DEVELOPMENTS IN OUR CONCEPTS OF DIFFRACTION PHENOMENA (ON THE 130TH ANNIVERSARY OF THE DEATH OF THOMAS YOUNG)*

G. D. MALYUZHINETS

Usp. Fiz. Nauk **69**, 321-334 (October, 1959)

1. INTRODUCTION

THE simplest laws of propagation of electromagnetic, acoustic or other waves in homogeneous or non-homogeneous media are obtained in the region of extremely small wavelengths λ, that is, for the case described by geometrical optics. The wave fields of geometrical optics are localized planes in the sense that over small regions they obey the same laws as plane waves propagating in homogeneous media in well defined directions. We will refer to these directions as rays. In the case of isotropic media the aggregate of the rays represents a family of lines orthogonal to the family of the propagating wave fronts.

On the basis of the ray concept, we can consider the geometric-optical field as decomposed into arbitrarily small tubular sections (referred to as ray tubes), the propagation within each of which is entirely independent of the rest of the field. This leads to a phenomenon characteristic of geometrical optics. Thus, if an obstruction is placed along the path of propagation such that some of the rays terminate on it while others pass by, a sharp boundary line is formed between the bright and dark regions.

However, in case of a small but finite wavelengths λ the sharp boundary between the bright and dark regions becomes blurred. This represents the ordinary phenomenon of diffraction. It seems as if a splitting of the edge rays takes place and therefore the field penetrates partially into the region of geometrical shadow. The word diffraction (from the Latin diffringere — to divide, split) related to this apparent splitting of rays was first proposed by Grimaldi (1665) to describe the broadening he observed of a narrow beam of sunlight passing through a small aperture and the smearing out of a shadow of a rod illuminated by a narrow light beam. Therefore, diffraction in the narrow sense of the word is used to describe these relatively small deviations in the propagation of waves from the laws of geometrical optics when the wave field or each of its individual components closely resemble locally-plane waves, although new directions of propagation may appear which cannot be explained by geometrical optics.

The term diffraction in the broad sense is used to describe arbitrary deviations from the laws of geometrical optics. This includes almost all wave phenomena associated with propagation in regions of arbitrary shape.

2. HISTORICAL DEVELOPMENT OF THE DIFFRACTION CONCEPT

The first wave treatment of the diffraction phenomenon was given by Young (1800) and the second by Fresnel (1815). The characteristic features of these treatments are easily followed in the simple example of diffraction by a straight edge.

Young* considers the wave field behind an obstruction to be the result of two phenomena, diffraction proper and interference. To explain the diffraction phenomenon Young introduced, in addition to the geometric-optical principle of propagation of locally-plane waves in the direction of rays, the notion of transverse transmission of the oscillation amplitudes directly along the wave fronts. He has shown that the transmission velocity, i.e., the amplitude flux, is proportional to the wavelength and to the amplitude gradient along the wave front. Figure 1, taken from Young's article, illustrates how the diffracted wave is formed behind the obstruction owing to the ray amplitude transmission along the cylindrical wave fronts from the boundary of the geometrical shadow. The formation of the diffracted waves is thus, according to Young, of local character since it takes place in the vicinity of the boundary of the shadow behind the edge of the obstruction. A similar diffracted wave is also formed in the bright region (see Fig. 2). Cylindrical wave fields are thus present, which appear to be emitted from the edge of the aperture. Interference between the diffracted wave and the portion of the incident wave not blocked by the obstruction explains the presence of interference fringes on

*This article is a revision of the article "Wave Diffraction" written for the new edition of the Physics Dictionary.

*Thomas Young, A Course of Lectures on Natural Philosophy and Mechanical Arts, Vol. I–II, London, 1807.

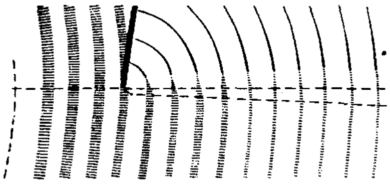

FIG. 1

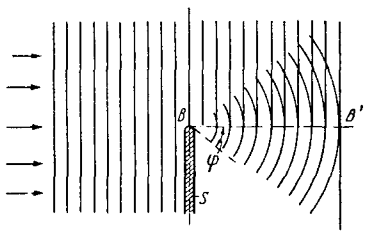

FIG. 2

the screen B′ above the boundary BB′ of the geometrical shadow and the absence of fringes below the boundary.

Unable to explain with Young's method the fact that near the boundary line of the geometrical shadow ($|\varphi| \ll 1$)* sharp and rounded edges lead to the same diffraction pattern on a screen, Fresnel (1815) rejected the treatment of diffraction as a local phenomenon that takes place in the vicinity of the aperture edge. Using the concepts embodied in Huygens' principle, he visualized the diffraction phenomenon as a consequence of interference of an infinitely large number of waves originating from virtual sources distributed over the entire aperture surface (see Fig. 3). To visualize graphically the magnitude of the diffraction field at an arbitrary point A behind an aperture large compared with the wavelength λ, Fresnel subdivided the aperture plane into circular half-wave zones. The zones, starting with the first one, can be constructed with a compass. This is shown in Fig. 4 for the example of a plane wave incident on a circular aperture of radius a and for a point A located on the aperture axis. In the example shown, exactly five Fresnel zones fit in the aperture. The effects of the Huygens sources distributed in the adjacent zones 2, 3 and 4, 5 essentially cancel each other, because of the opposite phases of the oscillations that reach A from these zones. The resultant field at the point A is therefore determined primarily by the Huygens' sources of the first zone. When $OA = a^2/\lambda$, there is only one Fresnel zone on the aperture. Beyond this distance a beam of light, sound, or of any other radiation spreads (in width) relatively rapidly because of diffraction. The region of

*The region of large diffraction angles φ, where differences are found in the diffraction properties of sharp and rounded screen edges, was not investigated at that time.

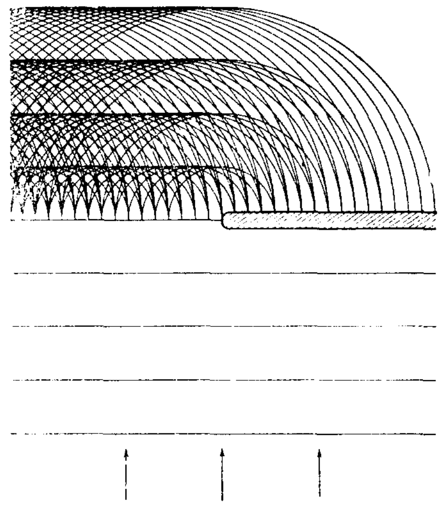

FIG. 3

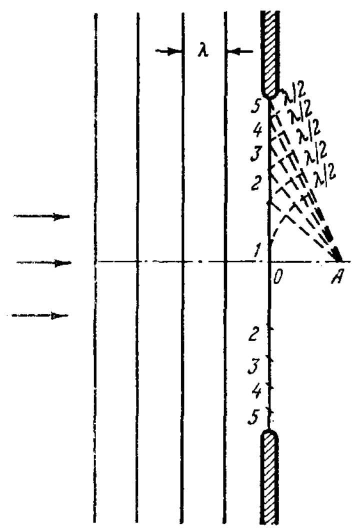

FIG. 4

space for which the distances from the aperture are much greater than a^2/λ is referred to as the Fraunhofer zone. In this region the diffracted wave regains its directivity and obeys essentially the laws of geometrical optics.

Fresnel's notions of diffraction, which he developed mathematically, soon became predominant and brought about the final victory of the wave theory of light over the Newtonian corpuscular theory. Although Fresnel's method is the more formal one, insofar as it introduces sources far removed from an edge in order to describe diffraction by the edge, Young's local approach, which is more representative of the physical nature of the diffraction phenomenon, was considered incorrect after Fresnel's time. Later Magie (1888) and Rubinowicz (1917, 1924) have shown that the results obtained by Fresnel's methods can be reduced by means of a mathematical transformation to the same form as predicted by Young.

In the more exact formulation of Helmholtz (1859) and Kirchhoff (1882), Huygens' principle leads to an integral equation that relates the value of the wave field at an arbitrary point of a given region to the value of the field and its derivatives at the boundary of the region. Huygens' principle is thus used only to formulate the mathematical problem as an integral equation. Generally this integral equation cannot be solved. It is, as a rule, possible to get an idea of the diffraction field and determine its values on the basis of Huygens' principle only if the aforementioned boundary conditions (i.e., the values of the field and its derivatives at the boundary) are known from other considerations. According to Kirchhoff's approximate solution the values of the field and its derivatives in the aperture plane are assumed to be the same as if the screen had no influence on the incident wave. Thus, in the shadow directly behind the screen, these values are assumed to be zero. This makes it possible to find approximate relations for the diffraction field of apertures of arbitrary form for both the near zone (Fresnel diffraction) and the far zone (Fraunhofer diffraction). A comparison with the exact solution of diffraction by a semi-infinite plane, obtained by Sommerfeld (1896), shows that although the approximate solution for the Fresnel diffraction leads to smaller values of the diffraction field in the shadow region for large diffraction angles, it gives a good picture of the diffraction field for small diffraction angles, i.e., near the boundary of the shadow.

Fresnel's method entails difficulties whenever it is impossible to guess beforehand, even approximately (as does Kirchhoff for large apertures), the distribution of the elementary sources on the boundary surfaces. This applies, for example, to diffraction by an absorbing surface when the wave is propagating along (i.e., parallel to) the surface, or to the bending of waves by a smoothly convex obstruction.

3. MODERN DEVELOPMENTS OF YOUNG'S NOTIONS OF DIFFRACTION[1-5]

According to the laws of geometrical optics, the propagation within each ray tube is independent of the rest of the field. The amplitude of the ray (a quantity defined such that the square of its magnitude is proportional to the energy flux along the ray tube) is assumed to remain constant along each ray tube. However, it can be different from zero in one group of ray tubes and equal to zero in adjacent ones. This corresponds to the presence of sharp boundary lines (between bright and dark regions) in the geometrical theory. According to more accurate ideas diffraction, in the narrow sense of the word, represents in the first approximation the effect of transverse diffusion of the ray amplitude from some ray tubes to adjacent ones, or, in other words, the diffusion of ray amplitudes along the wave fronts.

On the basis of this idea, all results of the simplified Fresnel theory for diffraction by apertures of arbitrary shape in a plane surface and small diffraction angles are obtained by considering the transverse diffusion of the ray amplitudes along the fronts of the nearly-plane waves. If we substitute the expression $U = A(x, y, z)\, e^{-i(\omega t - kx)}$ for a nearly plane wave propagating in the x direction into the wave equation $\partial^2 U/\partial t^2 = c^2 \Delta U$, we obtain for a continuously varying amplitude $A(x, y, z)$,

$$\frac{\partial A}{\partial x} - \frac{D}{c}\frac{\partial^2 A}{\partial x^2} = \frac{D}{c}\left(\frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2}\right),$$

where $D = \lambda c/4\pi i$. The second term in the left-hand side of the equation is small compared with the first one (because of the small value of the wavelength λ). Neglecting this term we obtain

$$\frac{\partial A}{\partial x} = \frac{D}{c}\left(\frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2}\right). \tag{1}$$

If we set $x = ct$, that is, if we place our coordinate system on a moving wave front which coincides at $t = 0$ with the plane $x = 0$ in which the screen with the aperture is located, Eq. (1) can be rewritten in the form of the two-dimensional diffusion equation for diffusion and heat conduction,

$$\frac{\partial A}{\partial t} = D\left(\frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2}\right). \tag{2}$$

When a plane wave of unit amplitude impinges upon a screen with an aperture (see Figs. 5 and 6) then, if we assume the amplitude directly behind the aperture also to be unity and the amplitude behind the screen to be zero, a spreading of the amplitude along the wave front is found, which increases with the advance of the latter — analogous to the conventional diffusion or heat conduction. This is illustrated graphically in Figs. 5 and 6 by means of vertical lines of varying thickness which is assumed to be proportional to $|A|$ on the wave

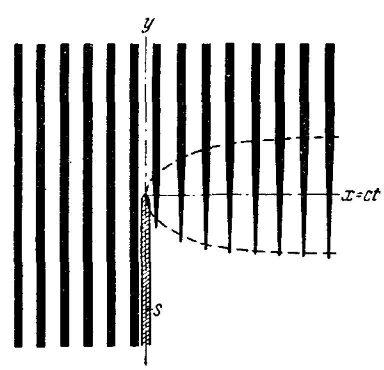

FIG. 5

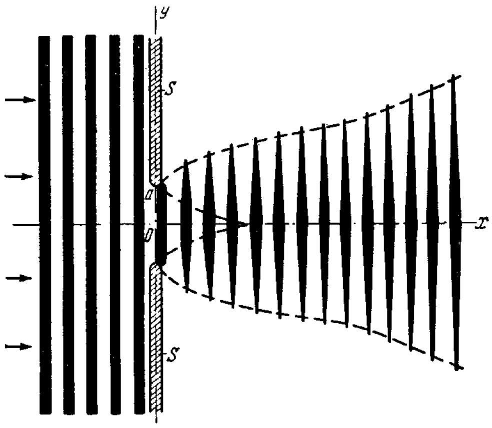

FIG. 6

front. The determination of such a spread in amplitude from Eqs. (1) and (2) leads to equations that agree with the approximate equations for Fresnel diffraction. As $\lambda \to 0$ the diffusion coefficient $D = \lambda c/4\pi i$ disappears, i.e., $\partial A/\partial x = 0$ (the case of geometrical optics with its sharp shadow boundaries).

The imaginary value of the coefficient D which leads to a similarity between Eq. (2) with Schrödinger's wave equation in quantum mechanics, implies that the diffusion of the complex amplitude A is accompanied with a phase shift, and as a result oscillations are possible in the distribution of the magnitude $|A|$ of the amplitude along the wave front. In most typical cases, however, the diffusion of the wave amplitude shows more similarity with the conventional diffusion and heat conduction processes than with quantum mechanics.

The method described permits the solution of problems that cannot be solved by Fresnel's method. One example is the problem of wave propagation along an absorbing surface $y = 0$ (characterized by a surface impedance $1/g$) where the boundary condition for the surface has the same form as the condition for "surface heat conductivity," i.e., $\partial A/\partial y = hA$, where $h = 2\pi g/i\lambda$. In Fig. 7 an example is illustrated where the wave is gliding first along an ideal reflecting surface

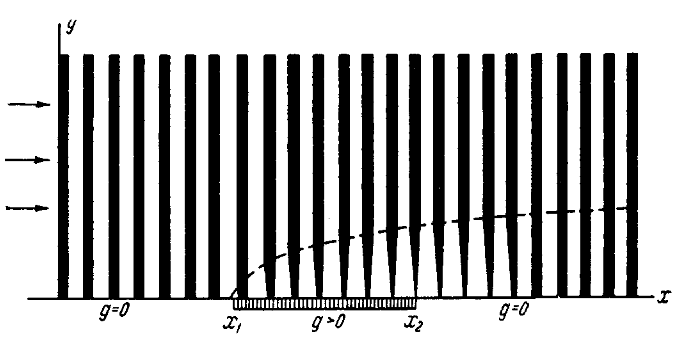

FIG. 7

($g = 0$), then passes over an absorbing region ($g > 0$), $x_1 < x < x_2$, after which it reaches again a nonabsorbing surface. The diffraction phenomenon manifests itself here in the fact that the amplitude of the wave A, determined from equations fully analogous to the equations for heat conduction and diffusion, decreases on the lower portion of the wave front as the wave advances along the absorbing section. The process taking place is similar to the cooling of an initially heated plate by outward heat conduction h from the lower end. When the wave again enters the non-absorbing region the inverse process of "warming up" of the lower end at the expense of the "warm" upper end starts. This is illustrated in Fig. 7.

The phenomenon of transverse amplitude diffusion along a wave front, like conventional diffusion or heat conduction, is of local character and is strongly pronounced only in the zones of effective diffusion, where the gradients of the complex amplitude of the wave front are relatively large. In Fig. 5 a similar zone is outlined by a parabola (dotted line). As the wavelength decreases this parabola approaches the boundary line of the geometrical shadow and coincides in the limit with it. In the case of an aperture (Fig. 6) the two parabolic zones of effective diffusion overlap at a distance a^2/λ. This distance appears also in Fresnel's investigation of diffraction. Further spreading of the amplitude maximum continues approximately linear with distance.

Examination of the transverse amplitude diffusion along the front of a plane wave is insufficient to yield a more exact idea of the diffraction phenomenon. An analogous investigation is necessary of the diffusion of the ray amplitude along curved wave fronts, which are obtained in accordance with the laws of geometrical optics for given shapes of the diffracting objects and given positions of the field sources. Let us limit ourselves to the two-dimensional case of a simple field described by the wave equation $\frac{\partial^2 U}{\partial t^2} = \frac{c^2}{n^2}\Delta U$, propagating in a nonuniform medium with a continuously varying refractive index n. It is then possible, within the approximation of geometrical optics, to introduce (see Fig. 8) a family of rays $\eta_1, \eta_2, \eta_3 \ldots$ and the orthogonal family of wave front surfaces $\xi_1, \xi_2, \xi_3 \ldots$, where ξ is the optical path length (eikonal). Introducing the specific transverse area of the ray tube, $\kappa = d\sigma/d\eta$, the expression for the wave field in terms of the coordinates ξ and η can be written

$$U = \frac{A}{\sqrt{n\kappa}} e^{-i(\omega t - k\xi)},$$

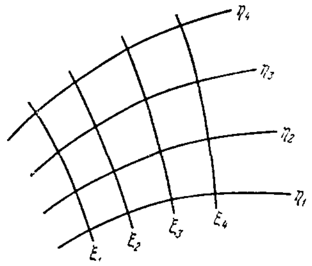

FIG. 8

where $A(\xi, \eta)$ is the ray amplitude, also referred to as the attenuation function. Substituting the above expression for U into the wave equation and neglecting terms similar to those neglected previously, an expression analogous to Eq. (1) is obtained for A,

$$\frac{\partial A}{\partial \xi} = \frac{D}{c}\frac{1}{\sqrt{n\varkappa}}\frac{\partial}{\partial \eta}\left(\frac{1}{n\varkappa}\frac{\partial}{\partial \eta}\frac{A}{\sqrt{n\varkappa}}\right) \tag{3}$$

Setting $\xi = ct$, this equation goes over into the diffusion heat conduction equation for a curved front that is being deformed with time.

The foregoing description of diffraction by aperture edges, based on the consideration of transverse diffusion along plane wave fronts, is only a poor approximation and is not more exact than Fresnel's. For a more exact description it is necessary to consider Young's more complete picture of the wave field and to take into consideration the fact that the transverse diffusion of the ray amplitude takes place along the fronts of a cylindrical wave that diverges from the aperture edge. To determine the diffusion of the ray amplitude, Eq. (3) must be used instead of Eq. (1). For a uniform medium $(n = 1)$ and the cylindrical case, Eq. (3) becomes.

$$\frac{\partial A}{\partial r} = \frac{D}{cr^2}\frac{\partial^2 A}{\partial \varphi^2} \tag{4}$$

Figure 9 shows a schematic diagram of the transverse diffusion for the case of diffraction of a plane wave by an ideal reflecting wedge with an arbitrary opening angle. The dotted parabolas show two zones of effective diffusion, which encompass the boundaries of the geometrical shadow of the incident and

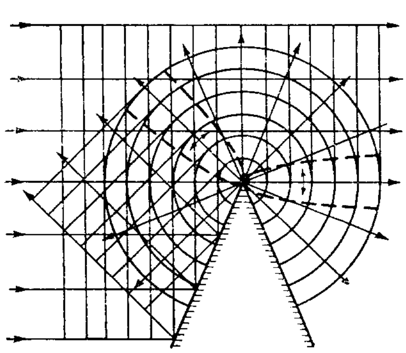

FIG. 9

reflected waves, respectively. The curved arrows within these zones indicate the direction of the transverse diffusion along the cylindrical fronts. The other arrows indicate the propagation direction of the wave fronts. Because of the spreading of the cylindrical wave fronts the gradients of the ray amplitude are very small in the regions outside the parabolas and the transverse diffusion is therefore very weak in these regions and can be practically neglected. The diverging wave in these regions, which can be referred to as the Fraunhofer zones, has the characteristics of a conventional cylindrical wave originating at the edge of the wedge and possessing a well defined directivity. Actually, the origin of this wave is not at the edge of the wedge but in the zone of effective diffusion. Strictly speaking, it is in this zone that the diffraction phenomenon takes place. Since the diffusion process is associated with a certain delay, the cylindrical wave in the Fraunhoffer zones lags in the conjugate plane wave on the boundary of the geometrical shadow by $\lambda/4$.

The application of Eq. (4) to the solution of diffraction by an ideally reflecting wedge leads to results that coincide asymptotically for large distances $r \gg \lambda$ with the exact solution found by Sommerfeld. Since the transverse diffusion approximation is not valid at the edge of the wedge, the asymptotic agreement shows that the sections in the direct vicinity of the edge have a negligible influence on the overall diffusion effect, which accumulates as the wave spreads away from the edge. In the small angular region $|\varphi| \ll 1$ near the boundary of the geometrical shadow the cylindrical wave differs very little from a plane wave and can be considered, together with the portion of the incident wave which is not cut off by the screen, as a single nearly-plane wave. This is essentially the justification for the foregoing approximation, in which amplitude diffusion was assumed to take place behind the aperture along nearly-plane wave fronts (see Figs. 5 and 6). Since the zones of effective diffusion, which is influential at small diffraction angles, also belong to the region $|\varphi| \ll 1$ results based on the assumption of diffusion along nearly-plane waves are valid for small diffraction angles. The error resulting from the incorrect assumption of the wave front direction is in part automatically corrected by the diffusion effect. The smaller the initial error the better, of course, is the correction.

The reasons for the incorrect results of the approximate calculations for large diffraction angles at both sides of the geometrical shadow become apparent when it is realized that the diffusion actually takes place along cylindrical rather than plane wave fronts. However, similar simplifications in the form of the wave fronts, which result in only small deviations in the zones of effective diffusion and their vicinity, can be conveniently used for an approximate quantitative analysis of more complex diffraction phenomena. This is also the approach used by Leontovich and Fock.[1-3]

In case of diffraction by rounded edges the phenomenon of transverse diffusion in the bright and shadow regions has its specific characteristics. These are more easily followed if we assume for simplicity that the normal derivative of the field is zero at the boundary and investigate the wave propagation along planes that terminate at the rounded edge either only in the rear or only in the front. In the presence of a convex obstacle (see Fig. 10) the ray from the source to an arbitrary point in the shadow region can be constructed in accordance with Fermat's principle. In a homogeneous medium it coincides with a thread stretched between these two points. In case of a rear rounding (see Fig. 11) the wave fronts in the shadow region are evolutes of this type of rays. The diffraction depends on the transverse diffusion of the ray amplitudes from the bright to the shadow region along these wave fronts. The zone of effective diffusion can be correspondingly divided into three regions, D_a, D_b and D_c, indicated in Fig. 11 by dotted lines. In the region D_a and in the neighboring angular region the diffraction picture is similar to the one obtained in the vicinity of the geometrical shadow in case of diffraction by a sharp screen edge or by a wedge (see Fig. 9). We observe that within the region D_c two ray tubes sufficiently far removed from each other (for example, two ray tubes located within the sections Δ' and Δ'' respectively) are not linked by a common wave front. The diffusion of the ray amplitude along the arc s can therefore take place only through a "cascading" process. That is to say, diffusion and ray propagation so alternate that the diffusion process in successive ray tubes starts only after the diffusion in the preceding ones is completed. As a result, if the radius of curvature r_0 of the convex body is constant, there is established in the zone D_c a diffusion process that decreases in strength along the arc s. The decrease of the ray amplitude in adjacent ray tubes along the arc proves to be proportional to the amplitude. Thus,

$$\frac{dA}{dS} = -aA, \qquad (5)$$

where more detailed calculations show that

$$-a = \alpha \frac{i-\sqrt{3}}{2} \sqrt[3]{\frac{\pi}{\lambda r_0^2}}.$$

The constant α depends on the material properties of the body and on the type of polarization of the incident wave. $\alpha \cong 1.02$ — the first root of the derivative of the Airy function — if the normal derivative of the field is zero at the boundary; if the field itself is zero at the boundary, then $\alpha \cong 2.34$.

Equation (5) is valid also for the case of varying radius of curvature, $r_0 = r_0(s)$, provided that the radius changes only slowly along the arc s in the interval which corresponds to the length of the section of the ray within the region D_c. Solving Eq. (5) we obtain

$$A(s) = A(s_0) \exp\left[\alpha \frac{i-\sqrt{3}}{2} \sqrt[3]{\frac{\pi}{\lambda}} \int_{s_0}^{s} r_0^{-\frac{2}{3}}(s)\right] ds,$$

or, for r_0 = constant,

$$A(s) = A(s_0) \exp\left[\alpha \frac{i-\sqrt{3}}{2} \sqrt[3]{\frac{\pi}{\lambda r_0^2}} (s-s_0)\right]. \qquad (6)$$

The real part in the exponent indicates an exponential decrease of the amplitude of the ray gliding along the arc s. The imaginary component in the exponential indicates a decrease in the propagation velocity. This is related to the delay resulting from the foregoing mechanism of sequential (step-by-step) diffusion in the zone D_c. When one of the rays branches off tangentially from the arc s, leaves the zone D_c, and enters the Fraunhofer zone further amplitude decrease and time retardation stops. The ray amplitude now remains constant and the decrease in the amplitude of the wave field is due only to the widening in the cross sections of the ray tubes. However, because of the substantial decrease of the ray amplitude in its passage through the diffusion zone, in the shadow region and at large diffraction angles the field be-

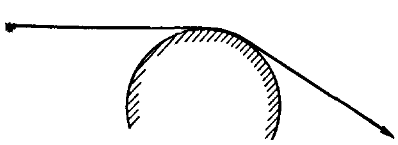

FIG. 10

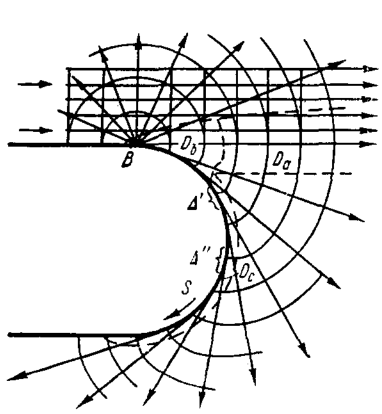

FIG. 11

hind a convex obstruction is much weaker than the field behind an equal screen with sharp edges. In the lower part of Fig. 11 the surface of the obstruction becomes a plane and the diffusion process discontinues if the surface is ideally reflecting with the boundary condition $\partial u/\partial n = 0$ (otherwise the diffusion process described in connection with Fig. 7 would take place).

Within the intermediate zone D_b near the point B the wave fronts are cylindrical and have relatively small radii. The diffracted wave directed from point B upwards and backwards toward the incident wave is formed at the expense of the ray amplitude diffusion along these cylindrical wave fronts. The greater the radius of curvature of the surface of the body to the right of the point B, the smaller near this point the disturbance that originates from the penetration of the field into the narrow region between the boundary of the geometrical shadow and the surface of the body. When the radius of curvature is large the diffusion along the cylindrical wave fronts of small radii is very weak and the dispersion in the upward and backward direction is therefore, in contrast to the case of a sharp edge, negligibly small.

A characteristic property of diffraction by an edge of an obstruction which is rounded in front (see Fig. 12) is the transverse diffusion of the ray amplitude directly from the incident to the reflected wave. To analyze this process it is necessary to generalize the geometric-optical notion of incident and reflected waves for a curved boundary surface.[6]

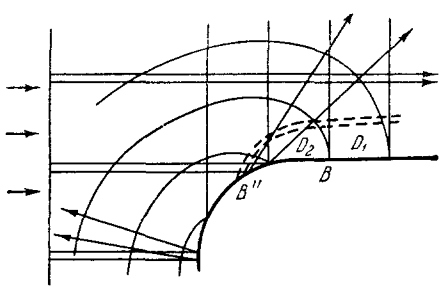

FIG. 12

When the normal derivative at the boundary of the body is zero ($\partial U_N/\partial n = 0$) the diffraction field, U_N, can be uniquely resolved into a sum of an incident wave U_1 and a reflected wave U_2, where $U_1 = 1/2 \times (U_N + U_D)$ and $U_2 = 1/2(U_N - U_D)$; U_D is a diffraction field analogous to U_N but satisfying the boundary condition $U_D = 0$. We can now imagine that two separate spaces surround the object, and assume that in one of the spaces only the incident wave U_1 is present and that in the other only the reflected wave U_2 is present. At a smooth boundary surface of the body the reflected wave is then found to be a continuous and smooth extension of the incident wave (if the reversed sign of the surface normal is taken into consideration). In the limit as $\lambda \rightarrow 0$, U_1 and U_2 go over into conventional incident and reflected waves of geometrical optics. At small but finite wavelengths λ the diffraction phenomena can be, as before, described by the process of ray amplitude diffusion along geometric-optical fronts continuously passing over from the region of the incident to the region of the reflected wave. In Fig. 12 where the incident wave is assumed to be a plane wave it is seen that as the point B is approached the transverse area of the ray tubes of the reflected wave increases relatively sharply. This results in a rapid attenuation of the field. It is also noticed that the rays squeeze more and more closely to the boundary of the body where the incident and reflected waves join. Owing to these two factors, considerable differences develop in the ray amplitude over small sections of the combined fronts of the incident and reflected waves. This causes the transverse diffusion.

Thus, in contrast to the reflection of a plane wave from a plane boundary surface, or from a smoothly convex boundary at steep incident angles, there is in this case a twofold mechanism of energy transfer from the incident to the reflected wave. In addition to the conventional reflection mechanism characteristic of geometrical optics, there is in this case the additional process of diffusion transfer.

The two overlapping zones D_1 and D_2 of effective diffusion, for the incident and reflected waves respectively, are shown dotted in Fig. 12. At the boundary of the body the ray amplitude diffuses from D_1 into D_2. The incident wave near the surface therefore becomes weaker as it approaches the point B, and this results at the same time in a diminishing of the resultant field. Behind the point B where, as shown in Fig. 12, the incident wave is assumed to glide along an ideally reflecting plane ($\partial U/\partial n = 0$) the reflection link between the reflected and incident waves is broken. However, the diffusion link remains preserved. As the wave moves away from the point B the resultant field gradually settles and coincides in the limit with the undisturbed incident plane wave.

If instead of the horizontal surface behind the point B the body is further curved as in Fig. 11, the resultant diffraction phenomena will be analogous to the ones described previously, with the formation of zones D_a, D_b, D_c of effective diffusion.

In the case of diffraction at the edge of a large smoothly convex body, the dispersion in opposite directions from the point B can be neglected (see Fig. 11). The fronts of the geometrical approximation for the incident, reflected, and diffracted (into the shadow) waves are then found to be almost normal to the surface of the body in the regions of effective diffusion (except for the far sections of the zone D_a) and their neighborhood, and are all found to propagate along the body in the same direction. By making a further approximation and replacing these wave fronts with fronts that are exactly normal to the boundary surface, Fock and Leontovich[3] developed an approximate theory for the propagation of radio waves around the earth.

The diffraction phenomena in nonuniform media are analogous to the diffraction phenomena in homogeneous media. For example, when a cylindrical wave in a medium having a refractive index varying with the height, $n(z) = e^{z/r_0}$, is reflected by a plane (see Fig. 13) the rays bend upwards and a shadow region is formed, into which the field penetrates as a result of the transverse diffusion between the ray tubes. As in the example shown in Fig. 11, zones of effective diffusion D_a, D_b, and D_c are formed. For the rays emerging from D_c into the Fraunhofer zone the field decreases from ray to ray in the direction of s, according to the same exponential law as given by Eq. (6) for the homogeneous case, and the propagation takes place with the same time delay. In fact, through a simple transformation, $s = r_0\varphi$ and $z = r_0 \ln(r/r_0)$, the wave field in the nonuniform medium in the semi-infinite region $z > 0$ is mapped into the region, $r > r_0$, outside a homogeneous cylinder $(n = 1)$.

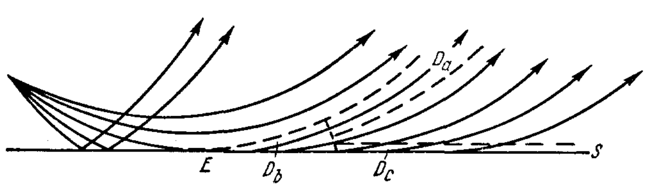

FIG. 13

This problem is thus exactly analogous to the problem of diffraction in a homogeneous medium by a cylindrical obstruction with a radius of curvature r_0.

4. DIFFRACTION IN THE BROAD SENSE OF THE WORD

Diffraction in the broad sense of the word is generally understood to be a study of excited wave fields in finite or infinite regions filled with homogeneous or nonhomogeneous media. The theoretical investigation of diffraction reduces to the solution of the mathematical problem of forced oscillations in the given region.

As an example of the formulation of the problem we consider the case of a scalar wave field $U(Q)e^{-i\omega t}$ excited by a point-source harmonic oscillator $e^{-i\omega t}$ (ω is the radian frequency) situated at the point $Q = Q_0$ (see Fig. 14). In the acoustic case the function $U(Q)$ can represent either the sonic pressure or the velocity potential. In the case of a two-dimensional electromagnetic problem this function can represent one of the components of the Hertz vector, or the field intensity of the electric or magnetic fields.

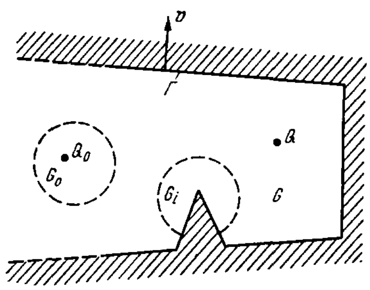

FIG. 14

As an illustration let us consider the case when the field in a given three-dimensional region with a piecewise-smooth boundary Γ is described by the equation

$$\Delta U + k^2 n^2(Q)\, u = -4\pi\delta(Q - Q_0), \tag{7}$$

where the refractive index $n(Q) > 0$ is a continuous function of the point Q within the region G, $n(Q_0) = 1$, k is a constant parameter, and $\delta(Q - Q_0)$ is a delta function.

At the boundary Γ the field is continuous and satisfies one of the boundary conditions:

$$U = 0;\quad \frac{\partial U}{\partial \nu} = 0;\quad \frac{\partial U}{\partial \nu} - ikgU = 0 \qquad (\operatorname{Re} g \geqslant 0), \tag{8}$$

where ν is the outward normal.

The third condition implies in the acoustic case that the normal impedance $1/g$ is known at the boundary. In the electromagnetic case it is referred to as the boundary condition of Leontovich.[7] In special cases the boundary conditions can also include derivatives of arbitrarily high order, as for example, in the case of diffraction of sound waves by elastic membranes.[8]

Equations (7) and (8) do not define the problem uniquely. It is necessary to add conditions that exclude, in the case of a lossless system, the possibility of exciting free oscillations of a natural frequency equal to the frequency ω of the source and have eigenfunctions orthogonal to the distribution function of the source. It is also necessary to add a condition that guarantees the absence of superfluous sources of forced oscillations at infinity or at the boundary of the region.

In order to exclude from the analysis the possible natural oscillations of the system, we use the fact pointed out by I. M. Gel'fand that the solution $U(Q; k)$ is an analytic function of k only for forced oscillations. The sources at infinity are excluded by making use of the stability condition. This condition is based on the fact that in a lossy medium the field strength must decay rather than increase with increasing distance from the source. Since the positive imaginary component of the wave number k represents the losses of the medium, the stability condition leads in conventional cases to the requirement that the solution $U(Q; k)$ be bounded for $\mathrm{Im}(k) > 0$ over the entire region G, except for the small neighborhood of the source. However, in certain regions a bounded solution may not exist when $\mathrm{Im}(k) < |\mathrm{Re}(k)|^{10}$ (because of concentrated effects). It is therefore generally convenient, following Fock,[9] to impose the requirement that the solution be bounded for pure imaginary values of k, that is, to require that in the region $G - G_0$

$$|U(Q; k)| < \infty \tag{9}$$

when the phase of k is $\pi/2$. The requirement that the solution be bounded also excludes the presence of superfluous sources at the boundary of the region. The problem of forced oscillations or of diffraction in case of a point source is thus reduced to the problem of finding, subject to the boundary conditions (8), a solution of Eq. (7) analytic in k and satisfying the stability condition (9).

The function $U(Q; k)$, which represents forced oscillations, is also referred to as the resolvent. The function $U = K(Q, Q_0)$, which is considered to be a function of the pair of observation and source points, is Green's function for the given equation, region and boundary conditions.

In more general diffraction problems the field can become infinitely large at certain points on the boundary (ribs, sharp edges). Then, to exclude superfluous sources on the boundary, we use instead of the requirement that the solution be bounded the requirement that the solution be square integrable (energy bound) in the neighborhood of the sharp tip G_i (see Fig. 14). The latter condition is equivalent to a limitation on the order of magnitude of the decrease in field with decreasing distance d from the singular element (i.e., from the sharp tip). For example, in case of a sharp edge on the boundary the field should increase not faster than $d^{-1/2}$ as the edge is approached.

If all diffracting objects are distributed at a finite distance $R < A$ from the origin of the coordinate system and all the sources are situated in the region $R > A$, it is convenient, for the case of a uniform medium $(n = 1)$, to separate the incident wave $U_0(Q)$ from the overall diffraction field $U(Q)$. (In the case of a single point source as assumed in Eq. (7), the incident wave is simply the spherical wave $U_0(Q) = \frac{1}{R_1} e^{ikR_1}$.) The stability condition (9) can then be written in the form:

$$|U - U_0| < \infty \tag{10}$$

for the region G where $\mathrm{Im}(k) > 0$.

This formulation of the problem permits us to solve diffraction problems involving not only spherical incident waves but waves of arbitrary shape. This formulation is in particular convenient when the incident wave is a plane wave arriving from infinity. The stability condition (10) can then be replaced by the Sommerfeld emission condition:[12]

$$\lim_{R\to\infty} R\left[\frac{\partial(U-U_0)}{\partial R} - ik(U-U_0)\right] = 0 \tag{11}$$

for $k > 0$.

As an example of an exact solution we consider the solution[13] of the wave equation $\nabla^2 U + k^2 U = 0$ for Sommerfeld's problem of diffraction of a plane wave

$$U_0 = e^{-ikr\cos(\varphi-\varphi_0)} \quad (|\mathrm{Re}\,\varphi_0| < \Phi)$$

in a wedge shaped region G $(-\Phi < \varphi < \Phi,\ r > 0)$ with the boundary conditions $U(r, \pm\Phi) = 0$. The general solution of the wave equation in the region G is given by the integral

$$U = \frac{1}{2\pi i}\int_\gamma e^{-ikr\cos\alpha}\, s(\alpha+\varphi)\, d\alpha$$

where the integration contour γ is shown in Fig. 15. The stability condition reduces in this case to the requirement that the function $s(\alpha) - (\alpha - \varphi_0)^{-1}$ be regular in the strip $|\mathrm{Re}(\alpha)| \le \Phi$. Substituting

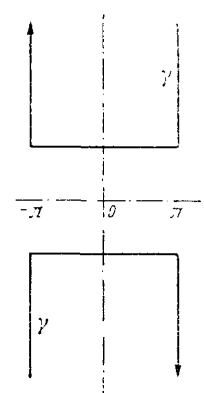

FIG. 15

the integral into the boundary conditions, we obtain the two identities:

$$\int_\gamma e^{-ikr\cos\alpha} s(\alpha \pm \Phi)\, d\alpha \equiv 0 \qquad (r > 0).$$

This requires[14] that the functions $s(\alpha \pm \Phi)$ be even. We thus obtain two simple functional relationships:

$$s(\alpha \pm \Phi) - s(-\alpha \pm \Phi) = 0.$$

The solution of these equations which satisfies the foregoing requirement of regularity, is of the form

$$s(\alpha) = \frac{\pi}{4\Phi}\left[\cot\frac{\pi}{4\Phi}(\alpha - \varphi_0) + \tan\frac{\pi}{4\Phi}(\alpha + \varphi_0)\right].$$

The solution of the diffraction problem is obtained by substituting this function into the original integral.

The solution for the case when the boundary conditions are of the third kind

$$\frac{\partial U}{r\,\partial\varphi} \mp ikg_{\pm} U = 0 \qquad (\varphi = \pm\Phi)$$

can be derived in an analogous way.[15] The functional relations, however, are more complicated.

[1] M. A. Leontovich, Izv. Akad. Nauk SSSR, Ser. Fiz. 8, 16 (1944).

[2] V. A. Fock, JETP No. 1, 693 (1945); Izv. Akad. Nauk SSSR, Ser. Fiz. 10, 171 (1946).

[3] M. A. Leontovich and V. A. Fock, JETP 557 (1946).

[4] G. D. Malyuzhinets, Paper at the All-Union Conference on Problems of Electrical Oscillations and Waves at the Gorkiĭ University (Dec. 12, 1946), Изв. Высш. учебн. завед., сер. радиофиз. (News of the Instit. of Higher Learning, Radiophysics Series) (1960).

[5] C. E. Schensted, IRE Trans. on Antennas and Propagation, No. 3, 240 (1956).

[6] G. D. Malyuzhinets, Some Generalizations of the Reflection Method in Diffraction Theory, Doctoral Thesis, U.S.S.R. Acad. Sci. M., 1950.

[7] M. A. Leontovich, Новейшие исследования по распространению радиоволн (New Investigations on the Propagation of Radio Waves), 2nd Collection, M.-L., 1948.

[8] G. D. Malyuzhinets, Dokl. Akad. Nauk SSSR **118**, 1099 (1958), Soviet Phys.-Doklady **3**, 52 (1958). IV Всес. акустич. конференция, Рефераты докладов (Fourth All-Union Acoustic Conference, Abstracts of Papers), U.S.S.R. Acad. Sci. M., 1958.

[9] V. A. Fock and V. R. Bursian, J. Russ. Phys. Chem. Soc. (Physics Part) **58**, 356 (1926). V. A. Fock, Ann. Physik **17**, 401 (1933).

[10] G. D. Malyuzhinets, Dokl. Akad. Nauk SSSR **78**, No. 3 (1951).

[11] H. Meixner, Z. Naturforsch. **7a**, 506 (1948); A. W. Mane, Z. Naturforsch. **7a**, 387 (1952); Z. angew. Math. u. Mech. **33**, No. 1-2 (1953).

[12] A. Sommerfeld, Jahresberichte d. Deutsch. Math. Vereinigung **21**, 309 (1912).

[13] F. Frank and R. Mises, Differential-and-Integral-Equations of Mathematical Physics (Russ. Transl.), Chapter XX, ONTI, 1937.

[14] G. D. Malyuzhinets, Dokl. Akad. Nauk SSSR **118**, 1099 (1958), Soviet Phys.-Doklady **3**, 52 (1958).

[15] G. D. Malyuzhinets, Dokl. Akad. Nauk SSSR **121**, 436 (1958), Soviet Phys.-Doklady **3**, 752 (1959).

Translated by L. E. Bergstein