to be diffracted under the image line is not zero, and the image involves an aperture distribution, infinitely wide, extending indefinitely downwards from the image line.

Fig. 6 illustrates diffraction of the Zenneck wave (34) and (35) under the image line. The associated shadow edge slopes downwards from the image line at the Brewster angle, and the whole of the region above the earth's surface, where the diffracted field must be calculated, is within the shadow. It is clear that in much of this region the edge-wave approximation to diffraction of the Zenneck wave under the image line may be used. But the edge-wave approximation is one in which the angular spectrum (32), or (12) with the value (33) for A, is merely interpreted as a polar diagram. This is simply equivalent to interpreting (31), and therefore (30), as a polar diagram (since the angular spectrum V of reflected waves for a perfectly conducting earth may always be interpreted as a polar diagram). Thus use of the edge-wave approximation for diffraction of the Zenneck wave (34) and (35) under the image line is merely equivalent to interpretation of the angular spectrum of reflected waves as a polar diagram, and this, as we saw at the beginning of this Section, is what leads to the ray theory. It follows that the ray theory of reflection from the earth is unsatisfactory under those conditions in which the edge-wave approximation is not an adequate description of diffraction of the Zenneck wave (34) and (35) under the image line.

Now, we know the conditions under which the edge-wave approximation may be used.[8] We draw, in Fig. 6, a parabola with focus on the image line, with axis along the shadow edge cast by diffraction of the Zenneck wave under the image line, and with semi-latus-rectum $\lambda/2\pi$. It is the region inside this parabola in which the edge-wave approximation is unsatisfactory. If, therefore, the parabola extends above the surface of the earth, there is a region above the surface and within the parabola where the edge-wave approximation, and consequently the ray theory, is unsatisfactory. In such a region, shown shaded in Fig. 6, diffraction of the Zenneck wave under the image line must be calculated using the formulae (14) and (15) involving the Fresnel integral (13), with the value (33) for A. Thus, radiation from a horizontal slot of voltage V above a flat imperfectly reflecting earth is obtained by adding to the field for a perfectly conducting earth the field

$$\left\{\begin{aligned} \mathscr{E} &= -\sqrt{(2)}jkV(S_0/C_0)F[(S+S_0)\sqrt{(\pi r/\lambda)}] \\ &\qquad (-S, C, 0)\exp\{-j[k(C_0x - S_0y) - \pi/4]\} \quad . \quad (36) \\ H &= -\sqrt{(2)}jk\eta V(S_0/C_0)F[(S+S_0)\sqrt{(\pi r/\lambda)}] \\ &\qquad (0, 0, 1)\exp\{-j[k(C_0x - S_0y) - \pi/4]\} \quad . \quad (37) \end{aligned}\right.$$

where the axes are as indicated in Fig. 6, and the Fresnel integral $F(u)$ is given by (13). But above the parabola drawn in Fig. 6 the argument of the Fresnel integral in (36) and (37) becomes large, and the field then becomes substantially identical with that given by the ray theory.

It is the necessity, in the shaded region of Fig. 6, of calculating the field produced by diffraction of the Zenneck wave under the image line by using the complete expressions (36) and (37) involving the Fresnel integral, instead of by means of the edge-wave approximation, that constitutes, from the point of view adopted in this paper, the Sommerfeld theory of propagation over a flat, imperfectly reflecting earth, as distinct from the ray theory.

(5) TRANSMITTER AND RECEIVER AT THE EARTH'S SURFACE

Let us examine the implications of (36) and (37) for the special case treated originally by Sommerfeld,[1] namely that for which transmitter and receiver are at the earth's surface. Both the primary line-source and the image line then coincide at the earth's surface; both lie along the z-axis at the earth's surface, and Fig. 6 is replaced by Fig. 7. For a receiver close to the

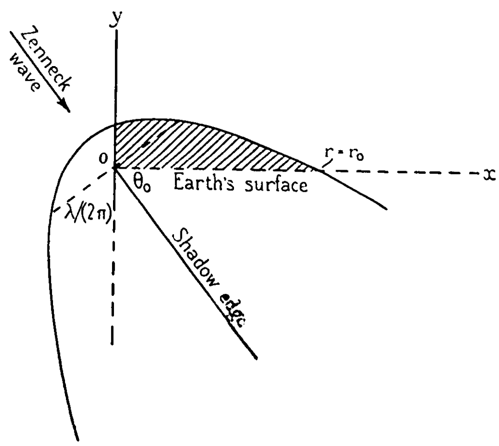

Fig. 7.—Diagram to illustrate the region (shaded) in which the ray theory is inapplicable, when the transmitter height is zero.

earth's surface the ray theory is inapplicable for all ranges up to a certain characteristic range r_0, at which the parabola (focus at 0, axis along shadow edge formed by diffraction of the Zenneck wave under 0, semi-latus-rectum $\lambda/(2\pi)$ intersects the earth's surface. Beyond the range r_0 the field begins to resemble that appropriate to the ray theory in a way that we shall investigate.

With transmitter and receiver at the earth's surface, we put

$$x = r,\ y = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)$$

$$S = 0,\ C = 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad (39)$$

in (36) and (37). The electric field (36) is then vertical, and its value is

$$\mathscr{E} = -\sqrt{(2)}jkV(S_0/C_0)F[S_0\sqrt{(\pi r/\lambda)}]\exp[-j(kC_0r - \pi/4)] \quad . \quad . \quad . \quad . \quad (40)$$

We also recall that S_0, the sine of the Brewster angle of elevation for the earth, is small, and so write

$$C_0 = 1 - \tfrac{1}{2}S_0^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (41)$$

in the exponential phase-factor in (40), and $C_0 = 1$ in the coefficient of the Fresnel integral. The contribution to the vertical electric field strength near the earth's surface, at range r, due to diffraction of the Zenneck wave under 0, is therefore

$$\mathscr{E} = -\sqrt{(2)}jkVS_0F[S_0\sqrt{(\pi r/\lambda)}] \exp[jS_0^2(\pi r/\lambda)]\exp[-j(kr-\pi/4)] \quad . \quad (42)$$

To this must be added the field appropriate to a slot of voltage V, at 0, in the presence of a perfectly reflecting earth. This is double the field in the absence of the earth, and is therefore, from (7),

$$\mathscr{E}_\infty = 2V\exp[-j(kr-\pi/4)]/\sqrt{(r\lambda)} \quad . \quad . \quad (43)$$

Adding (42) and (43) we see that the complete vertical electric field near the earth's surface at range r is

$$\mathscr{E} = \{2V\exp[-j(kr-\pi/4)]/\sqrt{(r\lambda)}\} \{1 - j\pi S_0\sqrt{(2r/\lambda)}F[S_0\sqrt{(\pi r/\lambda)}]\exp[jS_0^2(\pi r/\lambda)]\} \quad . \quad (44)$$

Dividing (44) by (43), we obtain the ratio of the vertical electric field, at range r, near the surface of an imperfectly reflecting earth, to that which would exist for a perfectly conducting earth, namely

$$\mathscr{E}/\mathscr{E}_\infty = 1 - j\pi S_0\sqrt{(2r/\lambda)}F[S_0\sqrt{(\pi r/\lambda)}]\exp[jS_0^2(\pi r/\lambda)] \quad . \quad (45)$$

In this form there is no need to retain the specialization of a line source along the earth's surface. By an application of the method of steepest descent[9, 10] we may extend the validity of (45) to almost any practical transmitter close to the earth's surface. This means that, if we substitute for $\mathscr{E}_\infty$ in (45) the field at range r (appreciably greater than λ) near a perfectly reflecting earth for the transmitter in which we are interested, then $\mathscr{E}$ will be the field for the same transmitter at the same range near the surface of an imperfectly reflecting earth, the sine of whose (small) Brewster angle of elevation is S_0, given by (23).

The parabola in Fig. 7 is the curve on which the argument of the Fresnel integral in (45) is unity, so that the range r_0 at which the parabola crosses the earth's surface is given by

$$S_0\sqrt{(\pi r_0/\lambda)} = 1 \quad . \quad . \quad . \quad . \quad . \quad (46)$$

Hence

$$r_0 = \lambda/(\pi S_0^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad (47)$$

$$= (\lambda/\pi)(\eta'/\eta)^2 \quad . \quad . \quad . \quad . \quad (48)$$

from (23). In terms of r_0 we may rewrite (45) as

$$\mathscr{E}/\mathscr{E}_\infty = 1 + \sqrt{(2\pi r/r_0)}F[\sqrt{(r/r_0)}]\exp\{j[(r/r_0) - \pi/2]\} \quad . \quad (49)$$

The first term of (49) corresponds to the field for a perfectly conducting earth, and the second term is the correction for an imperfectly reflecting earth arising from diffraction of the Zenneck wave under 0 in Fig. 7.

Points on the surface of the earth in Fig. 7 for which r is large compared with r_0 are well outside the parabola, and at these points the ray theory should be applicable. Now, for glancing incidence on the earth's surface, the Fresnel reflection coefficient is -1, from (22). Hence the direct wave and the wave reflected from the earth's surface are equal in amplitude and opposite in phase, and therefore the field on the earth's surface according to the ray theory is zero. Thus, to agree with the ray theory outside the parabola in Fig. 7, the second term in (49) should become -1 when r is large compared with r_0. To verify this we need the asymptotic formula for the Fresnel integral (13), applicable when the argument u is large compared with unity. This is[11]

$$F(u) \sim \frac{1}{\sqrt{(2\pi)}}\frac{\exp[-j(u^2+\pi/2)]}{u}\left[1 + \frac{j}{2u^2} + O\left(\frac{1}{u^4}\right)\right] \quad . \quad (50)$$

Using the first term only of (50), it is easily verified that the second term of (49) is practically -1 when r is large compared with r_0, thus agreeing with the ray theory outside the parabola in Fig. 7.

It should be noted that, in the region where r is large compared with r_0, the field is only zero at the earth's surface to the order in $1/r$ involved in normal calculations of distant field strength (for example, in the presence of a perfectly conducting earth, or in the absence of the earth). To a higher order in $1/r$ there is a residual field which is of some practical importance, and which may be calculated by using the second term in (50). The residual field at the earth's surface, when r is large compared with r_0, is given by

$$\mathscr{E}/\mathscr{E}_\infty = \tfrac{1}{2}(r_0/r)\exp(-j\pi/2) \quad . \quad . \quad . \quad (51)$$

For values of r small compared with r_0 we are inside the parabola in Fig. 7. Consequently (49) does not here agree with the ray theory, which would still predict a field strength small compared with that appropriate to a perfectly conducting earth. Indeed, for r small compared with r_0, the second term in (49) is small on account of the factor $\sqrt{(r/r_0)}$, so that the field is practically identical with that for a perfectly conducting earth, instead of small as suggested by the ray theory. The extent of the departure from the field for a perfectly conducting earth, at ranges small compared with r_0, is obtained by putting $r = 0$ in (49), except in the factor $\sqrt{(r/r_0)}$. Using the known value of $F(0)$, we obtain

$$\mathscr{E}/\mathscr{E}_\infty = 1 + \sqrt{(\pi r/r_0)}\exp(-j3\pi/4) \quad . \quad . \quad (52)$$

when r is small compared with r_0.

The complete formula (49), giving the transition from (52) —r small compared with r_0—to (51)—r large compared with r_0—may be put into the form originally obtained by Sommerfeld[1] by substituting for $F(u)$ the integral expression (13). We thus obtain

$$\mathscr{E}/\mathscr{E}_\infty = 1 + 2\sqrt{(r/r_0)}\exp\{j[(r/r_0) - \pi/2]\}\int_{\sqrt{(r/r_0)}}^{\infty}\exp(-jv^2)dv \quad . \quad . \quad . \quad . \quad (53)$$

and this leads, without difficulty, to the Sommerfeld formula in its usual form by employing

$$-j(r/r_0) \quad . \quad . \quad . \quad . \quad . \quad . \quad (54)$$

as a variable defining range.

The normalized range defined by (54) is what Sommerfeld called the numerical distance, and it may be wondered why he defined it as $-j(r/r_0)$ instead of simply r/r_0. The reason is that Sommerfeld arranged his notation to be convenient when the earth is acting as an imperfect conductor rather than as a "lossy" dielectric, whereas the reverse has been done in (49) and (53). Let us examine the expression (48) for r_0 as a function of the electrical properties of the earth. If κ and σ are the permittivity and conductivity of the earth, while κ_0 and zero are those of the atmosphere, then, at angular frequency ω, the ratio of the characteristic admittance η' of the earth to that for the atmosphere is given by[7]

$$\left(\frac{\eta'}{\eta}\right)^2 = \frac{\kappa - j(\sigma/\omega)}{\kappa_0} \quad . \quad . \quad . \quad . \quad . \quad (55)$$

The earth acts as a "lossy" dielectric or as an imperfect conductor according as ω is greater or less than σ/κ, which frequency is the reciprocal of the earth's time-constant. For an earth acting mainly as a "lossy" dielectric we have approximately

$$(\eta'/\eta)^2 = \kappa/\kappa_0 \quad (\omega \gg \sigma/\kappa) \quad . \quad . \quad . \quad . \quad (56)$$

and for an earth acting mainly as an imperfect conductor we have

$$(\eta'/\eta)^2 = -j\sigma/\kappa_0\omega \quad (\omega \ll \sigma/\kappa) \quad . \quad . \quad . \quad (57)$$

Substituting from (56) and (57) into (48) we have for dielectric behaviour of the earth ($\omega \gg \sigma/\kappa$)

$$r_0 = (\lambda/\pi)(\kappa/\kappa_0) \quad . \quad . \quad . \quad . \quad . \quad (58)$$

and for conducting behaviour ($\omega \ll \sigma/\kappa$)

$$r_0 = -j|r_0| \quad . \quad . \quad . \quad . \quad . \quad . \quad (59)$$

where $$|r_0| = (\lambda/\pi)[\sigma/(\kappa_0\omega)] \quad . \quad . \quad . \quad . \quad . \quad (60)$$

From (54) and (59) we see that, when the earth is acting mainly as a conductor, Sommerfeld's numerical distance (54) is real and equal to $r/|r_0|$, so that (51) and (52) are then more conveniently written

$$\mathscr{E}/\mathscr{E}_\infty = \tfrac{1}{2}(|r_0|/r)\exp(-j\pi) \quad (r \gg |r_0|) \quad . \quad . \quad (61)$$

$$\mathscr{E}/\mathscr{E}_\infty = 1 + \sqrt{(\pi r/|r_0|)}\exp(-j\pi/2) \quad (r \ll |r_0|) \quad . \quad (62)$$

We may notice that the above theory for an earth that acts mainly as a conductor is of considerably greater practical importance than for a dielectric earth. Consider the following numerical values:[7]

$$\textit{Land}\begin{cases}\kappa = 10\kappa_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (63)\\ \sigma = 10^{-3}\ \text{mho/metre} \quad . \quad . \quad . \quad (64)\end{cases}$$

$$\textit{Sea}\begin{cases}\kappa = 81\kappa_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (65)\\ \sigma = 4\ \text{mhos/metre} \quad . \quad . \quad . \quad . \quad (66)\end{cases}$$

Combining these with the value

$$\kappa_0 = 8{\cdot}854 \times 10^{-12}\ \text{farad/metre} \quad . \quad . \quad . \quad (67)$$

we see that the frequency σ/κ is roughly

Land 2 Mc/s . . . . . (68)

Sea 1 000 Mc/s . . . . . (69)

It is only above these respective frequencies that the earth may be regarded as a dielectric. Moreover, from (58), (63), (65) and (67), the appropriate value of r_0 is

Land 3λ . . . . (70)

Sea 25λ . . . . (71)

Combining (70), (71) with (68), (69) we see that r_0 is well under a kilometre whenever the earth is behaving as a dielectric, so that at ranges of practical interest the ray theory may nearly always be regarded as applicable, with the small residual field at the earth's surface being given by (51). On the other hand, at frequencies less than (68) and (69) respectively, the earth behaves mainly as a conductor and we deduce from (60), (64), (66) and (67) that

$$\textit{Land} \quad |r_0| = 2/f^2\ \text{kilometres} \quad . \quad . \quad . \quad . \quad (72)$$

$$\textit{Sea} \quad |r_0| = 6\,800/f^2\ \text{kilometres} \quad . \quad . \quad . \quad (73)$$

where f is frequency in megacycles per second. The values in (72) and (73) of $|r_0|$, at frequencies less than (68) and (69) respectively, can easily exceed distances of practical interest, so that, for a conducting earth, the whole range of r in formula (49), from values of r large compared with $|r_0|$ [equation (61)] down to values of r small compared with r_0 [equation (62)], is required.

(6) RÉSUMÉ OF SECTIONS 3, 4, AND 5

(*a*) To a good degree of approximation, radiation from a line source parallel to an imperfectly reflecting earth may be described in terms of radiation in the presence of a perfectly conducting earth, together with a field arising from diffraction of the Zenneck wave under the image line.

(*b*) When diffraction of the Zenneck wave under the image line may be satisfactorily described in terms of the edge-wave approximation, then the ray theory of reflection from an imperfectly reflecting earth is applicable, and this occurs above a parabola having focus on the image line, axis along the shadow edge, and semi-latus-rectum $\lambda/(2\pi)$.

(*c*) When diffraction of the Zenneck wave under the image line cannot be satisfactorily described in terms of the edge-wave approximation, the full diffraction-theory employing the Fresnel integral must be used, and this is equivalent to the Sommerfeld formula for propagation over a flat imperfectly reflecting earth.

(*d*) There is no difficulty in making the appropriate modifications of interpretation when the earth is not a pure dielectric and when the source is not a line source.

(7) SOME FURTHER CONSIDERATIONS

There are some further points that should be made in connection with the theory outlined in Sections 3, 4 and 5:

(*a*) We have deliberately concentrated upon vertical polarization. There would be a corresponding theory for horizontal polarization if the permeabilities of air and earth were as different as their permittivities. In fact, however, the Fresnel reflection coefficient of the earth for horizontally-polarized waves may nearly always be taken as -1, and the ray theory applied.

(*b*) Equation (49) gives the vertical electric field close to the earth's surface, but there is also a small horizontal component in the direction of propagation. It is approximately $1/\eta'$ times the horizontal magnetic field strength, since the field in the earth is practically a plane wave travelling vertically downwards and abstracting energy sideways from the wave travelling almost horizontally above the earth's surface. Thus, by definition of characteristic admittance, the vertical and horizontal components of electric field strength just above the earth's surface are $1/\eta$ and $1/\eta'$ times the horizontal magnetic field strength. The small horizontal electric field at the earth's surface in the direction of propagation, expressed as a ratio to the vertical component, is therefore η/η', the value of which may be deduced from (55).

(*c*) In connection with diffraction of the Zenneck wave under the image line, there is a point of mathematical difficulty arising when the earth is not a pure dielectric. The wave to be diffracted under the image line is the wave above the earth's surface in the Zenneck theory, and this increases exponentially downwards when the conductivity of the earth is not zero.[6] The aperture distribution therefore increases exponentially with distance downwards from the image line, and this leads to certain mathematical difficulties of convergence. These may be avoided, however, by an application of Babinet's principle: we think of diffraction of the Zenneck wave over the image line, and combine this suitably with the undiffracted Zenneck wave, thereby arriving at the same answer without having encountered a divergent integral. It will be remembered that an error of sign crept into Sommerfeld's original paper.[1] Though not present in a later work by the same author,[12] this error caused much misunderstanding,[13, 14, 15, 16] and it is perhaps worth mentioning that it corresponds to diffracting the Zenneck wave on the wrong side of the image line. This explains why the incorrect curves of Rolf,[17] based on the formula with the wrong sign, showed oscillations of field strength, these oscillations corresponding to the well-known oscillations on the illuminated side of the edge of a shadow, arising from interference between the edge wave and the undiffracted wave.

(*d*) In Sections 4 and 5 we employed the approximate form (22) for the Fresnel reflection coefficient of the earth. This form is sufficiently accurate for most practical purposes, and it simplifies the discussion. But the methods, in fact, apply to the accurate form of the Fresnel reflection coefficient. The accurate form contains a simple pole of the type (16) which corresponds to diffraction of the Zenneck wave under the image line. On separating the pole from the remainder of the Fresnel reflection coefficient [cf. equation (31)], we find that the remainder does not correspond simply to the image in a perfectly conducting earth, as was the case for (22). The corresponding angular spectrum may, however, be satisfactorily interpreted as a polar diagram, and this is all that is necessary for arriving at a useful answer.

(*c*) Production of curves exhibiting propagation over a flat imperfectly reflecting earth involves tabulation of the Fresnel integral (13). For a dielectric earth we merely require values of the Fresnel integral for positive real values of the argument u, and these are completely given by the Cornu spiral. To deal with earth of any conductivity, however, we require values of the Fresnel integral throughout the region $0 \leqslant \arg u \leqslant \frac{1}{4}\pi$ in the complex plane. Tables of the Fresnel integral in this region have been prepared for values of $|u|$ between 0 and 1. Norton's curves of propagation over a flat imperfectly reflecting earth[18] constitute a way of plotting the Fresnel integral convenient for the problem concerned. In particular his curves for propagation over a flat, purely dielectric earth are simply a re-plot of the Cornu spiral, well known in optics.

(8) PROPAGATION PATHS PARTLY OVER LAND AND PARTLY OVER SEA

The foregoing investigation of the Sommerfeld theory was undertaken with a view to finding a solution of the problem of propagation along a path lying partly over one type of surface of the earth and partly over another, land and sea being the obvious practical examples. This apparently simple problem has so far eluded satisfactory solution;[19, 20] plausible answers that have been suggested often fail to satisfy even approximately the law of reciprocity, which states that the answer must be unaffected if transmitter and receiver are interchanged. It is not intended to describe in this paper the solution of the problem that has been obtained, but sufficient will be said in this Section to indicate the advantage of the treatment of the Sommerfeld theory given in the preceding Sections.

We suppose, for simplicity, that the land and sea are at the same level (though there is no difficulty in dealing with a cliff edge if desired). This is indicated in Fig. 8, the coast line C

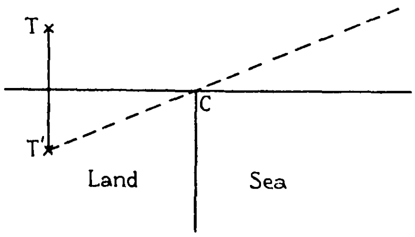

Fig. 8.—Radiation from a transmitter over land and sea.

being straight and normal to the diagram. Let T be a transmitter which we suppose is located over the land, but which might equally well be over the sea. Let us suppose for simplicity that T is a horizontal line-source parallel to the coast line; there is no particular difficulty in converting to a point source. T′, in Fig. 8, is the image line of T in the surface of the earth.

The first point to notice about the problem depicted in Fig. 8 is that refraction below the earth's surface at the interface between land and sea is unlikely to be of importance in practice: the refracted waves in both land and sea are travelling almost vertically downwards and are rapidly attenuated. It is obviously reasonable to assume that electrons vibrating near the surface of the land are doing so in substantially the same way as if the land extended indefinitely in a horizontal direction. The same assumption may (with less certainty) be made about the sea. This is equivalent to saying that the wave reflected from the earth is a combination of:

(*a*) A wave from the image in the land (as though the land extended indefinitely) diffracted over the coast line.

(*b*) A wave from the image in the sea (as though the sea extended indefinitely) diffracted under the coast line.

These two waves, added to the direct wave from the primary source, should give a useful first approximation to the field existing above both land and sea.

If we can assume that ray theory is applicable to reflection both from the land and from the sea, application of the above method is straightforward. But the cases of greatest interest are those in which the ray theory is inapplicable and the Sommerfeld theory must be used. We need to know, in fact, the exact images in both land and sea, not the approximate ones on which the ray theory is based. These exact images are precisely what have been calculated in Section 3. It follows that the field produced above the surface of the earth in Fig. 8 is, to a first approximation, a combination of:

(i) The field for a perfectly conducting earth of infinite extent.

(ii) The field formed by the Zenneck wave appropriate to the land, diffracted under the image line T′ and over the coast line C.

(iii) The field formed by the Zenneck wave appropriate to the sea, diffracted under the image line T′ and under the coast line C. The amplitudes of the two Zenneck waves are determined by (33), using the appropriate Brewster angle of elevation in each case.

It is apparent that each Zenneck wave is diffracted by two straight edges in succession. This involves certain difficulties into which we shall not enter.

It may be noticed that the method outlined above for handling propagation across a coast line is quite different from that suggested by Ratcliffe.[21] Ratcliffe points out that for a low-level transmitter and receiver above a dielectric earth, equations (51) and (52) imply a phase slip of $\frac{1}{2}\pi$ in going from zero range to a range appreciably greater than r_0, while for a conducting earth the phase slip is π in accordance with (61) and (62). He interprets this as a small but significant reduction in phase velocity, and then goes on to regard coastal phenomena as ordinary optical refraction using different phase velocities over land and sea. In particular, for oblique propagation across the coast line, he interprets the phenomenon of so-called "coastal refraction"[22, 23] as arising from the difference of phase velocity on the two sides of the coast, making the assumption that the phase velocity along the coast must be the same on either side. But, according to the view outlined above, coastal phenomena have nothing to do with refraction and are, in fact, diffraction phenomena. A field appropriate to the land fades over continuously to one appropriate to the sea as the shadow edge, formed by T′C produced (Fig. 8), is crossed. There is no reason why the component of phase velocity parallel to the coast should not change smoothly from one value to another as this shadow edge is crossed, and this seems to destroy the basis of the refraction argument. It appears more likely that so-called coastal refraction is in reality a diffraction phenomenon, confined mainly to a region within $\lambda/(2\pi)$ of the coast, and arising in the manner suggested by Barfield.[24]

(9) REFERENCES

(1) SOMMERFELD, A.: *Annalen der Physik*, 1909, **28**, p. 665.
(2) WEYL, H.: *ibid.*, 1919, **60**, p. 481.
(3) VAN DER POL, B., and NIESSEN, K. F.: *ibid.*, 1930, **6**, p. 273.
(4) NIESSEN, K. F.: *ibid.*, 1933, **18**, p. 893.
(5) VAN DER POL, B.: *Physica*, 1935, **2**, p. 843.
(6) ZENNECK, J. A.: *Annalen der Physik*, 1907, **23**, p. 846.
(7) BOOKER, H. G.: *Journal I.E.E.*, 1947, **94**, Part III, p. 181.
(8) BOOKER, H. G., and CLEMMOW, P. C.: See page 11.
(9) JEFFREYS, H. and B. S.: "Methods of Mathematical Physics" (Cambridge University Press, 1946), p. 472.
(10) WATSON, G. N.: "Theory of Bessel Functions" (Cambridge University Press, 1944), p. 235.
(11) JAHNKE, E., and EMDE, F.: "Funktionentafeln" (B. G. Teubner, Leipzig, 1938,) p. 35.
(12) SOMMERFELD, A.: *Annalen der Physik*, 1926, **81**, p. 1135.
(13) NORTON, K. A.: *Nature*, 1935, **135**, p. 954.
(14) BURROWS, C. R.: *ibid.*, 1936, **138**, p. 284.
(15) BURROWS, C. R.: *Proceedings of the Institute of Radio Engineers*, 1937, **25**, p. 219.
(16) WISE, W. H.: *Bell System Technical Journal*, 1937, **16**, p. 101.
(17) ROLF, B.: *Proceedings of the Institute of Radio Engineers*, 1930, **18**, p. 391.
(18) NORTON, K. A.: *ibid.*, 1941, **29**, p. 623.
(19) GRÜNBERG, G.: *Physical Review*, 1943, **63**, p. 185.
(20) MILLINGTON, G.: *Proceedings I.E.E.*, 1949, **96**, p. 53.
(21) RATCLIFFE, J. A.: *Proceedings of the Institute of Radio Engineers*, 1947, **35**, p. 938.
(22) ECKERSLEY, T. L.: *Radio Review*, 1920, **1**, p. 421.
(23) SMITH-ROSE, R. L.: *Nature*, 1925, **116**, p. 426.
(24) BARFIELD, R. H.: *ibid.*, 1925, **116**, p. 498.

SOUTH MIDLAND RADIO GROUP: CHAIRMAN'S ADDRESS

By K. R. STURLEY, Ph.D., B.Sc., Member.

"PROBLEMS IN BROADCAST TRANSMISSION AND RECEPTION"

(ABSTRACT *of Address delivered at* BIRMINGHAM, *26th September*, 1949.)

Before considering the problems encountered in broadcasting it is necessary to have a clear picture of the chain of processes. The transmitting chain involves the studio, microphone, audio-frequency amplifiers and programme-distributing system, the recording apparatus, Post Office lines, transmitter and aerial. The receiving chain consists of the aerial, r.f. amplifier, frequency changer, i.f. amplifier, detector, a.f. amplifier and loudspeaker At the present time the weakest links are at the end of each chain, namely the studio and the loudspeaker.

When a programme originates from an enclosed space, the sound received by the microphone is modified by reflections from the boundaries of that space and from any large objects in it. In the early days, attempts were made to suppress studio acoustics by heavy draping, but the result was unsatisfactory for such studios had a large output at low frequencies due to the selective absorption of the draping. The present-day trend is to try to control studio acoustics so that artistic presentation of the programme is improved. Reverberation time has not proved an entirely satisfactory measure of studio performance, and it is being supplemented by examination with pulsed tone. Examples of acoustic effects are easily obtained by recording speech in the open air (little reverberation), in a studio (more or less controlled reverberation) and in a reverberant room (uncontrolled reverberation).

The microphone can introduce all forms of distortion. The early carbon types were lacking in bass response and had resonances in the middle-frequency range; a later development, the Reisz, was a great improvement but had a low signal/noise ratio and tended to emphasize sibilants. The moving-coil and ribbon microphones now used in broadcasting represent a considerable advance. The ribbon has a good frequency response and has the advantage, exploited in drama, of an approach to a figure-of-eight directional diagram with very much reduced response at the sides. The moving-coil microphone is practically omni-directional. Developments are taking place towards the production of a unidirectional condenser microphone, which has advantages when outside broadcasts from theatres are required.

Dr. Sturley is with the British Broadcasting Corporation.

The chief problem in the a.f. amplifiers and programme distribution system is to preserve good signal/noise ratio and to reduce attenuation and intermodulation distortion to negligible proportions. The majority of broadcast recordings are carried on discs, and one of the main troubles is to preserve the h.f. response as the centre of the disc is approached. Increased loading at the recording cutter point and failure of the recording head to track the reduced wavelength cause appreciable h.f. loss, which is only partially counteracted by equalization circuits in the recorder.

At the transmitter, constancy of carrier frequency to better than 1 part in 10^6 presents no difficulty. Over-modulation must not be allowed to occur because, apart from possible damage to transmitter apparatus, it causes harmonic sidebands to be produced and these spread over into adjacent channels. When a common aerial is used for two different programmes, very careful filtering is necessary to prevent one modulated carrier from entering the r.f. stages of the other transmitter, where cross-modulation can take place.

At the receiver, signal/noise ratio is an important factor, and a satisfactory value can generally be achieved by correct siting of the aerial and by including a r.f. amplifier before the frequency changer. In detection, the circuit constants (C and R) must be so proportioned that the voltage across the detector load resistance is able to follow the r.f. peaks of modulation when the r.f. envelope voltage is falling. A wrong value of a.c.-to-d.c. load can also cause distortion due to the d.c. charge on the coupling capacitor between detector load resistance and a.f. stage. Probably one of the weakest links in the chain of reception is the loudspeaker, which is rarely able to reproduce faithfully transient electric signals applied to its input.