A technique for solving the wave equation and prospects for physical applications arising therefrom

A. V. Kukushkin
(Submitted 18 February 1992; resubmitted 16 September 1992)
Usp. Fiz. Nauk **163**, 81–95 (February 1993)

A new technique is proposed for solving the Helmholtz equation, based (in contrast to the classical method of separation of variables) on combining the variables. As a result, the wave equation (a partial differential equation) is converted into a single second-order ordinary differential equation. One solution of the latter is the classical exponential with imaginary argument (a simple self-similar solution). The second (the compound self-similar solution) consists of two factors: an exponential of imaginary argument (the first solution) and a tabulated special function. In the two-dimensional case this function is a complementary error function, or, in a special case, the Fresnel integral in complex form. In three dimensions the second factor is the exponential integral function. The physical part of the work is concerned with utilizing the compound self-similar solutions in physical applications involving external problems of electrodynamics and acoustics. These include (in the two-dimensional case) the theory of open waveguide structures.

1. INTRODUCTION

As is well known, the classical method for solving partial differential equations in general and the wave equation in particular is by separation of variables. On account of its great generality this method predominates in the majority of physical applications. As for the object of our investigations, i.e., the wave equation itself, the search for new ways of solving it, aside from its methodological interest in view of the role which this equation plays in theoretical physics, would also be interesting in the practical sense, since new mathematical techniques often open a way to new physical results.

In connection with ways of achieving this goal it is difficult to imagine that completely new methods can arise out of thin air. Here what is most needed is a new way of looking at what is already familiar and a way to overcome prejudice in treating what is well known. The first fruitful ideas can arise, e.g., by rereading some classical result well known to everybody. For us this classical example is the familiar and very elegant (in the mathematical sense) solution of the Sommerfeld problem for the diffraction of a plane electromagnetic wave by an ideally conducting and infinitely thin half-plane.[1–3] With that as our starting mark and also using the results which we have recently published[4,5] generalizing the corresponding classical Sommerfeld formulas to the case of a half-plane with "soft" boundary conditions, we can find, as will be shown, a way of grouping the independent variables which transforms the two-dimensional Helmholtz equation written in polar coordinates into a single ordinary differential equation. This way of grouping the independent variables, which in contrast to the method of separation of variables combines them, is the key to solving the methodological problem. This is confirmed by extending the grouping technique to the case with three special variables. Here, just as in the two-dimensional case, the Helmholtz equation (written in spherical coordinates) is transformed into an ordinary differential equation, but only for axisymmetric wave functions.

Characterizing the physical content of this work, we note that it contains both purely methodological results, which apply entirely to the three-dimensional case and partly to the two-dimensional case, and also original results (which apply only to the two-dimensional problems). In particular, the methodological nature of the work applied to the solution of the three-dimensional Helmholtz equation is associated with finding a new method for an already well-known closed solution of the problem of diffraction of a plane acoustic wave by a paraboloid of revolution with "hard" boundary conditions for axisymmetric excitation, when the front of the incident wave is perpendicular to the axis of the paraboloid.[6] As regards the two-dimensional models, the purely mathematical results reduce here to the statement that in connection with the familiar Sommerfeld result the new method does not just simplify the process of finding this solution, but compared with the classical methods may be said to reduce this process to one which is almost trivial in its simplicity.

As for the new physical results obtained using the mathematical techniques developed here, this aspect of the problem is exhibited most clearly in the theory of open waveguide structures,[4,5] and touches on the most general aspects of the physical origins of the theory. This will also be treated in what follows, but from a formal viewpoint of clarifying the mathematical apparatus which we introduced in Ref. 5 by purely heuristic arguments.

 1063-7869/93/020081-13$10.00

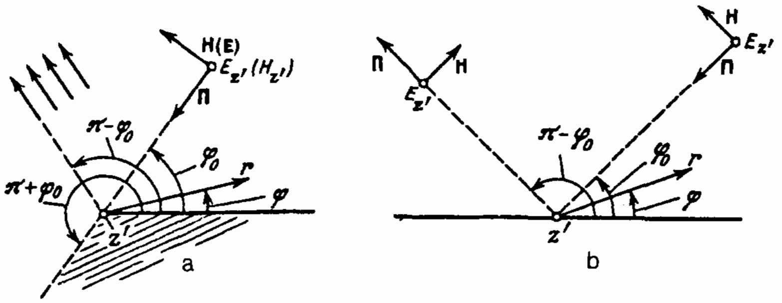

FIG. 1.

2. THE SOMMERFELD SOLUTION AND THE CHOICE OF A WAY OF GROUPING THE INDEPENDENT VARIABLES IN THE TWO-DIMENSIONAL HELMHOLTZ EQUATION

It is well known that the Sommerfeld solution for the diffraction of a plane electromagnetic wave by a perfectly conducting and infinitely thin half-plane (Fig. 1a) can be written in the form[2,3]

$$E_{z'},H_{z'}=(U(z,\varphi-\varphi_0)\mp U(z,\varphi+\varphi_0))e^{i\omega t}, \tag{1}$$

where we have written $z=kr$ and $k=2\pi/\lambda$, λ is the wavelength in free space, r, φ are polar coordinates, and φ_0 is the polar angle of incidence of the plane wave (Fig. 1a). If we introduce the notation $\psi=\varphi\pm\varphi_0$, the function U can be written[2] as a contour integral over the angular spectrum of the plane waves:

$$U=\frac{1}{4\pi}\int_C\left[1-e^{i/2(\theta+\psi)}\right]^{-1}e^{iz\cos\theta}d\theta, \tag{2}$$

where the two-branched contour C is shown in Fig. 2.

It is well known that the Sommerfeld solution (1) plays an extraordinarily important role as a model solution in diffraction theory. It is used as an example in order to study the fundamental behavior in the diffraction of plane waves by an obstacle. This is related to the fact that the integral (2) can be expressed in closed form in terms of tabulated functions:

$$U=\frac{1}{2}e^{iz\cos\psi}\text{erfc}\left[(2iz)^{1/2}\cos\frac{\psi}{2}\right], \tag{3}$$

where the complementary error function (or probability integral)

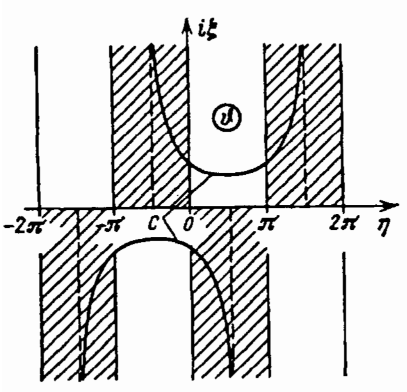

FIG. 2.

$$\text{erfc}w=\frac{2}{\sqrt{\pi}}\int_w^\infty e^{-t^2}dt \tag{4}$$

in this case is the complex Fresnel integral, in which

$$w\equiv(2iz)^{1/2}\cos\frac{\psi}{2}=e^{i\pi/4}(2z)^{1/2}\cos\frac{\psi}{2}, \tag{5}$$

so that the Sommerfeld solution (1) acquires the form of a closed mathematical expression which can be represented in terms of the function

$$U=\tfrac{1}{2}e^{-iz}e^{w^2}\text{erfc } w. \tag{6}$$

It is clear that the latter satisfies the two-dimensional Helmholtz equation

$$z^{-1}\frac{\partial}{\partial z}\left(z\frac{\partial U}{\partial z}\right)+z^{-2}\frac{\partial^2 U}{\partial\varphi^2}+U=0, \tag{7}$$

just as does the function describing a plane wave:

$$U\equiv e^{iz\cos\psi}=e^{-iz}e^{w^2}, \tag{8}$$

in terms of which the wave fields in two-dimensional boundary-value problems are usually expanded, as was done, e.g., in Eq. (2) and which (note carefully) enters into (6) as a factor in front of the erfc w function.

Besides the Helmholtz equation (7), the Sommerfeld solution (1) satisfies the associated boundary conditions on the half-plane. Specifically, in the case of TM polarization, when the electric field vector of the incident plane wave ($E_{z'}$) is parallel to the edge of the half-plane, i.e., for $\varphi=0.2\pi$, we have the condition

$$E_{z'}\Big|_{\varphi}=0.2\pi=0,$$

and in the case of TE polarization

$$\frac{\partial H_{z'}}{\partial\varphi}\Big|_{\varphi=0.2\pi}=0.$$

The special role played by the Sommerfeld solution in diffraction theory follows from the fact that a completely rigorous solution of the boundary-value problem can be represented here in terms of tabulated functions, the analyticity properties of which express all of the behavior of the phenomenon of wave diffraction by objects. In consequence of this, the fundamental analyticity properties of the complementary error function play an extremely important role here. We summarize the basic properties of this function.

First, from the theory of these functions[7] it is known that the following functional relation holds:

$$\operatorname{erfc} x+\operatorname{erfc}(-x)=2, \tag{9}$$

valid for any complex x. In contrast to this the formula

$$e^{x^2}\operatorname{erfc} x\approx(\pi x)^{-1}+o(x^{-3}) \tag{10}$$

for the asymptotic behavior of this function at large values $|x|\gg 1$ holds only under the condition

$$|\arg x|<\pi/2. \tag{11}$$

Essentially, these two formulas are completely sufficient to characterize all the most important properties of the fields undergoing diffraction in the half-plane problem. For example, if condition (11) is satisfied, this implies that the *field described by the function (6) contains no wave fields* other than the cylindrical wave.

Consequently, condition (11) is related to those which control the localization of the shadow regions of space adjacent to the obstacle (the half-plane), since the opposite condition

$$|\arg x|>\pi/2,$$

applied to relation (9) yields

$$\tfrac{1}{2}e^{-iz}e^{w^2}\operatorname{erfc} w=e^{-iz}[e^{w^2}-\tfrac{1}{2}e^{w^2}\operatorname{erfc}(-w)],$$

where $|\arg(-w)|<\pi/2$ already holds in the right-hand side; this obviously refers to the formulas specifying the localization of the illuminated parts of space, because, as follows from the previous relation, the field of interest to us contains an additive exponential (geometrical) part in the form of a uniform plane wave. Thus, the equation

$$|\arg w|=\pi/2, \tag{12}$$

which under these conditions reduces to the expression

$$\cos\frac{\varphi\pm\varphi_0}{2}=0, \tag{13}$$

controls the geometrical position in space of the boundary between light and shadow.

This boundary moves in a straight line (polar ray), which coincides with the farthest light ray out of the set of rays of a geometrical beam of reflected and transmitted rays (we are talking about a single extreme ray only because we are considering not the entire beam but its intersection with the plane). Recall that it is just condition (12) which plays the fundamental analytical role in specifying the most important features of the diffraction phenomenon studied using the half-plane model problem. It should be kept in mind that because the parameter φ_0 is real, Eq. (12) reduces here to Eq. (13), which leads to the linear relations $\varphi=\pi\pm\varphi_0$ connecting the angular coordinates of the boundary between light and shadow for the transmitted and reflected light beams with the polar angle of incidence of the plane wave. It is worth emphasizing again that these relations are linear because the parameter φ_0 which enters into the argument w of the complementary error function that vanishes at $\varphi=\pi\pm\varphi_0$ is real. Substituting this value in the function (6) we are concerned with, and recalling that erfc $(0)=1$ (Ref. 7), we can easily find its value at the boundary between light and shadow:

$$U\equiv\tfrac{1}{2}e^{-ikr}=\tfrac{1}{2}e^{-i\omega/cr}, \tag{14}$$

where c is the velocity of light. Hence we can conclude that the boundary between the light and shadow is determined in space by the farthest light beam (reflected from the extreme end of the half-plane, i.e., its edge; see Fig. 1a) out of the whole set of light rays which are emitted by the current induced in the half-plane by the incident plane wave. Here, in the case of a uniform plane wave, the direction of the reflected light rays determined by the relation (13) is the same as that of the energy they transport. To put it another way, the direction in which the phase velocity of the rays is equal to the velocity of light coincides with the direction in which the plane wave transports energy. Hence the definition of the concept of a light ray as a line along which energy is transported[10] is completely equivalent to the other definition which is possible here, namely, a line along which the phase velocity of the wave is equal to the velocity of light in the medium in which the wave process is taking place (here and below we will concern ourselves with vacuum).

The question now arises as to what specifically changes if Eq. (12) contains a complex quantity $\dot{\theta}$ instead of the real parameter φ_0. The function (6), just as in (8), will still be determined by the Helmholtz equation. This is related to the fact that the metric coefficients of a polar (or cylindrical) coordinate system are independent of the angular variable φ.

In order to answer this question we should keep in mind that the exponential function (8) which enters into (6) as a factor but nevertheless can be moved out by using Eq. (9) into an additive form, in this case is a nonuniform plane wave. As is well known,[9] in the theory of open waveguide structures these waves act as modes of an infinite waveguide when the structures are modeled as an incident impedance plane. Consequently, if we transform from the representation (8) to the representation (6) we can expect that for a complex parameter $\dot{\theta}$ the latter form will be related to the description of waves directed by an impedance half-plane, i.e., the waves of an open waveguide for which the two-dimensional semi-infinite model is a half-plane with two-sided impedance boundary conditions prescribed on it. There is reason to anticipate that, just as in the case of a perfectly conducting half-plane where the Sommerfeld solution plays a very important role as a model solution for diffraction theory generally, the corresponding solutions for an impedance half-plane will be similarly important in connection with the theory of open waveguides. In other words, just as before in diffraction theory, relation (12) must play a fundamental analytical role in the theory of open waveguides.

Let us see whether this is so, and what the physical content is of applications of Eq. (12) in the theory of open waveguides.

We first dealt with these questions in Refs. 4, 5, where solutions of the form

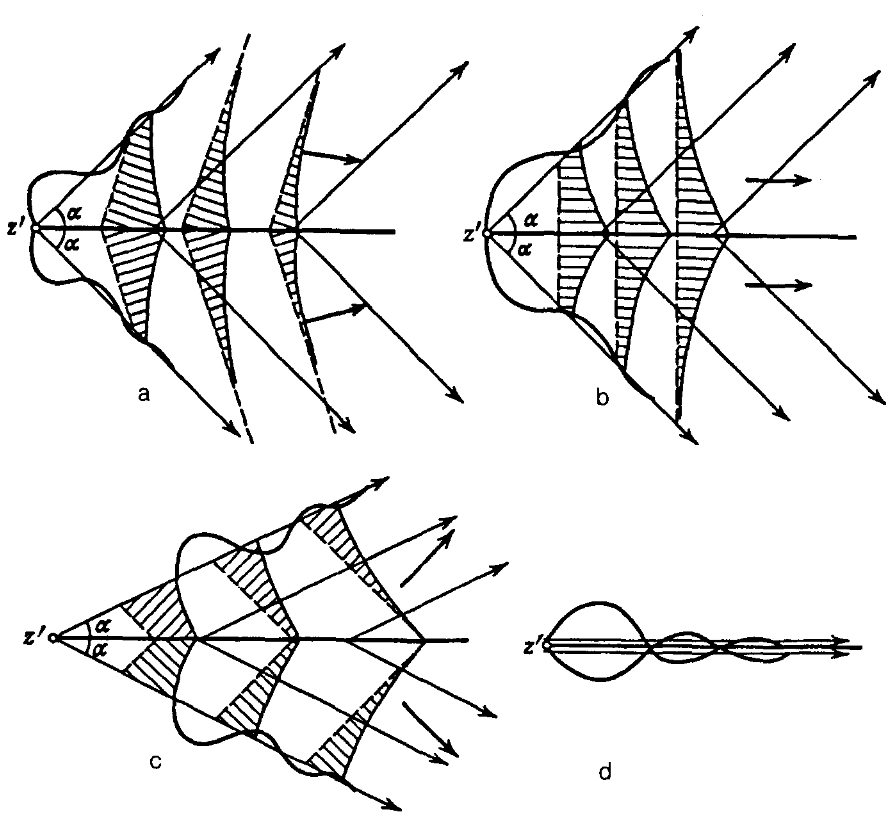

FIG. 3.

$$U = Ae^{-iz}(U(w_+) + U(w_-)), \quad (15)$$

$$U(w_\pm) = \tfrac{1}{2} e^{w^2} \pm \operatorname{erfc} \, w_\pm, \quad (16)$$

$$w_\pm = (2iz)^{1/2} \sin \frac{\varphi \pm \dot{\theta}}{2} = e^{i\pi/4}(2z)^{1/2} \sin \frac{\varphi \pm \dot{\theta}}{2}, \quad (17)$$

(where A is an arbitrary amplitude coefficient) were identified as quasimode functions of characteristic (and noncharacteristic) waves of an impedance half-plane corresponding to the slow characteristic (surface) waves for $\operatorname{Re} \dot{\theta} \equiv \eta \in [0, \arccos(\operatorname{ch}^{-1}\xi)]$, where $\xi = \operatorname{Im} \dot{\theta}$, and to fast noncharacteristic (propagating) modes for $\eta \in [\arccos(\operatorname{ch}^{-1}\xi), \pi/2]$. Here the same numbers $\dot{\theta}$ play the role of eigenvalues of the respective modes, being determined by the dispersion relation that follows from the impedance boundary conditions on the half-plane written for exponential (mode) components of the field,[4] since the latter enter adequately into the quasimode functions (15). We stipulate that here, as in Refs. 4, 5, only passive media are being treated.

Thus, if in the general case of a complex impedance a complex eigenvalue $\dot{\theta}$ is found, it is interesting to ascertain what the structure of the quasimode functions (15) is in this case and what specifically Eq. (12) means then, since the physical formulation of the problem has changed drastically compared with the classical Sommerfeld problem.

In the formal mathematical sense these changes reduce to the assertion that Eq. (12) can no longer be expressed as an equation of the form (13), since the complex nature of the parameter $\dot{\theta}$ means that it can now assume one of the two forms

$$\arg w_{+,-} = \pi/2, \quad (18)$$

$$\arg w_{+,-} = 3\pi/2. \quad (19)$$

It is found[4] that if we consider only those waves which are excited at the edge of the half-plane, then Eq. (18) is satisfied only for surface waves when $\eta \in [0, \arccos \operatorname{ch}^{-1}\xi]$, holds, while Eq. (19) holds only for unconfined waves, i.e., for $\eta \in [\arccos \operatorname{ch}^{-1}\xi, \pi/2]$.

Now let us consider the physical content of Eqs. (18) and (19).

If the parameter $\dot{\theta}$ is known, then, as can be deduced from Eqs. (17), Eqs. (18) and (19) must be solved with respect to the angular coordinate φ, and the resulting values will correspond as before to the angular coordinates of the boundary between light and shadow for a nonuniform plane wave excited at the edge of the impedance half-plane and propagating from this edge along its surface. In other words, according to relations (18) and (19) a nonuniform plane wave directed by the half-plane is contained in some angular sector of space adjacent to the half-plane itself (Fig. 3).

The dimensions of this sector, labeled in Fig. 3 with the letter α, is determined immediately from relations (18) and (19). Specifically, the angular coordinates of the boundary between light and shadow for the surface wave propagating over the half-plane, are calculated from Eq. (18), in which the expression for w_+ has been substituted. By replacing w_+ with w_- in this equation we can calculate the angular coordinates of the same wave localized below the half-plane. Consequently, within the polar angle α

above the half-plane there exists a nonuniform plane wave of the form

$$U = Ae^{-iz\cos(\varphi+\dot{\theta})},$$

and within the same angle α, but below the half-plane, there is another one given by

$$U = e^{-iz\cos(\varphi-\dot{\theta})}.$$

If $\dot{\theta} = i\xi$ holds, then we are concerned with an undamped surface wave directed by a reactive half-plane. Plots of the field corresponding to these two cases are shown in Fig. 3a, b respectively. The same thing applies to the unconfined wave, except that Eq. (19) with the "+" sign now gives us the angular coordinate of the boundary between light and shadow in the lower half-plane and with the "−" sign in the upper half-plane. Consequently, in the upper part of the angular sector the exponential

$$U = e^{-iz\cos(\varphi-\dot{\theta})},$$

is localized, and in the lower it is

$$U = e^{-iz\cos(\varphi+\dot{\theta})}.$$

This change of signs is related to the fact that the amplitude of the unconfined wave inside the localization sector grows if the observation point is shifted perpendicular to the surface of the half-plane. A plot of the field for this type of wave is shown in Fig. 3c, where in particular the exponential amplitude distribution of the wave on the surface is shown as in Fig. 3a, b for different phases by alternating regions with and without crosshatching. The heavy arrow is also used to show the direction in which energy is transported by the wave perpendicular to the surfaces of the different phases. Figure 3 also shows instantaneous snapshots of the real part of the field profile along the light rays, indicated in the figure by means of thin lines with arrows indicating the direction of propagation.

In order to determine the field strength of the quasimode function on the boundary between light and shadow of a nonuniform plane wave, it suffices to substitute in Eq. (15) the corresponding values of the arguments w_+ and w_- from (18) and (19), equal to $\pi/2$ or $3\pi/2$. This means that the argument of the function erfc $w_\pm$ in (15) takes on purely imaginary values here, namely, $m \pm i\varkappa$, where $\varkappa$ is a real positive number. In what follows we will use one additional analytical property of the complementary error function with an imaginary argument, noted in Ref. 5:

$$\operatorname{erfc}(\pm i\varkappa) = 1 \pm i \operatorname{Im} \operatorname{erfc}(i\varkappa),$$

since we are interested in the asymptotic (for $\varkappa \gg 1$) behavior of its imaginary part:

$$\operatorname{Im} \operatorname{erfc}(i\varkappa) \approx e^{\varkappa^2}[(i\sqrt{\pi}\varkappa)^{-1} + o(\varkappa)^{-3}].$$

From this it is not difficult to calculate the "penumbral" (for $\varphi = \alpha$) value of the function (15) in question,

$$U = Ae^{-iz}(\tfrac{1}{2}e^{-\varkappa^2} + f_\pm(z,\alpha)), \tag{20}$$

where the function $f_\pm$ has the asymptotic form of the amplitude of a cylindrical function for $z \gg 1$, i.e., is proportional to $z^{-1/2}$, regardless of whether it is a surface wave or an unconfined wave. However, for us it is important here that since

$$\varkappa^2 = kr\beta(\eta,\xi),$$

where β is a function that depends on the real and imaginary part of the eigenvalue of this mode, the modal (exponential) part of expression (20)

$$U_M = \tfrac{1}{2}Ae^{-i(\omega/c)r}e^{-\chi r} \tag{21}$$

(here χ is a real positive number) clearly implies the existence of complex damped light rays, streaming away from the surface of the half-plane, so to speak (see Fig. 3).

As was done before in diffraction theory using Eq. (12), one can uniquely determine the direction of propagation of the whole set of light rays reflected from the half-plane, although this equation is related to only one of them (the one on the end), so here we can do the same thing: using the end ray we can uniquely reproduce the propagation direction of the parallel rays (propagating with the speed of light), but now damped and complex, as shown in Fig. 3.

It is clear that in both cases (with the complex parameter $\dot{\theta}$ and with real φ_0) the rays in the beam are collinear because the beam itself is related to a plane wave, and it does not matter whether that wave is uniform or nonuniform. It is just for this reason that the location of the boundary ray plays such an important role. But Eq. (12) does more than determine the propagation direction of this ray. It also allows us to make an unambiguous choice between two possible definitions of a light ray.

The first definition, as lines along which energy is propagated, is completely inappropriate here, since it does not agree at all with the principles described above, which are based on the analytical properties of the solutions of the Helmholtz equation written in terms of the functions (6) in the general case of a complex parameter $\dot{\theta}$. These principles are consistent with another definition of a light ray, proposed by myself in Ref. 5 as a line along which the phase velocity of the waves is equal to the velocity of light in the medium (in this case a vacuum) in which the wave process takes place. In fact, this definition, which is (we repeat) a simple consequence of interpreting the analytical properties of the corresponding solution of the Helmholtz equation, satisfies the condition of generality better than the first definition, since it includes the case of a complex parameter $\dot{\theta}$ in a very natural way, and the case of a real parameter φ_0 is simply a limiting case of this, obtained by taking $\chi \to 0$ in Eq. (21). But of course the question very naturally arises as to the basis for the principles which guided us above in applying a new definition to the concept of a light ray based on the analytical properties of the corresponding solution of the wave equation.

Consequently, this question reduces to that of the choice of basis in solving the wave equation. If the solution is the exponential

$$U = e^{-iz}e^{x^2}, \tag{22}$$

then generally speaking there is no a priori argument in favor of choosing one definition for the concept of light rays instead of the other. But if we use a solution in the form

$$U=e^{-iz}e^{x^2}\operatorname{erfc} x, \tag{23}$$

then here Eq. (12) leads to the unique possible definition of light rays as lines along which the phase velocity of a wave is equal to the velocity of light. The resolution of this dilemma in favor of the representation (23) (to which in part the present paper is devoted) plays a central role in the theory of open waveguides, as shown in Refs. 4, 5. It ultimately leads to changes in the physical basis of the current theory of open waveguides, converting it into a standard form in which the only requirement on the mode elements of the field is that they satisfy the ordinary boundary conditions on the interfaces separating the various media. However, in contrast to the theory of closed waveguides, here the concept of a mode refers not to an infinite waveguide but to a different idealized model, that of a semi-infinite waveguide.[4,5] We will come back to this point after proving the formal equivalence of the solutions (22) and (23); evidently we can only make a selection between them as a basis for expanding the field if they are equally valid.

It must be said that we are touching upon a very subtle aspect of the problem here. From a formal point of view all special solutions of the wave equation, of which there are an infinite number for a partial differential equation, are equivalent. Nevertheless, there are serious arguments regarding the degree of generality, based mainly on the properties of the space-time symmetry,[11] about why the exponential solution (22) should be chosen as the basis from among all the possible solutions. In principle, therefore, very weighty reasons must be found in order to replace this basis solution with another.

One of our main purposes here is to find this justification, because the change of basis from (22) to (23), with all the resulting changes in the formal mathematical apparatus, was already demonstrated in Ref. 5, but in a purely heuristic manner.

If we compare the solutions (22) and (23), we readily note that they have homogeneous structure, which is probably not just an accident. But the structure of Eqs. (22) and (23) itself is a clue to the correct solution of this problem.

First, it is quite apparent that the most important property of this structure is that it has a rapidly oscillating phase factor e^{-iz}, so that it is better to separate this right at the start in writing the wave function:

$$U=e^{-iz}V(z,\varphi), \tag{24}$$

where we only know now that it must satisfy the Helmholtz equation (7).

Before going on let us recall that we are interested not only in the analytical side of the matter, but also in the geometrical aspects, because we noticed in the previous discussions that the exponential solution (22) has geometrical properties such that it can be distinguished from among all the other solutions. Since questions of geometrical interpretation can be of particular importance, in treating the solutions of the Helmholtz equation (7) we will relax the usual restrictions on the domain of the independent variables: $r\in[0,\infty]$, $\varphi\in[0,2\pi]$, allowing the angular variable φ to range over the entire real axis: $\varphi\in[-\infty, +\infty]$. Thus, we are offering a broader geometrical interpretation of the solutions of Eq. (7) as purely mathematical entities. We can only compare the resulting solution by restricting the geometrical region of definition to the ordinary two-dimensional geometrical space in which $\varphi\in[0,2\pi]$ holds, and thus draw conclusions about the physical significance of these solutions.

Now substituting expression (24) in Eq. (7) we find an equation for the amplitude V:

$$z^2V_{zz}+z(1-2iz)V_z+V_{\varphi\varphi}-izV=0, \tag{25}$$

which we call the reduced Helmholtz equation. Note that the argument of the amplitude function V, whether taken from Eq. (22) or from Eq. (23) is

$$x=(2iz)^{1/2}\sin\frac{\psi}{2},$$

where we have written $\psi=\varphi+\theta$ and θ is an arbitrary parameter which without loss of generality can be set equal to zero, so that we can write immediately

$$x=(2iz)^{1/2}\sin\frac{\varphi}{2}. \tag{26}$$

We see now that the variable x contains both independent variables r and φ, since $z=kr$. We also see that the form of the group (26) is related to both solutions (22) and (23) equally, so that it thus possesses a certain degree of generality, which immediately suggests using this grouping to create a new independent variable by means of Eq. (26), in which the variables r and φ are combined.

In fact, this is a departure from the classical approach to separation of variables.

3. SELF-SIMILAR SOLUTIONS OF THE REDUCED HELMHOLTZ EQUATION (TWO-DIMENSIONAL CASE)

Thus, taking x [see Eq. (26)] to be the new independent variable for the amplitude function V in the reduced Helmholtz equation (25), we rewrite the terms as

$$z^2V_{zz}=-\tfrac{1}{4}xV_x+\tfrac{1}{4}x^2V_{xx},$$

$$z(1-2iz)V_z=\tfrac{1}{2}xV_x-(iz)xV_x,$$

$$V_{\varphi\varphi}=-\tfrac{1}{4}xV_x-\tfrac{1}{4}x^2V_{xx}+(iz)\tfrac{1}{2}V_{xx}.$$

Now substituting all this in (25) we find

$$iz(V_{xx}-2xV_x-2V)=0,$$

from which we arrive at the ordinary differential equation

$$V_{xx}-2xV_x-2V=0. \tag{27}$$

If we now substitute

$$V=e^{x^2}W,$$

then Eq. (27) goes over to

$$W_{xx}-2xW_x=0, \tag{28}$$

which has the general solution

$$W=A\ \mathrm{erfc}\ x+B, \tag{29}$$

of which we can readily convince ourselves by directly substituting Eq. (29) in (28), recalling that

$$\mathrm{erfc}\ x=\frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}\mathrm{d}t$$

from which it follows by differentiation that

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{erfc}\ x=-\frac{2}{\sqrt{\pi}}e^{-x^2}.$$

Taking erfc x to be an entire function, it is not difficult to see that generally speaking there exist two other ways of writing Eq. (28), both completely equivalent to (29):

$$W=A\ \mathrm{erfc}\ x+B\ \mathrm{erfc}(-x),$$

$$W=A\ \mathrm{erfc}(-x)+B,$$

also containing only entire functions. This property of the solutions of Eq. (28) is due to the existence of the functional property (9) of the complementary error function, noted previously. The way we choose to write the general solution of Eq. (28) has no significance in principle.

Thus, we see that the solution of the two-dimensional Helmholtz equation derived using the self-similar solutions[1)] of the reduced Helmholtz equation (25) assumes the form

$$U=Ae^{-iz\cos\varphi}\mathrm{erfc}\left[e^{i\pi/4}(2z)^{1/2}\sin\frac{\varphi}{2}\right]+Be^{-iz\cos\varphi}, \tag{30}$$

where $(2z)^{1/2}$ is taken to be the absolute value of the square root. Thus the solutions (22) and (23) of the Helmholtz equation are actually set apart among all the other possible solutions because they are written in terms of the corresponding self-similar solutions of the reduced Helmholtz equation (25). In this sense the two solutions are equally interesting to us. And since one of them (the exponential) plays an exceptionally important role in mathematical physics, we now have some grounds to investigate more completely the possibility of extending the role of the second solution. We have in mind using it as a basis solution of the waveguide in terms of which the field can be expanded in two-dimensional problems arising in the same areas of mathematical physics. In order to make explicit the conditions under which the basis solution (22) can be replaced by (23), we must first perform a comparative analysis of their most general geometrical properties. From this point of view we first consider the exponential solution, i.e., the second term in expression (30), which for brevity we will call the general self-similar solution of the Helmholtz equation.

Since the exponential function is an entire function of its argument with period equal to 2π, it has the same period in the variable φ. In a certain context this is its main analytical property, since it enables us to identify the interval $\varphi\in[0,2\pi]$ as the nontrivial domain of definition of this function in the sense that to extend it further introduces no new information. The geometrical interpretation of this property is therefore that the natural geometrical space of this self-similar solution of the wave equation is the whole plane. Note that in this case the natural geometrical space of the solution defined above by purely analytical means coincides with the usual two-dimensional geometrical space, although it is not necessary in general that they coincide. This can easily be seen by treating the first term in Eq. (30) in the same way.

Here we are dealing with a product of entire functions which thus is itself an entire function of φ. As can be seen from expression (30), the argument of each of the factors of this product has its own natural periodicity in φ. The period of the argument of the complementary error function, which is equal to 4π, is twice that of the exponential. Consequently, it is this which determines the period of the product of the functions. From this we deduce that the natural geometrical space of the second self-similar solution (23) is the Riemann surface of two sheets, the upper given by $\varphi\in[0,2\pi]$ and the lower by $\varphi\in[2\pi,4\pi]$, fastened together along the ray $\varphi=0,2\pi(4\pi)$. On the upper sheet of the Riemann surface the solution in question behaves in accordance with Eq. (10) for $z\gg 1$ like a cylindrical wave, diverging (with time dependence $e^{i\omega t}$) from the branch point $z=0$ of the natural geometrical space of this solution. On the lower sheet, as can be seen from combining Eqs. (9) and (10), to this cylindrical wave is added the uniform plane wave $\exp(-iz\cos\varphi)$. Thus, we see that the natural geometrical space of the second self-similar solution is not at all the same as the usual two-dimensional geometrical space where this solution, or rather its partial derivatives beginning in first order, has first-order discontinuities on the polar ray $\varphi=0(2\pi)$. Naturally, this property of the self-similar solution we are studying here may not be reflected in the conditions under which it is applicable physically as a basis solution of the wave equation; this will be discussed further below. For now, rather than summarizing we restrict ourselves to establishing the fact that the transformation (26) which combined the independent variables of the two-dimensional Helmholtz equation actually converts the reduced form (25) into a second-order ordinary differential equation, two particular solutions of which are distinguished among the infinite set of other Helmholtz partial differential equations by their self-similarity.

In order to see whether the transformation (26) also works in three dimensions or whether it is only a coincidence that it works in the two-dimensional case, we must try to extend it to the three-dimensional Helmholtz equation.

4. SELF-SIMILAR SOLUTIONS OF THE THREE-DIMENSIONAL HELMHOLTZ EQUATION

The Helmholtz equation for a scalar wave function for the Hertz vector in a spherical coordinate system has the well-known form

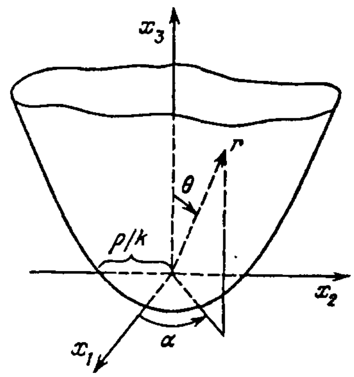

FIG. 4.

$$\frac{\partial}{\partial z}\left(z^2\frac{\partial U}{\partial z}\right)+\sin^{-1}\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial U}{\partial\theta}\right)+\sin^{-2}\theta\frac{\partial^2 U}{\partial\alpha^2}+z^2U=0, \tag{31}$$

where $z=kr$; r, θ, α are the corresponding spherical coordinates (Fig. 4). It is evident that the variable θ plays the role of the polar angular coordinate. Moreover, since the transformation (26), which here takes the form

$$x=(2iz)^{1/2}\sin\frac{\theta}{2}, \tag{32}$$

contains as before only two independent variables, if we want to ascertain its applicability to the three-dimensional differential form of the Helmholtz equation (31), we must restrict ourselves to axisymmetric wave functions U, which are independent of α. In this case the reduced form of the Helmholtz equation for the amplitude function $V(z,\theta)$ related to the wave function U by

$$U=e^{-iz}V(z,\theta),$$

takes the form

$$z^2V_{zz}+2z(1-iz)V_z+V_{\theta\theta}+\frac{\cos\theta}{\sin\theta}V_\theta-2izV=0. \tag{33}$$

Then, introducing a new independent variable x in accordance with Eq. (32), we reduce Eq. (33) in the same way as before to an ordinary differential equation

$$V_{xx}-(2x-x^{-1})V_x-4V=0.$$

After substituting

$$V=e^{x^2}W$$

we finally arrive at the equation

$$W_{xx}+(2x+x^{-1})W_x=0 \tag{34}$$

whose general solution is

$$W=A\mathrm{Ei}(-x^2)+B, \tag{35}$$

as follows readily by direct substitution of (35) in (34) if we recall that

$$\mathrm{Ei}(-x^2)=\int_{-x^2}^{\infty}e^tt^{-1}\mathrm{d}t.$$

Thus, the integral of Eq. (31) constructed using the general self-similar solution of the reduced equation (33), can be expressed as

$$U=Ae^{-iz\cos\theta}\mathrm{Ei}[-iz(1-\cos\theta)]+Be^{-iz\cos\theta}. \tag{36}$$

From this it is clear that the first term of expression (36) is the one of greatest interest to us; it contains a factor in the form of the exponential integral function $\mathrm{Ei}(-x^2)$, whose exponential factor has the same period in θ (equal to 2π) as does the argument $-x^2$ of this function. Everything else therefore follows from its intrinsic analytical properties.

As is well known,[12] in contrast to the complementary error function this function of x has an infinite number of sheets with a single finite branch point $x=0(\theta=0,\pm2\pi,\pm4\pi,...)$, where the function $\mathrm{Ei}(-x^2)$ itself has a logarithmic singularity:

$$\mathrm{Ei}(-x^2)\approx\ln\gamma x(|x^2|\ll1), \tag{37}$$

where γ is the Euler constant.

In the limit $|x^2|\gg1$, the principal value of the function $\mathrm{Ei}(-x^2)$, which we will distinguish from the multivalued function by replacing the letter E by the lower-case letter, has the asymptotic form[12]

$$\mathrm{ei}(-x^2)\approx x^{-2}e^{-x^2}\left(1-\frac{1!}{x^2}+\frac{2!}{x^4}-\frac{3!}{x^6}+...\right). \tag{38}$$

We now consider the behavior of the amplitude function $\mathrm{Ei}[-iz(1-\cos\theta)]$ on the θ axis.

According to (37), this function has a singularity at $\theta=0$. On the interval $\theta\in(0,\pi]$, it follows from Eqs. (36) and (38) that if we agree to take the principal value of the function here it is the amplitude of a spherical wave. On the interval $[\pi,2\pi)$ this function exactly repeats the variation on the reversed interval $(0,\pi]$. This is a consequence of the symmetry of the function $1-\cos\theta$ about the point $\theta=\pi$, and implies that extending the domain of definition of the angular coordinate θ in the spherical system of coordinates to $[0,2\pi]$ does not introduce anything new.[2)] When we intersect the coordinate $\theta=2\pi$, we also intersect the branch point of the function $\mathrm{Ei}(-x^2)$ and therefore wind up on a different sheet of this function for $\theta\in[2\pi,4\pi]$. The transition to the next sheet obviously occurs at $\theta=4\pi$, etc. Exactly the same thing happens at the points $\theta=-2\pi$, -4π,... . Hence the amplitude function, which has the cyclic constant $2\pi i$ (Ref. 12) can thus be represented in the one-dimensional θ continuum broken up into intervals $\theta\in[2\pi m,2\pi(m+1)]$, where $m=0,\pm1,\pm2,\pm3,...$, as

$$\mathrm{Ei}[-iz(1-\cos\theta)]=\mathrm{ei}[-iz(1-\cos\theta)]+2m\pi i. \tag{39}$$

It follows that the nontrivial domain of definition of the first term of the self-similar solution (36) is the whole θ axis. The geometrical interpretation of this is obviously that the natural geometrical space of the self-similar solution of the Helmholtz equation which is of interest to us is the dissected space consisting of a denumerably infinite set of three-dimensional spaces, fastened together at the polar rays $\theta=0, \pm 2\pi, \pm 4\pi, \ldots$.

This dissected space (a kind of Riemann volume) is the three-dimensional analog of a Riemann surface with an infinite number of sheets. The corresponding self-similar solution is continuous in precisely this space, together with all its partial derivatives of arbitrary order. This is the main difference in a nutshell between this self-similar solution and the second exponential form, which has a natural geometrical space that coincides with the ordinary three-dimensional geometrical space. What links them is that they comprise a pair of self-similar solutions of a second-order differential equation, which are thereby set apart among the infinite set of other solutions of the three-dimensional Helmholtz differential equation, regarded as a partial differential equation.

Ideas like this suggest the possibility of finding some other grouping, different from (32), which would combine all three independent variables. This would convert the three-dimensional Helmholtz equation into an ordinary differential equation, differing from Eq. (34) and having different self-similar solutions. Be that as it may, we are interested here in the question of the conditions for the applicability of these self-similar solutions in physical applications.

5. PROPERTIES OF THE TWO-DIMENSIONAL SELF-SIMILAR SOLUTIONS IN APPLICATIONS

In what follows we will refer to the exponential solution as the simple self-similar solution, and the nonexponential solution as the compound solution, bearing in mind that it consists of factors.

It is clear that utilization of both these solutions in physical applications presumes above all the presence of a three-dimensional geometrical space (the geometrical space of the observer). In this case the two-dimensional self-similar solutions should be regarded as solutions of the three-dimensional Helmholtz equation written in a cylindrical coordinate system such that the wave function U is taken (for some reason) to be independent of the coordinate z' in the system of coordinates z', r, φ (see Fig. 1a). All the other properties when these solutions are used will probably be determined by the analytical (and by the geometrical) properties of each.

We first note a property which is common to both solutions. This is the fact that solutions of the form

$$U_{1,2}=A_{1,2}(\gamma)e^{-iz\cos(\varphi\pm\gamma)}\operatorname{erfc}\left[(2iz)^{1/2}\sin\frac{\varphi\pm\gamma}{2}\right], \tag{40}$$

$$\widetilde{U}_{1,2}=\widetilde{A}_{1,2}(\gamma)e^{-iz\cos(\varphi\pm\gamma)} \tag{41}$$

with a nonzero complex parameter γ in the general case are the same self-similar solutions as those with $\gamma=0$. As was already emphasized, this is related to the fact that the metric coefficients of a polar or cylindrical coordinate system do not depend on the angular variable.

But with this the properties that the simple and compound solutions share in common end. Everything else will evidently depend precisely on the differences in the analytical and geometrical properties of these two solutions, and these differences, as already mentioned, are quite marked.

If the simple self-similar solution in the space of the observer, where $\varphi\in[0,2\pi]$ holds, together with its partial derivatives is continuous, then the compound solution, or rather its partial derivatives, exhibits first-order discontinuities in the coordinate half-plane $\varphi=0(2\pi)$. As noted previously, this is related to the structural features of the natural geometrical space of this solution. Of course, this property of the compound solution must be taken into account when it is utilized in applications.

The most natural step now would be to impose some boundary conditions on this half-plane. These would introduce an element of physical justification associated with the properties of the wave function to the corresponding discontinuities that manifest themselves on this half-plane. After these boundary conditions are imposed, it can be used to model some material object (e.g., a "thin" semiinfinite sheet of metal or insulator).

Thus, the main prerequisite for using the compound self-similar solution as the basic solution of the wave equation in applications is a need to formulate some sort of boundary-value problem on the half-plane. Here the idea is that, exactly as in the most general case of arbitrary boundary-value geometry, the full system of functions is constructed using the simple self-similar solution in the form of a mixed wave spectrum:

$$\begin{aligned}\widetilde{U}= & \sum_k A_k\exp[-iz\cos(\varphi+\gamma_k)]\\ & +\sum_k B_k\exp[iz\cos(\varphi+\delta_k)]+\int A(\gamma)\\ & \times\exp[-iz\cos(\varphi+\gamma)]d\gamma+\int B(\delta)\\ & \times\exp[iz\cos(\varphi+\delta)]d\delta,\end{aligned} \tag{42}$$

Now we have at least formal grounds to construct a full system of boundary-value functions on the half-plane using the compound self-similar solution

$$\begin{aligned}U= & \sum_k A_k\exp[-iz\cos(\varphi+\gamma_k)]\\ & \times\operatorname{erfc}\left[(2iz)^{1/2}\sin\frac{\varphi+\gamma_k}{2}\right]\\ & +\sum_k B_k\exp[iz\cos(\varphi+\delta_k)]\\ & \times\operatorname{erfc}\left[(2iz)^{1/2}\cos\frac{\varphi+\delta_k}{2}\right]+\int A(\gamma)\\ & \times\exp[-iz\cos(\varphi+\gamma)]\\ & \times\operatorname{erfc}\left[(2iz)^{1/2}\sin\frac{\varphi+\gamma}{2}\right]d\gamma\end{aligned}$$

$$+\int B(\delta)\exp[iz\cos(\varphi+\delta)]$$

$$\times\mathrm{erfc}\left[(2iz)^{1/2}\cos\frac{\varphi+\delta}{2}\right]d\delta. \quad (43)$$

The specific form of the representation (42) in any particular case depends on the details of the boundary-condition geometry, the mode of excitation, the spatial symmetry, and other physical factors. Thus, in the initial representation (42) one, two, or even three of the four terms can vanish. In the same way, the final form of the representation (43) can undergo similar transformations.

We present one very simple example from electromagnetic wave diffraction theory. Note that in the cylindrical coordinate system, when the potential U is independent of z', we can use it to represent either the electric field component $E_{z'}$ or the magnetic field component $H_{z'}$, depending on the polarization of the field.

Suppose now that a TM-polarized plane wave with unit amplitude is incident on a perfectly conducting infinite sheet of metal at some angle φ_0 (see Fig. 1b). If we leave out the time-dependent factor $e^{i\omega t}$ the solution of this problem in cylindrical coordinates will obviously have the form

$$E_{z'}=\exp[iz\cos(\varphi-\varphi_0)]-\exp[iz\cos(\varphi+\varphi_0)].$$

We see that it is composed of only two terms from the second sum of Eq. (42) for $\gamma_1=-\varphi_0$, $\gamma_2=+\varphi_0$, $B_1=1$, $B_2=-1$. Now we solve almost the same problem, but with half the metal plane "cut off" so that the electric vector of the incident plane wave is parallel to the edge thus produced.

We note that the four terms in expressions (42) and (43) are similar to one another in form, taken in pairs. Just as before, we make use of this by assuming that the solutions of the boundary-value problem require only the two corresponding terms of the sum from the second term of expression (43):

$$E_{z'}=\frac{1}{2}e^{iz\cos(\varphi-\varphi_0)}\mathrm{erfc}\left[e^{i\pi/4}(2z)^{1/2}\cos\frac{\varphi-\varphi_0}{2}\right]$$

$$-\frac{1}{2}e^{iz\cos(\varphi+\varphi_0)}\mathrm{erfc}\left[e^{i\pi/4}(2z)^{1/2}\cos\frac{\varphi+\varphi_0}{2}\right]. \quad (44)$$

There is one minor difference compared with the problem involving a plane treated above: here we have taken $B_1=1/2$, $B_2=-1/2$, which is simply related to the normalization of the complementary error function.

Thus, we see that the solution of (44) is the classical Sommerfeld solution.[1] However, here it is obtained in the form of a trivial set of discrete terms from the expansion in the compound self-similar solution. In general, this problem has to be solved (like any boundary-value problem) by starting with the complete solution in the form of a superposition of incident and scattered waves:

$$E_{z'}=e^{iz\cos(\varphi-\varphi_0)}$$

$$+B_1e^{iz\cos(\varphi-\varphi_0)}\mathrm{erfc}\left[(2iz)^{1/2}\cos\frac{\varphi-\varphi_0}{2}\right]$$

$$+B_2e^{iz\cos(\varphi+\varphi_0)}\mathrm{erfc}\left[(2iz)^{1/2}\cos\frac{\varphi+\varphi_0}{2}\right]. \quad (45)$$

It is not hard to see that when the expression (9) is taken into account, the boundary condition $E_{z'}(\varphi=0,2\pi)=0$, which applies to expression (45), yields the solution (44). The solution for the problem of TE-polarized plane waves is obtained in exactly the same way.

Of course, we can always write the scattered field in the form of an integral of plane waves, as is done everywhere.[1-3] That is, we can use the integral terms in the expansion in the simple self-similar solution (42). Then we can solve the integral equation by means of elaborate and fairly tedious mathematical techniques like the Wiener–Hopf method.[3] The alternative path, which we have demonstrated here, involves using an already available expansion in the compound self-similar solution (43). This reduces the process of obtaining a solution of the problem with a half-plane to a process which is literally trivial. To say more than this regarding the methodological advantages of the expansion (43) would be, in our opinion, completely superfluous.

We have already mentioned the discrete terms in the expansion (43) with complex parameters at the very beginning of this treatment. There is consequently some obscurity regarding the integral terms in Eq. (43). To dissipate it we must return again to the question of describing the characteristic waves of the semi-infinite "thin" unscreened waveguide. For this a natural model is the half-plane with the corresponding (two-sided) boundary conditions that we have indicated. In connection with the problem when this half-plane is a "thin" semi-infinite plane insulating waveguide, in Ref. 5 we have shown that the integral terms in the expansion (43) play the role of quasimode functions of the continuous wave spectrum of a semi-infinite waveguide in the theory of open waveguides. The modes of this spectrum are a continuum of light rays at grazing incidence along the surface (both upper and lower) of the half-plane. The edge where the half-plane is cut plays the role here of an exciting "inhomogeneity." The point is that, as shown in Ref. 5, the integral terms in (43) contain modes of the continuous wave spectrum of an open waveguide only when the angle α (the polar angle localizing the nonuniform plane wave directed by the half-plane; see Fig. 3) goes to zero. This case is illustrated graphically in Fig. 3d. There one of the light rays of the continuum (and an instantaneous snapshot of the real part of the field associated with this ray) is shown. It consists of light rays for which the direction of the field is described by Eq. (21). This is the integrand of an indefinite integral defined on $\chi\in[0,\infty]$, where χ is the variable of integration entering in expression (21). Note that the direction in which the light rays propagate, corresponding to grazing incidence along the surface of the half-plane, is generally the limiting value of all the possible directions shown in Fig. 3a–c. (This assumes that we have in mind deterministic relations between the currents of the half-plane and the light rays.$^{3)})^4$ An extraordinary property of the expansion (43) is precisely that in principle it does not contain modes with light rays "flowing in" to the surface of the half plane. This cannot, of course, be said of the representation (42). In the theory of open waveguides this difference in the representation plays a fundamental role. In fact, in deciding the dilemma of how to choose between the basic representations (42) and (43) in connection with the theory of open waveguides, we are essentially solving the problem of choosing an idealized physical model for an infinite or semi-infinite waveguide, respectively. If we choose the latter, then these waves (according to the current terminology[9] associated with the infinite waveguide model), both characteristic (exponentially damped at infinity in the transverse direction) and slow noncharacteristic (exponentially growing at infinity in the transverse direction), the light rays of which "flow into" the surface of the half-plane are automatically eliminated from consideration. We note in passing that the so-called Zenneck wave (the fast eigenmode) once predicted on the basis of the representation (42), has not been observed thus far, although its field satisfies all the conditions of the theory, for which the infinite waveguide serves as the basic model. Thus, the representation (43) contains from the start internal information reducing the number of additional possibilities allowed by the expansion (42). In particular, the presence of these extra possibilities implies that the theory of open waveguides formulated for infinite waveguide modes is also redundant. The physical basis of this theory, in addition to imposing the standard boundary conditions, contains conditions such as that at infinity (transverse or radial). Despite this (or rather, because of this) the theory lacks clear criteria according to which the spurious solutions, both characteristic and noncharacteristic, could be eliminated *a priori*. In contrast, with (43) chosen as the basic representation this problem is solved in a natural way irrespectively of what mode is being treated, characteristic or noncharacteristic. The only requirement imposed on the mode of the semi-infinite waveguide is the standard requirement that it satisfy the boundary conditions for this field. Thus, the semi-infinite waveguide model is preferable to that of the infinite waveguide as regards the derivation of an "economical" theory of open waveguides. Over and above this, as shown in Ref. 5, another consequence of reducing the physical basis of the theory to standard form is the relationship of the discrete and continuous spectral components of the scattered field. This means that the present theory of open waveguides is logically complete.

The only thing left to emphasize is that all these consequences stem from the simple analytical properties (9) and (10) of the compound self-similar solution of the Helmholtz equation. These are properties which increase the information content of the mathematical apparatus built on this solution in comparison with the conventional formalism, which is based on the simple (exponential) solution of the wave equation.

6. USE OF THE THREE-DIMENSIONAL SELF-SIMILAR SOLUTIONS IN APPLICATIONS

Although the self-similar solutions have been derived by a single technique without regard to the number of independent variables in the Helmholtz equation, the way these solutions are utilized in applications depends substantially on this point. For example, because the metric coefficients of the spherical coordinate system depend on the variable θ, functions constructed using formulas like (40) and (41) no longer have any meaning, since they do not satisfy the Helmholtz equation. Consequently, questions regarding the use in applications (in the theory of electromagnetic or acoustic waves) of the compound self-similar solution

$$U=Ae^{-iz\cos\theta}\,\mathrm{Ei}[-iz(1-\cos\theta)], \tag{46}$$

where $\mathrm{Ei}(-x^2)$ is the principal value of the exponential integral function, and $\theta\in(0,\pi]$ are resolved in a very different manner than in the two-dimensional analog.

Next, it makes no difference what the specific application of expression (46) is, whether as a Debye potential used in describing electromagnetic oscillations or as a velocity potential in the theory of sound waves. The question of the conditions under which this expression can be applied in the corresponding applications can probably be answered after one answers the question, what specifically are the surfaces of equal amplitudes of the wave described by the potential (46).

The relation (46) implies that these surfaces are specified by

$$z(1-\cos\theta)=\rho, \tag{47}$$

where ρ is a real parameter. It is not difficult to see that Eq. (47) describes the surface of a paraboloid of revolution (shown in Fig. 4) written in the spherical coordinate system. Here the parameter ρ acts as the focal parameter of the paraboloid (written in units of the wavelength).

From this it follows that Eq. (46) can be used most effectively to describe the space–time distribution of the field in acoustic or electromagnetic wave theory in boundary-value problems with the boundary condition applied to the surface of a paraboloid of revolution, except for the interior parabolic regions of the space containing the axis $\theta=0$, where the potential (46) has a singularity.

It is well known, however, that problems like these can be solved most readily in parabolic coordinates. These consist of a system of confocal paraboloids of revolution (coordinate surfaces) with focus at the origin of coordinates. Figure 5 shows the intersection of such a coordinate system with the x_2x_3 plane. It is easy to show that the parabolic coordinates are related to spherical coordinates by the following expressions:

$$\sigma^2=r(1-\cos\theta),\ \ \tau^2=r(1+\cos\theta),\ \ \alpha'=\alpha. \tag{48}$$

In addition we have

$$r\cos\theta=\tfrac{1}{2}(\tau^2-\sigma^2). \tag{49}$$

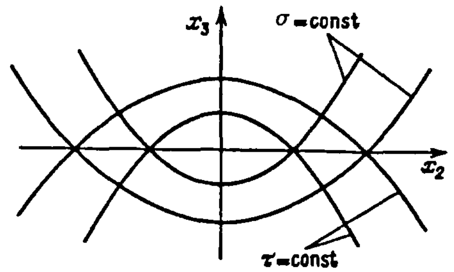

FIG. 5.

Substituting (48) and (49) into (46) we can easily obtain an expression for the potential U in the parabolic coordinate system:

$$U=e^{-i(k/2)(\tau^2-\sigma^2)}\mathrm{Ei}(-ik\sigma^2), \quad (50)$$

which, as can be seen by inspection, can be derived by the conventional method of separation of variables in the Helmholtz equation written in parabolic coordinates[13]

$$\sigma^{-1}\frac{\partial}{\partial\sigma}\left(\sigma\frac{\partial U}{\partial\sigma}\right)+\tau^{-1}\frac{\partial}{\partial\tau}\left(\tau\frac{\partial U}{\partial\tau}\right)+k^2(\sigma^2+\tau^2)U=0, \quad (51)$$

where we have omitted the dependence on α'.

The general theory of solutions of Eq. (51) was derived by Fok, who published some of his results in Ref. 13, where expression (50) was obtained as one of the particular solutions of Eq. (51). The same paper also indicated ways of using these solutions in the theory of electromagnetic wave diffraction by an ideal conducting paraboloid of revolution. However, Horton and Karal[6] gave the most direct approach to applications of the solution (50). In particular, just using the potential (50) they found a rigorous closed expression for a plane acoustic wave scattered by a paraboloid of revolution with a rigidly attached surface, where the wave is incident on the external (convex) surface of the paraboloid in the paraxial direction.

It is easy to see that the solution of this problem with boundary conditions

$$\left.\frac{\partial U}{\partial\sigma}\right|_{\sigma=(\rho')^{1/2}}=0,$$

where we have written $\rho'=\rho/k$ (the focal parameter of the paraboloids expressed in units of wavelengths) takes the form

$$U=\exp\left[-i\frac{k}{2}(\tau^2-\sigma^2)\right][1+\Gamma\mathrm{Ei}(-ik\sigma^2)],$$

$$\Gamma=i\rho/[2\exp(-i\rho)-i\rho\mathrm{Ei}(-i\rho)].$$

7. CONCLUSION

The general methodological significance of this work is obviously that it demonstrates a way of solving the Helmholtz equation which converts it into an ordinary differential equation. This gives rise to a pair of so-called self-similar solutions, which are distinguished from among the denumerably infinite set of other solutions of the wave equation, viewed as a partial differential equation, by precisely this fact. But this is only the superficial, formal side of the matter. The more important thing is that this implies consequences which are far from formal. This can be seen particularly clearly in the two-dimensional examples. The main result there reduces to demonstrating those analytical properties of the compound self-similar solution, used as the basic solution of the wave equation in order to model the wave processes in the theory of open waveguides, which result in quite general improvements (relative to the conventional form of this theory); this is practical evidence for the distinctiveness of the self-similar solutions.

We have shown that in the three-dimensional case the compound self-similar solution is not as general as its two-dimensional analog. This, however, is not a basis on which to draw final conclusions. It simply means, in our opinion, that further studies are necessary, whose direction may be very different from those carried out in this work, which is methodological in nature.

In this connection we note only that the direction of such studies cannot be entirely arbitrary. It must be determined from the properties of the three-dimensional solution; these, as shown here, differ very substantially from those of the two-dimensional analog.

[1)] The term "self-similar solution" is adopted from hydrodynamics[10] from the formal appearance of self-similarity, where the mathematical description of a physical process reduces to a problem with a single independent variable.

[2)] For spherical coordinates this extension simply means that if θ runs over the whole interval $[0,2\pi]$, then the end of a unit radius vector with $[0,2\pi]$ describes the whole sphere twice. The multivaluedness of the representation of the spatial distribution does not stem from the explicit symmetry of the factor $1-\cos\theta$ which appears in the argument of the function $\mathrm{Ei}(-x^2)$.

[3)] As shown by Kukushkin,[4,5] the existence of deterministic relations between the currents in the half-plane (if it is a semi-infinite "thin" insulating plate,[5] then the term currents here should be taken to mean the corresponding polarization currents in the insulator) and the light rays plays a central role in the development of the very concept of modes of a semi-infinite waveguide. In the theory of shielded waveguides, e.g., this concept is meaningless precisely because here the principle of determinacy applied to the modes of the field of an infinite waveguide never fails. This is because there are two planes, rather than one, which are constantly "exchanged" by the two reflected light rays. By virtue of this the presence of "incoming" rays here does not contradict the deterministic relations that hold in electrodynamics.

[1] A. Sommerfeld, F. Frank, and R. von Mieses, *Differential and Integral Equations of Mathematical Physics*, Springer, Berlin (1930). [Russ. transl., ONTI, L., M., 1937].

[2] P. Ya. Ufimtsev, *The Method of Boundary Waves in the Physical Theory of Diffraction* [in Russian], Sov. Radio, M., 1962.

[3] L. A. Vaĭnshteĭn, *Diffraction Theory and the Factoring Method* [in Russian], Sov. Radio, M., 1966.

[4] A. V. Kukushkin, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **33**, No. 10, 1138 (1990) [Sov. Radiophys. **33**, (1990)].

[5] A. V. Kukushkin, *ibid.* **33**, No. 11, 1242 (1990). [*ibid.*, **33** (1990)].
[6] C. W. Horton and F. C. Karal, J. Acoust. Soc. Am. **22**, No. 6, 855 (1950).
[7] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, GPO, Washington (1964). [Russ. transl., Nauka, M., 1979].
[8] Y. L. Luke, *Mathematical Functions and Their Approximations*, Academic Press, N.Y., 1975 [Russ. transl., Mir, M., 1980].
[9] V. V. Shevchenko, *Continuous Transitions in Open Waveguides*, Golem, Boulder, CO., 1971.
[10] *Encyclopedic Dictionary of Physics* [in Russian], Sov. Éntsiklop., M., 1983.
[11] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 2nd ed., Pergamon, Oxford, 1982.
[12] E. Jahnke, F. Emde, and F. Lösch, *Tables of Higher Functions*, 6th ed., McGraw-Hill, N.Y., 1960. [Russ. transl., Nauka, M., 1977].
[13] V. A. Fok, "Theory of diffraction from a paraboloid of revolution," in *Diffraction of Electromagnetic Waves by Some Bodies of Revolution* [in Russian], Sov. Radio, M., 1957.

Translated by D. L. Book