NEW YORK UNIVERSITY

WASHINGTON SQUARE COLLEGE OF ARTS AND SCIENCE

MATHEMATICS RESEARCH GROUP

RESEARCH REPORT No. EM-50

DIFFRACTION THEORY

A CRITIQUE OF SOME RECENT DEVELOPMENTS

by

C. J. BOUWKAMP

CONTRACT No. AF-19(122)-42

Declassified in Part - Sanitized Copy Approved for Release @ 50-Yr 2013/07/17 : CIA-RDP83-00423R000600590007-4

New York University

Washington Square College of Arts and Science

Mathematics Research Group

Research Report No. EM-50

DIFFRACTION THEORY

A CRITIQUE OF SOME RECENT DEVELOPMENTS

by

C. J. Bouwkamp

The research reported in this document has been made possible through support and sponsorship extended by the Geophysics Research Directorate of the Air Force Cambridge Research Center, under Contract No.AF-19(122)42. It is published for technical information only, and does not necessarily represent recommendations or conclusions of the sponsoring agency.

April 1953

50X1

Page Denied

Declassified in Part - Sanitized Copy Approved for Release @ 50-Yr 2013/07/17 : CIA-RDP83-00423R000600590007-4

ABSTRACT

A number of recent developments in the theory of diffraction of electromagnetic waves, particularly those dealing with apertures in plane conducting screens, are reviewed. The subjects treated include modifications of Kirchhoff's theory, the theory of small apertures, Babinet's principle for plane obstacles, variational principles, and singularities at sharp edges.

For completeness, a discussion from an alternative viewpoint of the problem of diffraction by an aperture by Professor N. Marcuvitz has been included in this report.

Table of Contents

I. Introduction

The theory of diffraction has three major fields of application: (1) optics, (2) radio-wave propagation, and (3) acoustics. In (2) and (3) the wavelengths considered are usually of the same order of magnitude as the diffracting obstacle, while in the case (1) the wavelengths are usually small compared to the obstacle. A further difference between these three fields is that (1) and (2) involve essentially vectorial problems, while the problems involved in (3) are mainly scalar. However, in most applications of diffraction theory to classical optics, light is considered as a scalar wave phenomenon (polarization effects are ignored). For example, calculations on diffraction by apertures are usually based on Kirchhoff's mathematical formulation of Huygens' principle. Experiments have shown that this is justifiable when the wavelength is small in comparison to the size of the aperture. Polarization cannot be ignored in radio-wave propagation, where the wavelength is of the same order of magnitude as the aperture. One way to cover this is by using an electromagnetic equivalent of the scalar Huygens-Kirchhoff formula. This scalar formula may be applied to any of the six rectangular components of the electric and magnetic vectors. In order that the six wave functions so obtained should satisfy Maxwell's equations we have to introduce certain contour integrals along the rim of the aperture (Kottler).

The theory of Kirchhoff and Kottler are poor substitutes for rigorous diffraction theory (wave equation plus boundary conditions) in the quasi-optical range because they do not correctly describe the field in the vicinity of the aperture and the edge. In the extreme case of very long waves they entirely fail to predict the correct order of magnitude of the field far from the aperture (Rayleigh).

The purpose of this report is to comment on some of the new developments in diffraction theory. Various modifications of the Kirchhoff theory have recently been proposed. Rayleigh's potential-approach has been extended to higher-order approximations. The integral-equation technique has been developed extensively, and variational methods have shown their usefulness in a great number of problems. Also, the rigorous form of Babinet's principle in plane obstacle diffraction theory has been obtained. New insight into the character of singularities at sharp edges has profoundly influenced many aspects of diffraction theory.

A number of important recent developments have not been treated in these lectures. Among these we can mention the exact solutions recently obtained for diffraction by circular apertures and disks, and the Wiener-Hopf technique which has proved its power in the solution of certain waveguide problems.

Only steady-state problems will be discussed. The time factor is understood to be exp(-iωt). For a general introduction into diffraction theory, which includes descriptions of the early work by Kirchhoff, Kottler and Rayleigh, see:

Baker and Copson, The Mathematical Theory of Huygens' Principle, Oxford, Clarendon Press, 1950.

Sommerfeld, Vorlesungen uber theoretische Physik, vol. 4, Optik, Wiesbaden, Dieterich Verlag, 1950.

II. Kirchhoff's Theory of Diffraction

Let Σ be a screen of vanishing thickness covering a finite part of the plane $z = 0$. Consider a system of sources in the left half-space $(z < 0)$. If the screen were absent, these sources would produce a wave field $u_o(P)$ at the point P. The actual field $u(P)$, produced when the diffracting screen is present, is the sum of $u_o(P)$ and $u_d(P)$, where $u_d(P)$ is the diffracted field due to the secondary sources on Σ. By Green's theorem

$$u_d(P) = \frac{1}{4\pi} \int \left\{ u \frac{\partial}{\partial \nu} \left(\frac{e^{ikr}}{r}\right) - \frac{e^{ikr}}{r} \frac{\partial u}{\partial \nu} \right\} d\Sigma ,$$

where the integration is over both faces of Σ, and r is the distance from P to $d\Sigma$. Further, $\partial/\partial\nu$ denotes differentiation with respect to the integration-point coordinates in the direction of the normal to Σ drawn into free space.

Kirchhoff made the following assumptions:

(i) $u = u_o, \ \frac{\partial u}{\partial \nu} = \frac{\partial u_o}{\partial \nu}$ on S (illuminated side of screen)

(ii) $u = 0 , \ \frac{\partial u}{\partial \nu} = 0$ (on dark side of screen).

The total field then becomes, in Kirchhoff's approximation,

$$(2.1) \qquad K_s(P) = u_o(P) + \frac{1}{4\pi}\int_S \left\{ \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} - u_o \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r}\right) \right\} d\Sigma ,$$

where now the integration is only over the illuminated part S of Σ, while n refers to the normal of S drawn into the shadow region $(z > 0)$.

Serious objections can be raised against Kirchhoff's theory[1]. In fact, as we let P approach a point Q on the screen Σ, equation (1) fails to reproduce the assumed values u_o and $\partial u_o/\partial n$. This can be shown with the use of the following theorem:

$$\frac{1}{4\pi} \int u_o \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r}\right) d\Sigma = - \frac{1}{4\pi} \frac{\partial}{\partial z} \int u_o \frac{e^{ikr}}{r} d\Sigma \rightarrow \pm \frac{u_o}{2} \begin{cases} P \rightarrow Q^+ \\ P \rightarrow Q^- \end{cases} .$$

Then the limiting values of $K_s(P)$ are

$$(2.2) \qquad Cu_o + \frac{1}{4\pi}\int_S \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} \, d\Sigma ,$$

where $C = \frac{1}{2}$ or $\frac{3}{2}$ according as P is on the dark or on the illuminated side of Σ. Consequently, only if

$$\frac{1}{2} u_o + \frac{1}{4\pi}\int_S \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} \, d\Sigma = 0 \qquad (P \text{ on } S)$$

will the limiting values of K_s be identical with the assumed values u_o. However, this condition cannot be fulfilled for arbitrary S and u_o, as can easily be seen if we take two screens, one inside the other, and subtract their field equations: Then, using the above condition, we obtain

$$\frac{1}{4\pi} \int_{\bar{S}} \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} \, d\Sigma = 0,$$

where $\bar{S}$ (the "difference" between the two original screens) and u_o are arbitrary. It would then follow that $\frac{\partial u_o}{\partial n} = 0$ on $\bar{S}$. But u_o was arbitrarily chosen and therefore $\frac{\partial u_o}{\partial n}$ would not necessarily be equal to zero. Hence we have a contradiction. This shows that Kirchhoff's procedure is not self-consistent. The same conclusion follows from a consideration of the limiting values of $\partial K_s/\partial n$, which are

$$(2.3) \qquad C\frac{\partial u_o}{\partial n} - \left(k^2 + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \frac{1}{4\pi} \int_S u_o \frac{e^{ikr}}{r} \, d\Sigma .$$

Generally it can be said that the reason for the inconsistency of Kirchhoff's theory is that u and $\partial u/\partial \nu$ cannot simultaneously be prescribed on Σ since the equation $\Delta u + k^2 u = 0$ is elliptic.

If Kirchhoff's boundary conditions on Σ were exact, then u and $\partial u/\partial n$ would jump by the amounts u_o and $\partial u_o/\partial n$ respectively across S. In fact, these jumps <u>are</u> produced by K_s, as may at once be verified from expressions (2) and (3) for the limiting values on S. This is in accordance with Kottler's interpretation of Kirchhoff's formula (1) in that K_s is the rigorous solution, not of a boundary-value problem, but of a saltus problem[2].

We shall now discuss the complementary problem of diffraction by an infinite plane screen with a finite aperture A. Assuming the same primary field u_o, we apply equation (1) where now S means the complement of A. To avoid the slight difficulty that S is no longer finite, it may be necessary to assume that the imaginary part of k is positive. Now the integral over S is equal to the integral over S + A minus the integral over A. The integral over S + A equals $-u_o(P)$ if $z > 0$ and equals zero if $z < 0$. Hence the total field behind the aperture, in Kirchhoff's sense, is

$$(2.4) \qquad K_a(P) = -\frac{1}{4\pi}\int_A \left\{ \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} - u_o \frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right) \right\} d\Sigma ,$$

and the total field in front of the aperture is

$$(2.5) \qquad K_a(P) = u_o(P) - \frac{1}{4\pi}\int_A \left\{ \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} - u_o \frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right) \right\} d\Sigma ,$$

where n is the normal of A drawn into the shadow region ($z > 0$).

The analytic continuation of the function (4) through the aperture is precisely the function (5), and vice versa. There are no discontinuities in the aperture, where we have

$$(2.6) \qquad \begin{aligned} K_a(P) &= \frac{1}{2} u_o(P) - \frac{1}{4\pi} \int_A \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} \, d\Sigma , \\ \frac{\partial K_a(P)}{\partial n} &= \frac{1}{2} \frac{\partial u_o(P)}{\partial n} + \left(k^2 + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) \frac{1}{4\pi}\int_A u_o \frac{e^{ikr}}{r} \, d\Sigma . \end{aligned}$$

Equation (4) is usually derived in a different manner, namely by applying Green's theorem for the half-space $z \geq 0$ and assuming that u and $\partial u/\partial n$ are zero on the dark face of the screen and that in the aperture they are equal to the unperturbed values. However, Kirchhoff's original method is preferable since it avoids the difficulty that (4) does not reproduce the values assumed in the aperture but rather the values (6).

For complementary problems (A = S), it follows from (1), (4) and (5) that

$$K_a + K_s = u_o \qquad \text{for } z > 0$$

$$K_a + K_s = 2u_o \qquad \text{for } z < 0.$$

This is Babinet's principle in the sense of Kirchhoff's theory [3].

It has often been suggested [4] that Kirchhoff's formula gives the first term of an accurate solution of a boundary-value problem by successive approximations. This was disproved by Franz [5] and Schelkunoff [6]. In addition, Franz derived equations (4) and (5) as follows. (A similar interpretation was given by Schelkunoff.) A zero-order approximation is $u = u_o$ $(z < 0)$, $u = 0$ $(z > 0)$. This agrees with the boundary condition at the black screen, but the wave equation is violated in the aperture. The wave equation is then restored by a correction term arising from the secondary sources in the aperture. This term equals the right-hand side of equation (4) for <u>all</u> P $(z \gtrless 0)$. By elaborating the new interpretation, Franz [7] devised an interesting diffraction theory which is applicable for all wavelengths. In the author's opinion, however, it is still an open question whether Franz's theory will eventually become of value in practice.

III. Modified Kirchhoff Theory

Modifications of Kirchhoff's theory have recently been proposed for planar diffraction problems*. One aim of this modified Kirchhoff's theory was to provide a means for distinguishing between the two principal boundary-value problems: (I) scalar wave function vanishing on a <u>soft</u> screen and (II) normal derivative of scalar wave function vanishing on a <u>rigid</u> screen (in the terminology of acoustics). The modified theory makes use of the two principal Green's functions of the half-space, which are known explicitly.

---

* See, for example, Bouwkamp, Thesis, Groningen, 1941.

Let u(P) denote a wave function that is regular for $z \geq 0$. Then, in the half-space $z \geq 0$, we have Rayleigh's formulas:

$$(3.1) \qquad R_1: \; u(P) = -\frac{1}{2\pi} \int \frac{e^{ikr}}{r} \frac{\partial u}{\partial n} \, d\Sigma ,$$

$$(3.2) \qquad R_2: \; u(P) = \frac{1}{2\pi} \int u \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r}\right) d\Sigma ,$$

where n points into the half-space $z \geq 0$, and the integration is over $z = +0$. Equations (1) and (2) follow at once from an application of Green's theorem using the two principal Green's functions of the wave equation for the half-space.

In applying Rayleigh's formulas to diffraction problems we shall assume, as in the usual version of Kirchhoff's theory, that u and $\partial u/\partial n$ may be replaced by the corresponding geometrical-optics values: immediately behind the screen u and $\partial u/\partial n$ are taken to be zero, and in the aperture they are replaced by u_o and $\partial u_o/\partial n$. We have previously denoted Kirchhoff's solutions by K_a and K_s, where the subscripts refer to diffraction by a finite aperture and by a finite screen respectively. It is convenient to introduce a similar notation for the wave functions derived from Rayleigh's formulas. For example, $R_{a1}(P)$ will denote the wave function for the diffraction by the aperture A based on Rayleigh's first formula (1). For obvious reasons, the corresponding modified Kirchhoff solutions will be termed "Rayleigh solutions".

The complete set of Rayleigh solutions for the diffracted field in the right half-space $(z > 0)$ then becomes

$$(3.3) \qquad R_{s1}(P) = u_o(P) + \frac{1}{2\pi} \int_S \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} \, d\Sigma ,$$

$$(3.4) \qquad R_{a1}(P) = -\frac{1}{2\pi} \int_A \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} \, d\Sigma ,$$

$$(3.5) \qquad R_{s2}(P) = u_o(P) - \frac{1}{2\pi} \int_S u_o \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r}\right) d\Sigma ,$$

$$(3.6) \qquad R_{a2}(P) = \frac{1}{2\pi} \int_A u_o \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r}\right) d\Sigma .$$

Unlike Kirchhoff's theory, the modified theory is self-consistent[8]. The reason for this is that in the modified theory it was sufficient to assume boundary values for either u (in the case of R_2) or $\partial u/\partial n$ (in the case of R_1). In fact, all values assumed to exist at $z = +0$, whether behind the screen or in the aperture, are exactly reproduced by the Rayleigh solutions when P approaches the plane $z = 0$ from the right.

The analytic continuation of the Rayleigh solutions into the illuminated half-space are easily obtained. Equations (3) and (5) remain valid for points P to the left of the plane $z = 0$. On the other hand, equations (4) and (6) are to be replaced by

$$(3.7) \qquad R_{a1}(P) = u_o(P) - u_o(-P) - \frac{1}{2\pi}\int_A \frac{e^{ikr}}{r} \frac{\partial u_o}{\partial n} \, d\sum ,$$

and

$$(3.8) \qquad R_{a2}(P) = u_o(P) + u_o(-P) + \frac{1}{2\pi}\int_A u_o \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r}\right) d\sum$$

respectively. Here $u_o(-P)$ means the value assumed by u_o at the reflection of P in the plane $z = 0$; $u_o(P) \mp u_o(-P)$ is the zero-order reflected field in the sense of geometrical optics.

Like Kirchhoff's solution, the Rayleigh solutions are exact solutions of saltus problems. The functions R_{s2} and R_{a2} jump from $2u_o$ on the illuminated face of the screen to zero on the dark face. Similarly the normal derivatives of R_{s1} and R_{a1} jump across the screen from $2\,\partial u_o/\partial n$ to zero. Further, the Kirchhoff solution is just the average of the two corresponding Rayleigh solutions, viz.,

$$K_a = \frac{1}{2}(R_{a1} + R_{a2}) \quad , \; K_s = \frac{1}{2}(R_{s1} + R_{s2}),$$

while Babinet's principle now assumes either the form

$$R_{a1} + R_{s1} = u_o \;(z > 0),\; R_{a1} + R_{s1} = 2u_o(P) - u_o(-P) \;(z < 0),$$

or the form

$$R_{a2} + R_{s2} = u_o \;(z > 0),\; R_{a2} + R_{s2} = 2u_o(P) + u_o(-P) \;(z < 0).$$

As we mentioned before, one aim of the modified Kirchhoff theory was to furnish a method for distinguishing between the two principal boundary-value problems previously noted. Accordingly, since $\partial R_{a1}/\partial n = 0$ on the dark face of the screen, R_{a1} of equation (4) was proposed [9] as an approximate expression for the diffracted field behind the aperture in the acoustically rigid screen for the incident wave u_o (i.e., $R_{a1} \approx \phi_2$). (ϕ_1 and ϕ_2 are defined in equations (9) and (10).) Similarly, R_{a2} was suggested for the diffracted field behind the acoustically soft screen (i.e., $R_{a2} \approx \phi_1$). It is obvious that these approximations will be accurate immediately behind the screen but poor in the vicinity of the aperture. The approximation is a complete failure if it is extended to the respective analytic continuations through the aperture into the illuminated space, because on the lit face of the screen we have $\partial R_{a1}/\partial n = 2\, \partial u_o/\partial n$ and $R_{a2} = 2u_o$. In fact, the reflected-field terms in equations (7) and (8) suggest the opposite correspondence between the Rayleigh solutions and the solutions of the boundary-value problems. We shall now discuss this correspondence in more detail.

Let $u_{a1}(x,y,z)$ denote the wave function for the diffraction of the primary field $u_o(x,y,z)$, incident from the left ($z < 0$), through a finite aperture in a <u>perfectly soft</u> plane screen. Then [10]

$$(3.9) \qquad u_{a1} = \begin{cases} u_o(x,y,z) - u_o(x,y,-z) + \phi_1(x,y,-z) & (z \leq 0) \\ \phi_1(x,y,z), & (z \geq 0) \end{cases}$$

where ϕ_1, defined for $z \geq 0$ only, has the following properties: (i) ϕ_1 is a solution of the wave equation; (ii) $\phi_1 = 0$ on the dark face of the screen; (iii) ϕ_1 is regular at infinity (Sommerfeld's radiation condition); (iv) $\partial\phi_1/\partial z = \partial u_o/\partial z$ in the aperture; (v) ϕ_1 is uniformly bounded, and $|\text{grad}\ \phi_1|^2$ is integrable over any finite part of three-dimensional space, including the rim of the aperture.

Let $u_{a2}(x,y,z)$ denote the corresponding wave function for an aperture in a <u>perfectly rigid</u> screen. Then

$$(3.10) \qquad u_{a2} = \begin{cases} u_o(x,y,z) + u_o(x,y,-z) - \phi_2(x,y,-z) & (z \leq 0) \\ \phi_2(x,y,z), & (z \geq 0) \end{cases}$$

where ϕ_2 is also defined for $z \geq 0$ and has similar properties as ϕ_1 except that (ii) and (iv) should be replaced by: (ii') $\partial\phi_2/\partial z = 0$ on the screen, and (iv') $\phi_2 = u_o$ in the aperture. Existence theorems and questions of uniqueness will not be discussed for the time being. Let it suffice to mention that the property (v) ensures that no energy is radiated by the singularities at the rim.

It is difficult, if not impossible, to determine the functions ϕ_1 and ϕ_2 for an aperture of arbitrary shape. The trouble is that either ϕ_1 or ϕ_2 solves a <u>mixed</u> boundary-value problem: ϕ and $\partial\phi/\partial n$ are given on mutually complementary parts of the plane $z = 0$ (see (ii) and (iv)). However, by virtue of (ii) and (ii') and Rayleigh's formula, we have for any aperture A the following relations:

$$(3.11) \qquad \phi_1 = \frac{1}{2\pi}\int_A \phi_1 \frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right) d\sum, \quad \phi_2 = -\frac{1}{2\pi}\int_A \frac{\partial\phi_2}{\partial n}\frac{e^{ikr}}{r} d\sum .$$

Now if we assume that the unknown values of ϕ_1 and $\partial\phi_2/\partial n$ in the aperture may be replaced by the respective unperturbed values of the incident wave, we find that $\phi_1 \approx R_{a2}$, $\phi_2 \approx R_{a1}$ $(z > 0)$. If this is substituted into equations (9) and (10) we obtain

$$(3.12) \qquad \begin{matrix} u_{a1} \approx R_{a2} \ (z > 0), & u_{a1} \approx 2u_o - R_{a2}, \ (z < 0) \\ u_{a2} \approx R_{a1} \ (z > 0), & u_{a2} \approx 2u_o - R_{a1}, \ (z < 0). \end{matrix}$$

For $z > 0$ the approximation (12) is identical with that discussed previously. It should be noted that the approximate solutions do satisfy the correct boundary conditions at the screen. However, they are not analytic functions: either the normal derivative or the approximate solution itself is discontinuous in the aperture.

In deriving the approximations (12) the properties (ii) and (ii') have been used. Yet almost equally characteristic for planar boundary-value problems are the properties (iv) and (iv'). The latter express that for problem I ($u = 0$ on the screen) $\partial u/\partial n$ is unperturbed in the aperture. Starting from this point of view we can derive a different set of approximations [11]. From Rayleigh's formula (1) and property (iv) it follows that

$$(3.13) \qquad \phi_1 = -\frac{1}{2\pi}\int_A \frac{e^{ikr}}{r}\frac{\partial u_o}{\partial n}\, d\sum - \frac{1}{2\pi}\int_S \frac{e^{ikr}}{r}\frac{\partial \phi_1}{\partial n}\, d\sum .$$

Unfortunately, the values of $\partial\phi_1/\partial n$ are not known on the infinite screen S. If we assume that they are approximately equal to zero, we arrive at $\phi_1 \approx R_{a1}$. By a similar reasoning in the case of problem II, we get $\phi_2 \approx R_{a2}$. It is not difficult to verify that this ultimately gives

$$(3.14) \qquad u_{a1} \approx R_{a1}, \quad u_{a2} \approx R_{a2}$$

everywhere in free space. The approximate solutions are now analytic and they produce the correct (unperturbed) values of $\partial u_{a1}/\partial n$ and u_{a2} in the aperture, but they violate the correct boundary conditions at the screen. Insofar as an accurate approximation to the field is more important in the aperture than in the vicinity of the screen, the approximations (14) seem preferable to the opposite relations (12).

An alternative way of showing the close relation between the original and modified Kirchhoff solutions is the following. Consider, for example, the diffracted field behind an aperture. Noting that $\partial/\partial n = \partial/\partial z' = -\partial/\partial z$ for any function of r, we have from equation (4) that

$$(3.15) \qquad 2K_a(P) = f(P) + \frac{\partial}{\partial z}\, g(P),$$

where f(P) and g(P) are both even functions of z, viz.,

$$(3.16) \qquad f(P) = -\frac{1}{2\pi}\int_A \frac{e^{ikr}}{r}\frac{\partial u_o}{\partial n}\, d\sum, \quad g(P) = -\frac{1}{2\pi}\int_A \frac{e^{ikr}}{r}\, u_o\, d\sum.$$

Comparison with (4) and (6) shows that

$$R_{a1}(P) = f(P), \quad R_{a2}(P) = \frac{\partial}{\partial z}\, g(P). \tag{3.17}$$

Consequently, the functions $R_{a1}(P)$ and $R_{a2}(P)$ are simply twice the even and odd parts respectively of $K_a(P)$.

In addition, the function $f(P)$ is equal to the velocity potential of a membrane vibrating in a rigid baffle with velocity distribution $-\partial u_o/\partial n$; the same holds for $g(P)$ and $-u_o$. Methods and numerical results of the theory of acoustic radiation are therefore valuable in diffraction theory also. Various authors have used the modified Kirchhoff theory in one way or another. Bremmer[12] applied Rayleigh's second formula to the diffraction theory of Gaussian optics. Various mathematical aspects of equation (2) were discussed by Luneberg[13], and Scheffers[14] emphasized the usefulness of this equation (in the Fourier form) for the calculation of Fraunhofer patterns. Durand[15] applied the Fourier equivalent of (6) to a circular aperture and to a half-plane for the case of a plane wave at normal incidence. Spence[16] compared R_{a1} for a circular aperture (plane wave at normal incidence) with the corresponding exact solution of the boundary-value problem II. Experimental results on the diffraction of sound around a circular disk were discussed in connection with the Kirchhoff approximations by Primakoff, Klein, Keller and Carstensen[17].

In concluding this section, we shall briefly discuss the Kirchhoff solutions for the diffraction of a plane wave normally incident on a circular aperture. Let <u>a</u> denote the radius of the aperture, and let the incident wave $u_o = e^{ikz}$ impinge from the left. Choosing the origin of coordinates at the center of the aperture, we have in the shadow region

$$K_a = -\frac{1}{2}\left(ikU + \frac{\partial U}{\partial z}\right), \quad R_{a1} = -ikU,$$

$$U = \frac{1}{2\pi}\int_A \frac{e^{ikr}}{r}\, d\textstyle\sum, \quad R_{a2} = -\frac{\partial U}{\partial z}, \tag{3.18}$$

where U is the velocity potential of Rayleigh's piston for unit velocity distribution.

The integral is easy to evaluate if P lies on the z axis. The result is

$$R_{a1} = e^{ikz} - e^{ik\sqrt{z^2+a^2}} , \tag{3.19}$$

$$R_{a2} = e^{ikz} - \frac{z}{\sqrt{z^2+a^2}} e^{ik\sqrt{z^2+a^2}} , \tag{3.20}$$

$$K_a = e^{ikz} - \frac{1}{2} e^{ik\sqrt{z^2+a^2}} - \frac{1}{2} \frac{z}{\sqrt{z^2+a^2}} e^{ik\sqrt{z^2+a^2}} . \tag{3.21}$$

It should be noticed that these expressions are equally valid if P is on the negative z axis. If a tends to infinity, R_{a1} and K_a do not reduce to the incident field e^{ikz}, unless the medium is assumed to be slightly absorbing (i.e., imaginary part of k is positive).

The respective Fraunhofer patterns are also easy to calculate. Let ρ, θ denote spherical coordinates with the positive z axis as polar axis. Then at large distances from the aperture

$$U \sim A(\theta)\ \rho^{-1}\ e^{ik\rho} ,$$

where $A(\theta)$ is the amplitude of the spherical wave. We find

$$(i/a)A_{a1} = \frac{J_1(ka\sin\theta)}{\sin\theta} , \tag{3.22}$$

$$(i/a)A_{a2} = \frac{J_1(ka\sin\theta)}{\tan\theta} , \tag{3.23}$$

$$(i/a)A_K = \frac{J_1(ka\sin\theta)}{2\tan(\theta/2)} , \tag{3.24}$$

where we use an obvious notation for the amplitude, and where $0 \leq \theta \leq \frac{\pi}{2}$, and J_1 is a Bessel function.

The amount of energy transmitted through the aperture can be computed by integrating $|A|^2$ over half of the unit sphere. It is convenient to introduce the <u>transmission coefficient</u>, which is the ratio of the transmitted energy to the energy incident on the aperture in the sense of geometrical optics. In the problem under discussion this coefficient is

(3.25) $$T = \frac{1}{\pi a^2} \int |A|^2 \, d\Omega = \frac{2}{a^2} \int_0^{\pi/2} |A(\theta)|^2 \sin\theta d\theta \, .$$

The relevant expressions for the three different cases mentioned above are

(3.26) $$T_1 = 1 - \frac{J_1(2ka)}{ka} \, ,$$

(3.27) $$T_2 = 1 + \frac{J_1(2ka)}{ka} - \frac{1}{ka} \int_0^{2ka} J_o(t)dt,$$

(3.28) $$T_K = 1 - \frac{1}{2} \left[J_o^2(ka) + J_1^2(ka) + \frac{1}{2ka} \int_0^{2ka} J_o(t)dt \right] \, .$$

Equation (26) is a classical result obtained by Lord Rayleigh.

For very small values of ka we have

(3.29) $$T_1 \sim \frac{1}{2}(ka)^2 \, , \; T_2 \sim \frac{1}{6}(ka)^2 \, , \; T_K \sim \frac{7}{24} (ka)^2.$$

These values of T are in complete disagreement with the results for the exact boundary-value problems. This is not surprising, since the Kirchhoff approximation holds for small wavelengths. For very large values of ka we have

$$T_1 \sim 1 - \frac{1}{\sqrt{\pi}} \; \frac{\sin(2ka)-\pi/4)}{(ka)^{3/2}} \, ,$$

(3.30) $$T_2 \sim 1 - \frac{1}{ka} + \frac{1}{2\sqrt{\pi}} \; \frac{\cos(2ka-\pi/4)}{(ka)^{5/2}} \, ,$$

$$T_K \sim 1 - (\frac{1}{4} + \frac{1}{\pi}) \frac{1}{ka} - \frac{1}{4\sqrt{\pi}} \; \frac{\sin(2ka)-\pi/4)}{(ka)^{3/2}} \, .$$

IV. Braunbek's Modification of the Kirchhoff Theory.

An attempt to improve the modified Kirchhoff solution was made by Braunbek[18] who observed that the Kirchhoff solution does not constitute the main term (in a series of powers of $1/ka$) of the exact solution. However, before entering into a discussion of Braunbek's theory, we shall first discuss the solution constructed by Sommerfeld for the diffraction of a plane wave by a half-plane:

$$\Phi = Ae^{-ik\rho\cos(\theta-\theta_0)}\int_0^{\sqrt{2k\rho}\cos[(\theta-\theta_0)/2]} e^{i\tau^2}\,d\tau$$

$$+ Be^{-ik\rho\cos(\theta+\theta_0)}\int_0^{\sqrt{2k\rho}\cos[(\theta+\theta_0)/2]} e^{i\tau^2}\,d\tau ,$$

on which Braunbek's theory is based.

A simple derivation of Sommerfeld's formulas is implicit in an interesting paper by J. Brillouin[19]; this proof we now present.

Consider a plane wave incident on a screen which cuts the plane of drawing in the upper half of the y axis (fig. 1, direction of incidence normal to edge of screen).

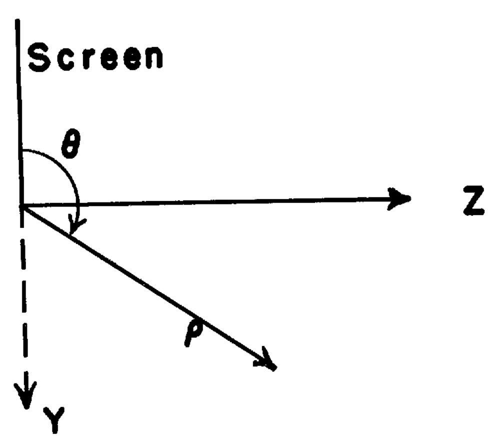

Fig. 1

(The restriction to normal incidence causes no lack of generality since the proof can be generalized to arbitrary oblique incidence.) Introduce new "semi-parabolic" coordinates u, v such that

(4.1)
$$v = y \qquad\qquad y = v$$
$$u = -y + \sqrt{y^2+z^2} \qquad\qquad z = \sqrt{u^2+2uv}$$

(i.e., one of the parabolic coordinates is rejected in favor of a rectangular coordinate), which can be written in terms of polar coordinates as well:

(4.2)
$$v = -\rho \cos \theta$$
$$u = 2\rho \cos^2 \frac{\theta}{2} ,$$

where $0 \leq \rho < \infty$ and $0 \leq \theta \leq \pi$.

Now, from (1) we see that

(4.3)
$$\frac{\partial}{\partial z} = \frac{z}{u+v} \frac{\partial}{\partial u} \; ; \qquad \frac{\partial}{\partial y} = \frac{\partial}{\partial v} - \frac{u}{u+v} \frac{\partial}{\partial u} .$$

With the use of (3), $\Delta_1 = \dfrac{\partial^2}{\partial y^2} + \dfrac{\partial^2}{\partial z^2}$ becomes

(4.4)
$$\Delta = \frac{\partial^2}{\partial v^2} + \frac{2u}{v+u} \left\{ \frac{\partial^2}{\partial u^2} - \frac{\partial^2}{\partial v \partial u} \right\} + \frac{1}{v+u} \frac{\partial}{\partial u} .$$

The wave equation $\Delta \Phi + k^2 \Phi = 0$ is not separable in these coordinates; however, we follow the same procedure as we would if it were separable and assume that $\Phi = F(v)G(u)$. The wave equation then becomes

(4.5)
$$G[F''+k^2F] + \frac{2u}{u+v} \left[F\left\{ G'' + \frac{1}{2u} G' \right\} - F'G' \right] = 0$$

which can be satisfied if we take F and G such that the relations

(4.6a)
$$F'' + k^2 F = 0$$

(4.6b)
$$F\left\{ G'' + \frac{1}{2u} G' \right\} - F'G' = 0$$

are satisfied.

From (6a) it follows that

(4.7)
$$F = \alpha \, e^{ikv}$$

and when this result is substituted into (6b), the latter becomes

$$\frac{G''}{G'} + \frac{1}{2u} = \frac{F'}{F} = ik = \frac{d}{du}\left[\log G'\sqrt{u}\right] . \qquad (4.8)$$

It is seen from (8) that

$$G' = \frac{\beta e^{iku}}{\sqrt{u}}$$

and finally that

$$G = \beta \int^{u} \frac{e^{ikt}}{\sqrt{t}}\, dt . \qquad (4.9)$$

Using the results of (7) and (9) the solution $\Phi = F(v)G(u)$ becomes

$$FG = (\text{const.})\, e^{ikv} \int^{u} \frac{e^{ikt}}{\sqrt{t}}\, dt. \qquad (4.10)$$

Transforming to polar coordinates, we obtain

$$FG = (\text{const.})\, e^{-ik\rho\cos\theta} \int^{2\rho\cos^2(\theta/2)} \frac{e^{ikt}}{\sqrt{t}}\, dt , \qquad (4.11)$$

and moreover, if we let $kt = \tau^2$, (11) becomes

$$FG = (\text{const.})e^{-ik\rho\cos\theta} \int^{\sqrt{2k\rho}\cos(\theta/2)} e^{i\tau^2}\, d\tau . \qquad (4.12)$$

If we had started with $\theta \pm \theta_o$ instead of θ (this just means a rotation of the coordinate system) we would have obtained two other solutions, and by taking a linear combination of these two, we would have as our final result

$$FG = Ae^{-ik\rho\cos(\theta-\theta_o)} \int_{o}^{\sqrt{2k\rho}\cos\frac{(\theta-\theta_o)}{2}} e^{i\tau^2}\, i\tau \qquad (4.13)$$

$$+ Be^{-ik\rho\cos(\theta+\theta_o)} \int_{o}^{\sqrt{2k\rho}\cos\frac{(\theta+\theta_o)}{2}} e^{i\tau^2}\, d\tau$$

which is exactly the form of Sommerfeld's solution and hence verifies it. (The boundary conditions can be met by an adjustment of the constants A and B.)

We shall now treat the two special cases referred to in Section III (i.e., $\phi_1 = 0$ on the screen, and $\partial\phi_2/\partial n = 0$ on the screen). However, we shall use henceforth a slightly different system of coordinates; see fig. 2.

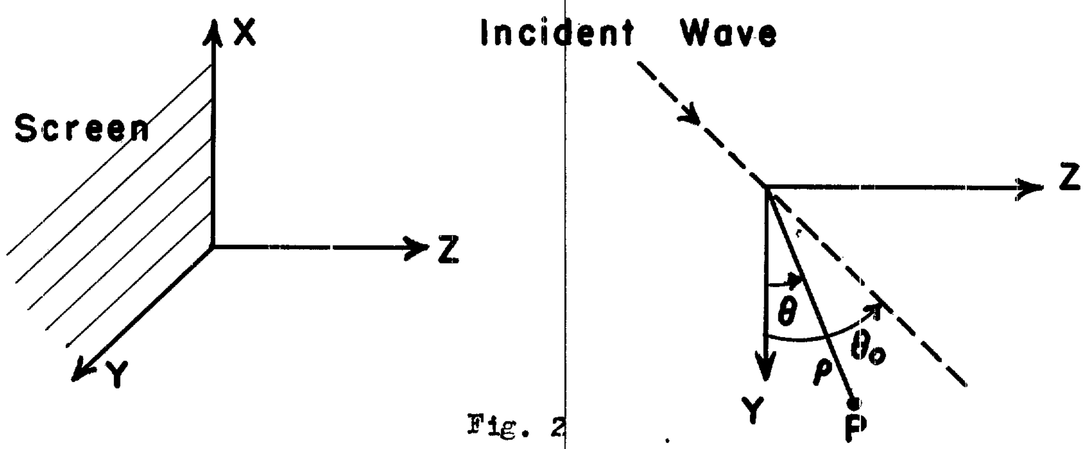

Fig. 2

Defining ϕ_1 and ϕ_2 to be

$$(4.14)\quad \left.\begin{matrix}\phi_1\\ \phi_2\end{matrix}\right\} = e^{ik\rho\cos(\theta-\theta_o)}\frac{e^{-i\pi/4}}{\sqrt{\pi}}\int_{-\infty}^{s_1} e^{i\tau^2}d\tau \mp e^{ik\rho\cos(\theta+\theta_o)}\frac{e^{-i\pi/4}}{\sqrt{\pi}}\int_{-\infty}^{s_2} e^{i\tau^2}d\tau$$

where $s_1 = \sqrt{2k\rho}\,\sin(\frac{\theta-\theta_o}{2})$, $s_2 = -\sqrt{2k\rho}\,\sin(\frac{\theta+\theta_o}{2})$, and where $u_o = e^{ik\rho\cos(\theta-\theta_o)}$ is the incident wave, it can be verified directly that ϕ_1 and $\partial\phi_2/\partial n$ both vanish on the screen.

There are other functions which satisfy the wave equation and, at the same time, the boundary conditions. Thus, for example, the functions $\partial^2\phi_1/\partial z^2$, $\partial^4\phi_1/\partial z^4$, etc., which obviously satisfy the wave equation, also vanish on the screen[20]. However, these functions are too singular at the edge and hence are not admissible solutions; ϕ_1 and ϕ_2 of equation

(14) <u>are</u> admissible[20].

For the purpose of discussing Sommerfeld's theory, two electromagnetic fields can be constructed from the solutions of the scalar wave equation previously discussed. The first field is

$$\vec{E}_{||} = (\phi_1, 0, 0), \tag{4.15a}$$

which vanishes on the screen, and for which $\vec{H}$, determined by Maxwell's equations, is given by

$$ik\,\vec{H} = (0, \frac{\partial\phi_1}{\partial z}, \frac{-\partial\phi_1}{\partial y});$$

and the second field is

$$\vec{H}_{||} = (\phi_2, 0, 0), \tag{4.15b}$$

the normal derivative of which vanishes on the screen, and for which

$$-ik\,\vec{E} = (0, \frac{\partial\phi_2}{\partial z}, \frac{-\partial\phi_2}{\partial y}).$$

The explicit expressions for all field components for the Sommerfeld half-plane problem are

$$E_x = \phi_1$$

$$H_y = \phi_2 \sin\theta_o - \cos\frac{1}{2}\theta_o \cos\frac{1}{2}\theta \sqrt{\frac{2}{\pi k\rho}}\; e^{i(k\rho+\pi/4)}$$

$$H_z = \phi_1 \cos\theta_o - \cos\frac{1}{2}\theta_o \sin\frac{1}{2}\theta \sqrt{\frac{2}{\pi k\rho}}\; e^{i(k\rho+\pi/4)}$$

in the case of incident field polarized parallel to the edge, and

$$H_x = \phi_2$$

$$E_y = -\phi_1 \sin\theta_o - \sin\frac{1}{2}\theta_o \sin\frac{1}{2}\theta \sqrt{\frac{2}{\pi k\rho}}\; e^{i(k\rho+\pi/4)}$$

$$E_z = \phi_2 \cos\theta_o + \sin\frac{1}{2}\theta_o \cos\frac{1}{2}\theta \sqrt{\frac{2}{\pi k\rho}}\; e^{i(k\rho+\pi/4)}$$

for the incident magnetic vector parallel to the edge. For the definition of coordinates, see Fig. 2, not Fig. 1.

Braunbek's theory makes use of the values of the scalar fields and their normal derivatives on the x,y plane (i.e., the plane determined by the screen). The non-vanishing derivatives of ϕ_1 and ϕ_2 necessary to compute the field quantities are

(4.16)

$$\frac{\partial\phi_1}{\partial z} = (ik\sin\theta_o)\phi_2 + \sqrt{\frac{2k}{\pi}}\, e^{-i\pi/4}\cos\tfrac{1}{2}\theta_o \left\{\cos\tfrac{1}{2}\theta\, \frac{e^{ik\rho}}{\sqrt{\rho}}\right\}$$

$$\frac{\partial\phi_2}{\partial z} = (ik\sin\theta_o)\phi_1 - \sqrt{\frac{2k}{\pi}}\, e^{-i\pi/4}\sin\tfrac{1}{2}\theta_o \left\{\sin\tfrac{1}{2}\theta\, \frac{e^{ik\rho}}{\sqrt{\rho}}\right\}$$

$$\frac{\partial\phi_1}{\partial y} = (ik\cos\theta_o)\phi_1 - \sqrt{\frac{2k}{\pi}}\, e^{-i\pi/4}\cos\tfrac{1}{2}\theta_o \left\{\sin\tfrac{1}{2}\theta\, \frac{e^{ik\rho}}{\sqrt{\rho}}\right\}$$

$$\frac{\partial\phi_2}{\partial y} = (ik\cos\theta_o)\phi_2 - \sqrt{\frac{2k}{\pi}}\, e^{-i\pi/4}\sin\tfrac{1}{2}\theta_o \left\{\cos\tfrac{1}{2}\theta\, \frac{e^{ik\rho}}{\sqrt{\rho}}\right\}.$$

To calculate the fields on the x,y plane, we let $\theta = 0$ and $\theta = \pi$. We confine ourselves to the case of normal incidence and hence take $\theta_o = \frac{\pi}{2}$. This gives us

(4.17)

$$\theta = 0 \begin{cases} \phi_1 = 0 \\ \dfrac{\partial\phi_2}{\partial z} = 0 \\ \dfrac{\partial\phi_1}{\partial z} = \dfrac{2ik}{\sqrt{\pi}}\, e^{-i\pi/4}\displaystyle\int_{\sqrt{k\rho}}^{\infty} e^{i\tau^2}d\tau + \sqrt{\frac{k}{\pi}}\, e^{-i\pi/4}\, \frac{e^{ik\rho}}{\sqrt{\rho}} \\ \phi_2 = \dfrac{2}{\sqrt{\pi}}\, e^{-i\pi/4}\displaystyle\int_{\sqrt{k\rho}}^{\infty} e^{i\tau^2}d\tau\ , \end{cases}$$

and

(4.18)

$$\theta = \pi \begin{cases} \dfrac{\partial\phi_1}{\partial z} = ik \\ \phi_2 = 1 \\ \phi_1 = 1 - \dfrac{2}{\sqrt{\pi}}\, e^{-i\pi/4}\displaystyle\int_{\sqrt{k\rho}}^{\infty} e^{i\tau^2}d\tau \\ \dfrac{\partial\phi_2}{\partial z} = ik - \dfrac{2ik}{\sqrt{\pi}}\, e^{-i\pi/4}\displaystyle\int^{\infty} e^{i\tau^2}d\tau - \sqrt{\frac{k}{\pi}}\, e^{-i\pi/4}\, \frac{e^{ik\rho}}{\sqrt{\rho}} \end{cases}$$

We shall now discuss Braunbek's attempt to improve the modified Kirchhoff solution for a circular aperture. As was mentioned in the beginning of this section Braunbek observed that Kirchhoff's solution does not constitute the main term of the exact solution. To obtain the correct main term, we must estimate the effect of the second integral of (3.13) on ϕ_1. (The problem for ϕ_2 is analogous.) Braunbek replaced $\frac{\partial \phi_1}{\partial n}$ on the screen by the value derived from Sommerfeld's theory for the half-plane, as if S had a locally straight edge. This is a plausible assumption, because $\partial\phi_1/\partial n$ is expected to be rapidly decreasing from the edge over a distance of a few wavelengths (confirmed experimentally in the case of ϕ_2 by Severin and Starke [21]), and the wavelengths here considered are small with respect to the radius of the aperture. For the circular aperture and a normally incident plane wave, Braunbek's approximation becomes

$$(4.19)\quad \phi_1 \approx B_1 = -\frac{ik}{2\pi}\int_A \frac{e^{ikr}}{r}\, d\Sigma - \frac{ik}{2\pi}\int_S \frac{e^{ikr}}{r}\left[\Phi(k\xi) - \psi(k\xi)\right] d\Sigma ,$$

where ξ is the distance to the rim and

$$(4.20)\quad \Phi(x) = \frac{2}{\sqrt{\pi}}\, e^{-i\pi/4}\int_{\sqrt{x}}^{\infty} e^{i\tau^2}\, d\tau \; ; \; \psi(x) = \frac{e^{i\pi/4}}{\sqrt{\pi x}}\, e^{ix} .$$

The evaluation of B_1 on the axis of the aperture is comparatively simple. Integration by parts and some trivial transformations yield

$$(4.21)\quad B_1(0,0,z) = e^{ikz} - \sqrt{\frac{2}{\pi}}\, e^{-i\pi/4} e^{ik\sqrt{z^2+a^2}} \int_0^{\infty} \left[\frac{\tau + ka + k\sqrt{z^2+a^2}}{\tau + 2k\sqrt{z^2+a^2}}\right]^{1/2} \frac{e^{i\tau}}{\sqrt{\tau}}\, d\tau .$$

The integral is elementary if $z = 0$. At the center of the aperture Braunbek's function is exactly equal to

$$(4.22)\quad B_1(o,o,o) = 1 - \sqrt{2}\, e^{ika}.$$

When z is greater than zero, the integral in (21) can be expanded asymptotically.

In a first approximation, the square root in the integrand may be replaced by its value of the integrand at the lower limit of integration. Then except for terms of order $1/ka$,

$$(4.23) \qquad B_1(0,0,z) = e^{ikz} - e^{ik\sqrt{z^2+a^2}}\left[1 + \frac{a}{\sqrt{z^2+a^2}}\right]^{1/2},$$

which is Braunbek's result (obtained somewhat differently). His computations showed that (22) and (23) are in excellent agreement with Meixner and Fritze's [22] exact values. The corresponding diffraction pattern was also evaluated by Braunbek.

In studying Braunbek's paper, the author encounters some difficulties in connection with a second type of approximation to ϕ_1. In starting from the equation

$$\phi_1 = \frac{1}{2\pi}\int_A \phi_1 \frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right) d\Sigma$$

given in Section III, Braunbek replaces ϕ_1 in the aperture by the value $1 - \Phi(k\xi)$ derived from Sommerfeld's theory. Then

$$(4.24) \qquad \phi_1 \approx B_1^* = \frac{\partial}{\partial z}\left[-\frac{1}{2\pi}\int_A \left\{1 - \Phi(k\xi)\right\}\frac{e^{ikr}}{r}\, d\Sigma\right].$$

Braunbek claims that, except for terms of order $1/ka$, the functions B_1 and B_1^* are identical on the axis of the aperture. The present author believes this to be incorrect, because from the values at the center of the aperture we can see, without any calculation at all, that

$$B_1^*(0,0,0) = 1 - \Phi(ka) \sim 1 - \frac{e^{i\pi/4}}{\sqrt{\pi ka}}\, e^{ika},$$

which differs from (22).

V. Variational Formulation of Scalar Diffraction Problems.

A variational formulation of planar diffraction problems, which permits accurate numerical evaluation of the diffracted amplitude and the transmission cross-section for a wide range of frequencies, was given by Levine and Schwinger[23]. They illustrated the utility of the variational method by applying it to the circular aperture for a normally incident plane wave. The analysis was criticized by Copson[24] since many of the integrals involved appeared to diverge. Copson, however, in deriving what he calls "Levine and Schwinger's variational principle" in a mathematically sound way, confined himself to the problem for ϕ_2, while his criticism concerns that for ϕ_1. Fortunately the divergent integrals that occur in Levine and Schwinger's paper are easy to eliminate without affecting the numerical results. However before we treat these problems let us first discuss the integral formulation of scalar diffraction problems with application to small apertures.

As is seen from the second equation of (3.11), the wave function ϕ_2 is uniquely determined in space by the values of its normal derivative in the aperture. Let the unknown values of $\partial\phi_2/\partial n$ in the aperture be denoted by $f(x,y)$. Recalling that $\phi_2 = u_o$ in the aperture, we find the integral equation[25]

$$(5.1a) \qquad \int_A f(x',y')G(x,x',y,y')dx'dy' = -u_o(x,y,o),$$

where the kernel G is symmetric and singular, viz.

$$G = (1/2\pi s)e^{iks} \qquad , \qquad s^2 = (x-x')^2 + (y-y')^2 ,$$

and where (x,y,o) is any point of the aperture.

Similarly, the first equation of (3.11) shows that ϕ_1 is uniquely determined by its value in the aperture:

$$(5.1) \qquad \phi_1 = \frac{\partial}{\partial z}\left[- \frac{1}{2\pi}\int_A \phi_1 \frac{e^{ikr}}{r} \, d\Sigma \right]$$

In this case we have the further condition that $\partial\phi_1/\partial z = \partial u_o/\partial z$ in the aperture. Since in equation (1) the term in brackets is a solution ψ, say, of the wave equation, we have

$$\frac{\partial\phi_1}{\partial z} = \frac{\partial^2\psi}{\partial z^2} = -\left(k^2 + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\psi,$$

so that the following relation is obtained[26]:

$$(5.2) \qquad \frac{\partial u_o(x,y,o)}{\partial z} = \left[k^2 + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right]\int_A \phi_1(x',y',o)G(x,x',y,y')dx'dy',$$

where (x,y,o) is in the aperture. This relation is not a pure integral equation; it is a differential-integral equation. It should be noted that the differential operator may not be applied to G under the integration sign since the resulting kernel would not be integrable. Maue[27] gave an equivalent form of equation (2), namely,

$$(5.3) \qquad \frac{\partial u_o}{\partial z} = k^2\int \phi_1 G d\sum - \int(\text{grad}\ \phi_1 \cdot \text{grad}\ G)d\sum,$$

where both gradients are to be taken with respect to the coordinates of integration x', y'. Equation (3) follows from (2) by a process of differentiation and integration by parts, and use of the condition $\phi_1 = 0$ at the edge of the aperture. The second integral in (3) is a Cauchy's principle value (small circle around x,y,o of radius $\epsilon \to 0$). A second integration by parts is impossible because of the singularities at the edge.

Only in a few simple cases can the differential-integral equation (2) be transformed into a pure integral equation. If the incident field is a plane wave, $u_o = \exp[ik(\alpha x+\beta y+\gamma z)]$, $\gamma = (1-\alpha^2-\beta^2)^{1/2}$, we have

$$\frac{\partial u_o(x,y,o)}{\partial z} = \left[k^2 + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right](i/k\gamma)\, e^{ik(\alpha x+\beta y)},$$

so that (2) becomes

(5.4) $$\int_A \phi_1(x',y',o)G(x,x',y,y') = \frac{i}{k\gamma} e^{ik(\alpha x+\beta y)} + \chi(x,y) ,$$

where χ is a solution of the two-dimensional wave equation in the aperture. The function χ can be uniquely determined (except for a constant multiplier) for a normally incident wave ($\alpha = \beta = o$) in the cases of the circular aperture [28] and the infinite slit [29]. The resulting equation can be transformed into a pure integral equation with a non-symmetric kernel.

Let us now return to the discussion of the variational formulation of planar diffraction problems as given by Levine and Schwinger. As was mentioned before, the divergent integrals that occur in their paper are easy to eliminate without affecting the numerical results. In what follows, the notation is suggested by that of Levine and Schwinger.

Let $A_1(n,n')$ denote the amplitude of the diffracted wave, where n and n' are unit vectors in the direction of observation and of propagation of the incident plane wave respectively. Further, let θ and θ' be the angles between the positive z axis and these unit vectors. From equation (1) it follows that

(5.5) $$A_1(n,n') = -(ik/2\pi)\cos\theta \int \phi_{n'}(\rho')e^{-ikn\rho'}dS',$$

where $\phi_{n'}(\rho')$ is the value of ϕ_1 in the aperture. The integral equation in Levine and Schwinger's paper [their equation (2.9)] contains a non-integrable kernel and should be interpreted in the sense of one of the differential-integral equations (2) or (3). Let us choose Maue's equation (3). Then

(5.6) $$2\pi\ ik\ \cos\theta' e^{ikn'\rho} = k^2\int\phi_{n'}(\rho')G(\rho,\rho')dS' - \int\nabla'\phi_{n'}(\rho')\nabla'G(\rho,\rho')dS',$$

where ρ is in the aperture and

(5.7) $$G(\rho,\rho') = \frac{e^{ik|\rho-\rho'|}}{|\rho-\rho'|} .$$

If we multiply through in equation (6) by $\phi_{n''}(\rho)$ (where $\phi_{n''}(\rho)$ is the solution for a plane wave in the direction n"), and integrate over the aperture, there results

$$(5.8)\quad 2\pi ik\cos\theta' \int \phi_{n''}(\rho)e^{ikn'\rho}dS = k^2\int \phi_{n''}(\rho)G(\rho,\rho')\phi_{n'}(\rho')dSdS' - \int \nabla\phi_{n''}(\rho)\nabla'\phi_{n'}(\rho')G(\rho,\rho')dSdS'.$$

The last integral appears after an integration by parts; the integrated term drops out because $\phi_{n''}$, like $\phi_{n'}$, vanishes at the edge of the aperture. The right-hand side of equation (8) is symmetrical in n' and n" and consequently so is the left-hand side. If we divide (8) by this left-hand side and by a similar term in which n' and n" are interchanged, and then use (5) and invert, we obtain

$$(5.9)\quad A_1(n'',n') = A_1(-n',-n'')$$

$$= \frac{\cos\theta'\cos\theta''\int \phi_{n'}(\rho)e^{-ikn''\rho}dS\int \phi_{-n''}(\rho)e^{ikn'\rho}dS}{\int\left\{\phi_{n'}(\rho)\phi_{-n''}(\rho')-k^{-2}\nabla\phi_{n'}(\rho)\nabla'\phi_{-n''}(\rho')\right\}G(\rho,\rho')dSdS'}.$$

It can be proved that the equation (9) is stationary with respect to small independent variations of $\phi_{n'}$ and $\phi_{-n''}$ about their <u>correct</u> values; those variations which do not violate the condition $\phi_1 = 0$ at the edge are admissible.

The stationary character of the expression (9) is of importance for approximate calculations. In fact, a judicious choice of aperture distributions in equation (9) may result in a reasonably correct value for A_1 without the necessity of solving the original differential-integral equation. Scale factors are of no account since (9) is homogeneous of degree zero in ϕ. The same remarks apply to the plane-wave transmission cross-section σ of any aperture in a plane screen (perfectly rigid or soft) which, according to Levine and Schwinger [30], is related to the amplitude A of the spherical wave at large distances behind the aperture in the direction of the incident wave by

$$(5.10) \qquad \sigma = -\frac{2\pi}{k} \operatorname{Im} A.$$

From equation (10) it follows that

$$(5.11) \quad \sigma_1(n) = -\frac{2\pi}{k}\cos^2\theta \operatorname{Im} \frac{\int \phi_n(\rho)e^{-ikn\rho}dS \int \phi_{-n}(\rho)e^{ikn\rho}dS}{\int\left\{\phi_n(\rho)\phi_{-n}(\rho')-k^{-2}\nabla\phi_n(\rho)\,\nabla'\phi_{-n}(\rho)\right\}G(\rho,\rho')dSdS'}$$

Levine and Schwinger discussed the limiting form of this equation for low and high frequencies. In the static case as k approaches zero

$$(5.12) \quad G = \frac{e^{ik|\rho-\rho'|}}{|\rho-\rho'|} = \left[\frac{1}{|\rho-\rho'|} + ik - \frac{k^2}{2}|\rho-\rho'| - \frac{ik^3}{6}|\rho-\rho'|^2 + o(k^4)\right].$$

Using equation (12), we obtain the following for the denominator of (11):

$$(5.13) \int\left\{-\frac{\nabla\phi_n\nabla'\phi_{-n}}{|\rho-\rho'|} - ik\nabla\phi_n\nabla'\phi_{-n} + k^2\left(\frac{1}{2}|\rho-\rho'|\,\nabla\phi_n\nabla'\phi_{-n} + \frac{\phi_n\phi_{-n}}{|\rho-\rho'|}\right) + ik^3\left(\phi_n\phi_{-n} + \frac{1}{6}|\rho-\rho'|^2\nabla\phi_n\nabla'\phi_{-n}\right) + o(k^4)\right\}dSdS'.$$

At this point we shall prove the theorem

$$(5.14) \quad \int_S dxdy \int_S dx'dy'\,\nabla\phi_1(x,y)\cdot\nabla'\phi_2(x',y')F(x,x',y,y') = -\int dx'dy'\phi_2(x',y')\int dxdy\phi_1(x,y)\,\Delta F,$$

which is necessary for the further calculation of expression (13). We assume that ϕ_1 and ϕ_2 are arbitrary functions defined in the region S and equal to zero on the boundary of S. The left-hand side of equation (13) can be written in the form

$$(5.15) \quad \int dxdy\,\nabla\phi_1(x,y)\cdot\int dx'dy'\left\{\nabla'\left\{\phi_2 F\right\} - \phi_2\nabla' F\right\}.$$

However, it follows from partial integration and the condition that $\phi_2 = 0$ on the boundary of S that

$$\int dx'dy' \, \nabla' \left\{\phi_2 F\right\} = \int_{\text{Boundary}} \phi_2 F (\text{direction cosine}) = 0.$$

Hence, expression (15) becomes

$$-\int_S dxdy \quad \nabla\phi_1(x,y) \cdot \int dx'dy' \phi_2(x',y') \quad \nabla' F,$$

which can be written as

$$(5.16) \qquad -\int dx'dy' \phi_2(x',y') \int dxdy \; \nabla \, \phi_1(x,y) \quad \nabla' F.$$

Using a similar procedure on the integral of $\nabla\phi_1(x,y) \quad \nabla' F$, the left side of equation (14) can be shown to be equal to

$$+\int dx'dy' \; \phi_2(x',y') \int dxdy \; \phi_1(x,y) \quad \nabla \; \nabla' F.$$

If we assume that F is a function of an argument of the form $(x-x')^2+(y-y')^2$ then $\nabla\nabla' = -\nabla = -\nabla'$, and thus the theorem is proved. Therefore if F is of the form $|\rho-\rho'|^2$ then we find, making use of the theorem (14), that

$$\int \nabla\phi_1 \cdot \nabla'\phi_2 |\rho-\rho'|^2 dSdS' = -4\int \phi_1(\rho)\phi_2(\rho')dSdS'.$$

Now, to continue our discussion of the limiting form of equation (11), we can see that the expression for the denominator becomes

$$(5.17) \qquad \left[-\int \frac{\nabla\phi_n \cdot \nabla'\phi_{-n}}{|\rho-\rho'|} + k^2 \int \frac{\phi_n\phi_{-n}}{|\rho-\rho'|} + \frac{1}{3} ik^3 \int \phi_n\phi_{-n} + 0(k^4)\right]$$

$$= \left[P + k^2 Q + \frac{1}{3} k^3 R\right].$$

The numerator of equation (11) becomes

$$(5.18) \left[\int \phi_n\phi_{-n} + ikn \cdot \int (\rho-\rho')\phi_n\phi_{-n} - \frac{1}{2} k^2 \int \phi_n\phi_{-n} \left[n \cdot (\rho-\rho')\right]^2\right] dSdS'$$

However, for small apertures (with respect to the wavelength) $\phi_n = \phi_{-n} = \phi$ since in this case ϕ is not dependent on a special n. Therefore the leading term of the transmission cross-section is given by

(5.19) $$\sigma_1(n) \sim \frac{2\pi}{3} k^4 \cos^2\theta \frac{\left[\int \phi(\rho) dS\right]^4}{\left[\int \frac{\nabla\phi(\rho)\nabla'\phi(\rho')}{|\rho-\rho'|} dS dS'\right]^2} \text{ as } k \to 0$$

which is an illustration of Rayleigh's λ^{-4} law for small scatterers.

The analogous problem for ϕ_2, discussed by Copson[31] and by Levine[32], is easier since it is based on the pure integral equation (1a). From this equation it follows that

(5.20) $$\int \Psi_{n'}(\rho') G(\rho,\rho') dS' = -2\pi e^{ikn'\rho},$$

where $\Psi_{n'}(\rho')$ is the value of $\partial\phi_2/\partial z$ in the aperture due to a plane incident wave traveling in the direction n'. The amplitude can be found from the second equation (3.11):

(5.21) $$A_2(n,n') = -\frac{1}{2\pi} \int \Psi_{n'}(\rho') e^{-ikn\rho'} dS',$$

and the analogue of (9) becomes

(5.22) $$A_2(n'',n') = A_1(-n',-n'') = \frac{\int \Psi_{n'}(\rho) e^{-ikn''\rho} dS \int \Psi_{-n''}(\rho) e^{ikn'\rho} dS}{\int \Psi_{n'}(\rho) G(\rho,\rho') \Psi_{-n''}(\rho') dS dS'},$$

which expression, again, is stationary with respect to small variations of $\Psi_{n'}$ and $\Psi_{-n''}$ about their correct values. In this case the variations are not restricted by a condition at the edge; the correct aperture fields Ψ are infinite there.

The corresponding formula for the transmission cross-section is

(5.23) $$\sigma_2(n) = -\frac{2\pi}{k} \operatorname{Im} \frac{\int \Psi_n(\rho) e^{-ikn\rho} dS \int \Psi_{-n}(\rho) e^{ikn\rho} dS}{\int \Psi_n(\rho) G(\rho,\rho') \Psi_{-n}(\rho') dS dS'},$$

while the leading term in the static case is given by

$$(5.24)\qquad \sigma_2(n) \sim 2\pi \frac{\left[\int \Psi(\rho)dS\right]^4}{\left[\int \frac{\Psi(\rho)\Psi(\rho')}{|\rho-\rho'|}\, dSdS'\right]^2} \qquad (k \to 0).$$

This leading term, it may be noted, is equal to $2\pi C^2$ where C is the electrostatic capacity of a metal disk of the same shape and size as the aperture.

The variational principle for ϕ_1 was successfully applied by Levine and Schwinger [33] for the diffraction of a normally incident plane wave through a circular aperture. The correct aperture distribution is assumed in the form [34]

$$(5.25)\qquad \phi_1 = \sum_{n=1}^{\infty} a_n(1-\rho^2/a^2)^{n-1/2} ,$$

where the coefficients a_n have yet to be determined. The denominator of equation (9),

$$(5.26)\qquad F = \int \left[k^2\phi_n(\rho)\phi_{n'}(\rho') - \nabla\phi_n(\rho)\nabla'\phi_{n'}(\rho')\right] G(\rho,\rho')dSdS',$$

can be written in the form

$$(5.27)\qquad F = \frac{1}{2\pi}\int_0^a \rho d\rho \int_0^a \rho' d\rho' \int_0^{2\pi} d\Psi' \left[k^2\phi_n(\rho)\phi_{n'}(\rho') - \frac{\partial\phi_n(\rho)}{\partial\rho}\frac{\partial\phi_{n'}(\rho')}{\partial\rho'}\cos(\Psi-\Psi')\right]$$

$$\times \frac{e^{ik\sqrt{\rho^2-2\rho\rho'\cos(\Psi-\Psi')+\rho'^2}}}{\sqrt{\rho^2-2\rho\rho'\cos(\Psi-\Psi')+\rho'^2}}$$

where we used the relationships $\nabla\phi_n(\rho)\cdot\nabla'\phi_{n'}(\rho') = \frac{\partial\phi_n}{\partial\rho}\ \frac{\partial\phi_{n'}}{\partial\rho'}\cos(\Psi-\Psi')$ and

$$G = \frac{e^{ik\sqrt{\rho^2-2\rho\rho'\cos(\Psi-\Psi')+\rho'^2}}}{\sqrt{\rho^2-2\rho\rho'\cos(\Psi-\Psi')+\rho'^2}}.$$

Now G can be written in the form

$$(5.28)\qquad \frac{e^{ik\sqrt{\rho^2-2\rho\rho'\cos(\Psi-\Psi')+\rho'^2}}}{2\pi\sqrt{\rho^2-2\rho\rho'\cos(\Psi-\Psi')+\rho'^2}} = \frac{1}{2\pi}\int_0^{\infty}\frac{\lambda d\lambda}{\sqrt{\lambda^2-k^2}}\, J_0(\lambda\sqrt{\rho^2-2\rho\rho'\cos(\Psi-\Psi')+\rho'^2}).$$

When we substitute equation (28) into equation (27) and use the addition theorem for Bessel functions, the denominator becomes

$$(5.29) \qquad F = 2\pi \int_0^\infty \frac{\lambda d\lambda}{\sqrt{\lambda^2-k^2}} \int_0^a \rho d\rho \int_0^a \rho' d\rho' \Big[k^2 \phi_n(\rho)\phi_{n'}(\rho') J_o(\lambda\rho) J_o(\lambda\rho') - \frac{\partial\phi_n}{\partial\rho}\frac{\partial\phi_{n'}}{\partial\rho'} J_1(\lambda\rho) J_1(\lambda\rho')\Big].$$

However, the expression $\int_0^a \rho d\rho \frac{\partial\phi_n}{\partial\rho} J_1(\lambda\rho)$ when integrated by parts gives us

$$\int_0^a \rho d\rho \frac{\partial\phi_n}{\partial\rho} J_1(\lambda\rho) = \rho\phi_n J_1(\lambda\rho)\Big|_o^a - \int \phi_n \frac{\partial}{\partial\rho}\Big\{\rho \; J_1(\lambda\rho)\Big\} d\rho$$
$$= \int_0^a \phi_n \frac{\partial}{\partial\rho}\Big\{\rho \; J_o'(\lambda\rho)\Big\} d\rho$$
$$= \int_0^a \phi_n \Big\{J_o' (\lambda\rho) + \lambda\rho J_o'' (\lambda\rho)\Big\}$$
$$= -\int_0^a \phi_n \lambda\rho J_o(\lambda\rho) d\rho.$$

From this we see that

$$\int_0^a \rho d\rho \int_0^a \rho' d\rho' \frac{\partial\phi_n}{\partial\rho}\frac{\partial\phi_{n'}}{\partial\rho'} J_1(\lambda\rho) J_1(\lambda\rho') = \lambda^2 \int_0^a \rho d\rho \int_0^a \rho' d\rho' \phi_n(\rho)\phi_n(\rho') J_o(\lambda\rho) J_o(\lambda\rho').$$

Hence equation (29) becomes

$$(5.30) \quad F = 2\pi \int_0^\infty \frac{\lambda(k^2-\lambda^2)}{\sqrt{\lambda^2-k^2}} d\lambda \int_0^a \rho d\rho \int_0^a \rho' d\rho' J_o(\lambda\rho) J_o(\lambda\rho') \phi_n(\rho)\phi_{n'}(\rho'),$$

which can be written finally as

$$(5.31) \quad F = -2\pi \int_0^\infty \lambda \sqrt{\lambda^2-k^2} d\lambda \int_0^a \rho d\rho \int_0^a \rho' d\rho' \phi_n(\rho)\phi_{n'}(\rho') J_o(\lambda\rho) J_o(\lambda\rho').$$

Equation (31) is the same as that obtained by Levine and Schwinger. However, they used a non-integrable kernel in the integral equation, and hence the validity of their procedure is somewhat doubtful. Although the present method is more difficult than Levine and Schwinger's, it is a valid procedure.

Now, if the series expansion (25), where the individual terms are of the form

$$(1 - \rho^2/a^2)^{n-1/2},$$

is substituted in (31), we get a double sum of terms:

$$(5.32)\quad F_{nm} = -2\pi\int_0^\infty \lambda\sqrt{\lambda^2-k^2}\,d\lambda\left[\int_0^a \rho d\rho(1-\rho^2/a^2)^{n-1/2}J_o(\lambda\rho)\right] \times\left[\int_0^a \rho d\rho(1-\rho^2/a^2)^{m-1/2}J_o(\lambda\rho)\right] .$$

However

$$\int_0^a \rho d\rho(1-\rho^2/a^2)^{n-1/2}J_o(\lambda\rho) = a^2\int_0^{\pi/2} J_o(\lambda a\sin\theta)\sin\theta\cos^{2n}\theta d\theta = \frac{a^2 J_{n+1/2}(\lambda a)2^{n-1/2}\Gamma(n+1/2)}{(\lambda a)^{n+1/2}} .$$

Substituting this into equation (32) we have

$$(5.33)\quad F_{nm} = -\pi 2^{n+m}\Gamma(n+\tfrac{1}{2})\Gamma(m+\tfrac{1}{2})a^3\int_0^\infty \frac{J_{n+1/2}(\lambda a)J_{m+1/2}(\lambda a)}{(\lambda a)^{n+m}}\sqrt{\lambda^2-k^2}\,d\lambda,$$

and then writing $\lambda = kv$ we obtain finally

$$(5.34)\quad F_{nm} = -2\pi a(2/ka)^{n+m-2}\,\Gamma(n+1/2)\Gamma(m+1/2) \cdot\int_0^\infty (v^2-1)^{1/2}v^{-(n+m)}J_{n+1/2}(kav)J_{m+1/2}(kav)dv.$$

Moreover, using the equation $\phi_n(\rho) = (1-\rho^2/a^2)^{n-1/2}$, we have

$$\int \phi_n(\rho)dS = 2\pi\int_0^a \rho d\rho(1-\rho^2/a^2)^{n-1/2} = 2\pi\int_0^{\pi/2}\sin\theta\cos^{2n}\theta d\theta = \pi a^2\frac{\Gamma(1)\Gamma(n+1/2)}{\Gamma(n+3/2)} = \frac{\pi a^2}{n+1/2} ,$$

and we can now write A_1 (the amplitude in the forward direction) as

$$(5.35)\quad \frac{-A_1}{2a} = \frac{\left[\sum_{n=1}^{\infty}\frac{a_n}{2n+1}\right]^2}{\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}c_{mn}a_m a_n} ,$$

where the coefficients c_{mn} are defined by

(5.36) $$c_{mn} = (2/ka)^{m+n}\,\Gamma(m+1/2)\,\Gamma(n+1/2)f_{mn}(ka)$$

and

(5.37) $$f_{mn}(ka) = \int_0^\infty (v^2-1)^{1/2} v^{-(m+n)} J_{m+1/2}(kav)J_{n+1/2}(kav)dv.$$

Let us now apply small variations δa_n to the <u>true</u> values of a_n. In view of the stationary character of (35) we then find

$$\frac{1}{2m+1}\sum_{n=1}^{\infty}\frac{a_n}{2n+1} = \frac{-A_1}{2a}\sum_{n=1}^{\infty} c_{mn}a_n \quad (m=1,2,\cdots).$$

On the other hand, it follows from equation (5) that

(5.38) $$A_1 = -ika^2\sum_{n=1}^{\infty}\frac{a_n}{2n+1}.$$

Elimination of A_1 thus gives the following infinite system of linear equations for the unknown a_n†:

(5.39) $$\sum_{n=1}^{\infty} c_{mn}a_n = \frac{1}{ika(m+1/2)} \quad (m = 1,2,\ldots).$$

An approximate solution of (39) was obtained by Levine and Schwinger by assuming $a_n = 0$ if $n > N$ and solving $a_1,\cdots,a_N$ from the first N equations of (39). The corresponding approximate value of the transmission coefficient becomes

(5.40) $$t_1^{(N)} = \mathrm{Re}\sum_{n=1}^{\infty}\frac{2}{2n+1}\,a_n^{(N)}.$$

It seems worth noting that the integrals (37) can be expressed in terms of $F_n = J_n + i\mathbf{H}_n$ where $\mathbf{H}_n$ is Watson's notation for the Struve function. The symmetry between the real and imaginary parts of f_{mn} were not recognized by Levine and Schwinger, although their expressions are easily transformed into the symmetrical form by a partial integration.

† Levine and Schwinger apparently overlooked the fact that their coefficients A_n and D_n are simply related by $A_n = ikD_n$. The factor C_o of Magnus [35] is thus equal to ik, so that his Table I provides at once the first terms of the power series for A_n.

One has

$$f_{11}(\alpha) = -\frac{i}{\pi}\left(\frac{\alpha}{2} - \frac{1}{4\alpha}\right) + \frac{1}{4\alpha}\left(1 + \frac{1}{4\alpha^2}\right)\int_0^{2\alpha} F_0(t)dt - \frac{1}{8\alpha^2} F_0(2\alpha) - \frac{1}{4\alpha} F_1(2\alpha),$$

$$f_{12}(\alpha) = -\frac{i}{\pi}\left(\frac{\alpha^2}{9} - \frac{1}{4} - \frac{3}{8\alpha^2}\right) + \frac{1}{8\alpha^2}\left(1 + \frac{3}{4\alpha^2}\right)\int_0^{2\alpha} F_0(t)dt - \frac{3}{16\alpha^3} F_0(2\alpha) - \frac{3}{8\alpha^2} F_1(2\alpha),$$

(5.41)

$$f_{12}(\alpha) = -\frac{i}{\pi}\left(\frac{\alpha^3}{36} - \frac{\alpha}{12} + \frac{3}{32\alpha} - \frac{45}{64\alpha^3}\right) + \frac{1}{16\alpha}\left(1 + \frac{3}{2\alpha^2} + \frac{45}{16\alpha^4}\right)\int_0^{2\alpha} F_0(t)dt$$

$$+ \frac{9}{32\alpha^2}\left(1 - \frac{5}{4\alpha^2}\right)F_0(2\alpha) - \frac{1}{16\alpha}\left(1 + \frac{45}{4\alpha^2}\right) F_1(2\alpha).$$

The infinite system of linear equations (39) was thoroughly studied by Magnus [36]. He showed that, for sufficiently small values of ka, the solution a_n is unique and can be expanded in a convergent power series in ka. It was also shown that in the N^{th} approximation the exact values of the first N coefficients of the power series for $a_1, \dots, a_N$ are obtained. Explicit recurrence formulae were given. Special attention was paid to the limiting form of (39) for $ka \to \infty$. Owing to an error of sign in the definition of the imaginary part of c_{mn}, the conjugate values should be taken in Magnus's table. The same error (and others) occurred in the paper by Sommerfeld [37], who took the wrong definition of $(\lambda^2-k^2)^{1/2}$ when $0 \leq \lambda \leq k$. The author checked Magnus's table of coefficients and found complete agreement with his own calculations, except for the last column. The following values are taken from the author's paper [38]; we give one term in addition to Magnus:

$$a_1 = -\frac{2ika}{\pi}\Big[1 + \frac{1}{6}(ka)^2 + \frac{2i}{9\pi}(ka)^3 + \frac{1}{120}(ka)^4 + \frac{13i}{225\pi}(ka)^5$$

$$+ \left(\frac{1}{5040} - \frac{4}{81\pi^2}\right)(ka)^6 + \frac{323i}{44100\pi}(ka)^7 + \dots\Big],$$

$$a_2 = -\frac{i(ka)^3}{9\pi}\Big[1 + \frac{3}{10}(ka)^2 + \frac{2i}{5\pi}(ka)^3 + \frac{1}{56}(ka)^4 + \frac{19i}{175\pi}(ka)^5 + \dots\Big],$$

$$a_3 = -\frac{i(ka)^5}{300\pi}\Big[1 + \frac{5}{14}(ka)^2 + \frac{10i}{21\pi}(ka)^3 + \dots\Big],$$

$$a_4 = -\frac{i(ka)^7}{17640\pi}\Big[1 + O(k^2a^2)\Big].$$

For the aperture distribution, an alternative expansion, which is somewhat simpler† than the Levine-Schwinger-Sommerfeld form (39), is in terms of Legendre functions, viz.,

(5.43) $$\phi_1 = \sum_{n=o}^{\infty} b_n P_{2m+1}(\sqrt{1-\rho^2/a^2}).$$

The diffracted amplitude in the forward direction is then determined by the first coefficient alone:

(5.44) $$A_1 = -\frac{1}{3}\, ika^2 b_o.$$

Application of the variational principle leads to the following infinite system for the unknowns b_n:

(5.45) $$\sum_{n=o}^{\infty} d_{mn} b_n = \frac{6}{ika}\,\delta_{mo} \quad (\delta_{oo} = 1,\ \delta_{mo} = o \text{ if } m > o),$$

where

(5.46) $$d_{mn} = (6/ka)^2\,\frac{\Gamma(m+3/2)\,\Gamma(n+3/2)}{m!\;n!}\; g_{mn}(ka);$$
$$g_{mn}(\alpha) = \int_o^{\infty} (v^2-1)^{1/2} v^{-2} J_{2m+3/2}(\alpha v) J_{2n+3/2}(\alpha v)\,dv.$$

In this case the various approximations to the transmission coefficients are given by the ratio of two determinants, viz.,

(5.47) $$t_1^{(N+1)} = \frac{4ka}{9\pi}\ \text{Imagin. part of}\ \frac{\begin{vmatrix} g_{11} & \cdots & g_{1N} \\ \vdots & & \\ g_{N1} & \cdots & g_{NN} \end{vmatrix}}{\begin{vmatrix} g_{oo} & \cdots & g_{oN} \\ \vdots & & \\ g_{No} & \cdots & g_{NN} \end{vmatrix}}$$

in which for N = 0 the upper determinant should be interpreted as unity. It may be verified that equation (47) gives exactly the same approximations as equation (39). The advantage of (47) is in its explicit analytical form, which invites a detailed study along the lines of Magnus's paper.

† It is also simpler than the Legendre-function expansion of Levine and Schwinger [39] obtained by direct integration of the differential-integral equation (5.4).

Only slight changes in the preceding analysis are necessary in order to cover the second boundary-value problem. First, the series analogous to that of (25) represents the aperture values of $\partial\phi_2/\partial z$; in this series we include a term with $n = 0$. Secondly, $+A_2/2a$ is given by equation (26) modified to include the terms $m,n = 0$. Thirdly, $f_{m\,n}(\alpha)$ should be replaced by

$$h_{mn}(\alpha) = \int_0^\infty (v^2-1)^{-1/2} v^{-(m+n)} J_{m+1/2}(\alpha v) J_{n+1/2}(\alpha v)\,dv, \tag{5.48}$$

which function is related to f_{mn} by [40]

$$h_{mn}(\alpha) = (m+n-2) f_{mn} - \alpha \frac{d}{d\alpha} f_{mn}(\alpha). \tag{5.49}$$

For example,

$$h_{oo}(\alpha) = \frac{1}{2\alpha} \int_0^{2\alpha} F_o(t)\,dt,$$

$$h_{01}(\alpha) = \frac{i}{\pi} + \frac{1}{4\alpha^2} \int_0^{2\alpha} F_o(t)\,dt - \frac{1}{2\alpha} F_o(2\alpha), \tag{5.50}$$

$$h_{11}(\alpha) = \frac{i}{\pi}\left(\frac{\alpha}{2} + \frac{3}{4\alpha}\right) + \frac{1}{4\alpha}\left(1 + \frac{3}{4\alpha^2}\right)\int_0^{2\alpha} F_o(t)\,dt - \frac{3}{8\alpha^2} F_o(2\alpha) - \frac{3}{4\alpha} F_1(2\alpha).$$

The corresponding first-order approximation of the transmission coefficient, $t_2^{(1)}$, was calculated by Miles [41] although this was obtained by a less powerful variational principle. Miles introduced an impedance parameter $Z = R-iX$, and the admittance $Y = Z^{-1} = G-iB$, which were evaluated for constant and static-field aperture values of $\partial\phi_2/\partial z$ and compared with the rigorous and Kirchhoff values of the transmission coefficient ($t_2 = \mathrm{Re}Z$). His curves are represented by

$$(P^2 + 4Q^2)Y = P^2\left[1 - \frac{F_1(2ka)}{ka}\right] - \frac{\pi}{2}\, i\, Q^2 \int_0^{2ka} F_o(t)\,dt,$$

where

$(P,Q) = (1,0)$, $(0,1)$ or $(1,2/ka\pi)$.

Again, the simplest analytical approximations of the transmission coefficient t_2 (circular aperture, plane wave at normal incidence) are obtained if we start from an expansion in Legendre functions. We now assume that in the aperture

(5.51) $$\frac{\partial \phi_2}{\partial z} = \sum_{n=o}^{\infty} B_n \frac{P_{2n}(\sqrt{1-\rho^2/a^2})}{\sqrt{1-\rho^2/a^2}} \quad .$$

Then by equation (21), the scattered amplitude becomes $A_2 = -a^2 B_o$, while insertion of (51) in (22) gives

(5.52) $$\frac{A_2}{2a} = \frac{B_o^2}{\sum_{m=o}^{\infty} \sum_{n=o}^{\infty} D_{mn} B_m B_n} \quad ,$$

where

(5.53) $$D_{mn} = \frac{\Gamma(m+1/2)\,\Gamma(n+1/2)}{m!\; n!} G_{mn}(ka); \; G_{mn}(\alpha) = \int_o^{\infty} (v^2-1)^{\frac{1}{2}} J_{2m+1/2}(\alpha v) J_{2n+1/2}(\alpha v) dv.$$

Application of the variational principle for ϕ_2 gives the following infinite system of equations for the unknowns B_n:

(5.54) $$\sum_{n=o}^{\infty} D_{mn} B_n = -\frac{2}{a}\,\delta_{mo} \quad (m = o,1,2,\ldots).$$

The successive approximations of the transmission coefficient are then simply given in close analogy with equation (47), by

$$(5.55) \qquad t_2^{(N+1)} = -\frac{4}{\pi ka} \operatorname{Im} \frac{\begin{vmatrix} G_{11} & \cdots & G_{1N} \\ \vdots & & \\ G_{N1} & \cdots & G_{NN} \end{vmatrix}}{\begin{vmatrix} G_{oo} & \cdots & G_{oN} \\ \vdots & & \\ G_{No} & \cdots & G_{NN} \end{vmatrix}}$$

It should be noted that in the integrals (37), (46), (48), and (53) the roots are understood in the sense

$$(v^2-1)^{1/2} = -i(1-v^2)^{1/2}, \; (v^2-1)^{-1/2} = +i(1-v^2)^{-1/2} \quad \text{when } 0 \leq v < 1.$$

All these integrals can be expressed in terms of f_{mn} and, therefore, in terms of F_n and the indefinite integral of F_o.

VI. Rigorous Form of Babinet's Principle in Electromagnetic Diffraction Theory.

On several occasions we have discussed Babinet's principle in one form or another. Only recently has it been possible to extend this principle so as to be applicable to rigorous electromagnetic diffraction theory. Our representation is essentially that of Copson[42,43] . In what follows, the time factor $\exp(-i\omega t)$ is omitted; as before, k denotes the wave number.

Let $(\vec{f},\vec{g})$ denote any arbitrary incident field, where $\vec{f}$ stands for the electric field vector and $\vec{g}$ for the magnetic field vector. It is assumed, therefore, that $\vec{f}$ and $\vec{g}$ satisfy Maxwell's equations. Later on we shall use the term "complementary" incident field. This is the field defined by $(-\vec{g},\vec{f})$, in the order (electric, magnetic) vector. As is well known, this complementary field also satisfies Maxwell's equations. The complementary field of the complementary field is identical with the original field except for sign.

First of all we consider the diffraction of the field $(\vec{f},\vec{g})$ by a perfectly conducting plane screen (finite or infinite) of zero thickness. Secondly, we consider the diffraction of the complementary field $(-\vec{g},\vec{f})$ by an aperture in a perfectly conducting plane screen of zero thickness;

the aperture in the second problem is of the same size and shape as the screen in the first problem. For simplicity we call these two diffraction problems complementary diffraction problems. The rigorous form of Babinet's principle asserts that the solution of one of these apparently different problems gives, at once, the solution of the remaining problem. We now turn to the proof of this statement, and to its precise form.

In the first problem the total field everywhere in space is given by $(\vec{f} + \vec{E}^s,\ \vec{g} + \vec{H}^s)$, where the scattered field $(\vec{E}^s\ ,\ \vec{H}^s)$ can be derived from the vector potential $\vec{A}^s$ of the currents induced in the screen by the incident flow. Let $\vec{I}$ denote the surface current density vector. Then

$$\vec{A}^s = \frac{1}{c}\int \vec{I}\,\frac{e^{ikr}}{r}\,dS, \tag{6.1}$$

$$\vec{H}^s = \text{curl}\,\vec{A}^s, \qquad -ik\vec{E}^s = k^2\vec{A}^s + \text{grad div}\,\vec{A}^s.$$

The unknown two-component vector $\vec{I}$, defined only on the screen, has to satisfy certain integro-differential equations in order that the well-known boundary conditions shall be satisfied at the surface of the screen. The superscript s will be omitted for all quantities evaluated at the screen's surface. For example, the normal component of the total magnetic field must vanish at the screen. This requires that (z = 0 in the plane of the screen)

$$\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} = -g_z \text{ for all points } (x,y,o) \text{ on the screen.} \tag{6.2}$$

Similarly the x-component of the total electric field must vanish at the screen. Thus

$$\begin{aligned} ikf_x(x,y,o) &= k^2A_x + \frac{\partial}{\partial x}\left[\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y}\right] \\ &= k^2A_x + \frac{\partial^2 A_x}{\partial x^2} + \frac{\partial}{\partial y}\left[\frac{\partial A_x}{\partial y} - g_z\right] \\ &= k^2A_x + \Delta A_x - \frac{\partial g_z}{\partial y} \\ &= k^2A_x + \Delta A_x + ikf_x - \frac{\partial g_y}{\partial z}(x,y,o), \end{aligned}$$

where Δ is the two-dimensional Laplace operator and where we have used the facts that $\vec{A}^s$ has a zero z-component and $(\vec{f},\vec{g})$ is a solution of Maxwell's equations. The last equation can be simplified to

$$k^2A_x + \Delta A_x = \frac{\partial g_y}{\partial z} \quad \text{for all points } (x,y,o) \text{ on the screen.} \tag{6.3}$$

Also, from the condition that the y-component of the total electric field vanish at the screen, we have

(6.4) $$k^2 A_y + \Delta A_y = - \frac{\partial g_x}{\partial z} \quad \text{for all points } (x,y,o) \text{ on the screen.}$$

Finally, bearing in mind that on the screen we have

(6.5) $$A_x = \frac{1}{c} \int I_x \frac{e^{ikr}}{r} \, dS, \qquad A_y = \frac{1}{c} \int I_y \frac{e^{ikr}}{r} \, dS,$$

where $r^2 = (x-x')^2 + (y-y')^2$, $dS = dx'dy'$, $\vec{I} = \vec{I}(x',y')$, we see that equations (2) through (5) constitute a set of integro-differential relations for the unknown current density $\vec{I}$.

By physical intuition we expect these relations to have at least one "admissible" solution $\vec{I} = (I_x, I_y)$ satisfying all physical requirements as to singularities possibly occurring at the edge of the screen. It is not known whether the assumption of absolute integrability of $\vec{I}$ over the screen would entail a unique admissible solution. The integrals in (5) cannot be proper Riemann integrals; they are improper because of singularities of I_x and I_y at the edge. Recent work of Meixner [44], Maue [45], and others makes it plausible that the component of $\vec{I}$ tangential to the edge becomes infinitely large as $D^{-1/2}$ and that the component of $\vec{I}$ normal to the edge vanishes as $D^{1/2}$, where D is the distance to the edge. Similar properties hold for the field vectors themselves, although it is not clear at present what conditions are necessary and/or sufficient for a unique physically acceptable solution.

We now consider the complementary diffraction problem. There is need for a proper distinction between the fields in front of and behind the aperture. Let $(\vec{E}_o, \vec{H}_o)$ denote the total field in the illuminated space if there is no hole in the screen. For example

$$\left. \begin{aligned} E_{ox} &= -g_x(x,y,z) + g_x(x,y,-z) \\ H_{oy} &= f_y(x,y,z) + f_y(x,y,-z) \end{aligned} \right\} \quad z \leq 0.$$

Let $(\vec{E}_1, \vec{H}_1)$ denote the total field for $z \leq 0$ in the presence of the hole, and let $(\vec{E}_2, \vec{H}_2)$ denote the total field for $z \geq 0$ (that is, behind the aperture). Further let $\vec{K}$ be any two-component vector defined in the opening. Define the vector $\vec{F}^d$ by

(6.6) $$\vec{F}^d = -\frac{1}{c}\int \vec{K}\,\frac{e^{ikr}}{r}\,dS.$$

Then we will show that when $\vec{K} = (K_x, K_y)$ satisfies the proper conditions, the fields can be expressed as follows:

(6.7) $$\begin{cases} \vec{E}_1 = \vec{E}_o - \operatorname{curl}\vec{F}^d \\ -ik\vec{H}_1 = -ik\vec{E}_o + k^2\vec{F}^d + \operatorname{grad}\operatorname{div}\vec{F}^d \end{cases} \quad (z \leq 0)$$
$$\begin{cases} \vec{E}_2 = \operatorname{curl}\vec{F}^d \\ +ik\vec{H}_2 = k^2\vec{F}^d + \operatorname{grad}\operatorname{div}\vec{F}^d \end{cases} \quad (z \geq 0).$$

First of all, so long as $\vec{K}$ is integrable, the field defined by the preceding equations in terms of $\vec{K}$ via $\vec{F}^d$ satisfies Maxwell's equations in the half-spaces $z < 0$ and $z > 0$, and it satisfies the appropriate boundary conditions at the screen (i.e., that the tangential electric and normal magnetic fields vanish). Furthermore, no matter what $\vec{K}$ is, the tangential electric and normal magnetic fields are automatically continuous in the aperture.

The only conditions that are not automatically satisfied are that the normal $\vec{E}$ component and the tangential $\vec{H}$ component be continuous in the aperture. In fact these components are easily seen to be identical with the corresponding components of the undisturbed incident field. Again dropping the superscript d when we refer to values on the aperture, we find that these conditions are

(6.8) $$\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} = -\varepsilon_z, \quad \text{for all points } (x,y,o) \text{ on the opening.}$$

$$(6.9) \qquad k^2 F_x + \Delta F_x = \frac{\partial g_y}{\partial z},$$

$$(6.10) \qquad k^2 F_y + \Delta F_y = -\frac{\partial g_x}{\partial z}$$

for all points (x,y,o) on the opening.

It is remarkable that these equations are precisely the same as those for A_x, A_y in the first problem. Assuming that either system has a unique solution, we see that the two mutually complementary problems are simply identical: $\vec{F} \equiv \vec{A}$, $\vec{K} \equiv -\vec{I}$; hence $\vec{F}^d \equiv \vec{A}^s$. Moreover, if we introduce the notation $(\vec{E}_2, \vec{H}_2) \equiv (\vec{E}^d, \vec{H}^d)$ for the diffracted field behind the aperture, we have

$$(6.11) \qquad \vec{E}^d = \vec{H}^s \quad \text{and} \quad \vec{H}^d = -\vec{E}^s \qquad (z \geq 0),$$

and this is Babinet's principle in its rigorous form for diffraction by plane perfectly conducting screens and apertures.

By analogy to the first problem, $\vec{K}$ is called the magnetic current density. Its components show the same behavior at the edge as does the electric current density in the complementary problem. The vector $\vec{F}^d$ is called the magnetic vector potential.

In concluding this section, it may be noted that taking the complementary field [transformation $(\vec{f},\vec{g}) \rightarrow (-\vec{g},\vec{f})$] of a plane wave is equivalent to rotating the plane of polarization through a right angle counterclockwise, looking in the direction of propagation.

VII. Diffraction by a Small Aperture in a Perfectly Conducting Plane Screen

In this section we shall comment upon Bethe's [46] theory of electromagnetic diffraction by small apertures. It appears [47] that Bethe's first-order approximation is fundamentally incorrect with respect to the field near the aperture.

Let $\vec{E}^i$, $\vec{H}^i$ denote the incident field, and let $\vec{E}_o$, $\vec{H}_o$ denote the field in the illuminated half-space ($z \leq 0$) if there is no aperture in the screen ($z = 0$). Then

$$E_{ox} = E_x^i(z) - E_x^i(-z)$$

$$E_{oy} = E_y^i(z) - E_y^i(-z)$$

$$E_{oz} = E_z^i(z) + E_z^i(-z)$$

$$H_{ox} = H_x^i(z) + H_x^i(-z)$$

$$H_{oy} = H_y^i(z) + H_y^i(-z)$$

$$H_{oz} = H_z^i(z) - H_z^i(-z),$$

where $z \leq 0$ and where we have omitted explicit reference to the x,y-coordinates.

As we have seen in the preceding section, the diffracted field can be derived from fictitious magnetic currents (and charges) in the aperture:

$$(7.1)\quad \begin{cases} \vec{E}_1 = \vec{E}_o - \text{curl}\,\vec{F} & (z \leq 0) \\ \vec{H}_1 = \vec{H}_o + ik\vec{F} + (i/k)\,\text{grad div}\,\vec{F} & \text{(in front of aperture)} \\ \vec{E}_2 = \text{curl}\,\vec{F} & (z \geq 0) \\ \vec{H}_2 = -ik\vec{F} - (i/k)\,\text{grad div}\,\vec{F} & \text{(behind aperture)}, \end{cases}$$

where $\vec{F}$ is the magnetic vector potential given in terms of the currents $\vec{K}$ by means of

(7.2) $$\vec{F} = -\frac{1}{c}\int_{\text{aperture}} \vec{K}\,\frac{e^{ikr}}{r}\,dS.$$

The magnetic charge density, η , is found from

(7.3) $$\operatorname{div}\vec{K} = \frac{\partial K_x}{\partial x} + \frac{\partial K_y}{\partial y} = ikc\eta,$$

and the scalar magnetic potential ψ satisfies the equation

(7.4) $$\psi = \int_{\text{aperture}} \eta\,\frac{e^{ikr}}{r}\,dS = \frac{1}{k}\operatorname{div}\vec{F}.$$

In order that these formulae hold it is necessary that the component of $\vec{K}$ normal to the edge of the screen vanishes at the edge(cf. Bouwkamp [48]).

For arbitrary $\vec{K}$ the tangential electric and normal magnetic components are automatically continuous in the aperture. We have for these components

(7.5) $$\vec{K} = \frac{c}{2\pi}\,[\vec{\mathcal{E}}\times\vec{n}], \qquad \eta = \frac{1}{2\pi}(\vec{\mathcal{H}}\cdot\vec{n}),$$

where script letters refer to values in the aperture, and $\vec{n}$ is the unit vector in the positive z-direction.

Conditions for $\vec{K}$ are obtained by requiring that the normal electric and tangential magnetic components are continuous in the aperture. Thus we have

$$\mathcal{E}_z = \lim E_{1z} = \lim E_{2z} = \frac{1}{2}E_{oz}$$

$$\mathcal{H}_x = \lim H_{1x} = \lim H_{2x} = \frac{1}{2}H_{ox} \quad \text{(and same for y-comp.)}.$$

In the opening, therefore,

$$\vec{\mathcal{E}}\cdot\vec{n} = \tfrac{1}{2}\vec{\mathcal{E}}_0\cdot\vec{n} = \vec{\mathcal{E}}^i\cdot\vec{n}$$

$$\vec{\mathcal{H}}\times\vec{n} = \tfrac{1}{2}\vec{\mathcal{H}}_0\times\vec{n} = \vec{\mathcal{H}}^i\times\vec{n}.$$

Consequently, as a by-product we get the theorem: in the aperture the values of the tangential magnetic and normal electric field components are exactly equal to the values of the corresponding components of the undisturbed incident field.

As we have seen in the preceding section, the above requirements lead to the system of integral-differential equations

$$(7.7)\quad \left\{\begin{aligned} \mathcal{F}_x &= -\frac{1}{c}\int K_x \frac{e^{ikr}}{r}\,dS \\ \mathcal{F}_y &= -\frac{1}{c}\int K_y \frac{e^{ikr}}{r}\,dS \\ k^2\mathcal{F}_x + \Delta\mathcal{F}_x &= -\frac{\partial \mathcal{E}^i_y}{\partial z} \\ k^2\mathcal{F}_y + \Delta\mathcal{F}_y &= +\frac{\partial \mathcal{E}^i_x}{\partial z} \\ \frac{\partial \mathcal{F}_y}{\partial x} - \frac{\partial \mathcal{F}_x}{\partial y} &= \mathcal{E}^i_z \end{aligned}\right\}$$

$\vec{\mathcal{F}} = \vec{\mathcal{F}}(x,y)$

$\vec{K} = \vec{K}(x',y')$

$dS = dx'dy'$

$r^2 = (x-x')^2 + (y-y')^2$

(x,y,o) any point in aperture

(x',y',o) same

$\frac{\partial \mathcal{E}^i_y}{\partial z} = \left[\frac{\partial}{\partial z} E^i_y(x,y,z)\right]_{z=0}$

So far our formulation is general. We next consider the case of long waves ($k \to 0$), and assume that $\vec{K}$ can be expanded in a power series of k. Let $\vec{K}^o$ and $\vec{K}^1$ denote the parts of $\vec{K}$ of relative order zero and one respectively. It appears that relative order is the same as absolute order. Thus assume

$$\vec{K} = \vec{K}^o + \vec{K}^1 + O(k^2),$$

where $\vec{K}^o = O(1)$, $\vec{K}^1 = O(k)$. Then, expanding the exponential function in powers of k, we have

$$\vec{\mathcal{F}} = -\frac{1}{c}\int \frac{\vec{K}^o}{r}\,dS - \frac{1}{c}\int \frac{\vec{K}^1}{r} - \frac{1}{c}\int ik\vec{K}^o dS + O(k^2).$$

The third integral, of order k, apparently is independent of x and y. The result of differentiation of this term with respect to x or y is identically zero. We need not retain this term in the expressions for the fields in our order of approximation. Therefore $\vec{\mathcal{F}} = \vec{\mathcal{F}}^o + \vec{\mathcal{F}}^1 + \text{const.}k + O(k^2)$, where

$$(7.8a)\quad \begin{aligned} \vec{\mathcal{F}}^o &\equiv -\frac{1}{c}\int \frac{\vec{K}^o}{r}\,dS && \text{(order 1)} \\ \vec{\mathcal{F}}^1 &\equiv -\frac{1}{c}\int \frac{\vec{K}^1}{r}\,dS, && \text{(order k)} \end{aligned}$$

with the same convention as to the superscripts zero and one as before.

We have $\Delta\vec{\mathcal{F}} = \Delta\vec{\mathcal{F}}^{o} + \Delta\vec{\mathcal{F}}^{1} + O(k^2)$. Therefore, except for terms of order k^2, we may replace $k^2\vec{\mathcal{F}} + \Delta\vec{\mathcal{F}}$ by $\Delta\vec{\mathcal{F}}$. Equations (7) then reduce to (8a) and the following

$$
(7.8b)\begin{cases}
\Delta\mathcal{F}_x^{o} = \left[-\dfrac{\partial\mathcal{E}_y^{i}}{\partial z}\right]_{k=o}; & \Delta\mathcal{F}_x^{1} = \left[\dfrac{\partial}{\partial k}\left\{-\dfrac{\partial\mathcal{E}_y^{i}}{\partial z}\right\}\right]_{k=o} \\
\Delta\mathcal{F}_y^{o} = \left[+\dfrac{\partial\mathcal{E}_x^{i}}{\partial z}\right]_{k=o}; & \Delta\mathcal{F}_y^{1} = \left[\dfrac{\partial}{\partial k}\left\{+\dfrac{\partial\mathcal{E}_x^{i}}{\partial z}\right\}\right]_{k=o} \\
\dfrac{\partial\mathcal{F}_y^{o}}{\partial x} - \dfrac{\partial\mathcal{F}_x^{o}}{\partial y} = \left[\mathcal{E}_z^{i}\right]_{k=o}; & \dfrac{\partial\mathcal{F}_y^{1}}{\partial x} - \dfrac{\partial\mathcal{F}_x^{1}}{\partial y} = \left[\dfrac{\partial}{\partial k}\left\{\mathcal{E}_z^{i}\right\}\right]_{k=o}.
\end{cases}
$$

Let us evaluate the right-hand sides of equations (8b) for a plane-polarized obliquely incident electromagnetic wave. We choose the xz-plane as plane of incidence and call θ_o the angle between the z-axis and direction of incidence. The phase function of the incident wave therefore is $\exp[ik(x\sin\theta_o + z\cos\theta_o)]$. Further, let ϕ_o denote the angle between $\vec{E}^{i}$ and the xz-plane. Then, in an obvious component notation,

$$\vec{E}^{i} = (\cos\phi_o\cos\theta_o, \sin\phi_o, -\cos\phi_o\sin\theta_o)\exp[ik(x\sin\theta_o + z\cos\theta_o)]$$

$$\vec{H}^{i} = (-\sin\phi_o\cos\theta_o, \cos\phi_o, \sin\phi_o\sin\theta_o)\exp[ik(x\sin\theta_o + z\cos\theta_o)].$$

Accordingly, in this case we have

$$-\frac{\partial\mathcal{E}_y^{i}}{\partial z} = -ik\sin\phi_o\cos\theta_o\ \exp(ikx\sin\theta_o)$$

$$\frac{\partial\mathcal{E}_x^{i}}{\partial z} = ik\cos\phi_o\cos^2\theta_o\ \exp(ikx\sin\theta_o)$$

$$\mathcal{E}_z^{i} = -\cos\phi_o\sin\theta_o\ \exp(ikx\sin\theta_o),$$

and equations (8b) become

$$(7.9)\quad \begin{cases} \Delta\mathcal{F}_x^o = 0 & \Delta\mathcal{F}_x^1 = -i\sin\phi_o\cos\theta_o \\ \Delta\mathcal{F}_y^o = 0 & \Delta\mathcal{F}_y^1 = i\cos\phi_o\cos^2\theta_o \\ \dfrac{\partial\mathcal{F}_y^o}{\partial x} - \dfrac{\partial\mathcal{F}_x^o}{\partial y} = -\cos\phi_o\sin\theta_o & \dfrac{\partial\mathcal{F}_y^1}{\partial x} - \dfrac{\partial\mathcal{F}_x^1}{\partial y} = -ix\cos\phi_o\sin^2\theta_o . \end{cases}$$

Note that if we had replaced $-\dfrac{\partial\mathcal{E}_y^i}{\partial z}$, $+\dfrac{\partial\mathcal{E}_x^i}{\partial z}$, $\mathcal{E}_z^i$ in equations (7) by their corresponding constant values at the origin of coordinates (which may be the center of a circular aperture, for example) we would have obtained the approximate equations

$$(7.10)\quad \begin{cases} \Delta\mathcal{F}_x = -ik\sin\phi_o\cos\theta_o \\ \Delta\mathcal{F}_y = ik\cos\phi_o\cos^2\theta_o \\ \dfrac{\partial\mathcal{F}_y}{\partial x} - \dfrac{\partial\mathcal{F}_x}{\partial y} = -\cos\phi_o\sin\theta_o \end{cases} .$$

and consequently, bearing in mind that $\vec{\mathcal{F}} = \vec{\mathcal{F}}^o + \vec{\mathcal{F}}^1$, we would have obtained eqs. (9) except for the last one:

$$\frac{\partial\mathcal{F}_y^1}{\partial x} - \frac{\partial\mathcal{F}_x^1}{\partial y} = 0 \text{ instead of } = -ix\cos\phi_o\sin^2\theta_o .$$

Equations (10) are essentially those of Bethe [46] and Copson [42]. They can, therefore, only lead to a correct $\vec{\mathcal{F}}^o$ term; the term $\vec{\mathcal{F}}^1$ is necessarily wrong in their approximation.

It is possible to eliminate derivatives with respect to k by introducing derivatives with respect to x,y:

$$-\cos\phi_o\sin\theta_o\left[1+ikx\sin\theta_o\right] \approx \mathcal{E}_z^i(0,0,0) + x\,\frac{\partial\mathcal{E}_z}{\partial x}(0,0,0).$$

We then have expressed all constants in terms of center values. More generally, we assume

$$\mathcal{E}_z^i(x,y,o) \approx \mathcal{E}_z^i(0,0,0) + x\,\frac{\partial\mathcal{E}_z^i}{\partial x}(0,0,0) + y\,\frac{\partial\mathcal{E}_z^i}{\partial y}(0,0,0).$$

Our final equations thus become, for any plane-wave excitation, (7.11) $\Delta \mathcal{F}_x^o = 0$; (7.14) $\Delta \mathcal{F}_x^1 = -P$;

(7.12) $\Delta \mathcal{F}_y^o = 0$; (7.15) $\Delta \mathcal{F}_y^1 = Q$;

(7.13) $\dfrac{\partial \mathcal{F}_y^o}{\partial x} - \dfrac{\partial \mathcal{F}_x^o}{\partial y} = \gamma$; (7.16) $\dfrac{\partial \mathcal{F}_y^1}{\partial x} - \dfrac{\partial \mathcal{F}_x^1}{\partial y} = Rx + Sy$,

where $P = \dfrac{\partial E_y^1}{\partial z}$, $Q = \dfrac{\partial E_x^1}{\partial z}$, $R = \dfrac{\partial E_z^1}{\partial x}$, $S = \dfrac{\partial E_z^1}{\partial y}$, $\gamma = E_z^1$.

All these equations are evaluated at the center of coordinates. The constants have the following orders of magnitude:

$$\gamma = O(1); \quad P, Q, R, S, = O(k).$$

This theory will now be applied to the case of a circular aperture of small radius a ($ka \ll 1$). For that purpose we want some complicated integrals of the type

$$G(x,y) \equiv \frac{1}{\pi^2}\int \frac{f(x',y')}{\sqrt{a^2-x'^2-y'^2}} \frac{dx'dy'}{\sqrt{(x-x')^2+(y-y')^2}} \qquad (x'^2 + y'^2 \le a^2).$$

If

$f = 1$, then $G = 1$ and $\Delta G = 0$

$f = x'$, then $G = \frac{1}{2} x$ and $\Delta G = 0$

$f = x'y'$, then $G = \frac{3}{8} xy$ and $\Delta G = 0$

$f = x'^2$, then $G = \frac{1}{16}(4a^2+5x^2-y^2)$ and $\Delta G = \frac{1}{2}$.

(For more general results, see Bouwkamp[49,50].) In virtue of this we find that

$$-\frac{1}{c} K_x^o = \frac{p_o+p_1x'+p_2y'+p_3x'y'+p_4(x'^2-y'^2)}{\pi^2\sqrt{a^2-x'^2-y'^2}}$$

$$-\frac{1}{c} K_y^o = \frac{q_o+q_1y'+q_2x'+q_3x'y'+q_4(y'^2-x'^2)}{\pi^2\sqrt{a^2-x'^2-y'^2}}$$

solves (11) and (12) [with 10 arbitrary parameters (p and q)]. Furthermore, on evaluation of the relevant integrals we find that

$$\mathfrak{F}_x^o = p_o + \frac{1}{2} p_1 x + \frac{1}{2} p_2 y + \frac{3}{8} p_3 xy + \frac{3}{8} p_4 (x^2-y^2)$$

$$\mathfrak{F}_y^o = q_o + \frac{1}{2} q_1 y + \frac{1}{2} q_2 x + \frac{3}{8} q_3 xy + \frac{3}{8} q_4 (y^2-x^2)$$

and thus

$$\frac{\partial \mathfrak{F}_y^o}{\partial x} - \frac{\partial \mathfrak{F}_x^o}{\partial y} = \frac{1}{2}(q_2-p_2) + \frac{3}{8} \left[(q_3+2p_4)y - (p_3+2p_4)x\right].$$

This should equal γ, therefore $q_3 = -2p_4$, $p_3 = -2q_4$, $q_2-p_2 = 2\gamma$. Then we have

$$(7.17)\begin{cases} -\dfrac{1}{c} K_x^o = \dfrac{p_o+p_1x+p_2y-2q_4xy+p_4(x^2-y^2)}{\pi^2\sqrt{a^2-x^2-y^2}} \\ -\dfrac{1}{c} K_y^o = \dfrac{q_o+q_1y+q_2x-2p_4xy+q_4(y^2-x^2)}{\pi^2\sqrt{a^2-x^2-y^2}} \end{cases},$$

which contains 7 independent parameters (8 parameters p, q with one relation, $q_2-p_2 = 2\gamma$, between them), is a general solution of equations (11),(12), (13) and the first equation (8a). Equation (17) is presumably the general solution with the requirement that K be absolutely integrable over the aperture.

It is clear that additional conditions are necessary in order to find a unique solution.

Of course, the charge density should be integrable over the aperture. In our case, this means that $xK_x^o + yK_y^o$ should contain a factor $a^2-x^2-y^2$. Equivalently, the radial component K_ρ of $\vec{K}$ must vanish at $\rho = a$. If we assume

$$(xK_x^o+yK_y^o)\sqrt{a^2-x^2-y^2} \equiv (d_o+d_1x+d_2y)(a^2-x^2-y^2),$$

we find that all coefficients (also d_o, d_1, d_2) are to be zero, except for the one relation $p_2+q_2 = 0$. We are then left with a unique solution

$$\frac{1}{c} K_x^o = \frac{\gamma y}{\pi^2\sqrt{a^2-\rho^2}}, \quad \frac{1}{c} K_y^o = \frac{-\gamma x}{\pi^2\sqrt{a^2-\rho^2}}$$

or in vector notation

(7.18) $$\frac{1}{c}\vec{K}^{o} = \frac{1}{2\pi^2}\frac{\vec{\rho}\times\vec{F}_o(0,0,0)}{\sqrt{a^2-\rho^2}} \quad (= K_E \text{ of Bethe}).$$

Note that $\operatorname{div}\vec{K}^{o} = 0$.

In the same way we determine K^1. We split this vector into four simpler ones which can be more simply treated:

$$\vec{K}^1 = \vec{k}^{(1)} + \vec{k}^{(2)} + \vec{k}^{(3)} + \vec{k}^{(4)}.$$

First Part:

In what follows the notation $(a \rightarrow b)$ means "contribution to a is b".

$$\frac{1}{c}k_x^{(1)} = \frac{2P(-2a^2+2x^2+y^2)}{3\pi^2\sqrt{a^2-x^2-y^2}}, \qquad \frac{1}{c}k_y^{(1)} = \frac{2P\,xy}{3\pi^2\sqrt{a^2-x^2-y^2}};$$

$$\mathcal{F}_x \rightarrow \frac{P}{24}(20a^2-9x^2-3y^2), \qquad \mathcal{F}_y \rightarrow \frac{1}{4}P\,xy;$$

$$\Delta\mathcal{F}_x \rightarrow -P, \qquad \Delta\mathcal{F}_y \rightarrow 0;$$

$$\frac{\partial\mathcal{F}_y}{\partial x} - \frac{\partial\mathcal{F}_x}{\partial y} \rightarrow 0.$$

Second Part:

$$\frac{1}{c}k_x^{(2)} = -\frac{2Q\,xy}{3\pi^2\sqrt{a^2-x^2-y^2}}, \qquad \frac{1}{c}k_y^{(2)} = -\frac{2Q(-2a^2+x^2+2y^2)}{3\pi^2\sqrt{a^2-x^2-y^2}};$$

$$\mathcal{F}_x \rightarrow \frac{1}{4}Q\,xy, \qquad \mathcal{F}_y \rightarrow -\frac{Q}{24}(20a^2-3x^2-9y^2);$$

$$\Delta\mathcal{F}_x \rightarrow 0, \qquad \Delta\mathcal{F}_y \rightarrow Q;$$

$$\frac{\partial\mathcal{F}_y}{\partial x} - \frac{\partial\mathcal{F}_x}{\partial y} \rightarrow 0.$$

Third Part:

$$\frac{1}{c}\,k_x^{(3)} = \frac{4R\;xy}{3\pi^2\sqrt{a^2-x^2-y^2}}\;, \qquad \frac{1}{c}\,k_y^{(3)} = -\frac{2R(a^2+x^2-y^2)}{3\pi^2\sqrt{a^2-x^2-y^2}}\;;$$

$$\mathcal{F}_x \rightarrow -\frac{1}{2}R\;xy\;, \qquad \mathcal{F}_y \rightarrow \frac{R}{12}\;(8a^2+3x^2-3y^2)\;;$$

$$\Delta\mathcal{F}_x \rightarrow 0\;, \qquad \Delta\mathcal{F}_y \rightarrow 0\;;$$

$$\frac{\partial\mathcal{F}_y}{\partial x} - \frac{\partial\mathcal{F}_x}{\partial y} \rightarrow R\;x\;.$$

Fourth Part:

$$\frac{1}{c}\,k_x^{(4)} = \frac{2S(a^2-x^2+y^2)}{3\pi^2\sqrt{a^2-x^2-y^2}}\;, \qquad \frac{1}{c}\,k_y^{(4)} = -\frac{4S\;xy}{3\pi^2\sqrt{a^2-x^2-y^2}}\;;$$

$$\mathcal{F}_x \rightarrow -\frac{S}{12}\;(8a^2+3y^2-3x^2), \qquad \mathcal{F}_y \rightarrow \frac{1}{2}\,S\;xy\;;$$

$$\Delta\mathcal{F}_x \rightarrow 0, \qquad \Delta\mathcal{F}_y \rightarrow 0\;;$$

$$\frac{\partial\mathcal{F}_y}{\partial x} - \frac{\partial\mathcal{F}_x}{\partial y} \rightarrow Sy.$$

It can be verified that the corresponding charge density contributions are determined by

$$\frac{1}{c}\,\mathrm{div}\,\vec{k}^{(1)} = \frac{2Px}{\pi^2\sqrt{a^2-\rho^2}}$$

$$\frac{1}{c}\,\mathrm{div}\,\vec{k}^{(2)} = -\frac{2Qy}{\pi^2\sqrt{a^2-\rho^2}}$$

$$\frac{1}{c}\,\mathrm{div}\,\vec{k}^{(3)} = \frac{2Ry}{\pi^2\sqrt{a^2-\rho^2}}$$

$$\frac{1}{c}\,\mathrm{div}\,\vec{k}^{(4)} = \frac{-2Sx}{\pi^2\sqrt{a^2-\rho^2}}\;,$$

and hence

$$\text{(7.19)} \qquad \frac{1}{c}\,\mathrm{div}\,\vec{K}^1 = \frac{2}{\pi^2}\,\frac{(P-S)x+(R-Q)y}{\sqrt{a^2-\rho^2}}\,.$$

By virtue of Maxwell's equations for the incident field, we have

$$P - S = -ik\,H_x^i\,(0,0,0)$$

$$R - Q = -ik\,H_y^i\,(0,0,0).$$

Moreover $\mathrm{div}\,\vec{K}^o = 0$, so that the magnetic charge density becomes

$$\text{(7.20)} \qquad \eta = \frac{1}{ikc}\,\mathrm{div}(\vec{K}^o+\vec{K}^1) = -\frac{1}{\pi^2}\,\frac{\vec{\rho}\cdot\vec{H}_o(0,0,0)}{\sqrt{a^2-\rho^2}} \qquad \text{(Bethe)}.$$

Bethe's expression for $\frac{1}{c}\vec{K}^1$ (his K_H) is wholly incorrect, although $\mathrm{div}\,\vec{K}_H$ is correct. According to Bethe

$$K_H = \frac{ik}{\pi^2}\sqrt{a^2-\rho^2}\,\vec{H}_o(0,0,0),$$

but the correct value is

$$\frac{1}{c}K_x^1 = (K_H)_x + \frac{2}{3\pi^2}\left[\frac{(P-2S)(a^2-x^2-2y^2)+(2R-Q)xy}{\sqrt{a^2-x^2-y^2}}\right]$$

$$\frac{1}{c}K_y^1 = (K_H)_y + \frac{2}{3\pi^2}\left[\frac{(P-2S)xy+(2R-Q)(a^2-2x^2-y^2)}{\sqrt{a^2-x^2-y^2}}\right].$$

The correction is precisely a divergence-free vector. This difference can be written as the curl of a certain vector. In fact, one finds

$$\frac{1}{c}\vec{K}^1 = \vec{K}_H + \frac{2}{3\pi^2}\,\mathrm{curl}\left[\vec{n}\sqrt{a^2-x^2-y^2}\left\{(P-2S)y + (Q-2R)x\right\}\right]$$

$$= K_H + \frac{2}{3\pi^2}\,\mathrm{curl}\left[\vec{n}\sqrt{a^2-x^2-y^2}\left\{x\,\frac{\partial E_x^i}{\partial z} + y\,\frac{\partial E_y^i}{\partial z} - 2\left(x\,\frac{\partial E_z^i}{\partial x} + y\,\frac{\partial E_z^i}{\partial y}\right)\right\}\right].$$

Using vector notation we have the following expression, which can be applied to plane-wave excitation:

$$(7.21) \quad \frac{1}{c}\vec{K}^1 = \frac{1}{3\pi^2}\Big[\sqrt{a^2-\rho^2}\Big\{2ik\vec{H}_o+\vec{n}\times \mathrm{grad}(\vec{n}\cdot\vec{E}_o)\Big\} + \vec{\rho}\cdot\Big\{ik\vec{H}_o\times\vec{n}-\mathrm{grad}(\vec{n}\cdot\vec{E}_o)\Big\}\frac{\vec{n}\times\vec{\rho}}{\sqrt{a^2-\rho^2}}\Big],$$

where all quantities refer to values at the center of the aperture and the gradients are taken in the tangential direction of the aperture; $\vec{\rho}$ is the radius vector from the center of the circular aperture.

Equation (21) shows that K^1_ρ vanishes at the edge as $\sqrt{a^2-\rho^2}$ and that K^1_ϕ becomes infinitely large as $1/\sqrt{a^2-\rho^2}$, since $(\vec{n}\times\vec{\rho})$ has the direction of ϕ.

The complete expressions for our plane-wave-incidence case become

$$(7.22) \quad \begin{aligned} \frac{1}{c}K^1_\rho &= \frac{4ik}{3\pi^2}\sqrt{a^2-\rho^2}\Big[-\sin\phi_o\cos\theta_o\cos\phi+\cos\phi_o(1-\tfrac{1}{2}\sin^2\theta_o)\sin\phi\Big] \\ \frac{1}{c}K^1_\phi &= \frac{4ik}{3\pi^2}\frac{1}{\sqrt{a^2-\rho^2}}\Big[\sin\phi_o\cos\theta_o(a^2-\tfrac{1}{2}\rho^2)\sin\phi+\cos\phi_o\Big\{(a^2-\tfrac{1}{2}\rho^2) \\ &\quad + (\rho^2-\tfrac{1}{2}a^2)\sin^2\theta_o\Big\}\cos\phi\Big]. \end{aligned}$$

For the total field in the aperture we get

$$(7.23) \quad \begin{aligned} \mathcal{E}_\rho &= -\frac{2\cos\phi_o\sin\theta_o}{\pi}\frac{\rho}{\sqrt{a^2-\rho^2}} \\ &\quad -\frac{8ik}{3\pi}\frac{1}{\sqrt{a^2-\rho^2}}\Big[\sin\phi_o\cos\theta_o(a^2-\tfrac{1}{2}\rho^2)\sin\phi \\ &\quad + \cos\phi_o\Big\{(a^2-\tfrac{1}{2}\rho^2)+(\rho^2-\tfrac{1}{2}a^2)\sin^2\theta_o\Big\}\cos\phi\Big] + 0(k^2a^2) \\ \mathcal{E}_\phi &= \frac{8ik}{3\pi}\sqrt{a^2-\rho^2}\Big[-\sin\phi_o\cos\theta_o\cos\phi+\cos\phi_o(1-\tfrac{1}{2}\sin^2\theta_o)\sin\phi\Big] + 0(k^2a^2) \\ \mathcal{H}_z &= \frac{4\rho}{\pi\sqrt{a^2-\rho^2}}\Big[\sin\phi_o\cos\theta_o\cos\phi-\cos\phi_o\sin\phi\Big] + 0(ka) \end{aligned}$$

where

$\mathcal{E}_z$, $\mathcal{H}_\rho$, $\mathcal{H}_\phi$ are the undisturbed incident field.

For the case of normal incidence ($\theta_o=0$, $\phi_o=0$) this result reduces to the first few terms of an earlier expression (Bouwkamp[48]).

Once we know small-aperture approximations for $\vec{K}$, (see eqs. (18) and (21)) we can try to evaluate the vector potential $\vec{F}$ to the same order of approximation by means of (2) and ultimately to arrive at the corresponding field by means of (1). This process is easy for the field at large distances from the aperture (in the wave zone). Applying known methods we find the asymptotic representations

$$F_x \sim \frac{4a^3}{3\pi}\left[P-S+\frac{1}{2}ik\gamma\sin\theta\sin\phi\right]\frac{e^{ikr}}{r}$$

$$= -\frac{4a^3ik}{3\pi}\left[\alpha-\frac{1}{2}\gamma\sin\theta\sin\phi\right]\frac{e^{ikr}}{r}$$

$$F_y \sim -\frac{4a^3}{3\pi}\left[Q-R+\frac{1}{2}ik\gamma\sin\theta\cos\phi\right]\frac{e^{ikr}}{r}$$

$$= -\frac{4a^3ik}{3\pi}\left[\beta+\frac{1}{2}\gamma\sin\theta\cos\phi\right]\frac{e^{ikr}}{r},$$

where $\alpha = H_x^i(0,0,0)$, $\beta = H_y^i(0,0,0)$, and P, Q, R, S have the same meaning as before, while r, θ, ϕ are spherical coordinates.

In vector notation, one has

$$(7.24)\qquad F = -\frac{a^3ik}{3\pi}\left[2\vec{H}_o+\vec{E}_o\, x\vec{r}_o\right]\frac{e^{ikr}}{r}\left\{1+0\left(\frac{1}{r}\right)\right\},$$

where $\vec{r}_o$ is a unit vector in the r-direction; the values of $\vec{H}_o$ and $\vec{E}_o$ are those taken at the center of the aperture. Using $\mathrm{curl}(\vec{E}_o x\vec{r}_o) = 0(\frac{1}{r})$, we find that in the wave zone $\vec{E} = \mathrm{curl}\,\vec{F} + 0(\frac{1}{r})$, so that

$$(7.25)\qquad E_2 \sim \frac{k^2a^3}{3\pi}\vec{r}_o x\left[2\vec{H}_o+\vec{E}_o x\vec{r}_o\right]\frac{e^{ikr}}{r}$$

$$H_2 \sim \vec{r}_o x\vec{E} \sim \frac{k^2a^3}{3\pi}\vec{r}_o x\left[\vec{E}_o+\vec{r}_o x2\vec{H}_o\right]\frac{e^{ikr}}{r}.$$

These relations are those of Bethe. Although his expression for $\vec{K}$ was wrong, the far-field expressions turn out to be right. The explanation is that the correction term in $\vec{K}$ is a divergence-free vector.

The field (25) behind the aperture in the wave zone is exactly the sum of an electric dipole (moment $\vec{P}_e$) and a magnetic dipole (moment $\vec{P}_m$), both located at the center of the aperture:

(7.26)
$$\vec{P}_e = + \frac{a^3}{3\pi}\vec{E}_o \qquad \text{(normal to plane of the aperture)}$$
$$\vec{P}_m = - \frac{2a^3}{3\pi}\vec{H}_o \qquad \text{(in the plane of the aperture).}$$

Note that the sign of $\vec{P}_e$ is incorrect in Bethe's paper [46].

We now turn to the transmission cross-section of the small aperture. The procedure for calculating is well known:

$$|\vec{E}|^2 = k^2|\vec{r}_o x\vec{F}|^2 = k^2(F_\theta^2 + F_\phi^2)$$

$$F_\theta = \cos\theta[F_x\cos\phi+F_y\sin\phi] = -\frac{4a^3ik}{3\pi}\frac{e^{ikr}}{r}[\alpha\cos\phi+\beta\sin\phi]\cos\theta$$

$$F_\phi = -F_x\sin\phi+F_y\cos\phi = -\frac{4a^3ik}{3\pi}\frac{e^{ikr}}{r}[-\alpha\sin\phi+\beta\cos\phi+\frac{1}{2}\gamma\sin\theta]$$

$$\int_0^{2\pi}d\phi|\vec{E}|^2 = \frac{16a^6k^4}{9\pi^2}\frac{1}{r^2}[(\alpha^2+\beta^2)(1+\cos^2\theta)+\frac{1}{2}\gamma^2\sin^2\theta]\pi$$

$$\int_0^{2\pi}d\phi\int_0^{\pi/2}\sin\theta|\vec{E}|^2d\theta = \frac{16a^6k^4}{9\pi}\frac{1}{r^2}[\frac{4}{3}(\alpha^2+\beta^2)+\frac{1}{3}\gamma^2]$$

$$\int|r\vec{E}|^2\,d\Omega = \frac{16a^6k^4}{27\pi}[4(\alpha^2+\beta^2)+\gamma^2] = \frac{4a^6k^4}{27\pi}[4\vec{H}_o^2+\vec{E}_o^2].$$

Let S_{tot} denote the time-average (indicated by a bar) of the Poynting energy flux; then

(7.28)
$$S_{tot} = \frac{c}{4\pi}\overline{|Er|^2} = \frac{c}{27\pi^2}a^6k^4(4\overline{H_o^2}+\overline{E_o^2}).$$

This agrees with Bethe's expression (we have corrected a trivial misprint in his paper). In the plane-wave case one has

$$\overline{4H_o^2+E_o^2} = 2[4(\sin^2\phi_o\cos^2\theta_o+\cos^2\phi_o)+\cos^2\phi_o\sin^2\theta_o].$$

The right-hand side becomes

$$8\left[1 + \tfrac{1}{4}\sin^2\theta_o\right] \qquad \text{for } \phi_o = 0,$$

$$8\cos^2\theta_o \qquad \text{for } \phi_o = \tfrac{\pi}{2}.$$

Therefore

(7.29)

$$S_\perp = \frac{8c}{27\pi^2} a^6 k^4 \cos^2\theta_o$$

$$S = \frac{8c}{27\pi^2} a^6 k^4 \left[1 + \tfrac{1}{4}\sin^2\theta_o\right].$$

The energy flux of the incident wave per unit of cross-sectional area equals $c/8\pi$. Hence for the cross-sections,

(7.30)

$$\sigma_\perp = \frac{64}{27\pi} a^6 k^4 \cos^2\theta_o$$

$$\sigma_{\parallel} = \frac{64}{27\pi} a^6 k^4 \left[1 + \tfrac{1}{4}\sin^2\theta_o\right].$$

For the general case,

(7.31) $$\sigma = \frac{64a^6k^4}{27\pi}\left[1 + \left(\tfrac{1}{4}\cos^2\phi_o - \sin^2\phi_o\right)\sin^2\theta_o\right].$$

For unpolarized incident waves (natural "light"),

(7.32) $$\sigma' = \frac{64a^6k^4}{27\pi}\left[1 - \tfrac{3}{8}\sin^2\theta_o\right].$$

Equations (28) through (32) are due to Bethe.

A few words must be said about the field near the aperture (the quasi-static field). For this case we have to calculate F for small values of ka and ke. This is most effectively done in spheroidal coordinates u, v related to the cylindrical coordinates by

$$\rho^2 = a^2(1-u^2)(1+v^2), \; z = auv, \; 0 \le u \le 1, \; -\infty < v < \infty .$$

(See, for example, Bouwkamp [47].) Very tedious and complicated calculations are required, though. Therefore we shall give final results only. Note the branch $0 < \text{arc cot } v < \pi$.

Total electric field near aperture, for both $z < 0$ and $z > 0$:

$$E_x = \frac{2au}{\pi}\left[Q(v\,\mathrm{arccot}\,v-1) + \frac{2R-Q}{3(1+v^2)}\right]$$

$$+ \frac{2ux}{\pi a(u^2+v^2)(1+v^2)}\left[\gamma + \frac{4(Rx+Sy)-2(Qx+Py)}{1+v^2}\right];$$

$$E_y = \frac{2au}{\pi}\left[P(v\,\mathrm{arccot}\,v-1) + \frac{2S-P}{3(1+v^2)}\right]$$

$$+ \frac{2uy}{\pi a(u^2+v^2)(1+v^2)}\left[\gamma + \frac{4(Rx+Sy)-2(Qx+Py)}{1+v^2}\right];$$

$$E_z = \frac{2\gamma}{\pi}\left[\mathrm{arccot}\,v - \frac{v}{u^2+v^2}\right]$$

$$+ \frac{2}{\pi}(Rx+Sy)\left[\mathrm{arccot}\,v - \frac{v}{1+v^2} - \frac{4v}{3(u^2+v^2)(1+v^2)}\right]$$

$$+ \frac{4v(Qx+Py)}{3\pi(u^2+v^2)(1+v^2)}.$$

For the total magnetic field:

$$H_x = \frac{2\alpha}{\pi}\left[\mathrm{arccot}\,v - \frac{v}{1+v^2}\right] - \frac{4vx(\alpha x+\beta y)}{\pi a^2(u^2+v^2)(1+v^2)^2};$$

$$H_y = \frac{2\beta}{\pi}\left[\mathrm{arccot}\,v - \frac{v}{1+v^2}\right] - \frac{4vy(\alpha x+\beta y)}{\pi a^2(u^2+v^2)(1+v^2)^2};$$

$$H_z = -\frac{4u(\alpha x+\beta y)}{\pi a(u^2+v^2)(1+v^2)}.$$

Note that on the aperture $v = 0$, $u = \sqrt{1-\rho^2/a^2}$ and on the screen $u = 0$, $v = \pm\sqrt{\rho^2/a^2-1}$ (for $z = \pm 0$). By calculating the jump of $\vec{H}$ across $u = 0$ we are able to calculate the electric current density in the screen near the edge. This is left to the reader. For the case of normal incidence, see Bouwkamp [47].

VIII. On Copson's Theory of Diffraction

In a recent paper, Professor Copson [43] has commented upon the "note by the reviewer" that was added to my review [51] of Copson's earlier paper entitled "An integral-equation method of solving plane diffraction problems" [42]. As is obvious from Copson's answer as well as from two papers by Miles [52], my comments have led to some confusion as to the boundary conditions to be satisfied on the rim of the diffracting obstacle. In order that any possible misinterpretation of my criticism may be avoided, it seems worth while to discuss these questions in greater detail than is permitted in a short review. This also presents the opportunity to stress how carefully Copson's differential-integral equations must be handled in practical applications; we now know that his solutions for the small circular disk and aperture are in error*.

To present a clear picture of the problem, I shall quote the essence of Copson's theorem 4 as well as my comments published in Mathematical Reviews. The theorem in question is [42]**:

---

* It should be noted that Copson now agrees fully with the theory outlined in this section [54].

** For the discussion in this chapter we have preserved the original notation of Copson. His formulation in terms of scalar functions is essentially equivalent to that given in Section VII with the following exceptions:

(1) In Section VII, the time factor was taken as $e^{-i\omega t}$; Copson uses $e^{+i\omega t}$.

(2) In Section VII the exciting wave is incident from the left while here incidence from the right is assumed.

The functions v and u are the x and y rectangular components respectively of the vector potential F; w is the scalar potential ψ.

The functions e_y and $-e_x$ are proportional to the x and y rectangular components respectively of the magnetic current density K; h_z is proportional to the magnetic surface charge density η.

When these differences are taken into account the formulation in terms of vector functions given in Section VII is equivalent to that given here. The correspondence between the equations may be seen from the following table.

Equations in Section VII	Corresponding Equations in Section VIII
7.1	8.1 and 8.6
7.2 and 7.4	8.2 and 8.3
7.3	8.4
7.7	8.5

Let an electromagnetic field $\vec{E}^i$, $\vec{H}^i$ be incident in the half-space $z > 0$ on a perfectly conducting screen in the plane $z = 0$, the aperture in the screen being denoted by S. Then the total field in the half-space $z < 0$ is

$$(8.1)\qquad E_x = \frac{\partial u}{\partial z},\ E_y = \frac{\partial v}{\partial z},\ E_z = -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y},$$
$$H_x = ikv + \frac{\partial w}{\partial x},\ H_y = -iku + \frac{\partial w}{\partial y},\ H_z = \frac{\partial w}{\partial z},$$

where†

$$(8.2)\qquad (u,v,w) = \frac{1}{2\pi}\iint_S (e_x,e_y,h_z)\,\Phi\,dx'dy'$$

and

$$(8.3)\qquad ikw = -\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}.$$

The functions e_x, e_y, h_z which are connected by the relation

$$(8.4)\qquad ikh_z = -\frac{\partial e_y}{\partial x'} + \frac{\partial e_x}{\partial y'},$$

satisfy the differential-integral equations*

$$\frac{\partial}{\partial x}\iint_S h_z\,\Phi_o dx'dy' + ik\iint_S e_y\,\Phi_o dx'dy' = 2\pi H_x^i(x,y,o),$$
$$(8.5)\qquad \frac{\partial}{\partial y}\iint_S h_z\,\Phi_o dx'dy' - ik\iint_S e_x\,\Phi_o dx'dy' = 2\pi H_y^i(x,y,o),$$
$$\frac{\partial}{\partial x}\iint_S e_x\,\Phi_o dx'dy' + \frac{\partial}{\partial y}\iint_S e_y\,\Phi_o dx'dy' = -2\pi E_z^i(x,y,o),$$

when (x,y,o) is a point of S. If there were no aperture in the screen, the total field would be null in $z < 0$, but would be $\vec{E}^o$, $\vec{H}^o$, say, in $z > 0$. In the presence of the aperture, the total field in $z > 0$ is

$$(8.6)\qquad E_x = E_x^o - \frac{\partial u}{\partial z},\ E_y = E_y^o - \frac{\partial v}{\partial z},\ E_z = E_z^o + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y},$$
$$H_x = H_x^o - ikv - \frac{\partial w}{\partial x},\ H_y = H_y^o + iku - \frac{\partial w}{\partial y},\ H_z = H_z^o - \frac{\partial w}{\partial z}.$$

---

† $\Phi = e^{-ikR}/R$, $R^2 = (x-x')^2+(y-y')^2 + z^2$; $\Phi_o = \Phi$ when $z = 0$.

* Copson terms them "integral equations".

My comment upon this theorem was as follows[51]:

"There remains only one question not properly accounted for by this analysis, namely whether or not line charges along the rim of the screen are necessary*. For instance the proof of theorem 4 is incomplete. Suppose the equations (5) are solved rigorously, under the side-condition (4). It is not at once evident whether the wave functions u,v,w defined by (2) fulfill (3). Now it can be shown that

$$(8.7) \qquad -ikw - \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = \frac{1}{2\pi}\iint_S \left\{ \frac{\partial}{\partial x'} (e_y \Phi) - \frac{\partial}{\partial y'} (e_x \Phi) \right\} dx'dy'.$$

Maxwell's equations are satisfied if and omly if the right-hand integral vanishes. Because $e_x = E_x$, $e_y = E_y$ in the hole, the integral over S can be transformed into an integral along the rim of the hole

$$(8.8) \qquad \frac{1}{2\pi}\int \Phi E_s ds,$$

where E_s denotes the electric field tangential to the rim of the hole. Therefore the condition for Maxwell's equations to be satisfied is

$$(8.9) \qquad E_s = 0.$$

This is an extra condition with regard to the solution e_x, e_y, h_z. For a circular hole the condition is equivalent to $x'e_y - y'e_x = 0$ at the rim. This condition is satisfied in Copson's (and Bethe's) theory for the small circular hole, and Copson states that his approximate solution does not violate (3)."

The validity of Copson's assertions and his theorem has not been questioned; however, I claim that a necessary condition for Copson's theorem 4 to be self-consistent is

$$(8.10) \qquad e_s = 0 \quad \text{on the rim of S},$$

where e_s denotes the projection of the vector $\vec{e} = (e_x, e_y)$ at the rim upon the tangent to the rim.

The condition (10) does not imply that either e_x or e_y or both are finite on the rim. In fact, they generally become infinite on the rim of order $D^{-1/2}$, where D denotes the distance to the rim. Thus (10) does not exclude that e_n is

* The same question was raised by Bourgin in a review of Bethe's paper (Bethe [46])

infinitely large on the rim, where e_n means the projection of $\vec{e}$ at the rim upon the normal of the rim.

Incidentally, e_s and e_n should be considered as limiting values of the corresponding components of $\vec{e}$ when the field point tends to the rim from the interior of the aperture. The vector $\vec{e}$ is not defined outside S although, following Copson, we may take it to be zero there: in this case $(e_x,e_y)=(E_x,E_y)$ over the whole plane $z = 0$. Of course, the limiting values of e_x and e_y are zero when the field point is tending to the rim from outside S. In this sense, (10) states that e_s is continuous on the rim. On the other hand, e_n is in general discontinuous there.

To prove that (10) is a necessary condition it will be assumed that S is a finite part of the plane $z = 0$, with a simple closed boundary curve s. Moreover, it will be supposed that e_x, e_y, and h_z are absolutely integrable over S and have continuous first-order derivatives with respect to x' and y' in the interior of S.

Let S_o be any subdomain of S with boundary curve s_o such that s and s_o have no points in common. Then Stokes's theorem may be applied to the vector $\vec{e}$ in the domain S_o when the field point (x,y,z) is outside S. Thus

$$\iint_{S_o} \left\{ \frac{\partial}{\partial x'} (\Phi e_y) - \frac{\partial}{\partial y'} (\Phi e_x) \right\} dx'dy' = \int_{s_o} \Phi e_s ds,$$

where the integrand of the surface integral is the z component of curl $(\Phi\vec{e})$. Now using (4) and the fact that $(\partial/\partial x', \partial/\partial y')\Phi = -(\partial/\partial x, \partial/\partial y)\Phi$, we see that the integrand is equal to

$$-ikh_z \Phi - e_y \frac{\partial\Phi}{\partial x} + e_x \frac{\partial\Phi}{\partial y}.$$

It thus follows that

(8.11) $$-\frac{\partial}{\partial x}\iint_{S_o} e_y\Phi dx'dy' + \frac{\partial}{\partial y}\iint_{S_o} e_x\Phi\, dx'dy' - ik\iint_{S_o} h_z\Phi\, dx'dy' = \int_{s_o}\Phi e_s ds.$$

The change in the order of integration and differentiation is justified, since all integrals involved are uniformly convergent with respect to x,y,z and absolutely convergent with respect to e_x, e_y, h_z.

The same remark applies if in (11) the limiting case is taken where S_o becomes S, though it must be implicitly understood that certain regularity conditions as to the boundary curves s_o and s are satisfied. If so, then (11) leads to

$$(8.12) \qquad -\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} - ikw = \frac{1}{2\pi}\int_s \Phi \, e_s ds$$

and then (10) follows from (12) and (3).

As emphasized before, in the proof of (10) it has not been required that e_x and e_y are finite on s. Yet an explicit example may be useful.

If S is a circular aperture of radius a, the following functions are admissible (so far as integrability is concerned):

$$(8.13) \qquad \begin{aligned} e_x &= A\,\frac{2a^2-x'^2-2y'^2}{\sqrt{a^2-x'^2-y'^2}} + B\sqrt{a^2-x'^2-y'^2} + Cy', \\ e_y &= A\,\frac{x'y'}{\sqrt{a^2-x'^2-y'^2}} - Cx', \end{aligned}$$

where A, B, and C are constants. It is easy to verify that the corresponding function h_z is given by

$$(8.14) \qquad ikh_z = -(3A+B)\,\frac{y'}{\sqrt{a^2-x'^2-y'^2}} + 2C.$$

Clearly these functions are absolutely integrable when $x'^2+y'^2 \leq a^2$, and are continuously differentiable when $x'^2+y'^2 < a^2$. Unless A = 0, they are all singular on the rim. Notwithstanding this, e_s is finite on the rim, as is readily verified if polar coordinates are introduced. Letting $x' = r' \cos \phi'$, $y' = r' \sin \phi'$ we have

$$(8.15) \qquad \begin{aligned} e_r &= \left\{ A\,\frac{2a^2-r'^2}{\sqrt{a^2-r'^2}} + B\sqrt{a^2-r'^2} \right\} \cos \phi' \\ e_\phi &= -Cr' - (2A+B)\sqrt{a^2-r'^2}\,\sin \phi', \end{aligned}$$

and so in this case $e_s = -Ca$ on the rim. Consequently, the condition (10) would rule out the case $C \neq 0$, but it does not require A = 0.

In view of this, I shall now quote* from Copson's answer: [43] "Bouwkamp's argument depends on differentiation under the sign of integration and the use of Green's theorem. If e_x is continuous it vanishes on the rim, whereas, if it is discontinuous, it becomes infinite on the rim, as Bouwkamp himself has point out. And similarly for e_y. Thus either e_x and e_y vanish on the rim, and (3) is a consequence of (2) and (4); or else at least one of the functions e_x and e_y is infinite on the rim, and Bouwkamp's argument fails".

As a matter of fact, my proof of (10) shows that Copson's assertion is simply incorrect, as is also borne out by the explicit example (13). What matters is that e_s should not be singular on the rim; this constitutes an intermediate case which has been overlooked by Copson.

Consequently, I fully maintain my "criticism" of Copson's paper. That is to say, if we intend to solve Copson's differential-integral equations (4), (5), we should look for those solutions e_x, e_y which satisfy the auxiliary condition (10) because then and only then will the electromagnetic field specified by (1) and (6) solve Maxwell's equations. This information concerning the solution of the differential-integral equations (5), which was omitted in Copson's paper and which I therefore added in my review, does not weaken the value of Copson's theorem. On the contrary, it was meant to be and is in fact a further step towards the practical application of the theorem, especially in the construction of approximate solutions.

In addition, it is now evident that in a rigorous formulation of plane diffraction problems there is never need of additional line integrals along the rim. This settles an old question concerning the "rigorous" extension of the Huygens—Kirchhoff principle to electromagnetic diffraction problems. Whereas fictitious boundary values of the field vectors on an open surface in general require these line integrals in order that Maxwell's equations be satisfied, the correct boundary values automatically make these equations vanish identically.

At the time of writing my review of Copson's paper, his approximate solutions for the small circular disk and aperture seemed to be correct, since the condi-

---

* Actually, Copson discusses the complementary problem, his theorem 5. I prefer to keep the argument in accordance with what I criticized in Copson's work.

tion (10) happened to be fulfilled and Bethe's earlier results were confirmed. However, mainly because of Meixner's investigations, I have since come to the conclusion that Bethe's as well as Copson's first-order solutions are correct in the wave zone, but that they fail in and near the aperture or disk.

The fact that Copson's approximate solution for the small circular aperture is incorrect will be shown by the simplest possible example, namely the diffraction of a plane-polarized wave impinging in the normal direction. In this case Copson's equations (6.4) - (6.6) [42] reduce to

$$(8.16) \qquad e_x = \frac{4ik}{\pi}\sqrt{a^2-r'^2}\ ; \quad e_y = 0; \quad h_z = -\frac{4}{\pi}\frac{y'}{\sqrt{a^2-r'^2}}\ ,$$

if it is assumed that the incident wave is polarized parallel to the x axis and of unit amplitude ($E_x^i = 1$, $H_y^i = 1$). In the limiting case $ka \rightarrow 0$, Copson's equations (5) then reduce to

$$(8.17) \qquad \begin{aligned} &\frac{\partial}{\partial x}\iint_S h_z \frac{dx'dy'}{\rho} + ik\iint_S e_y \frac{dx'dy'}{\rho} = 0, \\ &\frac{\partial}{\partial y}\iint_S h_z \frac{dx'dy'}{\rho} - ik\iint_S e_x \frac{dx'dy'}{\rho} = -2\pi, \\ &\frac{\partial}{\partial x}\iint_S e_x \frac{dx'dy'}{\rho} + \frac{\partial}{\partial y}\iint_S e_y \frac{dx'dy'}{\rho} = 0, \end{aligned}$$

where

$$\rho^2 = (x-x')^2 + (y-y')^2, \ x^2+y^2 \leq a^2.$$

If the expressions (16) are substituted in (17), the right-hand members become, in the same order,

$$0; \ -2\pi + \pi k^2(2a^2-x^2-y^2) = -2\pi+O(k^2a^2); \ -2\pi ikx = O(ka);$$

thus, as in Copson's paper, (16) is an approximate solution of (17).

However, a second solution is provided by

$$(8.18) \qquad e_x = \frac{4ik}{3\pi}\frac{2a^2-x'^2-2y'^2}{\sqrt{a^2-r'^2}}\ ; \quad e_y = \frac{4ik}{3\pi}\frac{x'y'}{\sqrt{a^2-r'^2}}\ ; \ h_z = -\frac{4}{\pi}\frac{y'}{\sqrt{a^2-r'^2}}\ ,$$

which, if substituted in (17), will make the right-hand members equal to

$$-\frac{1}{2}\pi k^2 xy = O(k^2a^2);\ -2\pi + \frac{1}{12}\pi k^2(20a^2-3x^2-9y^2) = -2\pi + O(k^2a^2);\ 0.$$

This clearly demonstrates that equations (4) and (5) of Copson's theory must be handled very carefully in order to determine an approximate solution e_x, e_y, h_z. In addition it is to be noted that the approximate solution (18) is better than Copson's (16) since in the first case equation (17) is satisfied up to terms of relative order $(ka)^2$ and in the second case only up to terms of relative order ka.

The condition (10) is not decisive as to the question whether (16) or (18) is the physical solution, since both (16) and (18) are consistent with (10). Thus (10) is not a sufficient condition, at least not for the purpose of finding the long-wave approximation.

The solution of this difficulty is simple. A detailed investigation of the electromagnetic field calculated on the basis of Copson's (that is, Bethe's) approximate solution (16) has revealed (Bouwkamp[47]) that the corresponding electric field, which is throughout of order ka compared to the magnetic field, is discontinuous in the aperture. There is no sense in retaining E_x in the aperture and ignoring E_z, as Copson did in his approximation, since these quantities are of the same order of magnitude.

As was shown elsewhere (Bouwkamp[44]) the approximate solution (18) does not lead to a discontinuity in the electric field in the aperture.

IX. Diffraction by Narrow Slits

In this section we shall be concerned with approximate solutions for the diffraction of plane waves by narrow slits. The x-axis is assumed to be parallel to the edges, and we shall consider only problems that are independent of x. That is, the direction of incidence of the plane wave is in a plane normal to the edge. The integral-equation formulation of these two-dimensional problems follows at once from the corresponding formulation in three dimensions† if we integrate with respect to the x-coordinate. If in three dimensions we have $r^2 = (x-x')^2 + (y-y')^2 + (z-z')^2$ and in two dimensions $\rho^2 = (y-y')^2 + (z-z')^2$, then

$$(9.1) \qquad \int_{-\infty}^{\infty} \frac{e^{ikr}}{r} \, dx' = \pi i \, H_0^{(1)}(k\rho).$$

The two principal boundary-value problems for the slit are to be formulated as follows:

Problem I, $\phi = 0$ on the screen:

$$\phi = \phi_0(y,z) - \phi_0(y,-z) + \phi_1(y,-z) \qquad (z \leq 0)$$

$$\phi = \phi_1(y,z) \qquad (z \geq 0).$$

Problem II, $\partial\phi/\partial n = 0$ on the screen:

$$\phi = \phi_0(y,z) + \phi_0(y,-z) - \phi_2(y,-z) \qquad (z \leq 0)$$

$$\phi = \phi_2(y,z) \qquad (z \geq 0),$$

where the wave functions ϕ_1 and ϕ_2 (defined for $z \geq 0$ only) can be represented in the form of integrals extended over the aperture, the integrands containing the aperture values of ϕ_1 and $\partial\phi_2/\partial n$ respectively. The two-dimensional analogue of Rayleigh's formulas are

† See chapter III.

$$(9.2)\quad \begin{cases} \phi_1(y,z) = -\frac{i}{2}\frac{\partial}{\partial z}\int_{-a}^{a}\phi_1(y',0)H_o^{(1)}(k\sqrt{(y-y')^2+z^2})\,dy' \\ \phi_2(y,z) = -\frac{i}{2}\int_{-a}^{a}\frac{\partial}{\partial z'}\phi_2(y',0)H_o^{(1)}(k\sqrt{(y-y')^2+z^2})\,dy', \end{cases}$$

where 2a is the width of the slit.

As in the case of the circular aperture, we obtain differential-integral relations for ϕ_1 and ϕ_2 by requiring that $\partial\phi/\partial z = \partial\phi_o/\partial z$ in the aperture in the case of problem I, and $\phi = \phi_o$ in the aperture for problem II. Of course, ϕ_o is the incident field. These differential-integral relations are

$$(9.3)\quad -2i\left[\frac{\partial}{\partial z}\phi_o(y,z)\right]_{z=o} = (k^2+\frac{d^2}{dy^2})\int_{-a}^{a}\phi_1(y',0)H_o^{(1)}(k|y-y'|)dy'$$

$$(9.4)\quad 2i\,\phi_o(y,0) = \int_{-a}^{a}\left\{\frac{\partial\phi_2(y',z')}{\partial z'}\right\}_{z'=o}H_o^{(1)}(k|y-y'|)dy',$$

where $-a \leq y \leq a$.

Henceforth we consider only the case of normal incidence; that is, we choose $\phi_o = \exp(ikz)$. We also introduce the notations

$$y = a\sin\theta,\ y' = a\sin\theta',\ ka = \epsilon$$

$$F_1(\theta) = -\frac{1}{ika}\cos\theta\,\phi_1(a\sin\theta,0)$$

$$F_2(\theta) = a\cos\theta\left|\frac{\partial}{\partial z}\phi_2(a\sin\theta,z)\right|_{z=0}$$

$$(-\frac{\pi}{2}\leq\theta\leq\frac{\pi}{2},\ -\frac{\pi}{2}\leq\theta'\leq\frac{\pi}{2}).$$

Equations (3) and (4) then reduce to

$$(9.5)\quad 2i = (\varepsilon^2 + a^2\frac{d^2}{dy^2})\int_{-\pi/2}^{\pi/2}F_1(\theta')H_o^{(1)}(\varepsilon|\sin\theta-\sin\theta'|)\,d\theta'$$

$$(9.6)\quad 2i = \int_{-\pi/2}^{\pi/2}F_2(\theta')H_o^{(1)}(\varepsilon|\sin\theta-\sin\theta'|)d\theta'.$$

We now try to find approximate solutions of the equations (5) and (6) for the case where ε is small. To this end we recall that we know the eigenfunctions of the corresponding static potential problems. More explicitly, the eigenfunctions $f(\theta)$ of the homogeneous integral equation

$$(9.7)\quad \int_{-\pi/2}^{\pi/2} \log \left\{ 2|\sin\Theta-\sin\Theta'| \right\} f(\Theta')\, d\Theta' = \lambda f(\Theta)$$

are $\sin(2n+1)\Theta$ and $\cos 2n\Theta$, where $n = 0,1,2,\ldots$. The relevant eigenvalues λ are equal to $-\pi/m$, where $m = 2n,\ 2n+1$ $(m \neq o)$ respectively; they are zero if $m = o$. Thus

$$(9.8)\quad -\frac{1}{2}\log|2(\sin\Theta-\sin\Theta')| = \sum_1^{\infty}\frac{1}{2n}\cos 2n\Theta\cos 2n\Theta' + \sum_1^{\infty}\frac{1}{2n+1}\sin(2n+1)\Theta\sin(2n+1)\Theta'.$$

Both Θ and Θ' are between $-\pi/2$ and $\pi/2$. Equation (8) follows from the known [55] expansion

$$-\frac{1}{2}\log|2(\cos\phi-\cos\phi')| = \sum_1^{\infty}\frac{1}{m}\cos m\phi\cos m\phi'$$

when we make a suitable change of variables.

We need an appropriate expansion of the Hankel-function kernel in our integral equations. This is a series expansion in powers of ε in which the coefficients depend on Θ, Θ' and on $\log\varepsilon$. This expansion is

$$H_o^{(1)}(\varepsilon|\sin\Theta-\sin\Theta'|) = \frac{2i}{\pi}\left[\Psi_o+\Psi_1\varepsilon^2+\Psi_2\varepsilon^4+\cdots+\Psi_n\varepsilon^{2n} + O(\varepsilon^{2n+2}\log\varepsilon)\right],$$

where

$$\Psi_o = p + \log|2(\sin\Theta-\sin\Theta')|$$

$(n>o)$

$$\Psi_n = \frac{(-1)^n}{2^{2n}(n!)^2}(\sin\Theta-\sin\Theta')^{2n}\left[p - 1 - \frac{1}{2}\cdots-\frac{1}{n} + \log|2(\sin\Theta-\sin\Theta')|\right]$$

and p is an abbreviation for

$$p = \gamma + \log(\tfrac{1}{4}\varepsilon) - \frac{\pi}{2}i \qquad (\gamma = 0.577\cdots = \text{Euler's constant}).$$

Let us first consider problem II. Assume that

$$F_2(\Theta) = f_o + f_1\varepsilon^2 + f_2\varepsilon^4 + \cdots .$$

Then

$$F_2(\Theta')H_o^{(1)}(\varepsilon|\sin\Theta-\sin\Theta'|) = \frac{2i}{\pi}\Big[f_oY_o + \varepsilon^2(f_oY_1 + f_1Y_o)$$

$$+ \varepsilon^4(f_oY_2 + f_1Y_1 + f_2Y_o) + \cdots\Big] .$$

Substituting this in equation (6) gives us an infinite set of integral equations of the potential type, viz.,

$$(9.9)\quad \begin{cases} \int f_oY_o d\Theta' = \pi \\ \int f_1Y_p d\Theta' = -\int f_oY_1 d\Theta' \\ \int f_2Y_o d\Theta' = -\int f_oY_o d\Theta' - \int f_1Y_1 d\Theta', \text{ etc.} \end{cases}$$

Since we know from (7) or (8) that

$$\int Y_o d\Theta' = \pi\, p,$$

we immediately get the zero-order approximation

$$(9.10)\qquad f_o = \frac{1}{p} .$$

If we insert this result in the right-hand side of the second equation of (9) we get an integral equation for f_1, namely

$$\int f_1Y_o d\Theta' = \frac{\pi}{4}\left\{1 - \frac{1}{4}\left(2+\frac{1}{p}\right)\cos 2\Theta\right\}.$$

To solve for f_1 we assume f_1 has the form $f_1 = c_1+c_2\cos2\Theta'$, where c_1 and c_2 are independent of Θ'. Then

$$\int f_1Y_o d\Theta' = c_1\int Y_o d\Theta' + c_2\int Y_2\cos2\Theta' d\Theta' = c_1\pi p - \frac{\pi}{2}c_2\cos2\Theta.$$

Therefore we find by comparing coefficients that

$$(9.11)\qquad f_1 = \frac{1}{8p}\left\{2 + (2p+1)\cos 2\Theta\right\}.$$

On inserting f_o and f_1 from (10) and (11) in the third equation of (9) and evaluating the necessary integrals, we find

$$\int f_2 \Psi_o d\Theta' = -\frac{1}{p}\int \Psi_2 d\Theta' - \frac{1}{4p}\int \Psi_1 d\Theta' - (\frac{1}{4} + \frac{1}{8p})\int \Psi_1 \cos 2\Theta' d\Theta' =$$

$$-\frac{\pi}{64}\left\{ p - \frac{3}{4} - \frac{1}{2p} + (\frac{1}{p} - \frac{2}{3})\cos 2\Theta + \frac{1}{8}(\frac{1}{3} + \frac{1}{4p})\cos 4\Theta \right\}.$$

This in turn leads, in the same way as before for the function f_1, to an explicit solution for f_2, viz.,

$$(9.12)\quad f_2 = -\frac{1}{64}\,(1 - \frac{3}{4p} - \frac{1}{2p^2}) + \frac{1}{32}\,(\frac{1}{p} - \frac{2}{3})\cos 2\Theta + \frac{1}{128}\,(\frac{1}{3} + \frac{1}{4p})\cos 4\Theta.$$

Equivalent representations of the results (10), (11), (12) are

$$(9.13)\quad \begin{cases} f_o = \frac{1}{p} \\ f_1 = \frac{1}{8p} - \frac{1}{4} + (\frac{1}{2} + \frac{1}{4p})\cos^2\Theta \\ f_2 = \frac{1}{128}(1 - \frac{9}{4p} + \frac{1}{p^2}) - \frac{1}{16}(1 - \frac{3}{4p}\cos^2\Theta) + \frac{1}{48}(1 + \frac{3}{4p})\cos^4\Theta. \end{cases}$$

REMARK. In deriving this result we have used various transformations of an essentially elementary character. For convenience we list them here:

$$[\sin\Theta - \sin\Theta']^2 = 1 - \frac{1}{2}\cos 2\Theta - \frac{1}{2}\cos 2\Theta' - 2\sin\Theta\sin\Theta'$$

$$[\sin\Theta - \sin\Theta']^4 = \frac{9}{4} - 2\cos 2\Theta + \frac{1}{8}\cos 4\Theta + \sin\Theta'(\sin 3\Theta - 6\sin\Theta)$$

$$+ \sin\Theta\sin 3\Theta' + \cos 2\Theta'(\frac{3}{2}\cos 2\Theta - 2) + \frac{1}{8}\cos 4\Theta'$$

$$\int \Psi_1 d\Theta' = \frac{\pi}{16}\left\{-4p + (2p+1)\cos 2\Theta\right\}$$

$$\int \Psi_o \cos 2n\Theta' d\Theta' = -\frac{\pi}{2n}\cos 2n\Theta \qquad (n > o)$$

$$\int \Psi_2 d\Theta' = \frac{\pi}{64}\left\{\frac{9}{4}p - \frac{3}{4} + (\frac{1}{3} - 2p)\cos 2\Theta + \frac{1}{8}(p + \frac{7}{12})\cos 4\Theta\right\}$$

$$\int \Psi_1 \cos 2\Theta' d\Theta' = \frac{\pi}{8}\left\{\frac{1}{2}p + \frac{1}{4} - \frac{1}{3}\cos 2\Theta - \frac{1}{24}\cos 4\Theta\right\}.$$

Finally, our result (13) can be brought into the following form: If we assume the field expansion in the aperture to be

(9.14) $\frac{\partial}{\partial z'} \phi_2(y,o) = \frac{1}{a} \sum_{o}^{\infty} C_n (1 - y^2/a^2)^{n-1/2}$

then

$$(9.15) \begin{cases} C_o = \frac{1}{p} - \frac{1}{4}(1 - \frac{1}{2p})(ka)^2 + \frac{1}{128}(1 - \frac{9}{4p} + \frac{1}{p^2})(ka)^4 + \cdots \\ C_1 = \frac{1}{2}(1 + \frac{1}{2p})(ka)^2 - \frac{1}{16}(1 - \frac{3}{4p})(ka)^4 + \cdots \\ C_2 = \frac{1}{48}(1 + \frac{3}{4p})(ka)^4 + \cdots \\ C_3 = \cdots \end{cases}$$

correct up to and including terms of order $(ka)^4/(\log ka)^2$. The error, represented by ..., is $O[(ka)^5]$.

It is to be noted that our result confirms that of Sommerfeld [56] who only gave terms up to but excluding $(ka)^4$. A small error of Sommerfeld must be noted, however: his A' should be replaced everywhere by $2A'/\pi i$ (see also Hönl and Broschwitz [57]).

We now turn to calculating the wave field at large distances behind the slit. Let ρ, Θ be polar coordinates with respect to the positive z-axis. Then

$$\phi_2 \sim \sqrt{\frac{2}{\pi k\rho}}\; e^{i(k\rho-\pi/4)} \frac{1}{2i} \int_{-a}^{a} \frac{\partial}{\partial z'} \phi_2(y',0) e^{-iky'\sin\Theta} dy'.$$

Inserting (14) and (15) we get

$$\phi_2 \sim \frac{1}{p}\sqrt{\frac{\pi}{2k\rho}}\; e^{i(k\rho-3\pi/4)} \Big[1 + \frac{1}{4}(ka)^2\cos^2\Theta + \frac{1}{64}(ka)^4 \Big\{ -p\cos 2\Theta + \sin^4\Theta - 3\sin^2\Theta + \frac{3}{4} + \frac{1}{2p} \Big\} + \cdots \Big];$$

hence

$$|\phi_2|^2 \sim \frac{1}{|p|^2} \frac{\pi}{2k\rho} \Big[1 + \frac{1}{2}\varepsilon^2\cos^2\Theta + \frac{1}{32}\varepsilon^4 \Big\{ -R\cos 2\Theta + 2\cos^4\Theta + \sin^4\Theta - 3\sin^2\Theta + \frac{3}{4} + \frac{R}{2|p|^2} \Big\} \Big],$$

where $\varepsilon = ka$ and R = real part of p.

Averaging over all angles, and using

$$\overline{\cos^2\Theta} = \overline{\sin^2\Theta} = \frac{1}{2}$$

$$\overline{\cos^4\Theta} = \overline{\sin^4\Theta} = \frac{3}{8},$$

we find that the transmission coefficient is

$$\tau_2 \sim \frac{\pi^2}{4\varepsilon|p|^2}\left[1+\frac{1}{4}\varepsilon^2+\frac{3}{256}\varepsilon^4\left\{1+\frac{4}{3}\,\mathrm{Re}\left(\frac{1}{p}\right)\right\}\right]$$

or, explicitly,

$$\tau_2 = \frac{\pi^2/4ka}{[\gamma+\log(\frac{1}{4}ka)]^2+\frac{\pi^2}{4}}\left[1+\frac{1}{4}(ka)^2+\frac{3}{256}(ka)^4\left\{1+\frac{4}{3}\,\frac{\gamma+\log(\frac{1}{4}ka)}{[\gamma+\log(\frac{1}{4}ka)]^2+\frac{\pi^2}{4}}\right\}+\cdots\right] . \tag{9.16}$$

Note that this τ_2 is also the transmission coefficient for the electromagnetic case if the incident wave is polarized perpendicular to the edge and the screen is a perfect conductor.

Problem I can be solved in a similar way. We assume $F_1(\theta')$ to be of the form

$$F_1(\theta') = c_1\cos^2\theta' + c_2\cos^4\theta' + c_3\cos^6\theta' + \cdots ,$$

where

$$c_1 = g_1 + g_2\,\varepsilon^2 + g_3\,\varepsilon^4 + \cdots$$

$$c_2 = g_4\,\varepsilon^2 + g_5\,\varepsilon^4 + \cdots$$

$$c_3 = g_6\,\varepsilon^4 + \cdots$$

$$c_4 = \cdots$$

and where g_n depend on ε (not on θ'), but only as some power of $\log\varepsilon$. Then

$$\frac{\pi}{2i}F_1(\theta')H_0^{(1)}(\varepsilon|\sin\theta-\sin\theta'|) = g_1\Psi_0\cos^2\theta' + \varepsilon^2[g_2\Psi_0\cos^2\theta'+g_4\Psi_0\cos^4\theta'+g_1\Psi_1\cos^2\theta'] + \varepsilon^4[g_3\Psi_0\cos^2\theta' + g_5\Psi_0\cos^4\theta' + g_6\Psi_0\cos^6\theta' + g_2\Psi_1\cos^2\theta' + g_4\Psi_1\cos^4\theta' + g_1\Psi_2\cos^2\theta'] + \cdots.$$

We thus require integrals of the form

$$\int \Psi_n \cos^{2m}\theta' d\theta'.$$

The following explicit results have been found:

$$\int \Psi_o \cos^2\theta' d\theta' = \frac{\pi}{4} (2p - \cos 2\theta)$$

$$\int \Psi_o \cos^4\theta' d\theta' = \frac{\pi}{8} (3p - 2\cos 2\theta - \frac{1}{4}\cos 4\theta)$$

$$\int \Psi_o \cos^6\theta' d\theta' = \frac{\pi}{32}(10p - \frac{15}{2}\cos 2\theta - \frac{3}{2}\cos 4\theta - \frac{1}{6}\cos 6\theta)$$

$$\int \Psi_1 \cos^2\theta' d\theta' = \frac{\pi}{32} \left\{ \frac{1}{2} - 3p + (2p + \frac{1}{3})\cos 2\theta - \frac{1}{12} \cos 4\theta \right\}$$

$$\int \Psi_1 \cos^4\theta' d\theta' = \frac{\pi}{32} \left\{ \frac{1}{2} - 2p + (\frac{3}{2}p + \frac{1}{16})\cos 2\theta - \frac{1}{10} \cos 4\theta - \frac{1}{240} \cos 6\theta \right\}$$

$$\int \Psi_2 \cos^2\theta' d\theta' = \frac{\pi}{512} \left\{ 5p - \frac{7}{3} + (\frac{41}{24} - 5p)\cos 2\theta + (\frac{1}{2p} + \frac{11}{120})\cos 4\theta - \frac{1}{120}\cos 6\theta \right\}.$$

In the process of evaluation all functions $\Psi_n \cos^{2m}\theta'$ were first transformed into a form containing the cosines of the even multiples of θ and θ' and the sines of the odd multiples of θ and θ' in a symmetrical way, and then (7) and (8) were applied.

To cope with the operator $\epsilon^2 + a^2 \frac{d^2}{dy^2}$, we note that

$$a^2 \frac{d^2}{dy^2} \cos 2\theta = -4$$

$$a^2 \frac{d^2}{dy^2} \cos 4\theta = 16(2 - 3\cos 2\theta)$$

$$a^2 \frac{d^2}{dy^2} \cos 6\theta = -12(9 - 16\cos 2\theta + 10\cos 4\theta);$$

thus we have

$$a^2 \frac{d^2}{dy^2} \int \Psi_o \cos^2\theta' d\theta' = \pi$$

$$a^2 \frac{d^2}{dy^2} \int \Psi_o \cos^4\theta' d\theta' = \frac{3\pi}{2} \cos 2\theta$$

$$a^2 \frac{d^2}{dy^2} \int \Psi_o \cos^6\theta' d\theta' = \frac{5\pi}{8} (2\cos 2\theta + \cos 4\theta)$$

$$a^2 \frac{d^2}{dy^2} \int \Psi_1 \cos^2\theta' d\theta' = \frac{\pi}{8} (-2p - 1 + \cos 2\theta)$$

$$a^2 \frac{d^2}{dy^2} \int \Psi_1 \cos^4\theta' d\theta' = \frac{\pi}{32}(-6p-3+4\cos 2\theta + \frac{1}{2}\cos 4\theta)$$

$$a^2 \frac{d^2}{dy^2} \int \Psi_2 \cos^2\theta' d\theta' = \frac{\pi}{128}\left\{9p - \frac{3}{4} - (6p + \frac{3}{2})\cos 2\theta + \frac{1}{4}\cos 4\theta\right\}.$$

Consequently

$$2i(\varepsilon^2 + a^2 \frac{d^2}{dy^2}) \int F_1(\theta') H_o^{(1)}(\varepsilon|\sin\theta - \sin\theta'|)d\theta'$$

$$= g_1 \left[1 + \frac{\varepsilon^2}{4}(2p - \cos 2\theta)\right]$$

$$+ g_2 \left[\varepsilon^2 + \frac{1}{4}\varepsilon^4(2p - \cos 2\theta)\right]$$

$$+ g_4 \left[\frac{3}{2}\varepsilon^2 \cos 2\theta + \frac{1}{8}\varepsilon^4(3p - 2\cos 2\theta - \frac{1}{4}\cos 4\theta)\right]$$

$$+ g_1 \left[\frac{1}{8}\varepsilon^2(-2p-1+\cos 2\theta) + \frac{1}{32}\varepsilon^4\left\{\frac{1}{2} - 3p + (2p + \frac{1}{3})\cos 2\theta - \frac{1}{12}\cos 4\theta\right\}\right]$$

$$+ g_3 \left[\varepsilon^4\right]$$

$$+ g_5 \left[\frac{3}{2}\varepsilon^4 \cos 2\theta\right]$$

$$+ g_6 \left[\frac{5}{8}\varepsilon^4(2\cos\theta + \cos 4\theta)\right]$$

$$+ g_2 \left[\frac{1}{8}\varepsilon^4(-2p-1+\cos 2\theta)\right]$$

$$+ g_4 \left[\frac{1}{32}\varepsilon^4(-6p-3+4\cos 2\theta + \frac{1}{2}\cos 4\theta)\right]$$

$$+ g_1 \left[\frac{1}{128}\varepsilon^4\left\{9p - \frac{3}{4} - (6p + \frac{3}{2})\cos 2\theta + \frac{1}{4}\cos 4\theta\right\}\right]$$

+ terms of higher order.

According to the integral-differential equation (5) this should be equal to 2i plus terms of order ε^6. Therefore we find

$$g_1 = 1 \qquad g_3 = \frac{1}{16}(p^2 - \frac{7}{8}p + \frac{7}{32})$$

$$g_2 = \frac{1}{8} - \frac{p}{4} \qquad g_6 = \frac{1}{320}$$

$$g_4 = \frac{1}{12} \qquad g_5 = \frac{1}{64}(1-2p).$$

We finally get the following result: let

$$(9.17) \qquad \phi_1(y,0) = -ika \sum_{n=1}^{\infty} C_n (1-y^2/a^2)^{n-1/2};$$

then

$$(9.18) \qquad \begin{cases} C_1 = 1 - \frac{1}{4}(p-\frac{1}{2})(ka)^2 + \frac{1}{16}(p^2 - \frac{7}{8}p + \frac{7}{32})(ka)^4 + \cdots \\ C_2 = \frac{1}{12}(ka)^2 - \frac{1}{32}(p-\frac{1}{2})(ka)^4 + \cdots \\ C_3 = \frac{1}{320}(ka)^4 + \cdots \\ C_4 = \cdots \end{cases}$$

REMARK. This result is at variance with that of Sommerfeld[58] and that of Groschwitz and Hönl[59]. Sommerfeld gave (in our notation)

$$C_1 = 1 - \frac{1}{4}(p - \frac{3}{2})(ka)^2 + \cdots$$

$$C_2 = -\frac{1}{12}(ka)^2 + \cdots$$

Our result (17), (18) in Groschwitz-Hönl notation would be

$$C_o = \frac{\varepsilon}{2i}\left[1 + \varepsilon^2 \left\{ -\frac{1}{4}\log\gamma\varepsilon + \frac{1}{2}\log 2 + \frac{3}{16} + \frac{\pi i}{8} \right\}\right]$$

$$C_1 = -\frac{1}{32} i\varepsilon^3.$$

The transmission coefficient for problem I turns out to be

$$(9.19) \qquad \tau_1 = \frac{\pi^2}{32}(ka)^3 \left[1 + \frac{5}{16}(ka)^2 \left\{ 1 - \frac{8}{5}(\gamma + \log\frac{1}{4}ka) \right\} + \cdots \right].$$

Note that τ_1 applies to electromagnetic waves if the incident electric field is parallel to the edge.

Section X

Diffraction by an Aperture in a Planar Screen

by

N. Marcuvitz

X. Diffraction by an Aperture in a Planar Screen by Nathan Marcuvitz

The circular aperture problem solved by Bouwkamp can be treated by a somewhat different, but equivalent, method that emphasizes the vector aspect of the diffraction problem. A dyadic Green's function formalism is employed to obtain a transverse* vector integro-differential equation for the transverse electric field in the aperture. The knowledge of the aperture field permits a direct calculation of the electromagnetic fields at any point. The desired aperture field is obtained by solving first an inhomogeneous transverse vector partial differential equation and then an inhomogeneous vector integral equation of the first kind.

The formulation of the problem is independent both of the nature of the sources and the shape of the aperture. The solution of the inhomogeneous steady state field equations**

$$(10.1) \qquad \nabla \times \underset{\sim}{E} = ik\underset{\sim}{H} - \underset{\sim}{M}, \qquad \nabla \times \underset{\sim}{H} = -ik\underset{\sim}{E}$$

is to be obtained subject to boundary conditions:

(10.2)

a) $\underset{\sim}{n} \times \underset{\sim}{E} = 0$ on the screen

b) $\underset{\sim}{E}_{scat.}, \ \underset{\sim}{H}_{scat.} \to 0$ as $\underset{\sim}{r} \to \infty$ $(\mathrm{Im}\ k = \omega\sqrt{\mu\varepsilon} > 0)$

c) $\underset{\sim}{E}_{tan} \sim \sqrt{s}$, $E_{normal} \sim \dfrac{1}{\sqrt{s}}$ in the aperture,

where s is the normal distance to the rim of the aperture. As Bouwkamp first emphasized, conditions a) and b) alone are not sufficient to insure a unique field solution. Condition c) defines the singularity of the electric field at the rim of the aperture and, together with a) and b), does insure such a solution. Condition b) refers to that portion of the far fields due to current sources $\underset{\sim}{M}(\underset{\sim}{r})$ (real or induced) at a finite distance $\underset{\sim}{r}$. As Bouwkamp discussed above, it is convenient to divide the field region into two parts, $z < 0$ and $z > 0$, to the left and right of the screen, with sources prescribed in the region $z < 0$. The total field in the region $z < 0$ is equivalent to that produced in a half-space, with a perfect conductor at $z = 0$, both by the prescribed sources and by the "induced" magnetic current source $\underset{\sim}{n} \times \underset{\sim}{E}$

---

* i.e., parallel to the plane of the screen.

** M.K.S. system but normalized so that intrinsic impedance $\sqrt{\mu/\varepsilon}$ of vacuum is unity.

in the aperture ($\underset{\sim}{n}$ = unit normal vector in positive z - direction). Similarly the fields in z > 0 may be regarded as produced solely by an induced magnetic current density $-\underset{\sim}{n} \times \underset{\sim}{E}$ in the aperture.

It is convenient to introduce a "half-space" dyadic Green's function $Y(\underset{\sim}{r},\underset{\sim}{r}')$ defined by (cf. [40])

$$(10.3) \qquad \nabla \times (\nabla \times Y) - k^2 Y = -ik\varepsilon\delta(\underset{\sim}{r}-\underset{\sim}{r}')$$

subject to boundary conditions

a) $Y \to 0$ as $|\underset{\sim}{r}-\underset{\sim}{r}'| \to \infty$ (Im k > 0)

b) $\underset{\sim}{n} \times (\nabla \times Y) = 0$ at z = 0,

where ε is the unit dyadic defined by $\varepsilon \cdot \underset{\sim}{A} = \underset{\sim}{A}$, and $\delta(\underset{\sim}{r}-\underset{\sim}{r}')$ is the three-dimensional delta function. Physically $-Y'\underset{\sim}{r},\underset{\sim}{r}') \cdot e$ is the magnetic field produced at $\underset{\sim}{r}$ by a delta-function magnetic current flowing in the direction $\underset{\sim}{e}$ at $\underset{\sim}{r}'$ in the half-space. The half-space Green's function can be defined in terms of a "free-space" dyadic Green's function $Y_f(\underset{\sim}{r},\underset{\sim}{r}')$ which obeys Eq. (3) with the omission of condition b). The free-space Green's function is given by (cf. |40|)

$$(10.4) \qquad Y_f(\underset{\sim}{r},\underset{\sim}{r}') = -ik(\varepsilon + \frac{\nabla\nabla}{k^2})\, g(\underset{\sim}{r},\underset{\sim}{r}')$$

where $g(\underset{\sim}{r},\underset{\sim}{r}')$ is the scalar Green's function defined by

$$(10.5) \qquad \begin{aligned} &(\nabla^2 + k^2)g = -\delta(\underset{\sim}{r}-\underset{\sim}{r}') \\ &g \to 0 \text{ as } |\underset{\sim}{r}-\underset{\sim}{r}'| \to \infty \text{ (Im } k > 0). \end{aligned}$$

Although

$$(10.6) \qquad g = \frac{e^{ik|\underset{\sim}{r}-\underset{\sim}{r}'|}}{4\pi|\underset{\sim}{r}-\underset{\sim}{r}'|}$$

is a simple closed form, a more convenient representation for subsequent applications will be considered below. The half-space dyadic Green's function $Y(\underset{\sim}{r},\underset{\sim}{r}')$ is obtained by additive or subtractive superposition of two free-space dyadic Green's functions $Y_f(\underset{\sim}{r},\underset{\sim}{r}')$, one corresponding to the source at $\underset{\sim}{r}'$ and the other to its image at $\underset{\sim}{r}' - 2\underset{\sim}{n}\,\underset{\sim}{n} \cdot \underset{\sim}{r}'$. In particular, for a transverse magnetic source on the z = 0 plane,

$$(10.7) \qquad Y(\underset{\sim}{r},\underset{\sim}{r}') = -2ik(\varepsilon + \frac{\nabla\nabla}{k^2})\, g(\underset{\sim}{r},\underset{\sim}{r}'),$$

where $\underset{\sim}{r}' = (x',y',0)$.

The magnetic field due to both the prescribed and induced sources can be expressed in terms of $Y(\underset{\sim}{r},\underset{\sim}{r}')$, the field of a "point" source. Since the prescribed sources are assumed to lie in z < 0, the magnetic field in z < 0 can be represented by means of superposition (or equivalently by use of the vector Green's theorem) as

$$(10.8) \qquad \underset{\sim}{H}(\underset{\sim}{r}) = \underset{\sim}{H}_o(\underset{\sim}{r}) - \iint_{ap} Y(\underset{\sim}{r},\underset{\sim}{r}')\cdot \underset{\sim}{n} \times \underset{\sim}{E}(\underset{\sim}{r}')dS', \quad z < 0$$

where the first term $\underset{\sim}{H}_o$ is the magnetic field produced by the prescribed sources in the absence of the aperture, and the second term is the field produced by the induced sources $\underset{\sim}{n} \times \underset{\sim}{E}$ in the aperture. Similarly, the magnetic field in z > 0 produced by the induced sources $-\underset{\sim}{n} \times \underset{\sim}{E}$ in the aperture is

$$(10.9) \qquad \underset{\sim}{H}(\underset{\sim}{r}) = + \iint_{ap} Y(\underset{\sim}{r},\underset{\sim}{r}') \cdot \underset{\sim}{n} \times \underset{\sim}{E}(\underset{\sim}{r}')dS', \quad z > 0.$$

It is evident from Eq. (2) that the field representations (8) and (9) satisfy the field equations (1) in $z \gtrless 0$ and the boundary conditions (2a) and (2b). Moreover, since $\underset{\sim}{n} \times (\nabla \times Y)$ has a jump discontinuity of value $-ik\varepsilon_t\delta(\underset{\sim}{\rho}-\underset{\sim}{\rho}')$ † at z = z' = 0, it follows that Eqs. (8) and (9) yield values of $\underset{\sim}{n} \times \underset{\sim}{E}(\underset{\sim}{r})$ that are continuous as $z \to \pm 0$ and equal to the value of $\underset{\sim}{n} \times \underset{\sim}{E}$ in the aperture. The requirement of continuity of the $\underset{\sim}{H}(\underset{\sim}{r}) \times \underset{\sim}{n}$ given by (8) and (9) in the aperture region (z = ± 0) imposes the condition

$$(10.10) \quad -\underset{\sim}{H}_o(\underset{\sim}{r}) \times \underset{\sim}{n} = 2\iint_{ap} \underset{\sim}{n} \times Y(\underset{\sim}{r},\underset{\sim}{r}') \times \underset{\sim}{n} \cdot \underset{\sim}{E}_t(\underset{\sim}{r}')dS', \; r \to (x,y,\pm 0) \text{ "in the aperture"}$$

on the transverse electric field $\underset{\sim}{E}_t$ in the aperture. The continuity of $\underset{\sim}{n} \times Y \times \underset{\sim}{n}$ at z = 0 should be noted; however z is not permitted to equal z' = 0 since the integral in (10) becomes divergent at $\underset{\sim}{r} = \underset{\sim}{r}'$. Because of this latter fact, as Bouwkamp has emphasized, Eq. (10) is <u>not</u> a true integral equation but may be called a <u>pseudo-integral equation</u> for $\underset{\sim}{E}_t(\underset{\sim}{r})$.

In view of the representation (7), Eq. (10) may be rewritten as

$$(10.11) \quad -\underset{\sim}{H}_o(\underset{\sim}{r}) \times \underset{\sim}{n} = 4ik\left(\varepsilon_t - \frac{\underset{\sim}{n}\times\nabla_t\nabla_t\times\underset{\sim}{n}}{k^2}\right) \cdot \iint_{ap} g(\underset{\sim}{r},\underset{\sim}{r}')\underset{\sim}{E}_t(\underset{\sim}{r}')dS', \; \underset{\sim}{r} \text{ in the aperture,}$$

---

† In a rectangular x , y coordinate system, $\delta(\underset{\sim}{\rho}-\underset{\sim}{\rho}') = \delta(x-x')\delta(y-y')$;
$\varepsilon_t = \underset{\sim}{x}_o\underset{\sim}{x}_o + \underset{\sim}{y}_o\underset{\sim}{y}_o$ = transverse unit dyadic.

where the interchange of differentiation and integration is permissible since the integral exists before and after interchange. Since $g(\underset{\sim}{r},\underset{\sim}{r}')$ is integrable even in the limit $z = z' = 0$, Eq. (11) is a true integro-differential equation. To emphasize this fact we may rewrite (11) in the form*:

$$\text{(10.12a)} \qquad \text{curl}_t^2\, \underset{\sim}{F} - k^2\underset{\sim}{F} = -ik\underset{\sim}{H}_o(\underset{\sim}{\rho}) \times \underset{\sim}{n}$$

$$\text{(10.12b)} \qquad \underset{\sim}{F}(\underset{\sim}{\rho}) = 4\iint_{ap} g(\underset{\sim}{\rho},\underset{\sim}{\rho}')E_t(\underset{\sim}{\rho}')dS'$$

where we have defined

$$\text{(10.13)} \qquad \text{curl}_t^2\, \underset{\sim}{F} = \underset{\sim}{n} \times \nabla_t\nabla_t \times \underset{\sim}{n} \cdot \underset{\sim}{F} = \nabla_t\nabla_t \cdot \underset{\sim}{F} - \nabla_t^2\, \underset{\sim}{F},$$

and where $\underset{\sim}{\rho} = (x,y,0)$ is the coordinate vector in the aperture. It is necessary first to obtain the general solution of the vector differential equation (12a) for $\underset{\sim}{F}$, and then to solve the integral equation (12b) for $\underset{\sim}{E}_t$. The arbitrary constants in the resulting solution for $\underset{\sim}{E}_t$ are determined by imposing the remaining boundary condition (2c).

A general solution to Eq. (12a) can be obtained in a variety of ways, depending on the nature of the excitation $\underset{\sim}{H}_o$ and the shape of the aperture. It thus appears desirable to specialize at this point.

Diffraction of a Plane Wave by a Circular Aperture

Since an arbitrary source distribution can be resolved into plane wave constituents, it is basic to consider a plane wave incident on the aperture. There are two independent types of vector plane waves: the E- and the H- waves**, distinguished by their polarization. If a rectangular coordinate system with origin at the center of the aperture is oriented so that the plane wave is incident in the xz-plane, the transverse field distribution of an E-mode wave is given by [55]

---

* Bouwkamp writes Eq. (12a) in a somewhat different form. From (12a) one notes that

$$\text{(i)} \qquad -k^2\nabla_t \cdot \underset{\sim}{F} = -ik\nabla_t \cdot \underset{\sim}{H}_o \times \underset{\sim}{n} = -k^2E_{on}$$

whence on expansion of curl_t^2 of (12a)

$$\text{(ii)} \qquad (\nabla_t^2 + k^2)\, \underset{\sim}{F} = ik(\varepsilon_t + \frac{\nabla_t\nabla_t}{k^2}) \cdot \underset{\sim}{H}_o \times \underset{\sim}{n} = \frac{\partial}{\partial z}\, \underset{\sim}{E}_{ot} \cdot$$

Eqs. (i) and (ii) which <u>together</u> are equivalent to Eq. (12a), are employed by Bouwkamp but with $\underset{\sim}{F}$ replaced by $\underset{\sim}{F} \times \underset{\sim}{n}$.

** Cf. [55]; Section 26 contains a description of a complete orthogonal set of vector plane waves.

(10.14a) $$\underset{\sim}{e}'(\underset{\sim}{\rho}) = \underset{\sim}{h}'(\underset{\sim}{\rho}) \times \underset{\sim}{n} = -\nabla_t \frac{e^{ik_x x}}{ik_x} = -e^{ik_x x}\underset{\sim}{x}_o$$

and that of an H-mode wave is given by

(10.14b) $$\underset{\sim}{e}''(\underset{\sim}{\rho}) = \underset{\sim}{h}''(\underset{\sim}{\rho}) \times \underset{\sim}{n} = -\nabla_t \times \underset{\sim}{n}\frac{e^{ik_x x}}{ik_x} = e^{ik_x x}\underset{\sim}{y}_o ,$$

where $\underset{\sim}{e}$ and $\underset{\sim}{h}$ denote, respectively, the transverse electric and magnetic field distributions of the mode in question. With Bouwkamp's choice of plane wave excitation, the unperturbed transverse magnetic field $\underset{\sim}{H}_o(\underset{\sim}{\rho}) \times \underset{\sim}{n}$ in the aperture is a superposition of an E- and H- mode:

(10.15) $$\underset{\sim}{H}_o(\underset{\sim}{\rho}) \times \underset{\sim}{n} = I'\underset{\sim}{h}'(\underset{\sim}{\rho}) \times \underset{\sim}{n} + I''\underset{\sim}{h}''(\underset{\sim}{\rho}) \times \underset{\sim}{n},$$

where I' and I" are twice the incident amplitudes of the transverse magnetic fields of the plane waves defined in Eqs. (14). For Bouwkamp's choice of plane wave excitation,*

(10.16) $$I' = -2\cos\phi_o, \quad I'' = 2\sin\phi_o\cos\theta_o, \quad k_x = k\sin\theta_o .$$

The solution of Eqs. (12) for a composite wave of the form (15) is unnecessarily complicated since the symmetry of the excitation is concealed. Accordingly it is desirable to represent $\underset{\sim}{F}(\underset{\sim}{\rho})$ as a superposition of an E- and H- mode component, viz:

(10.17) $$\underset{\sim}{F}(\underset{\sim}{\rho}) = \left[I'\underset{\sim}{F}'(\underset{\sim}{\rho}) + I''\underset{\sim}{F}''(\underset{\sim}{\rho}) \right] ;$$

then (12a) decomposes into

(10.17a) $$\mathrm{curl}_t^2\underset{\sim}{F}' - k^2\underset{\sim}{F}' = -ik\underset{\sim}{h}'(\underset{\sim}{\rho}) \times \underset{\sim}{n}$$

(10.17b) $$\mathrm{curl}_t^2\underset{\sim}{F}'' - k^2\underset{\sim}{F}'' = -ik\underset{\sim}{h}''(\underset{\sim}{\rho}) \times \underset{\sim}{n},$$

and correspondingly (12b) decomposes into

(10.18a) $$\underset{\sim}{F}'(\underset{\sim}{\rho}) = 4\iint_{ap} g(\underset{\sim}{\rho},\underset{\sim}{\rho}')\underset{\sim}{E}'(\underset{\sim}{\rho}')dS'$$

(10.18b) $$\underset{\sim}{F}''(\underset{\sim}{\rho}) = 4\iint_{ap} g(\underset{\sim}{\rho},\underset{\sim}{\rho}')\underset{\sim}{E}''(\underset{\sim}{\rho}')dS',$$

where

(10.18c) $$\underset{\sim}{E}_t(\underset{\sim}{\rho}) = I'\underset{\sim}{E}'(\underset{\sim}{\rho}) + I''\underset{\sim}{E}''(\underset{\sim}{\rho}).$$

---

* See Bouwkamp's definition of the angles ϕ_o, θ_o, ρ on p.47.

Solutions of the vector partial differential equations (17) can be obtained readily. In fact the general solution of Eqs. (17) may be expressed in the form:

(10.19a) $$\frac{F'(\rho)}{ik} = \frac{h'(\rho)\times n}{k^2} + \sum_{m=1}^{\infty} 2\alpha_m \nabla_t \times n\left(\frac{J_m(k\rho)}{k} \sin m\phi\right) + \sum_{m=0}^{\infty} 2\alpha'_m \nabla_t \times n\left(\frac{J_m(k\rho)}{k} \cos m\phi\right)$$

(10.19b) $$\frac{F''(\rho)}{ik} = \frac{h''(\rho)\times n}{k^2-k_x^2} + \sum_{m=0}^{\infty} 2\beta_m \nabla_t \times n\left(\frac{J_m(k\rho)}{k} \cos m\phi\right) + \sum_{m=1}^{\infty} 2\beta'_m \nabla_t \times n\left(\frac{J_m(k\rho)}{k} \sin m\phi\right) .$$

From Eqs. (13) and (14) we see that the first term in Eq. (19a or b) is a particular solution, while the remaining "H-mode terms", which alone can satisfy the homogeneous equations (17), represent the complementary solution. The ϕ-polarization† of the complementary solution is of the most general form; however, it may be delimited by using the symmetry properties of the field. In view of the rotational symmetry of the structure about the center of the circular aperture, the ϕ-dependence of the field is determined by the nature of the excitation. The symmetry of the excitation is evident when a solution in powers of ik is considered. As Bouwkamp has shown, only the E-mode component (19a) contributes to zero-order in ik, whereas both the E- and H- components of F contribute to first-order in ik.

To make explicit the perturbation solution in powers of ik one employs the scheme:

(10.20) $$h(\rho) = h_o(\rho) + ik\, h_1(\rho) + (ik)^2 h_t(\rho) + \cdots$$

$$F(\rho) = F_o(\rho) + ik\, F_1(\rho) + (ik)^2 F_2(\rho);$$

then Eqs. (17) decompose into

(10.21a) $$\mathrm{curl}_t^2 F_o = 0$$

(10.21b) $$\mathrm{curl}_t^2 F_1 = -h_o \times n$$

(10.21c) $$\mathrm{curl}_t^2 F_2 + F_o = -h_1 \times n$$

(10.21d) $$\mathrm{curl}_t^2 F_3 + F_1 = -h_2 \times n$$

---

† Note that $x = \rho \cos \phi$; $y = \rho \sin \phi$.

(the superscripts ' and " are omitted). It should be noted that the zero- and first-order solutions $\underset{\sim}{F}_o$ and $\underset{\sim}{F}_1$ are not determined solely by Eqs. (21a and b); to determine the transverse divergence of these two solutions, Eqs. (21c and d) respectively are required. Use of Eqs. (14) in Eqs. (21a and c) indicates that in zero-order both $\underset{\sim}{F}_o$ and the corresponding aperture field $\underset{\sim}{E}_o(\underset{\sim}{\rho})$ have no preferred dependence, i.e., they are radial. Furthermore, Eqs. (14) reveal that the electric field (or $\underset{\sim}{h} \times \underset{\sim}{n}$) of an incident E- or H- mode is polarized along the $\underset{\sim}{x}_o$ or $\underset{\sim}{y}_o$ directions respectively. In accordance with this symmetry Eqs. (21b and d) indicate that in the first order the corresponding aperture electric fields must have the following polar components†:

$$(10.22a) \qquad E'_\rho \sim \cos \phi \qquad E''_\rho \sim \sin \phi$$
$$E'_\phi \sim \sin \phi \qquad E''_\phi \sim \cos \phi .$$

From (22a) the rectangular components follow by the transformation

$$(10.22b) \qquad E_x = \cos \phi \; E_\rho - \sin \phi \; E_\phi$$
$$E_y = \sin \phi \; E_\rho + \cos \phi \; E_\phi .$$

It is convenient to deal with rectangular field components since Eqs. (18) imply* that the angular symmetry of the rectangular (but not polar) components of $\underset{\sim}{E}$ and $\underset{\sim}{F}$ are similar. Since the symmetry of the first-order aperture field is known, we can give a more detailed characterization of the rim singularity than that given in (2c). Equations (22a) state that in the first order the rim condition (2c) may be decomposed into

$$(10.23) \qquad \sqrt{s}\, E'_\rho \sim \cos \phi \qquad E'_\phi \sim \sqrt{s} \sin \phi$$
$$\sqrt{s}\, E''_\rho \sim \sin \phi \qquad E''_\phi \sim \sqrt{s} \cos \phi .$$

Bouwkamp has pointed out that the zero-order solutions obtained in the literature** have usually been correct, but that this is not the case for the first-order solutions. Since the difference between the method employed here and that of Bouwkamp is most evident in the first order, only this order of solution will be considered below.

---

† See note on previous page.

* Since $g(\underset{\sim}{\rho},\underset{\sim}{\rho}')$ is independent of the ϕ orientation of the coordinate axes.

** See section VII. For a derivation employing the methods herein , cf. the author's "Coupling of Waveguides by Small Apertures", p.68, report R157-47, PIB-106, Polytechnic Institute of Brooklyn.

First order solution in (ik).

The general solutions (19) permit a ready determination of the first-order form* of $\underset{\sim}{F}(\underset{\sim}{\rho})$. It is relevant therefore to consider at this point the first-order form of the integral equations (18). Let

(10.24)
$$g(\underset{\sim}{\rho},\underset{\sim}{\rho}') = g_o(\underset{\sim}{\rho},\underset{\sim}{\rho}') + (ik)g_1(\underset{\sim}{\rho},\underset{\sim}{\rho}') + \cdots = \frac{1}{4\pi|\underset{\sim}{\rho}-\underset{\sim}{\rho}'|} + \frac{ik}{4\pi} + \cdots$$
$$\underset{\sim}{E}(\underset{\sim}{\rho}) = \underset{\sim}{E}_o(\underset{\sim}{\rho}) + (ik)\underset{\sim}{E}_1(\underset{\sim}{\rho}) + \cdots;$$

Then in view of the corresponding expansion of $\underset{\sim}{F}(\underset{\sim}{\rho})$ in (20) and the stated radial nature of $\underset{\sim}{E}_o(\underset{\sim}{\rho})$, one has the first-order integral equation (for either the ' or " component):

$$\underset{\sim}{F}_1(\underset{\sim}{\rho}) = 4\iint_{ap} g_o(\underset{\sim}{\rho},\underset{\sim}{\rho}')\,\underset{\sim}{E}_1(\underset{\sim}{\rho}')dS',$$

where $g_o(\underset{\sim}{r},\underset{\sim}{r}')$ is the static form of the Green's function in (5), defined to within a constant by

$$\nabla^2 g_o(\underset{\sim}{r},\underset{\sim}{r}') = -\delta(\underset{\sim}{r}-\underset{\sim}{r}').$$

In an oblate spheroidal coordinate system θ, ϕ associated with an aperture of radius a,

(10.26)
$$x = a\cos\theta\cos\phi$$
$$y = a\sin\theta\sin\phi,$$

where $0 < \phi < 2\pi$ and $0 < \theta < \pi/2$. The static Green's function g_o can be represented diagonally in this coordinate system in terms of a complete set of orthogonal functions $\underset{\sim}{P}_n(\cos\theta)\,\begin{matrix}\cos m\phi\\ \sin m\phi\end{matrix}$ which are periodic in ϕ and whose derivative with respect to θ vanishes at $\theta = \frac{\pi}{2}$. The latter property implies that n and m are either both even or both odd integers, and that $m \leq n$. The desired diagonal representation** [60] [61] is

(10.27)
$$g_o(\underset{\sim}{\rho},\underset{\sim}{\rho}') = \frac{1}{8a}\sum_{n=o}^{\infty}\sum_{m=o}^{n}\varepsilon_m(2n+1)\left[\frac{(n-m)!}{(n+m)!}P_n^m(0)\right]^2 P_n^m(\cos\theta)P_n^m(\cos\theta')\cos m(\phi-\phi'),$$

where ε_m is the Neumann number and equals 1 or 2 depending on whether m = 0 or > 0. On substitution of (26) and (27) one obtains as the "diagonal" form of the integral equation (25):

---

* The first-order solution can likewise be obtained by means of Eqs. (21).

** Note that Eq. (27) represents a convenient form for some of Bouwkamp's integral theorems.

$$\underset{\sim}{F}_1(\underset{\sim}{\rho}) = \frac{a}{2}\int_0^{2\pi} d\phi' \int_0^{\pi/2} d\Theta' \sin\Theta' \, \mathcal{E}(\Theta') \left\{ \sum_{n,m} \epsilon_m (2n+1) \left[\frac{(n-m)!}{(n+m)!} P_n^m(0)\right]^2 P_n^m(\cos\Theta) P_n^m(\cos\Theta') \cos m(\phi - \phi') \right\}, \tag{10.28}$$

where

$$\mathcal{E}(\Theta) = \cos\Theta \, \underset{\sim}{E}_1(\underset{\sim}{\rho}).$$

It is evident that Eq. (28) can be solved immediately on representation of $\underset{\sim}{F}_1(\underset{\sim}{\rho})$ in a series of $P_n^m(\cos\Theta)\begin{matrix}\cos m\phi \\ \sin m\phi\end{matrix}$ functions. Two cases will be distinguished, one in which $\underset{\sim}{F}_1 = \underset{\sim}{F}_1'$ and the other in which $\underset{\sim}{F}_1 = \underset{\sim}{F}_1''$.

E-mode Solution

In view of Eqs. (14a) and (22a) we see that the only terms of the general E-mode solution (19a) that yield contributions to the first order in ik are

$$\underset{\sim}{F}_1'(\underset{\sim}{\rho}) = -\frac{1}{k^2}\left(1 - \frac{k_x^2\rho^2}{2}\cos^2\phi\right)\underset{\sim}{x}_o + 2\alpha_1 \nabla_t \times \underset{\sim}{n}\left(\frac{J_1(k\rho)}{k}\sin\phi\right) + 2\alpha_3 \nabla_t \times \underset{\sim}{n}\left(\frac{J_3(k\rho)}{k}\sin 3\phi\right) \tag{10.29}$$

where the omitted α_m, α_m' terms have the wrong symmetry and the omitted $ik_x x \underset{\sim}{x}_o$ term has been considered in the zero-order approximation. To the desired order Eq. (29) becomes

$$\underset{\sim}{F}_1'(\underset{\sim}{\rho}) = \left[\alpha_1 - \frac{1}{k^2} + (\sin^2\Theta_o - \alpha_1 k^2)\frac{\rho^2}{4} + ((\alpha_1+\alpha_3)k^2 + 2\sin^2\Theta_o)\frac{\rho^2}{8}\cos 2\phi\right]\underset{\sim}{x}_o + (\alpha_1 - \alpha_3)\frac{k^2\rho^2}{8}\sin 2\phi \, \underset{\sim}{y}_o \tag{10.30}$$

where we have put $k_x = k\sin\Theta_o$. Substituting $\rho = a\sin\Theta$ and noting that

$$\sin^2\Theta = \frac{2}{3}\left[1 - P_2(\cos\Theta)\right] = \frac{1}{3}P_2^2(\cos\Theta), \tag{10.31}$$

one obtains

$$\underset{\sim}{F}_1'(\underset{\sim}{\rho}) = \left[\alpha_1 - \frac{1}{k^2} + (\sin^2\Theta_o - \alpha_1 k^2)\frac{a^2}{6} + (\alpha_1 k^2 - \sin^2\Theta_o)\frac{a^2}{6}P_2(\cos\Theta) + ((\alpha_1+\alpha_3)k^2 + 2\sin^2\Theta_o)\frac{a^2}{24}P_2^2(\cos\Theta)\cos 2\phi\right]\underset{\sim}{x}_o + (\alpha_1 - \alpha_3)\frac{k^2a^2}{24}P_2^2(\cos\Theta)\sin 2\phi \, \underset{\sim}{y}_o. \tag{10.32}$$

The integral equation for the first-order aperture field $\underset{\sim}{E}_1'(\underset{\sim}{\rho})$ is given by (28) with (32) as the left-hand member. On equating coefficients of corresponding terms, one finds that

$$\cos\Theta \, \underset{\sim}{E}_1'(\underset{\sim}{\rho}) = \left[A_o + A_2 P_2(\cos\Theta) + A_2^2 P_2^2(\cos\Theta)\cos 2\phi\right]\underset{\sim}{x}_o + B_2^2 P_2^2(\cos\Theta)\sin 2\phi \, \underset{\sim}{y}_o, \tag{10.33}$$

where

$$\pi a A_o = \alpha_1 - \frac{1}{k^2} + (\sin^2\theta_o - \alpha_1 k^2)\frac{a^2}{6}$$

$$\frac{1}{4}\pi a A_2 = (\alpha_1 k^2 - \sin^2\theta_o)\frac{a^2}{6}$$

$$\frac{3}{8}\pi a A_2^2 = ((\alpha_1+\alpha_3)k^2 + 2\sin^2\theta_o)\frac{a^2}{24}$$

$$\frac{3}{8}\pi a B_2^2 = (\alpha_1-\alpha_3)\frac{k^2a^2}{24}.$$

The application of the rim boundary condition (23) to (33) yields

$$A_2^2 = B_2^2$$

$$A_o - \frac{A_2}{2} - 3A_2^2 = 0;$$

from this the two arbitrary constants α_1 and α_3 follow:

$$\alpha_3 k^2 = -\sin^2\theta_o$$

$$\alpha_1 k^2 = 1 \quad \text{(to first order)}.$$

Equation (33) then yields for the first-order E-mode component of the electric field in the aperture:

$$(10.34)\quad \underset{\sim}{E}_1'(\underset{\sim}{\rho}) = \frac{2}{3\pi}\left[\sqrt{a^2-\rho^2}\,(2-\sin^2\theta_o)\underset{\sim}{x}_o + \frac{\underset{\sim}{\rho}\underset{\sim}{\rho}\cdot\underset{\sim}{x}_o}{\sqrt{a^2-\rho^2}}(1+\sin^2\theta_o)\right].$$

H-mode solution

As in the case of Eq. (29) the only terms in the general H-mode solution (19b) which are of interest for the first-order solution **are**:

$$(10.35)\quad \underset{\sim}{E}_1''(\underset{\sim}{\rho}) = \frac{(1-k_x^2x^2/2)}{k^2-k_x^2}\underset{\sim}{y}_o + 2\beta_1\nabla_t\times\underset{\sim}{n}\left(\frac{J_1(k\rho)}{k}\cos\phi\right) + 2\beta_3\nabla_t\times\underset{\sim}{n}\left(\frac{J_3(k\rho)}{k}\cos 3\phi\right),$$

where the coefficients β_o and β_2 have been employed to remove in first order the **ikx** $\underset{\sim}{y}_o$ term of the particular solution, and where all β_m and β_m' terms with the wrong symmetry (cf. Eqs. 22) have been omitted. Using the relations (31), we can rewrite Eq. (36) to the first order as follows:

$$\underset{\sim}{E}_1''(\underset{\sim}{\rho}) = -(\beta_1+\beta_3)\frac{k^2a^2}{24}P_2^2(\cos\theta)\sin 2\phi\,\underset{\sim}{x}_o + \left[\frac{1}{k^2} - \beta_1 + (\beta_1 k^2 - \tan^2\theta_o)\frac{a^2}{6}\right.$$

$$(10.37)\quad \left. -(\beta_1 k^2 - \tan^2\theta_o)\frac{a^2}{6}P_2(\cos\theta) + ((\beta_1-\beta_3)k^2 - 2\tan^2\theta_o)\frac{a^2}{24}P_2^2(\cos\theta)\cos 2\phi\right]\underset{\sim}{y}_o.$$

As before, if we substitute (37) into the integral equation (28) for the aperture field $\underset{\sim}{E}_1''$ and equate the corresponding coefficients, we obtain

$$\cos\theta\, \mathbf{E}_1''(\rho) = C_2^2 P_2^2(\cos\theta)\sin 2\phi\, \underset{\sim}{x}_o + \left[D_o + D_2 P_2(\cos\theta) + D_2^2 P_2^2(\cos\theta)\cos 2\phi\right] \underset{\sim}{y}_o \,, \tag{10.38}$$

where

$$\frac{3}{8}\pi a C_2^2 = -(\beta_1+\beta_3)\frac{k^2a^2}{24}$$

$$\pi a D_o = \frac{1}{k^2-k_x^2} - \beta_1 + (\beta_1 k^2 - \tan^2\theta_o)\frac{a^2}{6}$$

$$\frac{1}{4}\pi a D_2 = (\tan^2\theta_o - \beta_1 k^2)\frac{a^2}{6}$$

$$\frac{3}{6}\pi a D_2^2 = ((\beta_1-\beta_3)k^2 - 2\tan^2\theta_o)\frac{a^2}{24} \,.$$

Furthermore, the imposition of the rim conditions (23) on (38) leads to

$$C_2^2 = -D_2^2$$

$$D_o - \frac{D_2}{2} + 3D_2^2 = 0 \,,$$

and hence the arbitrary constants β_1 and β_3 in (38) become

$$\beta_3 k^2 = -\tan^2\theta_o$$

$$\beta_1 k^2 = \sec^2\theta_o \quad \text{(to the first order).}$$

Equation (38) then yields for the H-mode component of aperture field in the first order:

$$\underset{\sim}{E}_1''(\underset{\sim}{\rho}) = -\frac{2}{3\pi}\left[2\sqrt{a^2-\rho^2}\, \underset{\sim}{y}_o + \frac{\underset{\sim}{\rho}\underset{\sim}{\rho}\cdot\underset{\sim}{y}_o}{\sqrt{a^2-\rho^2}}\right]. \tag{10.39}$$

In view of Eqs. (16) and (18c) we find by superposition that the total first-order aperture field for the Bouwkamp choice of incident wave is

$$ik\underset{\sim}{E}_1(\underset{\sim}{\rho}) = \frac{2ik}{3\pi}\left[\sqrt{a^2-\rho^2}\,(2I'\underset{\sim}{x}_o - I''\underset{\sim}{y}_o - I'\sin^2\theta_o \underset{\sim}{x}_o) + \frac{\rho}{\sqrt{a^2-\rho^2}}\,\underset{\sim}{\rho}\cdot(I'\underset{\sim}{x}_o - I''\underset{\sim}{y}_o + I'\sin^2\theta_o\underset{\sim}{x}_o)\right] \tag{10.40a}$$

or, as Bouwkamp obtained,

$$= \frac{2}{3\pi}\left[\sqrt{a^2-\rho^2}\,(-ik2\underset{\sim}{H}_o \times \underset{\sim}{n} - \nabla_t E_{on}) + \frac{\rho}{\sqrt{a^2-\rho^2}}\,\underset{\sim}{\rho}\cdot(-ik\underset{\sim}{H}_o \times \underset{\sim}{n} + \nabla_t E_{on})\right] \tag{10.40b}$$

where $\underset{\sim}{H}_o$ and $\nabla_t E_{on}$ are the unperturbed values evaluated at the center of the aperture.

References

1. Baker and Copson, The Mathematical Theory of Huygens' Principle, (Oxford, Clarendon Press), 1950, p. 71.

2. Ibid., p. 98.

3. Sommerfeld, Vorlesungen über theoretische Physik, vol. 4, Optik (Wiesbaden, Dieterich Verlag) 1950, p. 208 ff.

4. Born, Optik (Berlin, Springer Verlag) 1933, p. 152.

5. Franz, Z. Physik 125 (1949) 563.

6. Schelkunoff, Comm. Pure Appl. Math. 4 (1951) 43.

7. Franz, Z. Physik 128 (1950) 432; Proc. Roy. Soc. A63 (1950) 925.

8. Sommerfeld, op.cit., p. 203.

9. Bouwkamp, Thesis (Groningen, Wolters) 1941.

10. Ibid.

11. Braunbek, Z. Physik 127 (1950) 381.
    Meixner and Fritze, Z. Angew. Physik 1 (1949) 535.
    Severin and Starke, Acoustica (Akust. Beiheft, 2) 1952, p. 59.

12. Bremmer, Comm. Pure Appl. Math. 4 (1951) 61.

13. Luneberg, Mathematical Theory of Optics (Brown University) 1944.

14. Scheffers, Ann. Physik 42 (1942) 211.

15. Durand, C.R. Acad. Sci. Paris 226 (1948) 1440, 1593.

16. Spence, J. Acoust. Soc. Amer. 21 (1949) 98.

17. Primakoff et al., J. Acoust. Soc. Amer. 19 (1947) 132.

18. Braunbek, Z. Physik 127 (1950) 381; Ibid. 127 (1950) 405.

19. Brillouin, C. R. Acad. Sci. Paris 229 (1949) 513.

20. Bouwkamp, Physica 12 (1947) 467.

21. Severin and Starke, Acoustica-Akust. Beiheft 2 (1952) 59.

22. Meixner and Fritze, Z. Angew. Physik 3 (1951) 171.

23. Levine and Schwinger, Phys. Rev. 74 (1948) 958, 1212.

24. Baker and Copson, op.cit., p. 186.

25. Ibid., p. 160.

26. Ibid., p. 184.

27. Maue, Z. Physik 126 (1949) 601, equation 60.

28. Levine and Schwinger, Phys. Rev. 74 (1948) 958.

29. Sommerfeld, op .cit., p. 283.

30. Levine and Schwinger, Phys. Rev. 74 (1948) 958.

31. Baker and Copson, op .cit., p. 186.

32. Levine, Comm. Pure Appl. Math. 3 (1950) 355.

33. Levine and Schwinger, Phys. Rev. 74 (1948) 958.

34. Sommerfeld, Ann. Physik 42 (1942) 389.

35. Magnus, Res. Rep. No. EM-32, N.Y.U., Washington Sq. College, Math. Res. Group (1951).

36. Idem.

37. See ref. 34, also Am. Physik 2 (1948) 85.

38. Bouwkamp, Physica 16 (1950) 1.

39. Levine and Schwinger, Phys. Rev. 74 (1948) 958; appendix 1.

40. Levine and Schwinger, Comm. Pure Appl. Math. 3 (1950) 355.

41. Miles, J. Acoust. Soc. Amer. 21 (1949) 140, 434.

42. Copson, Proc. Roy. Soc. A 186 (1946) 110.

43. Copson, Ibid, A 202 (1950) 277.

44. Meixner, Ann. Physik 6 (1949) 2.

45. Maue. Z. Physik 126 (1949) 601.

46. Bethe, Phys. Rev. 66 (1944) 163.

47. Bouwkamp, Philips Res. Rep. 5 (1950) 321.

48. Bouwkamp, Philips Res. Rep. 5 (1950) 401.

49. Bouwkamp, Proc. Kon. Ned. Ak. Wet. Amsterdam 52,(1949) 987.

50. Bouwkamp, Ibid, 53 (1950) 654.

51. Bouwkamp, Math. Revs. 8 (1947) 179.

52. Miles, Quart. Appl. Math. 7 (1949) 45; J.Appl.Phys. 20 (1949) 760.

53. Bourgin, Math. Revs. 6 (1945) 165.

54. Copson, Math. Revs. 12 (1951) 774, and private correspondence.

55. cf. Marcuvitz, Waveguide Handbook (New York, McGraw Hill), 1949.

56. Sommerfeld, op.cit., page 294.

57. Groschwitz and Hönl, Z. Physik 131 (1952) 305.

58. Sommerfeld, op.cit., page 295.

59. See ref. 57.

60. MacRoberts, T.M., "Spherical Harmonics", page 218, 2nd edition, Dover.

61. Morse and Feshbach, Equations of Mathematical Physics, (McGraw Hill, 1953).