PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Solution of a Diffraction Problem I. The Wide Double Wedge

H. M. Nussenzveig

Phil. Trans. R. Soc. Lond. A 1959 **252**, 1-30
doi: 10.1098/rsta.1959.0012

SOLUTION OF A DIFFRACTION PROBLEM*
I. THE WIDE DOUBLE WEDGE

By H. M. NUSSENZVEIG
Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil

(*Communicated by M. J. Lighthill, F.R.S.—Received* 31 *January* 1959)

CONTENTS

The problem of radiation from a semi-infinite parallel-plate waveguide, terminated by an infinite plane flange ('double wedge') is investigated in the *k*-representation. The problem is reduced to an infinite system (*S*) of linear equations. Part I deals with the case of a 'wide double wedge' (the width of the waveguide much greater than the wavelength). This may be considered as a perturbation of the problem of diffraction by a single rectangular wedge, the rigorous solution of which is known (Reiche 1912). In the *k*-representation, this solution satisfies an integral equation, which is a limiting form of system (*S*). This leads to an approximate solution of the wide double-wedge problem, which appears to be a very good approximation, except in the neighbourhood of critical frequencies of the waveguide. The diffraction patterns and the reflexion coefficient are evaluated in this approximation. The accuracy and limits of applicability of classical diffraction theory are discussed. In the neighbourhood of critical frequencies, strong reflexion appears. This is a new limiting case of the problem, which is connected with the theory of quasi-stationary currents. Neumann's iteration method is applied to system (*S*), in order to investigate the possibility that Kirchhoff's approximation is the first step of an accurate solution by successive approximations. The convergence of Neumann's series is discussed.

1. Introduction

The classical theory of diffraction (Fresnel–Kirchhoff) is usually applied to the diffraction by large apertures (smallest aperture dimension $\gg \lambda$, where λ is the wavelength). The results are in excellent agreement with experiment, for not too large diffraction angles. How can this agreement be explained?

* This paper contains the results of a doctoral thesis submitted to the Faculty of Science of the University of São Paulo.

It is claimed, sometimes, that the closeness to geometrical optics justifies Kirchhoff's approximation. However, Kirchhoff's approximation may fail (in the neighbourhood of a sharp edge, for instance), without appreciably affecting the classical diffraction patterns (Sommerfeld 1954, pp. 264–5). As Sommerfeld remarks, 'It is amazing that the classical diffraction theory nevertheless yields for all practical purposes satisfactory results'.

It has also been suggested that Kirchhoff's formula might be the first step of an accurate solution by successive approximations (Born 1933, p. 152). A proposed iteration method, based on Helmholtz's formula, does not succeed, as was shown by Franz (1949) and Schelkunoff (1951). This does not rule out the possibility of an accurate solution by a different method of successive approximations.

In the few problems where the rigorous solution is known, the accuracy of classical theory may be determined by direct comparison. So far, the scope of rigorous theory has been restricted to the ideal case of perfect conductors. The first problem to be rigorously solved (Sommerfeld 1896) was that of a perfectly conducting wedge (in particular, a half-plane). This is not very convenient for comparison with classical theory, since no linear dimension of the diffracting object is finite.

Diffraction by objects having one typical finite dimension a is more interesting. We may define, in this case, a 'short-wavelength region' ($\lambda \ll a$), an 'intermediate wavelength region' ($\lambda \sim a$) and a 'long-wavelength region' ($\lambda \gg a$). The domain of classical theory is the short-wavelength region, where interference phenomena play an important role. In the long-wavelength region, quasi-static approximations may be employed; the phenomena which occur in this domain are of a much simpler nature.

Important cases of diffraction by objects having a finite dimension, which may be rigorously solved by the Wiener–Hopf technique, appear in the theory of radiation from semi-infinite waveguides (Vajnshtejn 1954). Although the solutions become very complicated in the short-wavelength region, a partial comparison with the results of Kirchhoff's theory is possible. A review of this work, as well as an excellent account of recent progress in diffraction theory, has been given by Bouwkamp (1954).

As may be seen by the preceding discussion, no satisfactory answer has yet been given to the following questions: (A) How can the success of classical diffraction theory be explained? (B) What is the accuracy, and what are the limits of applicability of this theory? (C) Can Kirchhoff's approximation be considered as the first step of an accurate solution by successive approximations? Since most known rigorous solutions are based on complex mathematical methods, which frequently obscure the physical meaning of the results, it would be desirable: (D) To find a physical picture of the phenomena which occur in the short- and in the long-wavelength regions, allowing us to establish a connexion between them.

The object of the present paper is to throw some light on the above questions. It has been shown in a previous paper (Nussenzveig 1959) that, by employing the 'k-representation', we may arrive at a qualitative understanding of the success of classical diffraction theory. In this paper, to obtain quantitative results, a typical diffraction problem, to be described in § 2, will be investigated in the k-representation. Part I deals with the short-wavelength region; the long-wavelength region will be considered in part II. The conclusions will be given at the end of part II.

2. Formulation of the problem

The simplest diffraction problem, involving one finite dimension, which may be treated in the k-representation, seems to be the 'double-wedge' problem. A double wedge is a pair of perfectly conducting rectangular wedges (figure 1). It is equivalent to a parallel-plate waveguide of width $d = 2a$, with an infinite plane flange at the mouth. The incident wave is one of the waveguide modes, travelling towards the open end.

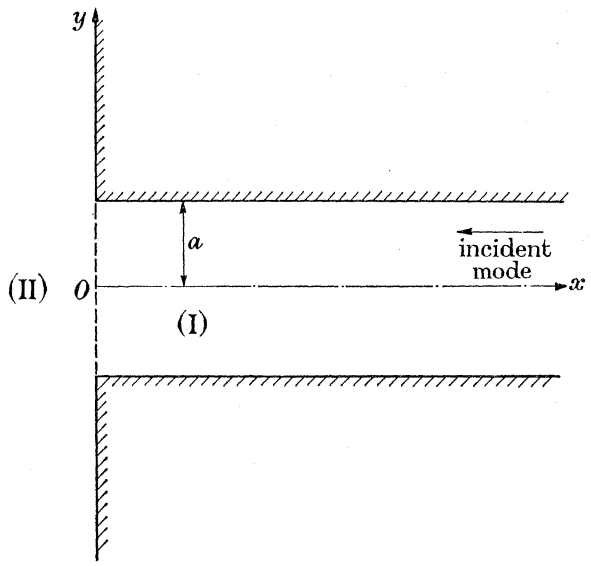

Figure 1. Co-ordinate system.

The co-ordinate system is shown in figure 1. The time factor $\exp(-\mathrm{i}\omega t)$ $(\omega = kc = 2\pi c/\lambda)$ will be omitted throughout. We shall consider only the case of transverse magnetic (TM) waves, which may be described by a single scalar wave function $u(x,y)$. The field components are

$$\mathbf{H} = (0, 0, u); \quad \mathbf{E} = \left(\frac{\mathrm{i}}{k}\frac{\partial u}{\partial y}, -\frac{\mathrm{i}}{k}\frac{\partial u}{\partial x}, 0\right). \tag{2·1}$$

The function $u(x,y)$ has the following properties (Bouwkamp 1954, p. 38):

$$(a) \quad (\partial^2/\partial x^2 + \partial^2/\partial y^2 + k^2)\, u(x,y) = 0, \tag{2·2}$$

$$(b) \quad \partial u/\partial y(x, \pm a) = 0 \quad (x > 0), \tag{2·3}$$

$$(c) \quad \partial u/\partial x(0, y) = 0 \quad (|y| > a), \tag{2·4}$$

$$(d) \quad u(0+, y) = u(0-, y) \quad (|y| < a), \tag{2·5}$$

$$(e) \quad \partial u/\partial x(0+, y) = \partial u/\partial x(0-, y) \quad (|y| < a), \tag{2·6}$$

(f) u satisfies Sommerfeld's radiation condition. This condition takes different forms in regions I and II (see figure 1): (f I) the only incoming wave in region I is the incident mode; (f II) there are no incoming waves in region II.

(g) ∇u is quadratically integrable over any domain of three-dimensional space, including the edges of the double wedge ('edge condition').

This boundary-value problem admits also an acoustical interpretation: it suffices to re-interpret $u(x,y)$ as the velocity potential of sound waves, assuming that the walls of the double wedge are perfectly rigid.

We shall take as incident wave a symmetric TM mode, so that the total wave function will be symmetric about Ox: $u(x,y) = u(x,-y)$. The symmetric TM modes of a parallel-plate waveguide may be classified according to the eigenvalues of k_y,

$$k_{y,n} = n\pi/a \quad (n = 0, 1, 2, \ldots). \tag{2·7}$$

For $n = 0$, we obtain the principal mode, the only mode which may travel within the waveguide for any value of the frequency; $k_{y,n} < k$ corresponds to travelling modes; $k_{y,n} > k$ leads to evanescent modes (Toraldo di Francia 1956, p. 245). A travelling mode may be regarded as arising from the superposition of a plane wave, travelling at an angle θ_n with respect to the x-axis, and its mirror image with respect to the plates, the 'angle of propagation' θ_n being given by $\sin\theta_n = k_{y,n}/k$. The incident mode will be denoted by $n = s$.

According to condition (f I), the wave function in region I may be written as

$$u_{\mathrm{I}}(x,y) = \cos(k_{y,s}y)\exp(-\mathrm{i}k_{x,s}x) + \sum_{n=0}^{\infty} a_n \cos(k_{y,n}y)\exp(\mathrm{i}k_{x,n}x), \tag{2·8}$$

where
$$k_{x,n} = (k^2 - k_{y,n}^2)^{\frac{1}{2}}, \quad \mathscr{I}(k_{x,n}) \geqslant 0. \tag{2·9}$$

Equation (2·8) represents the incoming mode (normalized to unit amplitude), plus a superposition of all 'reflected' travelling modes and all evanescent modes. The unknown coefficients a_n will be called 'mode amplitudes'.

Similarly, according to (f II), the wave function in region II may be represented by a superposition of outgoing travelling waves and evanescent waves:

$$u_{\mathrm{II}}(x,y) = \int_0^\infty A(k_y)\cos(k_y y)\exp(-\mathrm{i}k_x x)\,\mathrm{d}k_y, \tag{2·10}$$

$$k_x = (k^2 - k_y^2)^{\frac{1}{2}}, \quad \mathscr{I}(k_x) \geqslant 0. \tag{2·11}$$

This representation will be called 'k-representation' (Nussenzveig 1959).

The coefficients a_n and $A(k_y)$ may be expressed in terms of each other. In fact, it follows from (2·5) that

$$a_m + \delta_{m,s} = (1/a\zeta_m)\int_{-a}^{+a}\mathrm{d}y\int_0^\infty A(k_y)\cos(k_{y,m}y)\cos(k_y y)\,\mathrm{d}k_y, \tag{2·12}$$

where $\zeta_0 = 2$, $\zeta_m = 1\,(m \neq 0)$, and $\delta_{m,s} = 0$ if $m \neq s$, $\delta_{m,s} = 1$ if $m = s$. Similarly, it follows from (2·5) and (2·6) that

$$A(k_y) = -(1/\pi k_x)\sum_{n=0}^{\infty} k_{x,n}(a_n - \delta_{n,s})\int_{-a}^{+a}\cos(k_y y)\cos(k_{y,n}y)\,\mathrm{d}y. \tag{2·13}$$

Equations (2·12) and (2·13) may be combined in two different ways. If we eliminate a_n between them, we obtain an integral equation to determine $A(k_y)$; if we eliminate $A(k_y)$, the result is an infinite system of linear equations in the infinitely many unknowns a_n. We shall adopt the second procedure, which leads to a simpler treatment.

The infinite system of linear equations which results from the elimination may be reduced to the form

$$(S) \qquad a_m = \sum_{n=0}^{\infty} K_{m,n}a_n - K_{m,s}, \tag{2·14}$$

where

$$K_{m,n} = \tfrac{1}{2}\delta_{m,n} - (k_{x,n}/4\pi a\zeta_m)\int_{-a}^{+a}\mathrm{d}y\int_{-a}^{+a}\mathrm{d}y'\int_{-\infty}^{+\infty}\cos(k_{y,m}y)\cos(k_{y,n}y')\cos(k_y y)\cos(k_y y')\,\mathrm{d}k_y/k_x. \tag{2·15}$$

If we find a solution of system (S), the corresponding wave function (which must satisfy condition (g)) is determined by (2·8), (2·10) and (2·13). Our problem is therefore reduced to finding the solution of system (S).

3. The coefficients $K_{m,n}$

In this section, we shall give an outline of the methods for the evaluation of the coefficients $K_{m,n}$ defined in (2·15). Computational details may be found elsewhere (Nussenzveig 1957; henceforward quoted as A).

Introducing the dimensionless parameters

$$K = ka = 2\pi a/\lambda; \quad \gamma = k_y/k; \quad \gamma_n = k_{y,n}/k = n\pi/K \tag{3·1}$$

and employing the integral representation

$$H_0(k\,|y-y'|) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \cos\,[k_y(y-y')]\,\mathrm{d}k_y/k_x, \tag{3·2}$$

where H_0 denotes Hankel's function of the first kind of order zero,* we can reduce (2·15) to the form

$$K_{m,n} = \tfrac{1}{2}\delta_{m,n} + [(1-\gamma_n^2)^{\frac{1}{2}}/(2K\zeta_m)]\,c_{m,n}, \tag{3·3}$$

where
$$c_{m,n} = -\frac{1}{2}\int_{-K}^{+K}\mathrm{d}t\int_{-K}^{+K} H_0(|t-t'|)\cos(\gamma_m t)\cos(\gamma_n t')\,\mathrm{d}t' = c_{n,m}. \tag{3·4}$$

If we make in (3·4) the change of variables: $t = u-K$; $t'-t = v$, and integrate with respect to u, we find, for $m \neq n$,

$$c_{m,n} = (-1)^{m+n}(\gamma_m^2-\gamma_n^2)^{-1}[\gamma_m S_m(K) - \gamma_n S_n(K)], \tag{3·5}$$

where
$$S_n(K) = \int_0^{2K} H_0(v)\sin(\gamma_n v)\,\mathrm{d}v; \tag{3·6}$$

whereas, for $m = n$, we find

$$c_{n,n} = \tfrac{1}{2}\{[S_n(K)/\gamma_n] - 2KC_n(K) + D_n(K)\} \quad (\text{for} \quad n \neq 0), \tag{3·7}$$

$$c_{0,0} = -2KC_0(K) + D_0(K), \tag{3·8}$$

where
$$C_n(K) = \int_0^{2K} H_0(v)\cos(\gamma_n v)\,\mathrm{d}v, \tag{3·9}$$

$$D_n(K) = \int_0^{2K} H_0(v)\cos(\gamma_n v)\,v\,\mathrm{d}v. \tag{3·10}$$

Integrating (3·10) by parts in two different ways gives

$$D_n(K) = (1-\gamma_n^2)^{-1}[2KH_1(2K) + (2\mathrm{i}/\pi) + \gamma_n S_n(K)], \tag{3·11}$$

where H_1 is Hankel's function of the first kind of order one.

Substituting (3·5) to (3·11) in (3·3), we finally get

$$K_{m,n} = (-1)^{m+n}[(1-\gamma_n^2)^{\frac{1}{2}}/(2K\zeta_m)]\,(\gamma_m^2-\gamma_n^2)^{-1}[\gamma_m S_m(K) - \gamma_n S_n(K)] \quad (m \neq n), \tag{3·12}$$

$$K_{n,n} = \tfrac{1}{2}[1-(1-\gamma_n^2)^{\frac{1}{2}}C_n(K)] + [(1-\gamma_n^2)^{-\frac{1}{2}}/4K]\,[2KH_1(2K) + (2\mathrm{i}/\pi) + S_n(K)/\gamma_n] \quad (n \neq 0), \tag{3·13}$$

$$K_{0,0} = \tfrac{1}{2}[1 - C_0(K) + H_1(2K) + (\mathrm{i}/\pi K)]. \tag{3·14}$$

* Since we shall employ only Hankel functions of the first kind, the superscript (1) will be omitted.

These equations reduce the evaluation of the coefficients $K_{m,n}$ to that of the functions $S_n(K)$ and $C_n(K)$ defined in (3·6) and (3·9). To study the behaviour of these functions, we must distinguish between different wavelength regions. In the short-wavelength region ($K \gg 1$), we may employ asymptotic expansions; in the long-wavelength region ($K \ll 1$), expansions about the origin. The latter case will be deferred to part II.

Let us consider the case $K \gg 1$. We may write

$$S_n(K) = S(\gamma_n) - \mathscr{S}_n(K); \quad C_n(K) = C(\gamma_n) - \mathscr{C}_n(K), \tag{3·15}$$

where
$$S(\gamma_n) = \int_0^\infty H_0(v)\sin(\gamma_n v)\,\mathrm{d}v; \quad C(\gamma_n) = \int_0^\infty H_0(v)\cos(\gamma_n v)\,\mathrm{d}v, \tag{3·16}$$

$$\mathscr{S}_n(K) = \int_{2K}^\infty H_0(v)\sin(\gamma_n v)\,\mathrm{d}v; \quad \mathscr{C}_n(K) = \int_{2K}^\infty H_0(v)\cos(\gamma_n v)\,\mathrm{d}v. \tag{3·17}$$

According to known results (Erdélyi 1954, pp. 43, 47, 99, 103),

$$C(\gamma_n) = (1-\gamma_n^2)^{-\frac{1}{2}}, \tag{3·18}$$

$$S(\gamma_n) = \begin{cases} (2\mathrm{i}/\pi)\,(1-\gamma_n^2)^{-\frac{1}{2}}\sin^{-1}\gamma_n & (\text{for } 0 \leqslant \gamma_n < 1), \\ (\gamma_n^2-1)^{-\frac{1}{2}}[1-(2\mathrm{i}/\pi)\ln\sigma_n] & (\text{for } \gamma_n > 1), \end{cases} \tag{3·19}$$

where
$$\sigma_n = \gamma_n + (\gamma_n^2-1)^{\frac{1}{2}} \tag{3·20}$$

and, in agreement with (2·9), $\mathscr{I}[(1-\gamma_n^2)^{\frac{1}{2}}] \geqslant 0$.

To obtain an asymptotic expansion of the integrals defined in (3·17), it suffices to replace $H_0(v)$ by its asymptotic expansion and to integrate term by term (for details, see A).

Let us introduce the parameter

$$\delta_n = (K\,|1-\gamma_n|/\pi)^{\frac{1}{2}}. \tag{3·21}$$

This parameter is a measure of the distance between the nth mode and the *critical mode* $n = c$, which is defined by the condition that all modes with $n \leqslant c$ are travelling modes, whereas modes with $n > c$ are evanescent modes. If $\delta_n \lesssim 1$, we shall say that the nth mode belongs to the *critical strip*. It turns out that we must distinguish between the cases $\delta_n \gg 1$ and $\delta_n \lesssim 1$.

For $\delta_n \gg 1$, we find the asymptotic expansions

$$\mathscr{S}_n(K) = -(\pi K)^{-\frac{1}{2}}(1-\gamma_n^2)^{-1}\gamma_n \exp[\mathrm{i}(2K-\tfrac{1}{4}\pi)]\,[1+(\mathrm{i}/16K)\,(\gamma_n^2-9)/(1-\gamma_n^2) + O(\delta_n^{-4})] + O(K^{-\frac{5}{2}}), \tag{3·22}$$

$$\mathscr{C}_n(K) = (\pi K)^{-\frac{1}{2}}(1-\gamma_n^2)^{-1}\exp[\mathrm{i}(2K+\tfrac{1}{4}\pi)]\,[1-(\mathrm{i}/16K)\,(3\gamma_n^2+5)/(1-\gamma_n^2) + O(\delta_n^{-4})] + O(K^{-\frac{5}{2}}). \tag{3·23}$$

It follows, by comparing these results with (3·18) and (3·19), that

$$|\mathscr{S}_n(K)/S(\gamma_n)| \ll 1; \quad |\mathscr{C}_n(K)/C(\gamma_n)| \ll 1, \tag{3·24}$$

so that $\mathscr{S}_n(K)$ and $\mathscr{C}_n(K)$ may generally be neglected for $\delta_n \gg 1$.

For $\delta_n \lesssim 1$, we find

$$\mathscr{S}_n(K) = \{\mathrm{i}[2(1-\gamma_n)]^{-\frac{1}{2}} - (K/\pi)^{\frac{1}{2}}\exp(\tfrac{1}{4}\mathrm{i}\pi)\,F^{\pm}(2\delta_n)/\delta_n\}\,[1+\tfrac{1}{4}(1-\gamma_n)] + O(K^{-\frac{1}{2}}), \tag{3·25}$$

$$\mathscr{C}_n(K) = \{[2(1-\gamma_n)]^{-\frac{1}{2}} - (K/\pi)^{\frac{1}{2}}\exp(-\tfrac{1}{4}\mathrm{i}\pi)\,F^{\pm}(2\delta_n)/\delta_n\}\,[1+\tfrac{1}{4}(1-\gamma_n)] + O(K^{-\frac{1}{2}}), \tag{3·26}$$

where
$$F^{+}(w) = \int_0^w \exp(\tfrac{1}{2}\mathrm{i}\pi t^2)\,\mathrm{d}t \tag{3·27}$$

is the Fresnel integral and $F^-(w)$ is its complex conjugate. Wherever a double sign occurs, the upper sign refers to $\gamma_n < 1$, and the lower one refers to $\gamma_n > 1$. Expanding (3·18) and (3·19) in powers of $(1-\gamma_n)$, and taking into account (3·25) and (3·26), we finally get, for $\delta_n \lesssim 1$,

$$S_n(K) = (K/\pi)^{\frac{1}{2}} \exp\left(\tfrac{1}{4}\mathrm{i}\pi\right) \left[1+\tfrac{1}{4}(1-\gamma_n)\right] F^{\pm}(2\delta_n)/\delta_n - 2\mathrm{i}/\pi + O(K^{-\frac{1}{2}}), \tag{3·28}$$

$$C_n(K) = (K/\pi)^{\frac{1}{2}} \exp\left(-\tfrac{1}{4}\mathrm{i}\pi\right) \left[1+\tfrac{1}{4}(1-\gamma_n)\right] F^{\pm}(2\delta_n)/\delta_n + O(K^{-\frac{1}{2}}). \tag{3·29}$$

In particular, for $\delta_n \ll 1$, we find, employing the power series expansion of the Fresnel integral,

$$S_n(K) = 2(K/\pi)^{\frac{1}{2}} \exp\left(\tfrac{1}{4}\mathrm{i}\pi\right) - 2\mathrm{i}/\pi \pm 4\exp\left(\tfrac{3}{4}\mathrm{i}\pi\right) (\pi K)^{\frac{1}{2}}\, \delta_n^2/3 + \ldots + O(K^{-\frac{1}{2}}), \tag{3·30}$$

$$C_n(K) = 2(K/\pi)^{\frac{1}{2}} \exp\left(-\tfrac{1}{4}\mathrm{i}\pi\right) \pm 4\exp\left(\tfrac{1}{4}\mathrm{i}\pi\right) (\pi K)^{\frac{1}{2}}\, \delta_n^2/3 + \ldots + O(K^{-\frac{1}{2}}). \tag{3·31}$$

These results may be extended to the case in which K is not $\gg 1$, with the help of the exact integrals

$$f(z) = \int_0^z H_0(v) \sin(z-v)\, \mathrm{d}v = \frac{2\mathrm{i}}{\pi} \cos z + zH_1(z), \tag{3·32}$$

$$g(z) = \int_0^z H_1(v) \sin(z-v)\, v\, \mathrm{d}v = \frac{4\mathrm{i}}{3\pi} \cos z + \tfrac{1}{3}[2zH_1(z) - z^2 H_0(z)], \tag{3·33}$$

which may be derived (see A) by a method similar to that which is employed in the evaluation of Kapteyn's trigonometric integrals (Watson 1948, p. 380). For $n \neq 0$, we have

$$S_n(n\pi) = -f(2n\pi); \quad C_n(n\pi) = f'(2n\pi); \tag{3·34}$$

$$n\pi S_n'(n\pi) = -f(2n\pi) + g(2n\pi); \quad n\pi C_n'(n\pi) = f'(2n\pi) - g'(2n\pi), \tag{3·35}$$

where primes indicate derivatives with respect to the argument. With the help of (3·32) to (3·35), $S_n(K)$ and $C_n(K)$ may be expanded in Taylor series about the point $K = n\pi$. The results are as follows:

$$S_n(K) = -2n\pi H_1(2n\pi) - \frac{2\mathrm{i}}{\pi} \mp \tfrac{2}{3}\pi\delta_n^2[2n\pi H_0(2n\pi) + H_1(2n\pi) + (\mathrm{i}/n\pi^2)] + O(\delta_n^4), \tag{3·36}$$

$$C_n(K) = 2KH_0(2n\pi) \mp \tfrac{4}{3}\pi K\delta_n^2 H_1(2n\pi) + O(\delta_n^4). \tag{3·37}$$

Equations (3·36) and (3·37) may be applied in the neighbourhood of $K = n\pi$, for any $n \neq 0$. For $K \gg 1$, they reduce to (3·30) and (3·31).

The above results suffice for the evaluation of the coefficients $K_{m,n}$ in all cases which will be required in part I. The final results, obtained with the help of (3·12) to (3·14), are listed below.

First case: $K \gg 1$, $\delta_m \gg 1$, $\delta_n \gg 1$,

$$K_{m,n} = (-1)^{m+n}[(1-\gamma_n^2)^{\frac{1}{2}}/(2K\zeta_m)]\,(K'_{m,n} + \Delta K'_{m,n}) \quad (m \neq n), \tag{3·38}$$

$$K'_{m,n} = (\gamma_m^2 - \gamma_n^2)^{-1}\,[\gamma_m S(\gamma_m) - \gamma_n S(\gamma_n)] = K'_{n,m}, \tag{3·39}$$

$$\Delta K'_{m,n} = (\pi K)^{-\frac{1}{2}} \exp\left[\mathrm{i}(2K - \tfrac{1}{4}\pi)\right] [(1-\gamma_m^2)\,(1-\gamma_n^2)]^{-1}\,[1 + O(\delta_m^{-2}) + O(\delta_n^{-2})] = \Delta K'_{n,m}, \tag{3·40}$$

$$K_{n,n} = [(1-\gamma_n^2)^{\frac{1}{2}}/(2K\zeta_n)]\,(K'_{n,n} + \Delta K'_{n,n}), \tag{3·41}$$

$$K'_{n,n} = (1-\gamma_n^2)^{-1}\,[(\mathrm{i}/\pi) + S(\gamma_n)/2\gamma_n] = \lim_{m \to n} (K'_{m,n}). \tag{3·42}$$

$\Delta K'_{m,n}$ is a small correction to $K'_{m,n}$ in this region.

Second case: $K \gg 1$, δ_m or $\delta_n \lesssim 1$.

In this case, we must employ (3·28) and (3·29). We shall not consider all different possibilities, but only the most interesting ones, which occur when δ_m or δ_n is $\ll 1$. These are particular instances of the following case.

Third case: $K \gtrsim 1$, δ_m or $\delta_n \ll 1$.

$$K_{m,n} = (-1)^{m+n+1}[2m\pi H_1(2m\pi) + (2\mathrm{i}/\pi) + \gamma_n S_n(K) + O(\delta_m^2)]\,[2K(1-\gamma_n^2)^{\frac{1}{2}}]^{-1} \quad (\delta_m \ll 1, m \neq n), \qquad (3·43)$$

$$K_{m,n} = (-1)^{m+n+1}(\pm 2\pi/K)^{\frac{1}{2}}\,\delta_n[2n\pi H_1(2n\pi) + (2\mathrm{i}/\pi) + \gamma_m S_m(K) + O(\delta_n^2)]\,[2K\zeta_m(1-\gamma_m^2)]^{-1} \quad (\delta_n \ll 1, m \neq n), \qquad (3·44)$$

$$K_{n,n} = \tfrac{1}{2} - \tfrac{1}{3}(\pm 2\pi K)^{\frac{1}{2}}\,\delta_n[2H_0(2n\pi) + \frac{1}{K}H_1(2n\pi) + (1/\pi K^2)] + O(\delta_n^3) \quad (\delta_n \ll 1). \qquad (3·45)$$

To obtain from (3·43) to (3·45) the expressions for the case $K \gg 1$, it suffices to replace Hankel's functions by their asymptotic expansions.

4. The single-wedge problem

(a) *Transition to the single-wedge problem*

To obtain the rigorous solution of system (S) in all possible cases appears to be a problem of considerable difficulty. Our approach will be the following: we shall look for rigorous solutions in certain limiting cases, and we shall try to extend the range of these solutions by perturbation methods. It will be shown later that there are three important limiting cases of the double-wedge problem.

We shall begin with the case of the *wide double wedge*, i.e. $K \gg 1$. The corresponding limiting case is $K \to \infty$. However, this limit is not uniquely defined. One possibility, which corresponds to the *geometrical optics limit*, is to keep a fixed, while $k \to \infty$. Another possibility is to keep k fixed, while $a \to \infty$. This may be done as follows: let us keep one of the wedges fixed, while the other one is removed to infinity. Let this process be carried out in such a way that the 'angle of propagation' θ_s of the incident wave (see § 2) is unchanged in the limit. The limiting form of the double-wedge problem in this process should be the problem of diffraction of a plane wave incident at an angle θ_s on a single rectangular wedge. This problem, which we shall call the *single-wedge problem*, has been rigorously solved by Reiche (1912).

Which of these possibilities should be chosen as 'unperturbed problem'? The best choice is, of course, that which corresponds to the smallest perturbation. Clearly, the solution of the single-wedge problem must be closer to that of the wide double-wedge problem than the geometrical optics limit. In fact, while the latter gives only a rough picture of the intensity distribution, the former contains a detailed description of the field, including the correct singularity at the edge. Therefore, we shall take Reiche's problem as 'unperturbed problem'.

Let us find out what happens to system (S) in the transition to the single-wedge problem. For this purpose, let us shift the origin of co-ordinates to the edge of the fixed wedge (figure 2). In the new co-ordinates ($x' = x$, $y' = y - a$), (2·8) becomes

$$u_{\mathrm{I}}(x', y') = (-1)^s \cos(k_{y,s}y')\exp(-\mathrm{i}k_{x,s}x') + \sum_{n=0}^{\infty}(-1)^n a_n \cos(k_{y,n}y')\exp(\mathrm{i}k_{x,n}x'). \qquad (4·1)$$

As K increases, the γ_n spectrum becomes denser and denser, γ_n tending to the continuous variable γ, and the difference $\Delta\gamma = \gamma_n - \gamma_{n-1} = \pi/K$ tending to the differential $d\gamma$, when $K \to \infty$. We may expect that the mode amplitudes, in general, also become infinitesimal, so that

$$a_n = (-1)^{n+s}[a(\gamma_n) + \Delta a_n(K)]\,\Delta\gamma, \quad \text{where} \quad \lim_{K\to\infty}[\Delta a_n(K)] = 0. \tag{4·2}$$

The limiting form of (4·1) should then be

$$u_{\mathrm{I}}(x', y') = A\left\{\cos(ky'\gamma_s)\exp[-ikx'(1-\gamma_s^2)^{\frac{1}{2}}] + \int_0^\infty a(\gamma)\cos(ky'\gamma)\exp[ikx'(1-\gamma^2)^{\frac{1}{2}}]\,d\gamma\right\}, \tag{4·3}$$

where A is an amplitude factor.

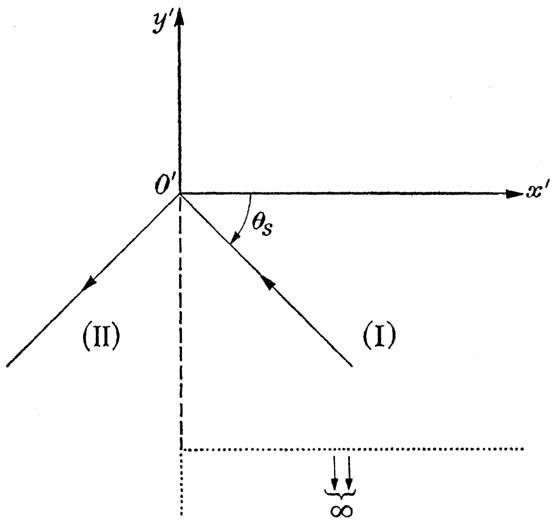

FIGURE 2. Single-wedge co-ordinate system.

On the other hand, according to (3·38) to (3·42), we may write, for $\delta_m \gg 1$, $\delta_n \gg 1$,

$$K_{m,n} = (-1)^{m+n}[K(\gamma_m, \gamma_n) + \Delta K_{m,n}(K)]\,\Delta\gamma, \tag{4·4}$$

where $K(\gamma_m, \gamma_n)$ and $\Delta K_{m,n}(K)$ are given by (3·38) to (3·42), and

$$\Delta K_{m,n}(K) = O(K^{-\frac{1}{2}}) \quad \text{for} \quad K \to \infty. \tag{4·5}$$

Substituting (4·2) and (4·4) in system (S), we get, for $\delta_s \gg 1$, $\delta_m \gg 1$,

$$a(\gamma_m) + \Delta a_m(K) = \sum_{\delta_n \gg 1}[K(\gamma_m, \gamma_n) + \Delta K_{m,n}(K)]\,[a(\gamma_n) + \Delta a_n(K)]\,\Delta\gamma$$
$$+ \sum_{\delta_n \lesssim 1} K_{m,n}[a(\gamma_n) + \Delta a_n(K)] - [K(\gamma_m, \gamma_s) + \Delta K_{m,s}(K)]. \tag{4·6}$$

As K increases, the region where $\delta_n \lesssim 1$ decreases, until, in the limit, it shrinks to the critical point, $\gamma = 1$. Since, according to (3·44), $K_{m,n}$ remains finite in this region, (4·6) shows that, excluding the case of 'exactly critical incidence' $\gamma_s = 1$, the limiting form of system (S) in the transition to the single-wedge problem is, at all points except $\gamma = 1$, the following linear integral equation of the second kind:

$$(R) \quad a(\gamma) = \int_0^\infty K(\gamma, \gamma')\,a(\gamma')\,d\gamma' - K(\gamma, \gamma_s), \tag{4·7}$$

where (see (4·4) and (3·38) to (3·42))

$$K(\gamma,\gamma') = [(1-\gamma'^2)^{\frac{1}{2}}/2\pi]\,(\gamma^2-\gamma'^2)^{-1}\,[\gamma S(\gamma)-\gamma' S(\gamma')] \quad (\gamma \neq \gamma'), \tag{4·8}$$

$$K(\gamma',\gamma') = [(1-\gamma'^2)^{-\frac{1}{2}}/2\pi]\,[(\mathrm{i}/\pi)+S(\gamma')/2\gamma'] = \lim_{\gamma\to\gamma'}[K(\gamma,\gamma')] \tag{4·9}$$

and $S(\gamma)$ is defined in (3·19).

Equation (R) is the integral equation of Reiche's problem in the k-representation. It follows from the above derivation that *the solution of the double-wedge problem must tend almost everywhere to the corresponding solution of the single-wedge problem when* $K\to\infty$ *by the process described above.* The phrase 'almost everywhere' refers to the previously mentioned exceptional cases: the result fails for all values of γ at exactly critical incidence ($\gamma_s = 1$), and it fails at the point $\gamma = 1$ for any other incidence. The critical mode, with $\gamma = 1$, which is reflected perpendicularly between the plates, is the only mode which can, so to speak, 'feel' the presence of another wedge, located at infinity. We shall discuss in § 5(g) what happens at exactly critical incidence.

(b) *Solution of the integral equation*

We shall now derive from Reiche's solution of the single-wedge problem the rigorous solution of equation (R). Taking $A = 1$ in (4·3), we get

$$a(\gamma) = (2k/\pi)\int_{-\infty}^{0}\cos(\gamma k y')\,[u(0-,y')-\cos(\gamma_s k y')]\,\mathrm{d}y'. \tag{4·10}$$

Therefore, to evaluate $a(\gamma)$, we need only the value of Reiche's solution on the half-plane $x' = 0$, $y' < 0$.

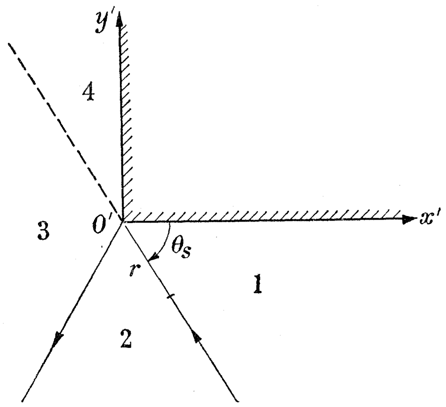

FIGURE 3. Division into regions for definition of geometrical optics field.

In polar co-ordinates $x' = r\cos\theta, y' = -r\sin\theta$ (figure 3), Reiche's solution, adapted to the TM case and to our notation, may be written as follows (Reiche 1912):

$$u(r,\theta) = u_g(r,\theta)+u_d(r,\theta), \tag{4·11}$$

where $u_g(r,\theta)$ is the geometrical optics field, given by (see figure 3)

$$u_g(r,\theta) = \begin{cases} \tfrac{1}{2}\{\exp[-\mathrm{i}kr\cos(\theta_s-\theta)]+\exp[-\mathrm{i}kr\cos(\theta_s+\theta)]\} & \text{(in regions 1 and 2)}, \\ \tfrac{1}{2}\exp[-\mathrm{i}kr\cos(\theta_s-\theta)] & \text{(in region 3)}, \\ 0 & \text{(in region 4)}, \end{cases} \tag{4·12}$$

and $u_d(r,\theta)$ is the diffracted field, given by

$$\begin{aligned}-4\sqrt{3}\,u_d(r,\theta) &= \exp[-\mathrm{i}kr\cos(\theta_s-\theta)]\,\{\mathrm{e}^{\frac{1}{3}\mathrm{i}\pi}\cos[\tfrac{1}{3}(\theta_s-\theta)]\,P_2+\mathrm{e}^{\frac{1}{6}\mathrm{i}\pi}\cos[\tfrac{2}{3}(\theta_s-\theta)]\,P_1\}\\ &\quad+\exp[-\mathrm{i}kr\cos(\theta_s+\theta)]\,\{\mathrm{e}^{\frac{1}{3}\mathrm{i}\pi}\cos[\tfrac{1}{3}(\theta_s+\theta)]\,P_2'+\mathrm{e}^{\frac{1}{6}\mathrm{i}\pi}\cos[\tfrac{2}{3}(\theta_s+\theta)]\,P_1'\}\\ &\quad+\exp[\mathrm{i}kr\cos(\theta_s-\theta)]\,\{-\mathrm{e}^{\frac{1}{3}\mathrm{i}\pi}\cos[\tfrac{1}{3}(\theta_s-\theta)]\,P_2''+\mathrm{e}^{\frac{1}{6}\mathrm{i}\pi}\cos[\tfrac{2}{3}(\theta_s-\theta)]\,P_1''\}\\ &\quad+\exp[\mathrm{i}kr\cos(\theta_s+\theta)]\,\{-\mathrm{e}^{\frac{1}{3}\mathrm{i}\pi}\cos[\tfrac{1}{3}(\theta_s+\theta)]\,P_2'''+\mathrm{e}^{\frac{1}{6}\mathrm{i}\pi}\cos[\tfrac{2}{3}(\theta_s+\theta)]\,P_1'''\},\end{aligned} \qquad (4\cdot13)$$

where

$$P_1=\int_{kr}^{\infty}\exp[\mathrm{i}u\cos(\theta_s-\theta)]\,H_{\frac{1}{3}}(u)\,\mathrm{d}u;\quad P_1'=\int_{kr}^{\infty}\exp[\mathrm{i}u\cos(\theta_s+\theta)]\,H_{\frac{1}{3}}(u)\,\mathrm{d}u;$$

$$P_1''=\int_{kr}^{\infty}\exp[-\mathrm{i}u\cos(\theta_s-\theta)]\,H_{\frac{1}{3}}(u)\,\mathrm{d}u;\quad P_1'''=\int_{kr}^{\infty}\exp[-\mathrm{i}u\cos(\theta_s+\theta)]\,H_{\frac{1}{3}}(u)\,\mathrm{d}u \qquad (4\cdot14)$$

and $P_2, P_2', \ldots$, are obtained from $P_1, P_1', \ldots$, by replacing $H_{\frac{1}{3}}(u)$ by $H_{\frac{2}{3}}(u)$.

Substituting Reiche's solution in (4·10), and making the change of variable $ky' = -v$, we get

$$\begin{aligned}-\pi\sqrt{3}\,a(\gamma) &= \mathrm{e}^{\frac{1}{6}\mathrm{i}\pi}\cos\tfrac{2}{3}\theta_s\int_0^{\infty}\mathrm{d}v\int_v^{\infty}\cos(\gamma v)\cos[\gamma_s(u-v)]\,H_{\frac{1}{3}}(u)\,\mathrm{d}u\\ &\quad-\mathrm{e}^{-\frac{1}{6}\mathrm{i}\pi}\sin\tfrac{1}{3}\theta_s\int_0^{\infty}\mathrm{d}v\int_v^{\infty}\cos(\gamma v)\sin[\gamma_s(u-v)]\,H_{\frac{2}{3}}(u)\,\mathrm{d}u.\end{aligned} \qquad (4\cdot15)$$

By inverting the order of integration, these integrals may be reduced to Fourier transforms of Hankel's functions. Making the substitution

$$\gamma=\sin\theta \quad (0\leqslant\gamma<1);\quad \gamma=\cosh\Theta \quad (1<\gamma<\infty), \qquad (4\cdot16)$$

we get (Erdélyi 1954)

$$a(\gamma) = -\mathrm{i}(\pi\sqrt{3})^{-1}\,[\cos\theta(\sin^2\theta-\sin^2\theta_s)]^{-1}\,[\cos\tfrac{4}{3}\theta_s\cos\tfrac{2}{3}\theta-\cos\tfrac{2}{3}\theta_s\cos\tfrac{4}{3}\theta] \quad (0\leqslant\gamma<1), \qquad (4\cdot17)$$

$$\begin{aligned}a(\gamma) &= -\tfrac{1}{2}(\pi\sqrt{3})^{-1}\,[\sinh\Theta\,(\cosh^2\Theta-\sin^2\theta_s)]^{-1}\,\{[\cos\tfrac{4}{3}\theta_s\cosh\tfrac{2}{3}\Theta+\cos\tfrac{2}{3}\theta_s\cosh\tfrac{4}{3}\Theta]\\ &\quad+\mathrm{i}\sqrt{3}\,[\cos\tfrac{4}{3}\theta_s\sinh\tfrac{2}{3}\Theta-\cos\tfrac{2}{3}\theta_s\sinh\tfrac{4}{3}\Theta]\} \quad (\gamma>1).\end{aligned} \qquad (4\cdot18)$$

Equations (4·17) and (4·18) give *the rigorous solution of the integral equation* (R). Comparing them with (4·11) to (4·14), we see that *a striking simplification is accomplished by employing the k-representation.*

(*c*) *Verification of the solution*

The verification that (4·17) and (4·18) satisfy equation (R), besides giving some insight into the structure of this equation, leads to a set of identities which will play an important role in later developments. However, the complete verification is very tedious, so that we shall not reproduce it here (see A). Its main steps are as follows:

Substituting (4·17) and (4·18) in equation (R), and separating real and imaginary parts, we get a set of relations, which may be expressed in terms of a small number of functions (defined below). These relations may then be reduced to trigonometric and hyperbolic identities, with the help of certain identities satisfied by the functions.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The functions which appear for $0 \leqslant \gamma < 1$ are

$$\mathscr{C}_\lambda(\theta,\theta_s) = \int_0^\infty \frac{\cosh(\lambda\phi)\,\mathrm{d}\phi}{(\cosh^2\phi - \sin^2\theta)\,(\cosh^2\phi - \sin^2\theta_s)},$$

$$\mathscr{S}_\lambda(\theta,\theta_s) = \int_0^\infty \frac{\sinh(\lambda\phi)\,\mathrm{d}\phi}{(\cosh^2\phi - \sin^2\theta)\,(\cosh^2\phi - \sin^2\theta_s)},$$

$$C_\lambda(\theta,\theta_s) = \mathscr{P}\int_0^{\frac{1}{2}\pi} \frac{\cos(\lambda\psi)\,\mathrm{d}\psi}{(\sin^2\psi - \sin^2\theta)\,(\sin^2\psi - \sin^2\theta_s)}, \tag{4·19}$$

where $\mathscr{P}$ denotes Cauchy's principal value. The functions which appear for $\gamma > 1$ are

$$\overline{\mathscr{C}}_\lambda(\Theta,\theta_s) = \mathscr{P}\int_0^\infty \frac{\cosh(\lambda\phi)\,\mathrm{d}\phi}{(\cosh^2\phi - \cosh^2\Theta)\,(\cosh^2\phi - \sin^2\theta_s)},$$

$$\overline{\mathscr{S}}_\lambda(\Theta,\theta_s) = \mathscr{P}\int_0^\infty \frac{\sinh(\lambda\phi)\,\mathrm{d}\phi}{(\cosh^2\phi - \cosh^2\Theta)\,(\cosh^2\phi - \sin^2\theta_s)},$$

$$\overline{C}_\lambda(\Theta,\theta_s) = \mathscr{P}\int_0^{\frac{1}{2}\pi} \frac{\cos(\lambda\psi)\,\mathrm{d}\psi}{(\sin^2\psi - \cosh^2\Theta)\,(\sin^2\psi - \sin^2\theta_s)}. \tag{4·20}$$

All these functions are defined for $-4 < \lambda < +4$. They satisfy a set of identities which we shall call the *R-identities*. These identities will be derived in the Appendix (§7).

(*d*) *Behaviour of the field near the edge*

Reiche's solution may be employed to study the behaviour of the field near the edge. For this purpose, it suffices to make $kr \ll 1$ in equations (4·13) and (4·14). The result (see A) is

$$u_d(r,\theta) = -\tfrac{1}{3} + (\sqrt{3}/2\pi)\,2^{\frac{1}{3}}\,\Gamma(\tfrac{1}{3})\,\mathrm{e}^{-\frac{1}{3}\mathrm{i}\pi}\cos\tfrac{2}{3}\theta_s\cos\tfrac{2}{3}\theta\,(kr)^{\frac{2}{3}}$$
$$+ (3\sqrt{3}/8\pi)\,2^{\frac{2}{3}}\,\Gamma(\tfrac{2}{3})\,\mathrm{e}^{-\frac{2}{3}\mathrm{i}\pi}\cos\tfrac{4}{3}\theta_s\cos\tfrac{4}{3}\theta\,(kr)^{\frac{4}{3}} + \dots \tag{4·21}$$

The components of the diffracted electric field follow from (2·1):

$$\begin{Bmatrix} E_{x,d}(r,\theta) \\ E_{y,d}(r,\theta) \end{Bmatrix} = -(2^{\frac{1}{3}}/\pi\sqrt{3})\,\Gamma(\tfrac{1}{3})\,\mathrm{e}^{\frac{1}{6}\mathrm{i}\pi}\cos\tfrac{2}{3}\theta_s \begin{Bmatrix} \sin\frac{1}{3}\theta \\ \cos\frac{1}{3}\theta \end{Bmatrix} (kr)^{-\frac{1}{3}}$$
$$+ (\sqrt{3}/2\pi)\,2^{\frac{2}{3}}\,\Gamma(\tfrac{2}{3})\,\mathrm{e}^{-\frac{1}{6}\mathrm{i}\pi}\cos\tfrac{4}{3}\theta_s \begin{Bmatrix} \sin\frac{1}{3}\theta \\ \cos\frac{1}{3}\theta \end{Bmatrix} (kr)^{\frac{1}{3}} + \dots \tag{4·22}$$

Equation (4·21) is related to Sommerfeld's representation of the branched wave functions as series of Bessel functions of fractional order (Frank & Von Mises 1935, p. 838). According to (4·21) and (4·22), the magnetic field is finite and continuous at the edge, whereas the electric field tends to infinity as $r^{-\frac{1}{3}}$.

The 'edge condition' (*g*) (§2) means, essentially, that the electric field must have the correct type of singularity at the edges. Therefore, we may now replace it by the following condition:

(*g'*) The electric field is everywhere finite, except at the edges, where it tends to infinity as $r^{-\frac{1}{3}}$.

This condition determines the asymptotic behaviour of the mode amplitudes. In fact, it follows from (2·1) and (4·3) that

$$\gamma a(\gamma) = -(2\mathrm{i}/\pi)\int_0^\infty E_{x,d}(0,-v)\sin(\gamma v)\,\mathrm{d}v, \quad \text{where} \quad v = -ky'. \tag{4·23}$$

The function $E_{x,d}(0, -v)$ is regular, except at the origin. Its behaviour near the origin is given by (4·22):

$$E_{x,d}(0, -v) = A(\theta_s)\, v^{-\frac{1}{3}} + O(v^{\frac{1}{3}}); \quad A(\theta_s) = -2^{\frac{1}{3}}\Gamma(\tfrac{1}{3})\, e^{\frac{1}{6}i\pi}\,(2\pi\sqrt{3})^{-1}\cos(\tfrac{2}{3}\theta_s). \tag{4·24}$$

Applying a general theorem on the asymptotic behaviour of Fourier integrals whose integrands have singularities at the end-points of the interval of integration (Erdélyi 1956, p. 48), it follows from (4·23) and (4·24) that

$$a(\gamma) \approx -(i\sqrt{3}/\pi)\,\Gamma(\tfrac{2}{3})A(\theta_s)\,\gamma^{-\frac{5}{3}}, \tag{4·25}$$

where ≈ means 'asymptotically equal'. Equation (4·25) also follows directly from (4·18). Thus, *the main term in the asymptotic expansion of the mode amplitudes is determined by the singularity at the edge*. Consequently, (g') may be replaced by the following condition, which applies directly to the k-representation:

(g'') The mode amplitudes a_n must behave asymptotically (i.e. for $\gamma_n \gg 1$) like $n^{-\frac{5}{3}}$.

5. The independent wedges approximation

(a) *Definition of the i.w.a.*

It follows from § 4 that, 'outside of the critical region', the solution of the wide double-wedge problem must be of the form

$$a_n = \alpha_n + \Delta\alpha_n, \tag{5·1}$$

where
$$\alpha_n = (-1)^{n+s}\, a(\gamma_n)\,\Delta\gamma \tag{5·2}$$

and
$$\lim_{K\to\infty}(\Delta\alpha_n/\alpha_n) = 0, \tag{5·3}$$

$a(\gamma_n)$ being the value taken by (4·17) and (4·18) for $\gamma = \gamma_n$.

'Outside of the critical region' means: if neither γ_s nor γ_n approaches the critical point. If we restrict ourselves to 'ordinary' (far from critical) incidence, (5·3) shows that, for large enough K, we may neglect $\Delta\alpha_n$ without committing a serious error, except near the critical point. We may expect that the largest error will occur for the value of n which is closest to the critical point, i.e. for $n = c$ or $n = c+1$; we shall assume, for definiteness, that this value is $n = c$. Then, (5·2) may be an unacceptable approximation for $n = c$, since it may lead, as will be seen later (§ 5 (f)), to an inadmissibly large value for the partial reflexion coefficient of this mode. To avoid this inconvenience, we shall employ the cth equation of system (S) to define α_c in terms of the remaining α_n:

$$(1 - K_{c,c})\,\alpha_c = \sum_{n=0}^{\infty}{}' K_{c,n}\,\alpha_n - K_{c,s}, \tag{5·4}$$

where the prime on the summation sign indicates the exclusion of $n = c$.

If we define α_n by (5·2) for $n \neq c$, and by (5·4) for $n = c$, we obtain an approximate solution of system (S), which will be called the *independent wedges approximation* (i.w.a.). The corresponding wave function follows from (2·8), (2·10) and (2·13). This wave function, by construction, satisfies *exactly* conditions (a), (b), (c), (e) and (f) (§ 2), as well as (g'') (§ 5). The remaining condition (d) is approximately satisfied, since the i.w.a. is an approximate solution of system (S).

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL A SOCIETY
PHILOSOPHICAL TRANSACTIONS OF

In so far as the integral of (4·3) may be replaced by the sum of (4·1), the wave function in the i.w.a. approximately corresponds (disregarding the correction for $n = c$) to a properly symmetrized superposition of two independent single-wedge wave functions, each of the wedges being excited only by the incident wave. This is the reason for the name 'independent wedges approximation'. In this sense, the i.w.a. may be likened to the first-order approximation in Schwarzschild's theory of diffraction by a slit (Schwarzschild 1902), for it does not take into account 'multiple' diffraction, i.e. diffraction by each wedge of waves diffracted by the other one. However, there is only an approximate correspondence, since Schwarzschild's first-order approximation would violate conditions (*b*) and (*c*) of §2, whereas the i.w.a. violates only condition (*d*).

(*b*) *The residuals of the i.w.a.*

In the previous section, we have inferred that the i.w.a. must be a 'good approximation' for 'ordinary incidence' and 'large enough' K. We shall now replace these qualitative ideas by more precise criteria.

The *absolute error* $\Delta\alpha_n$ of an approximate solution of system (S) satisfies the following system of linear equations:

$$(\Delta S) \quad \Delta\alpha_m = \sum_{n=0}^{\infty} K_{m,n}\Delta\alpha_n + R_{m,s}, \tag{5·5}$$

where
$$R_{m,s} = \sum_{n=0}^{\infty} K_{m,n}\alpha_n - (\alpha_m + K_{m,s}). \tag{5·6}$$

The system (ΔS) differs from (S) only by the replacement of the inhomogeneous term $-K_{m,s}$ by $R_{m,s}$. The quantities $R_{m,s}$ will be called the *absolute residuals* of the approximate solution α_n.

If the *relative residuals* $\epsilon_{m,s} = |R_{m,s}/K_{m,s}|$ are very small, it may be expected that the *relative error* $|\Delta\alpha_n/\alpha_n|$ will be correspondingly small. This is not always the case for finite systems of equations, but no simple practical criterion for recognizing exceptional cases (known as 'ill-conditioned systems') is available (Hartree 1952, p. 151). In the case of infinite systems, it seems that this problem has never been considered. We shall return to it later on.

We shall now try to estimate the order of magnitude of the residuals of the i.w.a. If we make

$$K_{m,n} = (-1)^{m+n}[K(\gamma_m, \gamma_n) + \Delta K_{m,n}]\Delta\gamma \quad (m, n, \neq c), \tag{5·7}$$

it follows from (5·2), (5·4), (5·6) and (4·7) that

$$R_{m,s} = (1-\delta_{m,c})\,[R_{m,s}(\delta) + R_{m,s}(\Delta K) + R_{m,s}(\alpha_c)], \tag{5·8}$$

where
$$R_{m,s}(\delta) = (-1)^{m+s}\Delta\gamma\left[\sum_{n=0}^{\infty}{}' K(\gamma_m, \gamma_n)\, a(\gamma_n)\,\Delta\gamma - \int_0^{\infty} K(\gamma_m, \gamma)\, a(\gamma)\,\mathrm{d}\gamma\right], \tag{5·9}$$

$$R_{m,s}(\Delta K) = (-1)^{m+s}\Delta\gamma\left[\sum_{n=0}^{\infty}{}' \Delta K_{m,n}\, a(\gamma_n)\,\Delta\gamma - \Delta K_{m,s}\right], \tag{5·10}$$

$$R_{m,s}(\alpha_c) = K_{m,c}\alpha_c, \tag{5·11}$$

the prime on the summation signs having the same meaning as in (5·4). Thus, three different sources contribute to the residuals: the difference (5·9) between integrals and sums of Reiche's solutions, the difference (5·10) between the kernel of equation (R) and the coefficients of system (S), and the 'α_c correction' (5·11).

The order of magnitude of (5·9) may be estimated with the help of the following numerical integration formula (Hennequin 1949):

$$\sum_{n=1}^{j} f(\gamma_n)\,\Delta\gamma - \int_{\gamma_i}^{\gamma_j} f(\gamma)\,\mathrm{d}\gamma = \tfrac{1}{2}\Delta\gamma\,[f(\gamma_i)+f(\gamma_j)] + \tfrac{1}{12}(\Delta\gamma)^2 \sum_{n=i+1}^{j} f''(\gamma'_n)\,\Delta\gamma \quad (\gamma_{n-1}<\gamma'_n<\gamma_n), \tag{5·12}$$

the function $f(\gamma)$ being differentiable up to the second order in the interval (γ_i, γ_j). If $f(\gamma_j)$ tends to zero for $j\to\infty$ and if the sum which appears in (5·12) converges, (5·12) may also be applied for $j\to\infty$. The order of magnitude of (5·10) and (5·11) may also be estimated, with the help of the expressions derived in § 3. After a lengthy calculation, which will not be reproduced here (see A), the following results are obtained:

The order of magnitude of α_c is given by

$$\alpha_c \sim (\Delta\gamma)^{\frac{1}{2}}/\cos\theta_s. \tag{5·13}$$

The order of magnitude of the absolute residuals is given by

$$R_{m,s} \sim [\cos\tfrac{1}{3}\theta_s/\cos\theta_s]\,(1-\gamma_m^2)^{-1}\,(\Delta\gamma)^{\frac{3}{2}} \quad (m \lesssim M), \tag{5·14}$$

$$R_{m,s} \sim [\cos\tfrac{1}{3}\theta_s/\cos\theta_s]\,\gamma_m^{-2}\ln\gamma_m(\Delta\gamma)^2 \quad (m \gtrsim M), \tag{5·15}$$

where M is defined by

$$\ln\gamma_M \sim (\Delta\gamma)^{-\frac{1}{2}}, \quad \text{so that} \quad \gamma_M \sim \exp(\surd K). \tag{5·16}$$

Dividing the absolute value of (5·14) and (5·15) by $|K_{m,s}|$, we obtain the order of magnitude of the relative residuals:

$$\epsilon_{m,s} \sim [\cos\tfrac{1}{3}\theta_s/\cos\theta_s]\,(\Delta\gamma)^{\frac{1}{2}} \quad \text{for} \quad \gamma_m \text{ not} \gg 1 \quad \text{and} \quad |1-\gamma_m^2|^{\frac{1}{2}} \text{ not} \ll 1 \tag{5·17}$$

$$\epsilon_{m,s} \sim 1 \quad \text{for} \quad \delta_m \lesssim 1, \tag{5·18}$$

$$\epsilon_{m,s} \sim [\cos\tfrac{1}{3}\theta_s/\cos^2\theta_s]\,\Delta\gamma \quad \text{for} \quad m \gtrsim M. \tag{5·19}$$

In intermediate regions, the order of magnitude of $\epsilon_{m,s}$ varies smoothly between the values given in (5·17) to (5·19).

It follows from (5·17) to (5·19) that: (*a*) All relative residuals become of the order of unity if and only if

$$\cos\tfrac{1}{3}\theta_s/(\surd K\cos\theta_s) \sim 1. \tag{5·20}$$

This happens when the incident wave belongs to the region $\delta_s \lesssim 1$, which will be called the *critical incidence region*. (*b*) For *ordinary incidence* (i.e. outside of the critical incidence region), the relative residuals are of the order of unity only in the *critical strip* $\delta_m \lesssim 1$. (*c*) Outside of the critical strip, all relative residuals are much smaller than unity, and tend to zero for $K\to\infty$, in which limit the critical incidence region and the critical strip are reduced to the points $\gamma_s = 1$, $\gamma = 1$.

Assuming that (S) is not an 'ill-conditioned system', we may infer that: (*a*) *The i.w.a. fails completely within the critical incidence region.* (*b*) *For ordinary incidence, the i.w.a. is a very good approximation for large K, except within the critical strip. For $K\to\infty$, the i.w.a. tends almost everywhere to the rigorous solution.* This agrees exactly with our previous conclusions.

It follows from Reiche's solution that the amplitude of the diffracted wave emanating from one wedge, near the edge of the other wedge, is of the order of magnitude of the first member of (5·20), whereas the amplitude of the incident wave is of the order of unity. Therefore, the condition for the validity of the i.w.a. is that the 'secondary' excitation due to multiple diffraction be negligible with respect to the 'primary' excitation, which arises from the incident wave. In this sense, the i.w.a. may be called a 'weak coupling' approximation, for it corresponds to a weak interaction between the wedges, whereas the case of critical incidence corresponds to 'strong coupling'.

(c) *The i.w.a. in region II*

According to (2·10) and (3·1),

$$u_{\mathrm{II}}(x,y) = \frac{1}{2}\int_{-\infty}^{+\infty} A(\gamma)\exp\{ik[\gamma y-(1-\gamma^2)^{\frac{1}{2}}x]\}\,d\gamma, \tag{5·21}$$

where $A(\gamma) = kA(k_y)$. It follows from (2·13) that

$$A(\gamma) = A^K(\gamma)+\bar{A}(\gamma), \tag{5·22}$$

where
$$A^K(\gamma) = (-1)^s\,(2/\pi)\,(1-\gamma_s^2/1-\gamma^2)^{\frac{1}{2}}\,(\gamma^2-\gamma_s^2)^{-1}\,\gamma\sin(\gamma K) \tag{5·23}$$

and
$$\bar{A}(\gamma) = -(2/\pi)\,(1-\gamma^2)^{-\frac{1}{2}}\,\gamma\sin(\gamma K)\sum_{n=0}^{\infty}(-1)^n\,(1-\gamma_n^2)^{\frac{1}{2}}\,(\gamma^2-\gamma_n^2)^{-1}\,a_n. \tag{5·24}$$

$A^K(\gamma)$ is the value taken by $A(\gamma)$ when all mode amplitudes are neglected. It corresponds, therefore, to *Kirchhoff's approximation*. $\bar{A}(\gamma)$ represents the correction to Kirchhoff's approximation.

If γ coincides with one of the γ_n, the series which appears in (5·24) diverges, but $\bar{A}(\gamma)$ remains finite, due to the factor $\gamma\sin(\gamma K)$. If γ_m is the closest neighbour to γ, it is convenient to separate from the sum the term $n = m$. The remaining series may be approximated by an integral (a Cauchy principal value). Replacing a_n by the i.w.a., this integral may be expressed in terms of the functions defined in (4·19) and (4·20), and may then be evaluated with the help of the R-identities, given in the Appendix. The effect of the difference between the sum and the integral may be estimated with the help of (5·12). The final result is (see A):

$$\pi\cos\theta\bar{A}(\gamma=\sin\theta) \simeq (-1)^s\,(\sin^2\theta-\sin^2\theta_s)^{-1}\{\sin(K\sin\theta)\,[\cos\tfrac{4}{3}\theta_s\sin\tfrac{2}{3}\theta+\cos\tfrac{2}{3}\theta_s\sin\tfrac{4}{3}\theta$$
$$-2\sin\theta\cos\theta_s]+(i/\surd 3)\cos(K\sin\theta)\,[\cos\tfrac{4}{3}\theta_s\cos\tfrac{2}{3}\theta-\cos\tfrac{2}{3}\theta_s\cos\tfrac{4}{3}\theta]\}, \tag{5·25}$$

for $\gamma < 1$. The expression for $\gamma = \cosh\Theta > 1$ is the analytic continuation of (5·25).

Comparing (5·25) with (5·23), it is easily seen that, in the neighbourhood of $\theta = \theta_s$, the correction to Kirchhoff's approximation is of the order of $(K\cos^2\theta_s)^{-1}$, which, for ordinary incidence, is very small. Therefore, *in the neighbourhood of* $\theta = \theta_s$, *Kirchhoff's approximation is the main term of the solution* (*for ordinary incidence*). It has been shown in a previous paper (Nussenzveig 1959) that this is a necessary condition for the validity of classical diffraction theory.

(d) *Classical optical patterns of the double wedge*

According to (5·21) to (5·23), the wave function in region II in Kirchhoff's approximation is given by

$$u^K_{\mathrm{II}}(x,y) = u^K_{\mathrm{II}}(x,y,\gamma_s)+u^K_{\mathrm{II}}(x,y,-\gamma_s), \tag{5·26}$$

where
$$u^K_{\mathrm{II}}(x,y,\gamma_s) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(\frac{1-\gamma_s^2}{1-\gamma^2}\right)^{\frac{1}{2}}\frac{\sin[(\gamma-\gamma_s)K]}{\gamma-\gamma_s}\exp\{ik[\gamma y-(1-\gamma^2)^{\frac{1}{2}}x]\}\,d\gamma. \tag{5·27}$$

Equation (5·26) corresponds to the decomposition of the incident mode in a couple of plane waves (§ 2): $u_0 = u_0(\gamma_s) + u_0(-\gamma_s)$, where $u_0(\gamma_s) = \frac{1}{2}\exp\{ik[\gamma_s y - (1-\gamma_s^2)^{\frac{1}{2}} x]\}$. The optical patterns of the double wedge in Kirchhoff's approximation are identical to those of an infinite slit of width $2a$ excited by this couple of plane waves. We may restrict ourselves to one of the components, e.g. (5·27), and to the region $y \geqslant 0$, since the wave function is symmetric with respect to $0x$.

The *geometrical optics pattern* corresponds to the limiting case $k \to \infty$; we shall assume that θ_s is unchanged in the limit (§ 4 (*a*)). It follows that

$$\lim_{k\to\infty}\left\{\frac{\sin[(\gamma-\gamma_s)K]}{\gamma-\gamma_s}\right\} = \pi\delta(\gamma-\gamma_s), \tag{5·28}$$

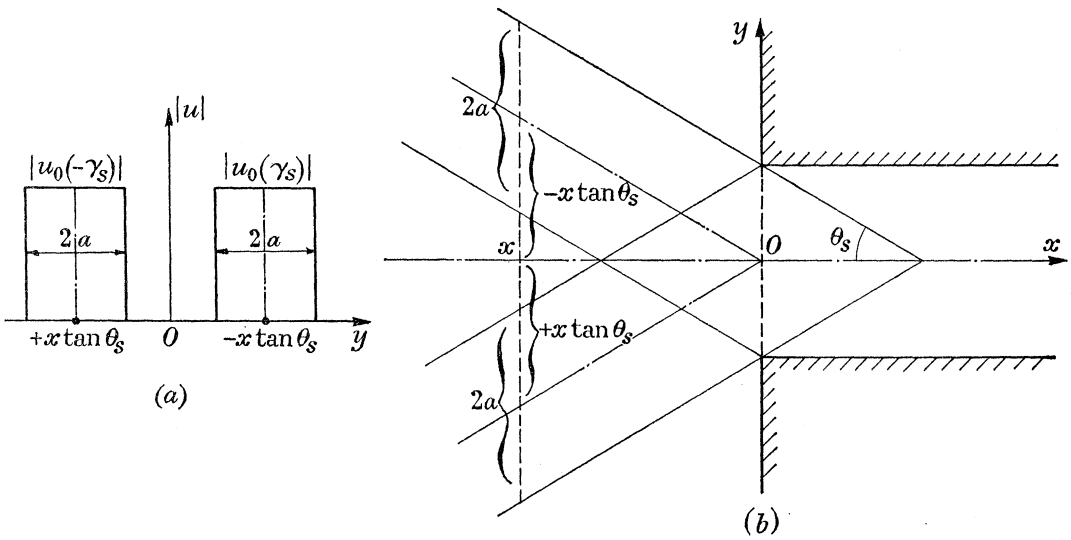

FIGURE 4. (*a*) Distribution in y for given x; (*b*) geometrical optics pattern.

where $\delta(x)$ is Dirac's delta function. Since k appears also in the exponent of the exponential in (5·27), we cannot simply replace γ by γ_s; we must take into account the variation of the exponential in the neighbourhood of $\gamma = \gamma_s$. Expanding the exponent in powers of $\gamma-\gamma_s$, we find

$$\lim_{k\to\infty} u_{\mathrm{II}}^K(x,y,\gamma_s) = \frac{1}{\pi} u_0(\gamma_s) \int_{-\infty}^{+\infty} \exp[\mathrm{i}t(y+x\,tg\theta_s)]\sin(at)\,\mathrm{d}t/t. \tag{5·29}$$

Introducing the function

$$R(y,y_0,a) = \begin{Bmatrix} 1 & \text{for } y_0-a<y<y_0+a \\ 0 & \text{for other values of } y \end{Bmatrix} = \frac{1}{\pi}\int_{-\infty}^{+\infty}\exp[\mathrm{i}(y-y_0)\,t]\sin(at)\,\mathrm{d}t/t, \tag{5·30}$$

we finally get

$$\lim_{k\to\infty} u_{\mathrm{II}}^K(x,y) = u_0(\gamma_s)\,R(y,-xtg\theta_s,a) + u_0(-\gamma_s)\,R(y,xtg\theta_s,a), \tag{5·31}$$

which is precisely equivalent to the geometrical optics pattern (figure 4). The important part of the domain of integration is the interval $|\gamma-\gamma_s| = \pi/K$. If the phase change of the exponential in this interval is less than 2π, contributions from different values of γ interfere constructively, so that, in the limit, we get the incident wave; if the phase change is greater than 2π, destructive interference sets in, and the integral vanishes in the limit.

The *Fraunhofer diffraction pattern* may be obtained by applying to (5·27) the general rule for the evaluation of the Fraunhofer pattern in the k-representation (Nussenzveig 1959). Adapted to the two-dimensional case of (5·21), this rule becomes

$$u_{\text{II}}(R,\theta) \approx \tfrac{1}{2}\pi\cos\theta\, A(\bar{\gamma})\, H_0(kR) \quad (R\to\infty), \tag{5·32}$$

where R, θ are polar co-ordinates ($x = -R\cos\theta$; $y = R\sin\theta$) and $\bar{\gamma}$ is the *stationary phase point*, given, in this case, by $\bar{\gamma} = \sin\theta$. Taking into account (5·23), we get

$$u^K_{\text{II}}(R,\theta,\gamma_s) \approx \tfrac{1}{2}ka\cos\theta_s\, H_0(kR)\,(\sin X)/X, \tag{5·33}$$

where
$$X = K(\sin\theta - \sin\theta_s). \tag{5·34}$$

This gives the well-known Fraunhofer diffraction pattern of a slit (Sommerfeld 1954).

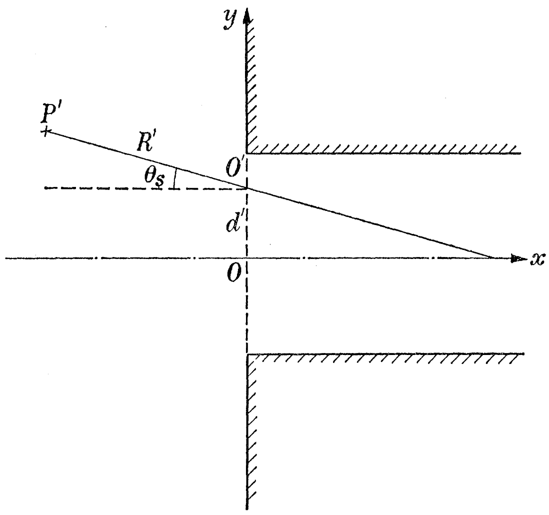

FIGURE 5. Co-ordinate system for evaluation of Fresnel pattern.

To find the *Fresnel diffraction pattern*, it is convenient to shift the origin of co-ordinates to the 'pole' O' of the incident wave with respect to the observation point (figure 5). Then, if $R' = \overline{O'P'}$, $d' = \overline{OO'}$, the phase of the exponential in (5·27) becomes

$$k[\gamma y - (1-\gamma^2)^{\frac{1}{2}}x] = kd'\gamma + kR'[\gamma\sin\theta_s + (1-\gamma^2)^{\frac{1}{2}}\cos\theta_s]. \tag{5·35}$$

In the Fresnel region ($kR' \gg 1, R' \gg d'$), the first term in the second member of (5·35) gives rise to a slowly varying factor, so that the stationary phase method (Nussenzveig 1959) may be applied just to the second term. Expanding the phase, up to the second order, about the stationary phase point $\bar{\gamma} = \sin\theta_s$, we get

$$u^K_{\text{II}}(x,y,\gamma_s) = (2\pi)^{-\frac{1}{2}}u_0(\gamma_s)\int_{w_1}^{w_2}\mathrm{d}v\int_{-\infty}^{+\infty}\exp[-\mathrm{i}(t^2+\surd(2\pi)\,vt]\,\mathrm{d}t = \tfrac{1}{2}(1-\mathrm{i})\,u_0(\gamma_s)[F(w_2)-F(w_1)], \tag{5·36}$$

where
$$\begin{Bmatrix} w_2 \\ w_1 \end{Bmatrix} = (k/\pi R')^{\frac{1}{2}}\cos\theta_s(\pm a - d') \tag{5·37}$$

and $F(w)$ is the Fresnel integral (3·27). Equation (5·36) gives the classical Fresnel diffraction pattern of a slit (Sommerfeld 1954).

It has been shown (Nussenzveig 1959) that all classical optical patterns depend only on a very small spectral region, centred on the direction of incidence. The above analysis allows us to determine the width of this region in the present case. Thus, in the geometrical optics pattern, we have seen that the width is given by $|\gamma-\gamma_s| \lesssim 1/K$. The same is true for the Fraunhofer pattern, since, according to (5·33), the intensity is appreciable only for $|X| \lesssim 1$. To determine the width for the Fresnel pattern, it suffices to study the contribution of a given interval, centred on γ_s, to the integral (5·27), with the help of the stationary phase method (Erdélyi 1956). The width is found to be given by

$$|\gamma-\gamma_s| \lesssim (kR')^{-\frac{1}{2}}\cos\theta_s, \tag{5·38}$$

which is very small in the Fresnel region ($R' \gg a$).

(e) The optical patterns in the i.w.a.

According to (5·23), (5·25) and (5·32), the wave function in the Fraunhofer region, in the i.w.a., is given by

$$u_{\text{II}}(R,\theta) \approx \tfrac{1}{2}(-1)^s(\sin^2\theta-\sin^2\theta_s)^{-1}\{\sin(K\sin\theta)\,[\cos\tfrac{4}{3}\theta_s\sin\tfrac{2}{3}\theta+\cos\tfrac{2}{3}\theta_s\sin\tfrac{4}{3}\theta] + (\mathrm{i}/\surd 3)\cos(K\sin\theta)\,[\cos\tfrac{4}{3}\theta_s\cos\tfrac{2}{3}\theta-\cos\tfrac{2}{3}\theta_s\cos\tfrac{4}{3}\theta]\}\,H_0(kR). \tag{5·39}$$

This result has a simple physical interpretation. It may be shown (see A) that (5·39) is equivalent to a superposition of two single-wedge wave functions (taken in the Fraunhofer region). The upper (lower) wedge diffracts the component of the incident wave which propagates in the direction $\theta_s(-\theta_s)$.

The mean energy current per unit length in the z-direction within the element of angle $\mathrm{d}\theta$, in the direction θ, is given by

$$\sigma(\theta)\,\mathrm{d}\theta = (c/8\pi)\,|u_{\text{II}}(R,\theta)|^2\,R\,\mathrm{d}\theta. \tag{5·40}$$

The quantity $\sigma(\theta)$ represents the intensity distribution in the Fraunhofer pattern (angular distribution). It contains one term due to the θ_s component of the incident wave, one term due to the $-\theta_s$ component, and an interference term. It suffices to discuss the first of these terms, which, according to (5·39) and (5·40), is given by

$$\sigma(\theta,\theta_s) = (cK^2/16\pi^2k)\,\{[(\cos\tfrac{4}{3}\theta_s\sin\tfrac{2}{3}\theta+\cos\tfrac{2}{3}\theta_s\sin\tfrac{4}{3}\theta)/2\sin\theta]^2\,(\sin^2X)/X^2 + [(\cos\tfrac{4}{3}\theta_s\cos\tfrac{2}{3}\theta-\cos\tfrac{2}{3}\theta_s\cos\tfrac{4}{3}\theta)/2\sin\theta]^2\,(\cos^2X)/(3X^2)\}, \tag{5·41}$$

where X has been defined in (5·34). Most of the intensity is concentrated in the region $|X| \lesssim 1$, where (5·41) may be expanded in powers of $\gamma-\gamma_s$:

$$\sigma(\theta,\theta_s) = (cK^2/16\pi^2k)\,\{\cos^2\theta_s\,(\sin^2X)/X^2-(2/K)\,(\sin^2\tfrac{1}{3}\theta_s/\sin\theta_s)\,(\sin^2X)/X + [(\cos\tfrac{4}{3}\theta_s\sin\tfrac{2}{3}\theta_s-2\cos\tfrac{2}{3}\theta_s\sin\tfrac{4}{3}\theta_s)/\sin\theta_s\cos\theta_s]^2\,(\cos^2X)/(27K^2)+\dots\}. \tag{5·42}$$

The first term represents Kirchhoff's approximation; the remaining terms give the corrections to classical diffraction theory. Let us introduce

$$w = (\surd K\cos\theta_s)^{-1}. \tag{5·43}$$

The main correction, given by the second term of (5·42), which is of the order of Xw^2 times smaller than Kirchhoff's approximation, is an odd function of X, which destroys the symmetry of the pattern. The other correction (third term of (5·42)) contains an additional factor of the order of Xw^2, and implies that the minima of the diffraction pattern are not zeros of the intensity. This conclusion also follows directly from (5·24). Both effects are very small for ordinary incidence.

Without explicitly evaluating the Fresnel diffraction pattern in the i.w.a., we can see that it differs very little from the classical pattern, for ordinary incidence. This follows directly from the physical interpretation of the i.w.a. To estimate the order of magnitude of the corrections, it suffices to study the coefficients $A(\gamma)$ in the region defined by (5·38). The main correction to the coefficients is of the order of $(kR)^{-\frac{1}{2}}/\cos\theta_s$, which is $\ll w$ in the Fresnel region.

Thus, *within the domain of validity of the i.w.a.* (*i.e. for ordinary incidence*), *the corrections to classical diffraction theory are very small.* Large corrections may appear only for large angles of diffraction (where the intensity is very weak), in the immediate neighbourhood of the aperture, or in the case of critical incidence.

There are two important factors which contribute to explain the success of classical diffraction theory. The first is the fact that the diffraction patterns depend only on a very small region of the spectrum (Nussenzveig 1959). The second factor is the closeness to geometrical optics in the short-wavelength region. As was shown in the preceding section, the geometrical optics pattern is determined by the same small region of the spectrum as the diffraction patterns. This explains why, under these conditions, Kirchhoff's approximation should be the main term of the solution in this spectral region. The second factor, by itself, does not suffice to explain the success of classical diffraction theory (§ 1). The reason why singularities near sharp edges do not affect the classical patterns is that they do not contribute appreciably to the Fourier coefficients, in the relevant region of the spectrum: they contribute mainly to the asymptotic region.

(*f*) *The reflexion coefficient*

The *reflexion coefficient* of the double wedge is defined as the ratio of the mean reflected energy current to the mean incident energy current (per unit length in the z-direction). It is given by

$$|R_s|^2 = \sum_{n=0}^{c} |r_{n,s}|^2, \tag{5·44}$$

where

$$|r_{n,s}|^2 = (\zeta_n \cos\theta_n/\zeta_s \cos\theta_s)\,|a_n^s|^2, \tag{5·45}$$

which may be called the partial reflexion coefficient of the nth mode when excited by the sth mode. The superscript s has been attached to the mode amplitudes a_n to indicate explicitly their dependence on the incident mode. Notice that evanescent modes do not contribute to (5·44).

It follows from (4·17), (5·2) and (5·45) that, for $n \neq c$, in the i.w.a.,

$$|r_{n,s}|^2 = (3K^2\zeta_n\zeta_s\cos\theta_n\cos\theta_s)^{-1}\left\{\frac{\sin\frac{1}{3}(\theta_n-\theta_s)}{\sin(\theta_n-\theta_s)}+\frac{\sin\frac{1}{3}(\theta_n+\theta_s)}{\sin(\theta_n+\theta_s)}\right\}^2 = |r_{s,n}|^2, \tag{5·46}$$

whereas, for $n = c$, according to (5·13),

$$|r_{c,s}|^2 \sim \cos\theta_c/(K\cos^3\theta_s). \tag{5·47}$$

If we had applied (5·46) for $n = c$, the partial reflexion coefficient would diverge for $\theta_c \to \frac{1}{2}\pi$, whereas (5·47) tends to zero in this case, and is at most of the same order as the partial reflexion coefficients of the neighbouring modes. This explains why the correction (5·4) was necessary.

Replacing the sum (5·44) by an integral, the reflexion coefficient in the i.w.a. may be evaluated, with the following result (see A):

$$|R_s|^2 \cong (2\pi\sqrt{3}\,\zeta_s K\cos^2\theta_s)^{-1}\left\{\cos\tfrac{1}{3}\theta_s - \left[\tfrac{1}{3}\cos\theta_s + \frac{\sin\frac{2}{3}\theta_s}{2\sin\theta_s}\right]\ln\left|\frac{4\cos^3\frac{1}{3}\theta_s}{\cos\theta_s}\right| + \frac{\sqrt{3}}{2}\frac{\cos\frac{2}{3}\theta_s}{\sin\theta_s}\ln\left|\frac{\cos\frac{1}{3}(\theta_s-\pi)}{\cos\frac{1}{3}(\theta_s+\pi)}\right|\right\}. \quad (5·48)$$

This may be replaced, with a mean error of less than 10%, by

$$|R_s|^2 \cong 0{\cdot}1\cos\tfrac{1}{3}\theta_s/(\zeta_s K\cos^2\theta_s). \quad (5·49)$$

The quantity $\zeta_s K\,|R_s|^2$ is plotted in figure 6 as a function of θ_s. According to (5·43) and (5·49), $|R_s|^2$ is of the order of w^2. Therefore, within the domain of validity of the i.w.a., the reflexion coefficient is very small. In fact, *the corrections to classical diffraction theory are precisely of the order of the reflexion coefficient.*

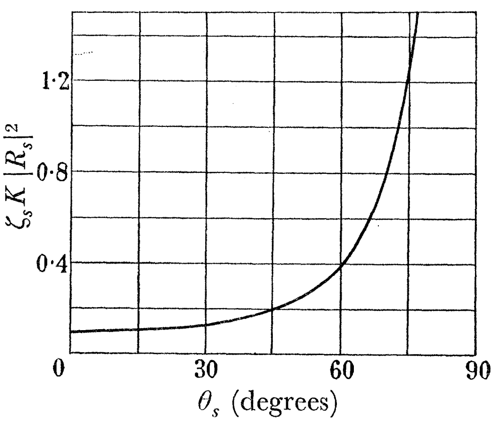

FIGURE 6. Reflexion coefficient of wide double wedge as a function of the angle of incidence.

The origin of reflexion is the distortion of the incident mode in the neighbourhood of the waveguide mouth. This occurs mainly within a distance of a few wavelengths from the edges. In fact, if t represents the order of magnitude of this distance, a fraction t/a of the incident current will be reflected, so that the reflexion coefficient will be of the order of $t/(a\cos\theta_s)$. Comparing this with (5·49), we see that $t\sim\lambda$. The additional factor $\cos\frac{1}{3}\theta_s/\cos\theta_s$ which appears in (5·49) must be related to the distortion of the incident wave by multiple diffraction. This discussion illustrates some effects that occur within a few wavelengths from a border, a region hitherto nearly unexplored in optics.

(*g*) *The case of critical incidence*

The reflexion coefficient increases monotonically with the angle of incidence (figure 6), for ordinary incidence. By extrapolation, this suggests the existence of strong reflexion in the critical incidence region. It will now be shown that *strong reflexion occurs in the neighbourhood of all critical frequencies.* If $\delta_c \ll 1$ (see (3·21)), the incident mode c is almost totally reflected, no matter what is the value of K. This includes, in particular, the narrow double wedge ($K \ll 1$), which will be studied in part II. Here we shall restrict ourselves to $K \gtrsim 1$.

According to (3·43) to (3·45), the coefficients $K_{m,n}$ are of the form:

$$K_{m,c} = K^{(1)}_{m,c}\,\delta_c + O(\delta_c^3)\ (m \neq c); \quad K_{c,n} = K^{(1)}_{c,n} + O(\delta_c^2) \quad (n \neq c); \quad K_{c,c} = \tfrac{1}{2} + K^{(1)}_{c,c}\,\delta_c + O(\delta_c^3). \tag{5·50}$$

In the limiting case of *exactly critical incidence* ($\delta_c = 0$), system (S) becomes

$$(S)\left\{\begin{array}{lll} (S') & a_m = \sum_{n=0}^{\infty}{}' K_{m,n}\,a_n & (m \neq c), \\ (C) & \tfrac{1}{2}a_c = \sum_{n=0}^{\infty}{}' K^{(1)}_{c,n}\,a_n - \tfrac{1}{2} & (m = c), \end{array}\right\} \tag{5·51}$$

where the prime on the summation sign means that $n = c$ is to be excluded. Thus, (S) is *separated* into a homogeneous system (S'), which does not contain a_c, and an equation (C). The *rigorous solution* of system (S) then is

$$a_m = 0 \quad (m \neq c); \quad a_c = -1. \tag{5·52}$$

This corresponds to *total reflexion of the incident mode*. According to (2·8), (2·10) and (2·13), the corresponding wave function identically vanishes everywhere. This means that there cannot be propagation in the waveguide at exactly critical incidence.

If $\delta_c \ll 1$, (5·52) suggests the following 'Ansatz' for the solution (to first order):

$$a_m = a^{(1)}_m\,\delta_c \quad (m \neq c); \quad a_c = -1 + a^{(1)}_c\,\delta_c. \tag{5·53}$$

Replacing in (S) and neglecting higher order terms, we get

$$(S)\left\{\begin{array}{lll} (S') & a^{(1)}_m = \sum_{n=0}^{\infty}{}' K_{m,n}\,a^{(1)}_n - 2K^{(1)}_{m,c} & (m \neq c), \\ (C) & a^{(1)}_c = 2\sum_{n=0}^{\infty}{}' K^{(1)}_{c,n}\,a^{(1)}_n - 4K^{(1)}_{c,c} & (m = c), \end{array}\right\} \tag{5·54}$$

so that (S) still separates into (S') and (C). Replacing (5·53) in (2·8), we find

$$u_{\mathrm{I}}(x,y) = -2\mathrm{i}\cos(k_{y,c}\,y)\sin(k_{x,c}\,x) + O(\delta_c). \tag{5·55}$$

Thus, interference between the incident and reflected modes gives rise to *quasi-stationary waves* within the waveguide. The terms which contain the 'perturbation parameter' δ_c may be called 'radiative corrections'.

For $\delta_c \sim 1$, the separation of (S) may no longer be effected. We may consider δ_c as a measure of the coupling between the critical mode and the other modes. For $\delta_c \ll 1$, the coupling is weak, and we may speak of 'radiative corrections'. This is the physical meaning of the separation of system (S). For $\delta_c \sim 1$, we have strong coupling, and this description becomes inadequate. This 'coupling between modes' is entirely different from the 'coupling between wedges' discussed in section *c*; for instance, in the neighbourhood of exactly critical incidence, the coupling is weak in the former sense, while it is strong in the latter.

It is usually assumed that a diffraction problem contains only two important limiting cases: the short-wavelength limit and the long-wavelength limit. We see now that *there exists a third important limiting case: the case of critical incidence*. Herein lies perhaps the essential difference between problems in which no linear dimension of the diffracting object is finite and problems in which the diffracting object has at least one finite dimension. This is the reason for the difference between the single-wedge problem and the limit of the double-wedge problem.

The importance of the third limiting case is that it provides a link between the theory of diffraction and the theory of quasi-stationary currents. The existence of strong reflexion in the neighbourhood of critical frequencies was established by Vajnshtejn (1954) for parallel-plate and circular semi-infinite waveguides, as well as for semi-infinite two-wire or co-axial lines. The importance of this result for linking the theory of diffraction with the theory of quasi-stationary currents seems, however, not yet to have received due attention.

To obtain quantitative results in the region $\delta_c \ll 1$, we must find the solution of system (S') in (5·54). This seems to be very difficult in the most general case. However, as we have already mentioned, the narrow double-wedge problem is a particular case of critical incidence. The solution of this problem, which will be discussed in part II, may be taken as an illustration of the effects which appear in this region.

6. The iteration method

(a) *Neumann's iteration method*

The problem which will be investigated in this section is related to question C of § 1: can Kirchhoff's approximation be considered as the first step of an accurate solution by successive approximations? Since Kirchhoff's approximation is a 'geometrical optics' approximation, we shall now take as 'unperturbed problem' the geometrical optics limit, and try to develop a method of successive approximations starting from this limit. This corresponds to the alternative approach suggested in § 4*a*.

It follows from (3·12) to (3·14) that

$$\lim_{k\to\infty} K_{m,s} = 0, \tag{6·1}$$

for finite m as well as for $m \to \infty$, except for $\gamma_m \to 1$, $\gamma_s \to 1$ (it is easily seen that this is an exceptional case, for $K_{c,c} = \frac{1}{2}$ for $\gamma_c = 1$, no matter what is the value of k). Therefore, except in the case of exactly critical incidence, system (S) becomes homogeneous when $k \to \infty$, so that Kirchhoff's approximation, $a_m = 0$, is valid almost everywhere in the geometrical optics limit.

According to (6·1), $K_{m,n}$ must be very small for $K \gg 1$, unless m and n belong to the critical strip. Ignoring the exceptions, for a moment, let us examine what would happen if we had

$$K_{m,n} = \epsilon \overline{K}_{m,n}, \quad \text{with } \overline{K}_{m,n} \text{ independent of } K \quad \text{and} \quad \epsilon = \epsilon(K) = O(K^{-\alpha}) \quad (\alpha \geqslant \delta > 0). \tag{6·2}$$

If this were the case, (S) would be nearly diagonal for $K \gg 1$. This suggests that we might apply *Neumann's iteration method*. Starting from Kirchhoff's approximation as zero-order approximation (N_0),

$$a_m^{N_0} = 0, \tag{6·3}$$

Neumann's first-order approximation (N_1) would be obtained by replacing (6·3) in (S):

$$a_m^{N_1} = -K_{m,s} \tag{6·4}$$

and higher-order approximations would be obtained by iteration:

$$a_m^{Nr} = -K_{m,s} + \sum_{p=0}^{\infty} K_{m,p} a_p^{Nr-1} = -K_{m,s} - (K^2)_{m,s} - \ldots - (K^r)_{m,s}, \tag{6·5}$$

where
$$(K^r)_{m,s} = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \ldots \sum_{v=0}^{\infty} K_{m,p} K_{p,q} \ldots K_{v,s} \quad (r \text{ factors}). \tag{6·6}$$

According to (6·2), this would lead to a Neumann series

$$a_m^{N\infty} = -\epsilon \overline{K}_{m,s} - \epsilon^2 \sum_{n=0}^{\infty} \overline{K}_{m,n} \overline{K}_{n,s} - \ldots \tag{6·7}$$

which might be expected to converge to the rigorous solution for small enough ϵ (large enough K).

Actually, (6·2) does not represent the coefficients $K_{m,n}$. According to § 3, the matrix $\|K_{m,n}\|$, for $K \gg 1$, is of the form

$$\|K_{m,n}\| = \left\| \begin{array}{c|c|c|c|c} m \downarrow \; n \rightarrow & N & C & N & A \\ N & \left(\frac{1}{K}\right) & \left(\frac{1}{K}\right) & \left(\frac{1}{K}\right) & \left(\frac{\ln \gamma_n}{K\gamma_n}\right) \\ C & \left(\frac{1}{\sqrt{K}}\right) & (1) & \left(\frac{1}{\sqrt{K}}\right) & \left(\frac{(\sqrt{K}) + (\ln \gamma_n)}{K\gamma_n}\right) \\ N & \left(\frac{1}{K}\right) & \left(\frac{1}{K}\right) & \left(\frac{1}{K}\right) & \left(\frac{\ln \gamma_n}{K\gamma_n}\right) \\ A & \left(\frac{\ln \gamma_m}{K\gamma_m^2}\right) & \left(\frac{(\sqrt{K}) + (\ln \gamma_m)}{K^{\frac{3}{2}}\gamma_m^2}\right) & \left(\frac{\ln \gamma_m}{K\gamma_m^2}\right) & \tilde{K}_{m,n} \end{array} \right\|. \tag{6·8}$$

The bracketed expressions give the order of magnitude of $K_{m,n}$ in different regions of the matrix. The order of magnitude changes continuously across the transition regions, indicated by dotted lines. The labelling of the rows and columns corresponds to a subdivision of the spectrum into three main regions, namely: *normal region* (N), for $\delta_n^2 \sim K$, and either $\gamma_n < 1$ or $\gamma_n > 1$, but not $\gamma_n \gg 1$; *critical strip* (C), for $\delta_n^2 \sim 1$; *asymptotic region* (A), for $\gamma_n \gg 1$ (the same applies to γ_m).

In the AA region of the matrix, according to § 3,

$$K_{m,n} \cong \tilde{K}_{m,n} = \frac{(-1)^{m+n}}{\pi^2} \frac{n \ln (m/n)}{(m^2 - n^2)} \quad (\gamma_m \gg 1, \gamma_n \gg 1). \tag{6·9}$$

Thus, in the asymptotic region of the matrix, the main term of $K_{m,n}$ does not depend on K. This is a very significant result; its physical interpretation will be given in part II.

There are two 'anomalous' regions in the matrix (6·8), where $K_{m,n}$ becomes independent of K, in disagreement with (6·2): the region CC, where $K_{m,n} \sim 1$, and the region AA. How does this affect the convergence of Neumann's method? It may be verified, with the help of (6·8), that, for ordinary incidence, the 'anomalous' regions give contributions of the same order to all iterations. Therefore, if Neumann's method converges in the critical strip and in the asymptotic region, it must converge very slowly. Even in the normal region, the convergence cannot be very rapid, for successive iterations give contributions of the same order.

(b) Neumann's first- and second-order approximations

For ordinary incidence, Neumann's method may be tested by comparing its results with the i.w.a. For this purpose, we shall compute the first- and second-order approximations.

Neumann's first-order approximation (N_1) is given by (6·4). Its value follows from the expressions given in §3.

According to (6·5) and (6·6), Neumann's second-order approximation (N_2) is given by

$$a_m^{N_2} = -K_{m,s} - \sum_{n=0}^{\infty} K_{m,n} K_{n,s}. \tag{6·10}$$

The series which appears in (6·10) may be evaluated by replacing it by an integral. The result may be expressed in terms of the functions defined in (4·19) and (4·20) and their first- and second-order derivatives with respect to λ. With the help of the R-identities, the following results are obtained (see A):

$$a_m^{N_2} = a_m^{N_1}\{1+\tfrac{1}{6}-(\theta_m^2+\theta_s^2)/6\pi^2+(\theta_m\theta_s/3\pi^2)\,[(\theta_m\tan\theta_s-\theta_s\tan\theta_m)/(\theta_m\tan\theta_m-\theta_s\tan\theta_s)]\} + O(K^{-\frac{3}{2}}), \quad \text{for} \quad \gamma_m = \sin\theta_m < 1, \tag{6·11}$$

$$\mathscr{R}(a_m^{N_2}) = \mathscr{R}(a_m^{N_1})\left[1+\tfrac{1}{8}+\frac{1}{2\pi^2}(\phi_m^2-\theta_s^2)+\frac{1}{\pi^2}\phi_m\theta_s\tanh\phi_m\tan\theta_s\right]+O(K^{-\frac{3}{2}});$$

$$\mathscr{I}(a_m^{N_2}) = \mathscr{I}(a_m^{N_1})\left\{1+\tfrac{1}{24}+\frac{1}{6\pi^2}(\phi_m^2-\theta_s^2)+\frac{\phi_m\theta_s}{3\pi^2}[(\phi_m\tan\theta_s-\theta_s\coth\phi_m)/(\phi_m\coth\phi_m+\theta_s\tan\theta_s)]\right\} + O(K^{-\frac{3}{2}}), \quad \text{for} \quad \gamma_m = \cosh\phi_m > 1. \tag{6·12}$$

These results may be applied when γ_m does not belong to the critical strip. Higher-order Neumann approximations may also be evaluated (up to terms of the order of $K^{-\frac{3}{2}}$) by a similar procedure.

Now let us compare N_1 and N_2 with the i.w.a. The following functions are plotted in figure 7: $(-1)^{m+s}\,\mathrm{i}K\zeta_m a_m(\gamma_m<1)$, for $\theta_s=0°$ and $\theta_s=87°$; for $\gamma_m>1$ and $\theta_s=0°$: $(-1)^{m+s+1}K\mathscr{R}(a_m)$ (curves I) and $(-1)^{m+s}K\mathscr{I}(a_m)$ (curves II). The i.w.a. is represented by full lines, N_1 by dash-and-dot lines, and N_2 by dashed lines. For other values of θ_s, similar curves are obtained.

The N_1 and N_2 curves always lie below the i.w.a. curves. The relative error $\epsilon_m^{N_1}(\epsilon_m^{N_2})$ of $N_1(N_2)$ with respect to the i.w.a. will be defined as the difference between corresponding results in the i.w.a. and in $N_1(N_2)$, divided by the i.w.a. result. Let us study its behaviour as a function of γ_m.

(1) *Normal region.* For $\gamma_m<1$, $\epsilon_m^{N_1}<20\%$; its mean value is $\sim 15\%$; it decreases for increasing γ_m or θ_s. For $\theta_s=0°$, N_1 may be obtained from the i.w.a. by the substitution: $\sin\tfrac{1}{3}\theta_m \rightarrow \sqrt{3}\,\theta_m/2\pi \cong \theta_m/3{\cdot}63$. For $\gamma_m>1$, let $\epsilon_m^{N_1}(\mathscr{R})$ and $\epsilon_m^{N_1}(\mathscr{I})$ be the relative errors of the real and imaginary part of $a_m^{N_1}$, respectively. Both errors increase with γ_m. For $1<\gamma_m<2$, $\epsilon_m^{N_1}(\mathscr{R})<20\ \%$ and $\epsilon_m^{N_1}(\mathscr{I})<10\ \%$. For $\gamma_m>2$, both errors become large.

Figure 7 shows that N_2 is considerably better than N_1. For $\gamma_m<1$, $\epsilon_m^{N_2}<5\ \%$; its mean value is $\sim 3\%$. For $\theta_s=0°$, N_2 may be obtained from the i.w.a. by the substitution:

$$\sin\tfrac{1}{3}\theta_m \rightarrow \frac{1}{3{\cdot}11}\theta_m - \frac{1}{215}\theta_m^3.$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The power series expansion of $\sin\frac{1}{3}\theta_m$ is:

$$\sin\tfrac{1}{3}\theta_m = \tfrac{1}{3}\theta_m - \tfrac{1}{162}\theta_m^3 + \dots.$$

Thus, Neumann's method is somewhat analogous to a power series development, but the distribution of the error is more uniform. For $1 < \gamma_m < 2$, $\epsilon_m^{N_2}(\mathscr{R}) \lesssim 5\%$ and $\epsilon_m^{N_2}(\mathscr{I}) \lesssim 1\%$. In contrast to N_1, N_2 remains a reasonable approximation even for $2 < \gamma_m < 10$, where $\epsilon_m^{N_2}(\mathscr{R}) \lesssim 15\%$ and $\epsilon_m^{N_2}(\mathscr{I}) \lesssim 3\%$.

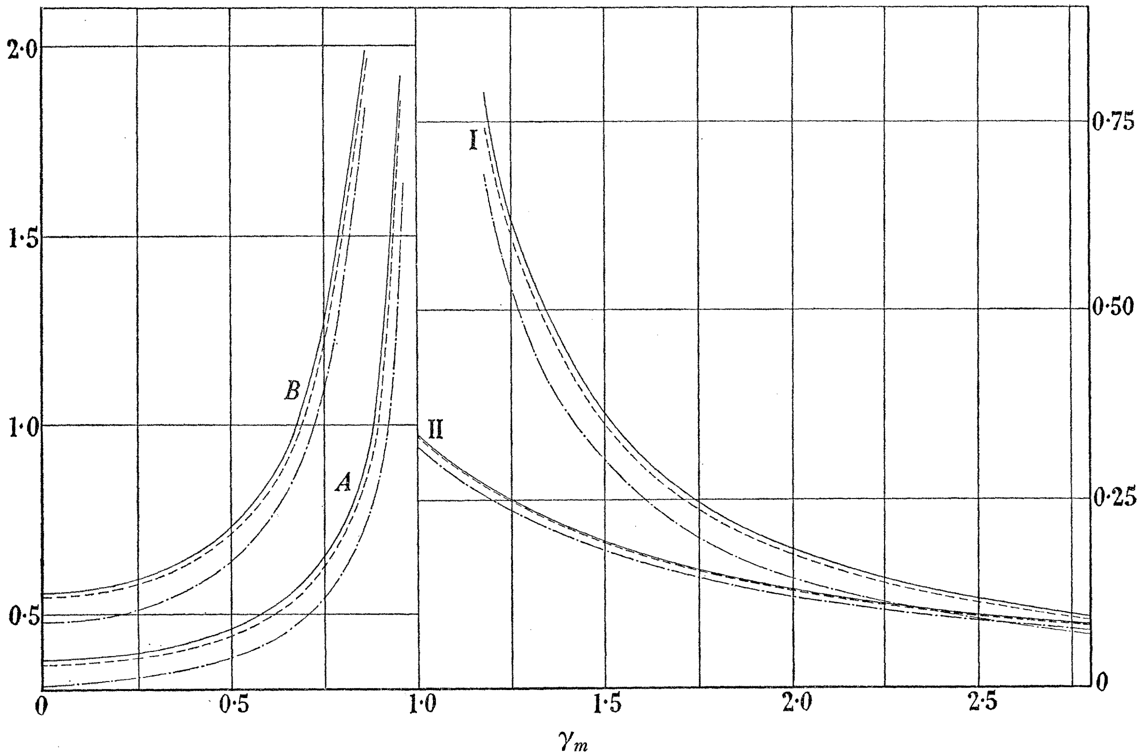

FIGURE 7. Comparison between the i.w.a. and Neumann's first- and second-order approximations; ——, i.w.a.; —·—·—, N_1; - - - -, N_2.

(2) *Critical strip.* Though the phase of N_1 is quite different from that of the i.w.a. in this region, the absolute value is of the same order of magnitude, even for $m = c$.

(3) *Asymptotic region.* In this region,

$$\alpha_m \approx (-1)^{m+s+1}(2^{\frac{1}{3}}/\surd 3\,K)\exp(-\tfrac{1}{3}\mathrm{i}\pi)\cos\tfrac{2}{3}\theta_s\,\gamma_m^{-\frac{5}{3}} \quad \text{(i.w.a.)}, \tag{6·13}$$

$$a_m^{N_1} \approx (-1)^{m+s}(\mathrm{i}/\pi K)\cos\theta_s\,\gamma_m^{-2}\ln\gamma_m. \tag{6·14}$$

Therefore, the asymptotic behaviour of N_1 is entirely different from that of the i.w.a. It follows that N_1 fails completely in the neighbourhood of the edges. N_2 has a larger domain of applicability. For instance, $\epsilon_m^{N_2}(\mathscr{I}) \lesssim 15\%$ for $\gamma_m < 10^2$. The asymptotic behaviour of N_2 recalls the expansion

$$\gamma_m^{-\frac{5}{3}} = \gamma_m^{-2}\exp(\tfrac{1}{3}\ln\gamma_m) = \gamma_m^{-2}[1 + \tfrac{1}{3}\ln\gamma_m + \tfrac{1}{18}(\ln\gamma_m)^2 + \dots]. \tag{6·15}$$

This supports the above-mentioned analogy with a power series development.

The following inferences are rendered *plausible* by these results: (*a*) The iteration method possibly converges for ordinary incidence; (*b*) If so, the convergence should be moderately rapid in the normal region, slowing down in the critical strip, and becoming increasingly slower in the asymptotic region.

From the practical point of view of *numerical convergence, the iteration method may be very useful in the normal region* (where N_2 is a good approximation), *but this is not the case in the asymptotic region.*

So far, the iteration method has been applied only for ordinary incidence. Let us consider now what happens in the extreme case of *exactly critical incidence*. It follows from (5·50) that $K_{m,c} = \frac{1}{2}\delta_{m,c}$ in this case, so that, according to (6·7),

$$a_m^{N\infty} = -(\tfrac{1}{2}+\tfrac{1}{4}+\tfrac{1}{8}+\ldots)\,\delta_{m,c} = -\delta_{m,c}. \tag{6·16}$$

Comparing (6·16) with (5·52), we see that, *for exactly critical incidence, Neumann's series converges to the rigorous solution.* This is a rather surprising result, for this is a case of 'strong coupling' (in the sense of § 5*b*). Neumann's zero-order approximation is a 'geometrical optics' approximation, and geometrical optics breaks down in this case. Notice, however, that (6·16) converges very slowly for $m = c$.

In view of the above, one would expect that Neumann's method still converges for $\delta_c \ll 1$. However, the convergence should be very slow, so that we shall not apply the method in this region, for $K \gtrsim 1$. An illustration of the behaviour of the iteration method in the neighbourhood of exactly critical incidence will be given in part II, where we shall apply it to the case $K \ll 1$.

7. Appendix. The *R*-identities

To derive the identities which are satisfied by the functions defined in (4·19), let us consider the complex integral

$$\int_\Gamma f(z,\beta,\theta,\theta_s)\,\mathrm{d}z = \int_\Gamma \frac{z^\beta\,\mathrm{d}z}{(z-\mathrm{e}^{2\mathrm{i}\theta})\,(z-\mathrm{e}^{-2\mathrm{i}\theta})\,(z-\mathrm{e}^{2\mathrm{i}\theta_s})\,(z-\mathrm{e}^{-2\mathrm{i}\theta_s})}, \tag{7·1}$$

taken over the contour Γ shown in figure 8. To render the integrand single valued, we take a branch cut along the negative part of the real axis, and define

$$z^\beta = \exp\left[\beta(\log|z| + \mathrm{i}\arg z)\right] \quad (-\pi \leqslant \arg z \leqslant \pi). \tag{7·2}$$

The contour Γ consists of two segments $\overline{AB}$ and $\overline{DE}$, on opposite sides of the branch cut, joined by a circle C_0 of radius δ ($\delta \to 0$), centred at the origin, and by a circle C_1, with radius unity and centre at the origin, indented by four half-circles of equal radii ϵ ($\epsilon \to 0$), centred on the four poles of the integrand. The integrand is analytic within Γ, so that

$$\int_\Gamma f\,\mathrm{d}z = 0. \tag{7·3}$$

It is easily seen that

$$\int_{\overline{AB}} f\,\mathrm{d}z + \int_{\overline{ED}} f\,\mathrm{d}z = -2\mathrm{i}\sin(\pi\beta)\int_0^1 \frac{x^\beta\,\mathrm{d}x}{(x+\mathrm{e}^{2\mathrm{i}\theta})\,(x+\mathrm{e}^{-2\mathrm{i}\theta})\,(x+\mathrm{e}^{2\mathrm{i}\theta_s})\,(x+\mathrm{e}^{-2\mathrm{i}\theta_s})}. \tag{7·4}$$

It may also be shown that

$$\lim_{\delta\to 0}\int_{C_0} f\,\mathrm{d}z = 0, \quad \text{for} \quad \beta+1>0. \tag{7·5}$$

The contribution of the four half-circles, for $\epsilon \to 0$, is

$$\lim_{\epsilon \to 0} \int_{4C} f \, \mathrm{d}z = \mathrm{i}\pi \Sigma \operatorname{res} f(z, \beta, \theta, \theta_s), \tag{7·6}$$

where the residues are to be taken at the four poles of the integrand. Finally, if $C_1 - 4C$ denotes the circle C_1, less the arcs cut off by the indentations,

$$\lim_{\epsilon \to 0} \int_{C_1 - 4C} f \, \mathrm{d}z = -\frac{\mathrm{i}}{8} \mathscr{P} \int_{-\frac{1}{2}\pi}^{+\frac{1}{2}\pi} \frac{\exp\left[2\mathrm{i}(\beta - 1)\,\psi\right] \mathrm{d}\psi}{(\sin^2\psi - \sin^2\theta)\,(\sin^2\psi - \sin^2\theta_s)}. \tag{7·7}$$

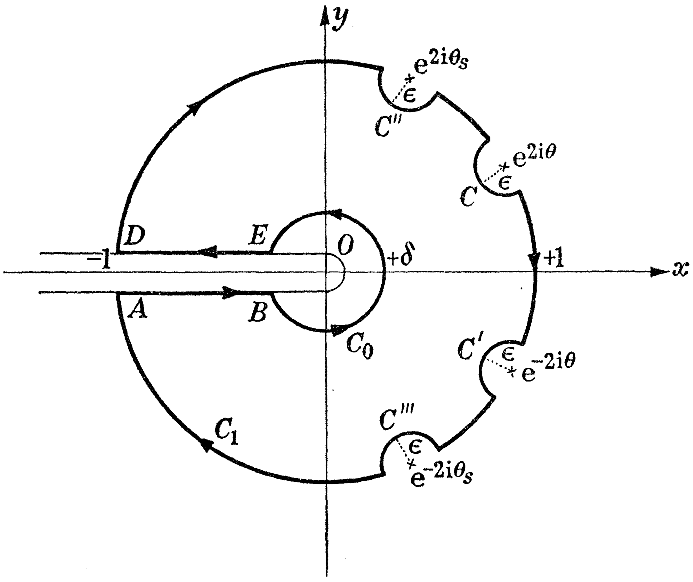

FIGURE 8. Contour of integration for equation (7·1).

If we substitute (7·4) to (7·7) in (7·3), making the substitution $x = \exp(-2\phi)$ in (7·4), and evaluating the sum over residues in (7·6), we find, taking $\beta = \frac{1}{2}\lambda + 1$,

$$\int_0^\infty \frac{\exp(-\lambda\phi)\,\mathrm{d}\phi}{(\cosh^2\phi - \sin^2\theta)\,(\cosh^2\phi - \sin^2\theta_s)} = \pi \operatorname{cosec} \tfrac{1}{2}\lambda\pi\,(\sin^2\theta - \sin^2\theta_s)^{-1} \left[\frac{\sin\lambda\theta}{\sin 2\theta} - \frac{\sin\lambda\theta_s}{\sin 2\theta_s}\right]$$
$$+ \tfrac{1}{2} \operatorname{cosec} \tfrac{1}{2}\lambda\pi\, \mathscr{P} \int_{-\frac{1}{2}\pi}^{+\frac{1}{2}\pi} \frac{\exp(\mathrm{i}\lambda\psi)\,\mathrm{d}\psi}{(\sin^2\psi - \sin^2\theta)\,(\sin^2\psi - \sin^2\theta_s)}. \tag{7·8}$$

Respectively by adding to (7·8) and by subtracting from (7·8) the same equation, with λ replaced by $-\lambda$, we get the identities

$$\mathscr{C}_\lambda(\theta, \theta_s) = \pi \operatorname{cosec} \tfrac{1}{2}\lambda\pi\,(\sin^2\theta - \sin^2\theta_s)^{-1} \left[\frac{\sin(\lambda\theta)}{\sin(2\theta)} - \frac{\sin(\lambda\theta_s)}{\sin(2\theta_s)}\right], \tag{7·9}$$

$$C_\lambda(\theta, \theta_s) = -\sin \tfrac{1}{2}\lambda\pi\, \mathscr{S}_\lambda(\theta, \theta_s), \tag{7·10}$$

which, according to (7·5), are valid for $-4 < \lambda < +4$.

To derive the identities satisfied by the functions defined in (4·20), we consider the integral

$$\int_{\Gamma'} g(z, \beta, \Theta, \theta_s)\,\mathrm{d}z = \int_{\Gamma'} \frac{z^\beta\,\mathrm{d}z}{(z + \mathrm{e}^{2\Theta})\,(z + \mathrm{e}^{-2\Theta})\,(z - \mathrm{e}^{2\mathrm{i}\theta_s})\,(z - \mathrm{e}^{-2\mathrm{i}\theta_s})}. \tag{7·11}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The contour Γ', shown in figure 9, differs from Γ by having two poles less on C_1, and one additional pole on the real axis, at $x = -e^{-2\Theta}$, surrounded by half-circles C and C' of radius ϵ, on each side of the branch cut.

Employing a method similar to that adopted in the previous case, we get, in the place of (7·8),

$$\mathscr{P}\int_0^\infty \frac{\exp(-\lambda\phi)\,\mathrm{d}\phi}{(\cosh^2\phi-\cosh^2\Theta)\,(\cosh^2\phi-\sin^2\theta_s)}$$

$$= -\pi\,\mathrm{cosec}\,\tfrac{1}{2}\lambda\pi\,(\cosh^2\Theta-\sin^2\theta_s)^{-1}\frac{\sin\lambda\theta_s}{\sin 2\theta_s}+\pi\cot\tfrac{1}{2}\lambda\pi\,(\cosh^2\Theta-\sin^2\theta_s)^{-1}$$

$$\times\frac{\exp(-\lambda\Theta)}{\sinh 2\Theta}+\tfrac{1}{2}\,\mathrm{cosec}\,\tfrac{1}{2}\lambda\pi\,\mathscr{P}\int_{-\frac{1}{2}\pi}^{+\frac{1}{2}\pi}\frac{\exp(\mathrm{i}\lambda\psi)\,\mathrm{d}\psi}{(\sin^2\psi-\cosh^2\Theta)\,(\sin^2\psi-\sin^2\theta_s)}. \quad (7·12)$$

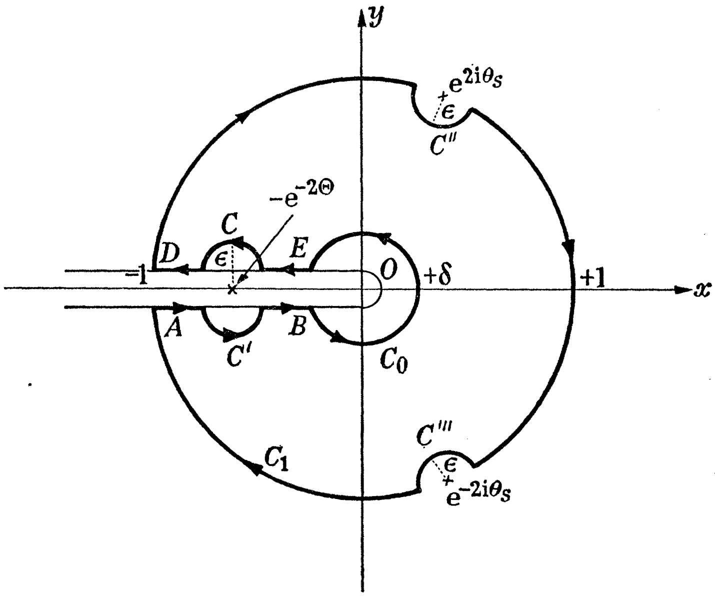

FIGURE 9. Contour of integration for equation (7·11).

Adding or subtracting the same equation, with λ replaced by $-\lambda$, we get

$$\mathscr{C}_\lambda(\Theta,\theta_s) = -\pi(\cosh^2\Theta-\sin^2\theta_s)^{-1}\left[\cot\tfrac{1}{2}\lambda\pi\frac{\sinh\lambda\Theta}{\sinh 2\Theta}+\mathrm{cosec}\,\tfrac{1}{2}\lambda\pi\frac{\sin\lambda\theta_s}{\sin 2\theta_s}\right], \quad (7·13)$$

$$\overline{C}_\lambda(\Theta,\theta_s) = -\sin\tfrac{1}{2}\lambda\pi\,\overline{\mathscr{S}}_\lambda(\Theta,\theta_s)-\pi\cos\tfrac{1}{2}\lambda\pi\,(\cosh^2\Theta-\sin^2\theta_s)^{-1}\frac{\cosh\lambda\Theta}{\sinh 2\Theta}. \quad (7·14)$$

These identities are valid for $-4<\lambda<+4$.

Equations (7·9), (7·10), (7·13) and (7·14) are the *R*-identities.

REFERENCES

Born, M. 1933 *Optik*. Berlin: Springer.
Bouwkamp, C. J. 1954 *Rep. Progr. Phys.* **17**, 35.
Erdélyi, A. 1954 *Tables of integral transforms*, vol. 1. New York: McGraw-Hill.
Erdélyi, A. 1956 *Asymptotic expansions*. New York: Dover Publications.
Frank, P. & Von Mises, R. 1935 *Differentialgleichungen der Physik*, vol. 2, 2nd ed. Braunschweig: Vieweg.

Franz, W. 1949 *Z. Phys.* **125**, 563.
Hartree, D. R. 1952 *Numerical analysis*. Oxford University Press.
Hennequin, A. 1949 *Calcul numérique et applications pratiques*. Paris: Les Cours de Sorbonne.
Nussenzveig, H. M. 1957 *Notas de Física*, Suppl. vol. 3 (quoted as A).
Nussenzveig, H. M. 1959 To appear in *An. Acad. Brasil. Cienc.*
Reiche, F. 1912 *Ann. Phys., Lpz.*, **37**, 131.
Schelkunoff, S. A. 1951 *Commun. Pure Appl. Math.* **4**, 43.
Schwarzschild, K. 1902 *Math. Ann.* **55**, 177.
Sommerfeld, A. 1896 *Math. Ann.* **47**, 317.
Sommerfeld, A. 1954 *Optics*. New York: Academic Press.
Toraldo di Francia, G. 1956 *Electromagnetic waves*. New York: Interscience.
Vajnshtejn, L. A. 1954 *Propagation in semi-infinite waveguides* (six papers translated from the Russian). New York University, Report no. EM-63.
Watson, G. N. 1948 *Theory of Bessel functions*, 2nd ed. Cambridge University Press.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES