Limits of applicability of the method of geometric optics and related problems[1)]

Yu. A. Kravtsov and Yu. I. Orlov

P. N. Lebedev Institute of Physics of the Academy of Sciences of the USSR and Moscow Energetics Institute

Usp. Fiz. Nauk **132**, 475–496 (November 1980)

We present on a heuristic basis the universal sufficient conditions for applicability of the method of geometric optics. In formulating the criteria, we make essential use of the concept of the "Fresnel volume" of a ray, whose boundary links the first Fresnel zones "threaded" on the ray. The fundamental criterion of applicability requires that the parameters of the medium and of the wave should vary little over the transverse section of the Fresnel volume. The second criterion, which stems from the first, requires that rays incident on the same given point should lie mostly outside the Fresnel volume of an adjacent ray. The effectiveness of these criteria has been demonstrated in many problems of electrodynamics and acoustics that allow a solution more precise than the ray solution. On the basis of the presented criteria, one can reveal the regions of inapplicability of the ray method (focal and caustic regions, penumbra regions in diffraction by screens and convex bodies, regions where lateral waves arise, etc.). If we know the dimensions of the regions of inapplicability, we can also solve a number of related problems. The most important of these problems are: the problem of determining the field in the neighborhood of caustics and foci and the problem of analyzing the wave pattern as a whole. The proposed criteria also allow a generalization to three-dimensional quantum-mechanical problems, while outlining the limits of applicability of the quasiclassical approximation.

PACS numbers: 42.10.Dy

CONTENTS

1. INTRODUCTION

a) Purpose of the article

The method of geometric optics is an effective instrument for finding wave fields under conditions of smoothly inhomogeneous and slowly nonstationary media. The geometric-optical approach rests substantially on rays, which play the role of a skeleton that bears the wave field.

In spite of the extremely broad application of the method of geometric optics, which sometimes appears as "geometric acoustics" or "geometric seismics", and which has a quantum-mechanical "twin brother" in the quasiclassical approximation, thus far the limits of applicability of this method in three-dimensional problems have not yet been elucidated.[2)] Below we shall try to formulate universal sufficient conditions for applicability of the method on a heuristic basis that rests on the Huygens-Fresnel picture of the formation of the wave field. The proposed heuristic criteria for applicability also enable one to solve a number of related problems (see Sec. 4 and the Conclusion).

b) Fundamental equations of the method of geometric optics

The method of geometric optics is based on the assumption that the properties of the medium, as characterized by the refractive index $n(\mathbf{r})$, and the parameters of the ray field:

$$u = Ae^{ik_0\Psi}, \quad k_0 = \frac{\omega}{c}, \tag{1.1}$$

vary smoothly on the scale of wavelengths in the medium, which is $\lambda = \lambda_0/n$.

[1)]This article is written from the materials of a lecture[1] given by the authors at the 5th All-Union School on Diffraction and Propagation of Waves (Chelyabinsk, 1979).

[2)]The conditions for applicability of the geometric-optical approximation for one-dimensional problems, i.e., essentially the conditions for applicability of the WKB approximation, have been studied in many articles. These conditions reduce to the requirement that the refractive index n should vary little over one wavelength in the medium: $\lambda\,(dn/dz) \ll n$, where $\lambda = \lambda_0/n$, and $\lambda_0 = 2\pi c/\omega$ is the wavelength in a vacuum. V. L. Ginzburg[34] has made a complete physical analysis of the problem with account taken of the zeros and poles of the refractive index.

 0038-5670/80/110750-13$01.10

The generally adopted expansion of the amplitude A in (1.1) in an asymptotic series in powers of $1/ik_0$ is[3]:

$$A=A_0+\frac{1}{ik_0}A_1+\frac{1}{(ik_0)^2}A_2+\ldots=\sum_{m=0}^{\infty}\frac{A_m}{(ik_0)^m}.$$

Substitution of this expansion into the Helmholtz equation $\nabla^2 u+k_0^2n^2u=0$ enables one to determine the sought amplitudes A_m and the eikonal ψ of the wave:

$$\psi=\psi^0+\int_0^{\tau}n^2\,d\tau=\psi^0+\int_0^{\sigma}n\,d\sigma. \quad (1.2)$$

$$A_m=\frac{A_m^0}{\sqrt{\mathscr{Y}}}-\frac{1}{2\mathscr{Z}(\tau)}\int_0^{\tau}\sqrt{\mathscr{D}(\tau')}\,\nabla^2 A_{m-1}\,d\tau. \quad (1.3)$$

Here we have $\psi^0\equiv\psi(0)$, $A_m^0\equiv A_m^0(0)$ as the initial values of the eikonal and the amplitudes at $\tau=0$. The integration in (1.2) and (1.3) is performed along the rays, as defined by the system of equations

$$\frac{d\mathbf{r}}{d\tau}=\mathbf{p},\quad \frac{d\mathbf{p}}{d\tau}=\frac{1}{2}\nabla n^2, \quad (1.4)$$

Here we have $\mathbf{p}=\nabla\psi$, while τ is a parameter of the ray, which is associated with the arc length σ of the ray by the equation $d\tau=d\sigma/n$.

The quantity $\mathscr{Y}$ in (1.3), which is called the divergence of the rays, is equal to $\mathscr{D}(0)/\mathscr{D}(\tau)$, where $\mathscr{D}(\tau)=\partial(x,y,z)/\partial(\xi,\eta,\tau)$ is the Jacobian of the transformation from the Cartesian coordinates x, y and z to the ray coordinates ξ, η, and τ. One can easily calculate the Jacobian $\mathscr{D}(\tau)$ if one knows the equations of the family of rays $\mathbf{r}=\mathbf{r}(\xi,\eta,\tau)$ (the variables ξ and η specify the "numbers" of the rays). The divergence $\mathscr{Y}$ of the rays is proportional to the ratio of the cross-sections ds of the elementary ray tube

$$\mathscr{Y}=\frac{n^0I^0}{nI}=\frac{n^0\,ds^0}{n\,ds},\quad I\equiv\frac{ds}{d\xi\,d\eta}. \quad (1.5)$$

Here the superscript "0" corresponds to $\tau=0$. Additional information on the fundamental equations of geometric optics is contained, e.g., in Refs. 4–6 and 35.

c) The necessary conditions for applicability of geometric optics

Usually one restricts the treatment solely to the zero-order approximation

$$u_0=A_0e^{ik_0\psi}=\frac{A_0^0e^{ik_0\psi}}{\sqrt{\mathscr{Y}}}. \quad (1.6)$$

Then one takes the limits of applicability of geometric optics to mean the limits of applicability of precisely the zero-order approximation (1.6).

The necessary conditions for applicability of geometric optics in (1.2) require the absence of sharp variations of the amplitude A, a sufficient smoothness of the phase fronts, and a sufficient slowness of variation of the refractive index n per wavelength:

$$\lambda\,|\nabla A_0|\ll A_0,\quad \lambda\,|\nabla p_j|\ll p,\quad \lambda\,|\nabla n|\ll n,\quad \lambda\equiv\frac{1}{k_0n}. \quad (1.7)$$

Owing to (1.5) and (1.6), the first of these conditions limits the rate of variation of the divergence of the rays $\lambda|\nabla I|\ll I$, while the first and second conditions imply the inequality

$$\lambda\ll|R_{1,2}|. \quad (1.8)$$

Here $R_{1,2}$ are the principal radii of curvature of the phase front.

d) Cumulative errors

The zero-order approximation of geometric optics satisfies the Helmholtz equation only approximately, to an accuracy of terms of the order of $\mu^2=(1/k_0nL)^2$. Let us assume that

$$u=u_0+\tilde{u}=A_0e^{ik_0\psi}+\tilde{u}. \quad (1.9)$$

Here $\tilde{u}$ is the correction to the geometric-optical field u_0. Then we can easily show that $\tilde{u}$ satisfies the equation

$$\nabla^2 u+k_0^2n^2\tilde{u}=-e^{ik_0\psi}\Delta A_0. \quad (1.10)$$

Here $\nabla^2 A_0$ is of the order of μ^2 with respect to $\nabla^2 u_0$ or to $k_0^2n^2u_0$.

In spite of the smallness of the closure error $\nabla^2 A_0e^{ik_0\psi}$, it can lead to cumulative errors at large distances that involve *diffraction effects*.[4] This is just why the inequality $\mu\sim 1/k_0L\ll 1$ and the conditions (1.7) and (1.8) associated with it prove to be only necessary, but not sufficient, conditions for applicability of geometric optics. Yet the sufficient conditions must in some way reflect the cumulative errors. In approaching the formulation of the sufficient conditions for applicability, we shall discuss the ways of constructing the Fresnel zones in an inhomogeneous medium, and shall introduce the Fresnel volume as a key concept for further constructions.

2. FRESNEL VOLUMES OF RAYS IN INHOMOGENEOUS MEDIA

a) The Fresnel volume. The physical content of the concept of a "ray"

In line with the Huygens-Fresnel principle, the field at the point of observation is formed by the interference of the secondary waves that are generated by the primary wave at each point of the phase front.[5] The decisive role here belongs to the first Fresnel zone, since the secondary waves from the first Fresnel zone differ mutually in phase by no more than π. Therefore they do not cancel one another, whereas the cumulative contribution from the higher Fresnel zones is very small, owing to the addition of oscillations opposed in phase.

We shall call the volume bounded by the envelope of

---

[3]We recall that one actually performs this expansion in the dimensionless small parameter $\mu\sim 1/k_0nL$ (L is the characteristic scale of the variation of A and n), as was first demonstrated by S. M. Rytov (see Refs. 2, 3, and also 4, 35).

[4]Properly speaking, all phenomena that lead to deviations from the laws of geometric optics are commonly called diffraction phenomena.

[5]Owing to the arbitrary choice of the phase front, the excitation of the secondary waves actually occurs at every point of space. As we know, the Huygens concept has been systematically realized by R. P. Feynman by using path-integrals.[7]

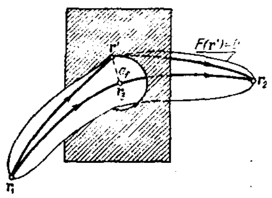

FIG. 1.

the first Fresnel zones the *Fresnel volume*.[6)]

The Fresnel volume marks out the portion of space that can be naturally considered to be the region of localization of a ray treated as a physical object. In fact, if we wish to single out a given ray by passing the wave through an aperture in a screen (Fig. 1), then the dimensions of the aperture must be greater than the cross-section of the Fresnel volume. If we narrow the aperture to dimensions smaller than the dimension of the first Fresnel zone, then this distorts the field associated with the given ray. Thus the Fresnel volume defines the region of space that gives rise to the wave field at a given point. It constitutes the physical content of the concept of a "ray". For wavelengths that are finite (though small, $\lambda \ll L$), the physical ray has a *finite thickness*, in contrast to the mathematical ray $\mathbf{r}=\mathbf{r}(\tau)$, which amounts to an infinitesimally thin line in space.

b) The equation of the boundary of the Fresnel volume

Let us construct the Fresnel volume for a point source in an inhomogeneous medium. Let the source lie at the point $\mathbf{r}_1$ and the observer at the point $\mathbf{r}_2$ (see Fig. 1). We shall call the ray $\mathbf{r}_1\to\mathbf{r}_2$ that joins the points $\mathbf{r}_1$ and $\mathbf{r}_2$ the *reference ray*. In the neighborhood of the reference ray $\mathbf{r}_1\to\mathbf{r}_2$, let us draw the *two-segment virtual* ray $\mathbf{r}_1\to\mathbf{r}'\to\mathbf{r}_2$, which consists of two segments of rays that satisfy the equations (1.4). Such a ray corresponds to the Huygens secondary waves excited at the point $\mathbf{r}'$. We denote by $\psi(\mathbf{r}_a, \mathbf{r}_b)$ the optical path (eikonal) along the ray that joins the arbitrary points $\mathbf{r}_a$ and $\mathbf{r}_b$:

$$\psi(\mathbf{r}_a, \mathbf{r}_b)=\int\limits_a^b n\,d\sigma. \tag{2.1}$$

The boundary of the Fresnel volume is the surface $F(\mathbf{r}')=0$, which includes the points $\mathbf{r}'$ such that the optical path $\psi_{\text{virt}}=\psi(\mathbf{r}_1,\mathbf{r}')+\psi(\mathbf{r}',\mathbf{r}_2)$ calculated along the virtual ray $\mathbf{r}_1\to\mathbf{r}'\to\mathbf{r}_2$ differs from the optical path $\psi_{\text{ref}}=\psi(\mathbf{r}_1,\mathbf{r}_2)$ along the reference ray by one half of the wavelength $\lambda_0=2\pi/k_0=2\pi c/\omega$ (the corresponding phase advances $k_0\psi_{\text{ref}}$ and $k_0\psi_{\text{virt}}$ differ by π). This surface amounts to the envelope of the first Fresnel zones threaded onto the reference ray. Its equation can be represented in the following form (see Fig. 1):

$$F(\mathbf{r}')=|\psi_{\text{virt}}-\psi_{\text{ref}}|-\frac{\lambda_0}{2}=|\psi(\mathbf{r}_1,\mathbf{r}')+\psi(\mathbf{r}',\mathbf{r}_2)-\psi(\mathbf{r}_1,\mathbf{r}_2)|-\frac{\lambda_0}{2}=0. \tag{2.2}$$

Similarly, one constructs the Fresnel volume for another formulation of the problem, namely, when the eikonal ψ^0 is given on some surface S. In this case the equation of the envelope of the first Fresnel zones is written in the form

$$\begin{aligned}F(\mathbf{r}')&=|\psi_{\text{virt}}-\psi_{\text{ref}}|-\frac{\lambda_0}{2}\\&=|[\psi^0(\mathbf{r}_1')+\psi(\mathbf{r}_1',\mathbf{r}')+\psi(\mathbf{r}',\mathbf{r}_2)]\\&\quad-[\psi^0(\mathbf{r}_1)+\psi(\mathbf{r}_1,\mathbf{r}_2)]|-\frac{\lambda_0}{2}=0.\end{aligned} \tag{2.3}$$

Here $\mathbf{r}_2$ is the point of observation, $\mathbf{r}_1\to\mathbf{r}_2$ is the reference ray, and $\mathbf{r}_1'\to\mathbf{r}'\to\mathbf{r}_2$ is the virtual ray (Fig. 2).

We note that the Fresnel volumes of the higher-order virtual rays (e.g., three- or four-segment) usually lie inside the surface $F(\mathbf{r}')=0$. Hence we can neglect them.

c) Approximate equation of the boundary of the Fresnel volume

In actually finding the Fresnel volume, we can employ the smallness of the wavelength as compared with all the characteristic distances in some given problem. This enables us to expand the eikonals in Eqs. (2.2) and (2.3) in terms of the deviations $\mathbf{r}'$ from the reference ray. If $\mathbf{r}_3$ is the point on the reference ray closest to $\mathbf{r}'$ (see Figs. 1 and 2), then we have the following expression, apart from quadratic terms in $\mathbf{r}'-\mathbf{r}_3$:

$$F(\mathbf{r}')\approx\frac{1}{2}|(\mathbf{r}'-\mathbf{r}_3,\nabla_3)^2[\psi(\mathbf{r}_1,\mathbf{r}_3)+\psi(\mathbf{r}_3,\mathbf{r}_2)]|-\frac{\lambda_0}{2}=0. \tag{2.4}$$

Here ∇_3 is the derivative with respect to $\mathbf{r}_3$. A linear term is lacking in this expansion, owing to the extremal properties of the reference ray.

Analogously we get the following for the surface in (2.3):

$$\begin{aligned}F(\mathbf{r}')\approx\frac{1}{2}|&(\mathbf{r}_1'-\mathbf{r}_1,\nabla_1)^2\psi^0(\mathbf{r}_1)\\&+(\mathbf{r}'-\mathbf{r}_3,\nabla_3)^2[\psi(\mathbf{r}_1,\mathbf{r}_3)+\psi(\mathbf{r}_3,\mathbf{r}_2)]|-\frac{\lambda_0}{2}=0.\end{aligned} \tag{2.5}$$

If the quadratic terms in (2.4) and (2.5) vanish, as happens at caustics, then we must take account of the cubic terms, etc.

d) The Fresnel volumes of a plane and a spherical wave in a homogeneous medium

In a homogeneous medium we have $\psi(\mathbf{r}_a,\mathbf{r}_b)=n|\mathbf{r}_a-\mathbf{r}_b|$. Therefore we can assign an explicit form to Eqs. (2.2)–(2.5). If the points $\mathbf{r}_1$ and $\mathbf{r}_2$ lie on the z axis, while $\rho'=\sqrt{x'^2+y'^2}$, then the approximate equation (2.4) describes the ellipsoid of rotation

$$F(\mathbf{r}')=\frac{n}{2}(\rho')^2\left(\frac{1}{L_1}+\frac{1}{L_2}\right)-\frac{\lambda_0}{2}=0. \tag{2.6}$$

Here we have $L_1=z'-z_1$, and $L_2=z_2-z'$. The cross-

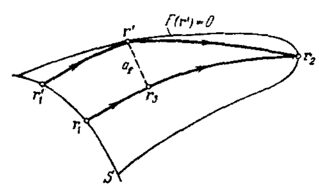

FIG. 2.

[6)]This name seems more suitable to us than the other terms that have been applied to various special cases: "region essential for diffraction",[8] "spatial Fresnel zone",[9] "three-dimensional Fresnel zone",[10] etc.

section of this ellipsoid in the plane $z' = \text{const}$ has the radius

$$a_f = \sqrt{\frac{\lambda_0/n}{L_1^{-1}+L_2^{-1}}}. \tag{2.7}$$

Analogously, in the case of a plane wave propagating along the z axis, we obtain from (2.5) the equation of the Fresnel paraboloid

$$F(\mathbf{r}') = \frac{n(\rho')^2}{2|z-z'|} - \frac{\lambda_0}{2} = 0. \tag{2.8}$$

The cross-section of the latter is

$$a_f = \sqrt{\frac{\lambda_0|z-z'|}{n}}. \tag{2.9}$$

3. HEURISTIC CRITERIA FOR APPLICABILITY OF GEOMETRIC OPTICS

Starting with the Huygens-Fresnel picture of the formation of the field in an inhomogeneous medium, let us formulate the following fundamental criteria for applicability of geometric optics:

Criterion I. The parameters of the medium, as well as the parameters of the wave (amplitude and phase gradient), must not vary appreciably over the cross-section of the Fresnel volume.

This condition presupposes that inequalities of the following form are satisfied:

$$a_f\,|\nabla_\perp A_0| \ll A_0,\quad a_f\,|\nabla p_j| \ll p,\quad a_f\,|\nabla n| \ll n. \tag{3.1}$$

Here $\nabla_\perp$ is the operator for differentiation transverse to the ray $p_j = \partial\psi/\partial x_j$, and a_f is the transverse dimension of the Fresnel volume.[7)] If we substitute the amplitude A_0, expressed in terms of the divergence $\mathscr{D}$, into (3.1), then we can obtain the bounds on the radii of curvature $R_{1,2}$ of the phase front from (3.1):

$$a_f \ll |R_{1,2}|. \tag{3.2}$$

If several rays arrive at the point of observation rather than one, the resultant field proves to be the sum of the fields associated with the individual rays. Near caustics, where the rays converge strongly, the first inequality of (3.1) breaks down, owing to the unbounded growth of the gradient of the amplitude. When the inequalities of (3.1) break down, one of the rays can be shown to lie inside the Fresnel volume of an adjacent ray. Therefore it is expedient to formulate another *auxiliary* condition for applicability, which is a consequence of criterion I, but which substantially facilitates the analysis of the problem in a number of cases:

Criterion II. The phase difference $k_0(\psi_1 - \psi_2)$ corresponding to rays arriving at the very same point must not be smaller than π:

$$k_0\,|\psi_1 - \psi_2| \geqslant \pi \tag{3.3}$$

(correspondingly, the path difference $|\psi_1 - \psi_2|$ must not be smaller than $\lambda_0/2$).

Upon somewhat simplifying the formulation of the problem, we can say that *a ray must not traverse an appreciable fraction of its path within the Fresnel volume of other rays arriving at the same point.* When criterion II breaks down, we risk taking double (or more than double) account of the contribution of the same rays to the resultant field.[8)]

The criteria discussed here in some form or other have been employed or have arisen in many studies, starting with the ground-breaking studies of Fresnel. In a number of cases they permit a rigorous substantiation. For example, in the diffraction of waves in homogeneous space (the initial conditions being assigned on some surface), the geometric-optical field corresponds to the stationary points of the Huygens-Kirchoff integral.[8] In this case the π-neighborhood of a stationary point (i.e., the neighborhood in which the phase differs from the stationary value by no more than π) is a cross-section of the Fresnel volume. Here the inequalities of (3.1) require constancy of the parameters of the wave and the medium within the bounds of such a π-neighborhood of the stationary point, while criterion II forbids the overlap of the π-neighborhoods of close-lying stationary points. We can also identify the cross-section of the Fresnel volume with the π-neighborhood of a stationary point in many other problems that allow an exact or approximate integral representation of the field. The general substantiation of the criteria being discussed in the case of a smoothly inhomogeneous medium stems from the condition of applicability of the stationary-phase method to the calculation of Feynman continual integrals, by means of which the Huygens principle can be most fully formulated.[32,33] In an extremely simplified formulation of the problem, one uses only two-segment virtual trajectories. Then one can derive criteria I and II by using the Huygens-Kirchoff integral with the approximate, rather than the exact, geometric-optical Green's function.[11]

In the cited cases and some others, criteria I and II arise from the stating of well-known facts. The new point that we are stressing is the *universality* and *sufficiency* of these criteria, even in inhomogeneous media. The sufficiency and universality of criteria I and II is confirmed by the fact that, in all of the numerous cases known to us, criteria I and II agree with the other methods of determining the limits of applicability of the method of geometric optics: from the first omitted

---

7)We note that criterion I is satisfied for a zone plate and for a circular aperture whose radius is an even number of Fresnel zones, but the field differs from the geometric-optical field. In order to rule out these degenerate situations, in the case of jumpwise variation of the parameters of the medium or of the field, we must take a_f in the inequalities of (3.1) to imply not necessarily the first Fresnel zone, but some characteristic Fresnel scale that is comparable with the radii of the first Fresnel zones.

8)L. A. Vaĭnshteĭn has pointed out a special case to the authors: in the problem of the radiation from a point source on an ideal conductive plane, the Fresnel volumes of the direct and of the specularly reflected rays with a phase difference less than π intersect over most of the path. Nevertheless, the ray description remains valid. The special character of this example involves the uniformity of the boundary conditions throughout the course of propagation. Consequently the fundamental criterion I is satisfied, and the geometric-optical approximation yields an exact solution of the problem.

term, by comparison with the exact, asymptotic, or numerical solutions, etc. We shall present the pertinent examples and comparisons in Secs. 5–9.

Another new point is the *pure ray* recipe proposed in Sec. 3 for constructing the Fresnel volume in an inhomogeneous medium. Thus the limitations I and II, which are diffractional in content, are expressed in *ray* language. Therefore it will not be a great exaggeration to say that geometric optics has acquired *internal* criteria of applicability.

4. RELATED PROBLEMS THAT CAN BE SOLVED BY EMPLOYING THE HEURISTIC CRITERIA

a) Estimates of the error of the zero-order approximation of geometric optics

The criteria formulated above enable one not only to answer the fundamental problem posed (of the sufficient conditions of applicability of the ray method), but also to illuminate a number of related problems that offer independent interest.

One of these problems is to estimate the errors of the zero-order approximation of geometric optics. We can hope that the parameters that figure in the inequalities (3.1) and (3.2) can serve as a heuristic measure of the inexactness of the field of (1.6):

$$\frac{|u_{\text{exact}}-u_0|}{|u_0|}\sim\gamma_{\text{heur}}\sim a_f^2\max\left\{\frac{|\nabla_\perp A_0|^2}{A_0^2},\ \frac{|\nabla_\perp n|^2}{n^2},\ \frac{|\nabla_\perp p_j|^2}{n^2}\right\}. \quad (4.1)$$

By itself, the estimate (4.1) cannot give reliable error levels throughout space. Yet it can suggest precisely where the error γ is known to be small, and where it is known to be large. In other words, these estimates permit one to outline both the region of applicability and that of inapplicability of geometric optics. The boundary of these regions can be defined roughly by the condition

$$\gamma_{\text{heur}}\sim 1. \quad (4.2)$$

b) Estimates of the wave field in focal, caustic, and other zones of inapplicability of geometric optics

At first glance it seems that geometric-optical calculations can be appropriate only in the region of applicability of the ray method, where its error is small: $\gamma_{\text{heur}}\ll 1$. Nevertheless, in a number of cases the ray approach can yield an estimate of the field that is correct in order of magnitude, though crude, for the field in a region of inapplicability of the ray method, in particular, in a penumbra zone or in the vicinity of foci and caustics. For example, let us examine the boundaries of caustic and focal zones. One can determine these boundaries by using (4.2), but one can also use criterion II (see Sec. 7).

While assuming the boundaries of the caustic and focal regions to be known, let us give estimates of the field inside these regions. First of all, we can do this by employing the values of the geometric-optical field at the boundary of a focal (or caustic) zone:

$$|u_{\text{foc}}|\sim\left|\frac{A^0}{\sqrt{\mathscr{J}}}\right|_{\text{at the boundary of the focal zone}}. \quad (4.3)$$

A method for estimating that differs in form, but is equivalent in essence, is based on the law of conservation of the energy flux in the ray tube, with the extra assumption that the initial energy flux $\Pi^0=n^0|A^0|^2\Delta S^0$ is uniformly smeared out over the focal (or caustic) zone:

$$\Pi^0=n^0|A^0|^2\Delta S^0=n_{\text{foc}}|u_{\text{foc}}|^2\Delta S_{\text{foc}}. \quad (4.4)$$

Here ΔS_{foc} is the width of the ray tube that corresponds to the focal (or caustic) zone (Fig. 3), and ΔS^0 is the initial cross-section of the ray tube. We obtain the following estimate from (4.5):

$$|u_{\text{foc}}|\sim|A^0|\sqrt{\frac{n^0\Delta S^0}{n_{\text{foc}}\Delta S_{\text{foc}}}}. \quad (4.5)$$

This is close to the estimate of (4.3), since $\mathscr{J}\approx n\Delta S/n^0\Delta S^0$. We shall demonstrate the effectiveness of these estimates below (in Secs. 5–9). A need for such estimates arises, e.g., in solving the problem of the possible occurrence of some particular nonlinear processes (harmonic generation, self-action, breakdown, etc., in a region of field concentration in foci and on caustics.

Estimates similar to (4.3) and (4.5) also are valid in other regions of inapplicability of geometric optics, e.g., in a penumbra region.

c) The problem of stability of the geometric-optical solution

There is another related problem—the behavior of geometric-optical solutions under small perturbations of the parameters of the medium, of phase boundaries, and/or of the initial conditions of the problem. We can explain the essence of the problem with the following simple example.

Let us treat a plane wave, which corresponds to a parallel beam of rays [Fig. 4a]. If we subject the plane initial wavefront to a weak, small-scale periodic perturbation, then the structure of the rays is cardinally distorted [Fig. 4b]. The field as calculated by geometric optics is also substantially distorted. At the same time, evidently, when $\delta\ll\lambda$, the true wave field is practically unchanged (only perturbations of the field of the order of δ/λ will appear). This example shows that the geometric-optical approximation is unstable (or more exactly, is very sensitive) with respect to small perturbations of the initial conditions. Instability of the same type also arises under analogous perturbations of the parameters of the medium and of the phase boundaries.

The resolution of the problem of the stability of geometric-optical solutions consists simply in the fact that the conditions for applicability of geometric optics break down under weak, small-scale perturbations. In particular, both the old and the new rays that have appeared in Fig. 4b, as a result of the perturbation of the

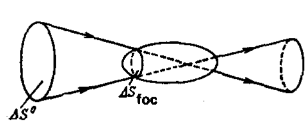

FIG. 3.

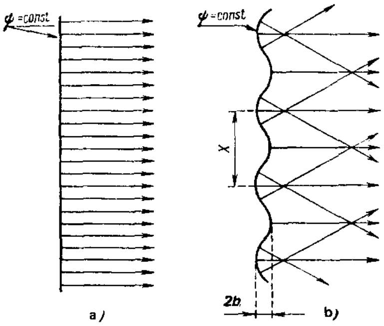

FIG. 4.

plane phase front become fictive and nonphysical. In fact, when the point of observation is set at the distance l from the screen, the Fresnel volumes surrounding the rays contain *many* inhomogeneities when $\sqrt{\lambda l} \gg X$. In this case, geometric optics is inapplicable.

d) Analysis of the wave pattern as a whole

This problem offers considerable interest under conditions in which one must obtain an estimate of the field in order of magnitude, or when one must compose a rough picture of the structure of the field before starting analytical or numerical calculations. The geometric-optical calculations and estimates in such a "rapid analysis" play a pilot role, since they enable one to estimate by simple means the value of the field over a considerable fraction of space, both in the region of applicability of the method and outside it [see Sec. 4b].

Thus the stated ideas allow not only a qualitative, but also a quantitative analysis of the structure of various high-frequency fields as a whole, based on geometric-optical notions alone. While somewhat exaggerating, we can say that this is equivalent to solving the wave problem while bypassing the solution of the wave equation. Such an analysis is especially valuable in engineering applications. In the overwhelming majority of such cases, it is important to know orders of magnitudes, rather than the exact values of the amplitudes of the wave field.

5. DIFFRACTION OF WAVES IN FREE SPACE

a) The form of the penumbra region in diffraction by a half-plane

As the first example of using the heuristic criteria, let us examine the very simple problem of the width of the penumbra region in the diffraction of a spherical (or cylindrical) wave by a half-plane (Fig. 5). The point of observation P lies on the boundary of the penumbra region (in Fig. 5 the umbra and penumbra regions are indicated by cross-hatching) if the Fresnel volume of the ray touches the edge of the half-plane. The radius of the first Fresnel zone $a_f = a_f(z)$ near the half-plane can be determined by Eq. (2.7) by setting $L_1 = z_0$ and $L_2 = z$ therein:

$$a_f = \sqrt{\frac{\lambda_0 z z_0}{(z+z_0)}}. \tag{5.1}$$

The distance of the point P from the axis is (see Fig.

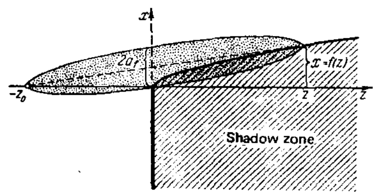

FIG. 5.

5):

$$\begin{aligned} x &= \frac{z+z_0}{z} a_f(z) \\ &= \sqrt{\frac{\lambda_0 z (z+z_0)}{z_0}} \equiv f(z). \end{aligned} \tag{5.2}$$

Thus the function $x = f(z)$ amounts to the equation of the boundary of the penumbra.

This boundary has the shape of a hyperbola having the rectilinear asymptote $x_a = [(z + (z_0/2)]\sqrt{\lambda_0/z_0}$. For all rays having a slope greater than the asymptote, geometric optics is valid at *infinitely large distances*. For these rays, the error of the method of geometric optics does not accumulate, in contrast to the rays that lie in the penumbra region. For a plane wave ($z_0 \to \infty$), the boundary of the penumbra region of (5.2) acquires the form of the parabola $x = \sqrt{\lambda_0 z}$. In this case all rays sooner or later fall into the penumbra region.

We note that we can apply Eq. (5.1) to a wave having an arbitrary phase front if we take z_0 to be the local radius of curvature of the phase front in the plane containing the normal to edge.

b) Formation of the ray field in the near and far zones of an antenna

This example explains why one can employ geometric optics in both the near and far zones of an antenna of large dimensions.

Near a cophased antenna with an aperture diameter $2b$, the geometric-optical approximation describes a beam of parallel rays (projector ray) with the same amplitude distribution as at the aperture itself. This approximation is applicable, in line with criterion I, as long as the radius of the first Fresnel zone $a_f = (\lambda_0)^{1/2} z$ is small in comparison with b, i.e., as long as $z \ll b^2/\lambda_0$.

In the far zone of the aperture, the field amounts to a directional spherical wave having a width of its directional diagram of the order of λ_0/b. Thus the characteristic scale of the variation of the amplitude of the field at the distance r from the antenna amounts to $l_\perp \sim \lambda_0 r/b$. According to (2.7), when the observation point lies at infinity, the dimension of the Fresnel zone on a sphere of radius r is $a_f = (\lambda_0 r)^{1/2}$. Hence the condition for applicability of geometric optics acquires the form

$$a_f \ll l_\perp, \quad \text{or} \quad r \gg \frac{b^2}{\lambda_0}.$$

Naturally, the condition derived here coincides with the ordinary criterion for the far zone of an aperture.

Thus, the geometric-optical description is applicable in both the far and near zones of the antenna. However, we should stress the difference between the initial conditions. In describing the field in the near zone, we assign the initial data at the aperture, whereas in the far zone we take as the initial data the field of a directional spherical wave, i.e., the far field of the antenna already formed. In the intermediate zone where $r \sim b^2/\lambda_0$, neither of the described approaches is applicable.

c) The field in the neighborhood of the focus of a lens

Let us study an ideal lens with an aperture of diameter $2b$ and a focal length F. Such a lens converts a plane wave into a convergent spherical wave with the radius of curvature F. Simple constructions yield the following expression for the radius of the first Fresnel zone in the plane of the lens:

$$a_f = \sqrt{\lambda_0 \frac{Fz}{F-z}}. \tag{5.3}$$

As we approach the focus ($z \to F$), the geometric-optical approximation loses force, since the radius a_f approaches infinity. Let $\zeta = F - z$ be the distance from the observation point z to the focus. Upon requiring that $a_f \ll b$, we obtain the following inequality for the quantity $\zeta = F - z$:

$$\zeta \gg \lambda_0 \left(\frac{F}{b}\right)^2 \equiv l_\parallel. \tag{5.4}$$

We can easily recognize the longitudinal dimension of the focal spot in the expression $l_\parallel = \lambda_0 (F/b)^2$. The ray AF proceeding from the edge of the lens to the focus (Fig. 6) intersects the plane $z = F - l_\parallel$ at the point B at the following distance from the axis of the lens:

$$l_\perp = b \frac{l_\parallel}{F} = \lambda_0 \frac{F}{b}. \tag{5.5}$$

This quantity characterizes the transverse dimension of the focal spot.

Further, let us extimate the field $|u_{foc}|$ at the focus of the lens by starting with the ideas expressed in Sec. 4. Here we equate the energy flux $\Pi^0 = (A^0)^2 \pi b^2$ through the lens to the energy flux $|u_{foc}|^2 \pi l_\perp^2$ through the focal spot (we are assuming that the energy flux at the focus of the lens is uniformly blurred over a circle of radius $l_\perp$). Then we obtain the following result for the three-dimensional problem:

$$|u_{foc}| \sim |A^0| \frac{b^2}{\lambda_0 F}. \tag{5.6a}$$

At the same time, at a two-dimensional focus we have:

$$|u_{foc}| \sim |A^0| \sqrt{\frac{b^2}{\lambda_0 F}}. \tag{5.6b}$$

These estimates of the focal field agree rather well with the exact maximal values of the field at the focus:

$$|u_{foc}|_{3\text{-D}} = \pi |A^0| \frac{b^2}{\lambda_0 F}, \qquad |u_{foc}|_{2\text{-D}} = 2 |A^0| \sqrt{\frac{b^2}{\lambda_0 F}}. \tag{5.7}$$

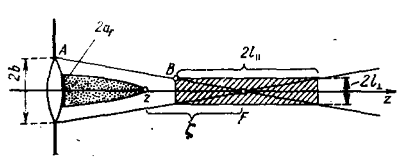

FIG. 6.

If we average the exact values over the focal spot, the agreement in (5.6) will be even better.

6. REFLECTION AND REFRACTION OF WAVES AT CURVILINEAR PHASE BOUNDARIES OF TWO HOMOGENEOUS MEDIA

a) Conditions of applicability of the reflection formulas

When a wave is incident on the curvilinear phase boundary of two homogeneous media, both reflected and refracted waves arise. In the geometric-optical approximation, the amplitudes A_{reflec} and A_{refrac} of these waves at the phase boundary S are associated with the amplitude A_{inc} of the incident wave by the local relationships

$$A_{reflec}|_S = \Gamma A_{inc}|_S, \; A_{refrac}|_S = D A_{inc}|_S. \tag{6.1}$$

Here Γ and D are respectively the coefficients of reflection and of transmission. Owing to the geometric principle of locality, they are defined by the formulas for a plane wave incident on the tangent plane (Fig. 7). Upon treating the amplitudes in (6.1) as the initial values in the formulas of geometric optics, we can calculate the field far from the phase boundary.

The conditions (3.1) for applicability require smooth variation of A_{inc}, Γ, and D within the limits of the cross-section b_f of the Fresnel volume at the phase boundary. Here we have $b_f = a_f/\cos\theta$, where θ is the angle of reflection or refraction of the ray. Moreover, Eq. (3.2) implies that the radii of curvature of the surface $a_{s1,2}$ must be large in comparison with b_f: $a_{s1,2} \gg b_f$. The value of a_f is calculated by Eq. (2.7) for a spherical wave, wherein L_1 is the radius of curvature R_{ph} of the phase front of the reflected (or refracted) wave immediately after reflection (refraction), and L_2 is the distance from the point of reflection (refraction). At the same time, the value of R_{ph} is calculated from the given values of the angle of incidence and the radii of curvature of the surface and of the phase front of the incident wave.[6,12]

Let us examine some effects that arise when the smoothness of variation of Γ and D breaks down inside the limits of the Fresnel volume.

b) The region of breakdown of the reflection formulas near a light-shadow boundary created by a convex object

The reflection formulas lose force at the angle of reflection $\theta = \theta_{min}$ at which the Fresnel volume of the reflected ray touches the light-shadow boundary at the surface of the convex object (Fig. 8), since the ampli-

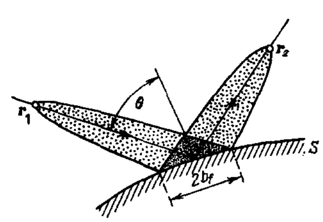

FIG. 7.

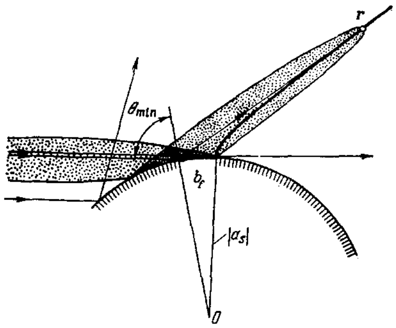

FIG. 8.

tude of the incident wave suffers a discontinuity at this boundary. An elementary calculation gives the following value for $\cos\theta_{\min}$:

$$\cos\theta_{\min} = \left(\frac{\pi}{k_0 |a_s|}\right)^{1/3}. \tag{6.2}$$

Here a_s is the local radius of curvature of the surface at the light-shadow boundary.[1,8,9] This value agrees very well with the values $\cos\theta_{\min} = (2/k_0|a_s|)^{1/3}$, which Fok derived by using the method of the parabolic equation.[12]

c) The region of inapplicability of the reflection formulas in the neighborhood of the angle of total reflection

The inapplicability of geometric optics near the critical angle of incidence $\theta = \theta_{cr}$ that corresponds to total reflection involves the fact that the phase of the coefficient of reflection $\alpha = \arg\Gamma$ varies rapidly here. Consequently $|\nabla\Gamma|$ becomes infinite at $\theta = \theta_{cr}$.

We can easily find the region of inapplicability from the condition that the Fresnel volume of the reflected ray should not include the point of reflection of the critical ray. The instant of contact is shown in Fig. 9 for the case of a plane surface, for which one can employ the mirror image of the source. The zone of inapplicability (cross-hatched in Fig. 9) obtained from these considerations agrees well with the asymptotics of the exact solution derived by L.M. Brekhovskikh (cf. Refs. 1, 13).

From the physical standpoint, the inapplicability of the method of geometric optics in this region involves the fact that a new type of ray arises here: diffraction rays, which describe the field of the lateral wave.[13]

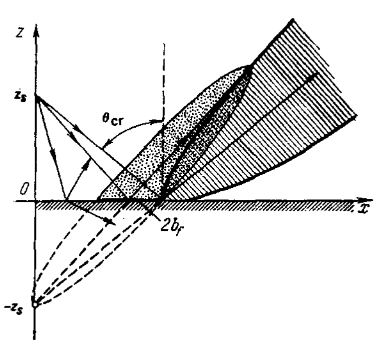

FIG. 9.

The neighborhood of the critical ray is a sort of penumbra region for the lateral wave, i.e., for the diffractive rays, which must be taken into account for $\theta > \theta_{cr}$ together with the ordinary reflected rays.

Similar restrictions pertain to the neighborhood of the angle of total refraction (Brewster angle) and also to the neighborhood of a ray that bears a zero field (i.e., a ray corresponding to a zero of the emission diagram of the source).[1]

d) Multiple reflection of rays from a concave mirror

If a source $\mathbf{r}_0$ lies on a concave phase boundary, then rays arrive at a point of observation $\mathbf{r}$ on the same surface that have undergone multiple reflection from the surface.

The multiplicity of reflection at which the rays lose their individuality has been determined[5]: if ψ_m is the optical path of the wave for m-fold reflection of the ray, then geometric optics loses force when $\psi_{m+1} - \psi_m \leq \lambda_0/2$. From the standpoint of the conditions that we have proposed for inapplicability of the ray method, this inequality implies the breakdown of criterion II: the $(m+1)$-fold reflected ray lies inside the Fresnel volume of the m-fold reflected ray.

7. PERICAUSTIC REGIONS OF INAPPLICABILITY OF GEOMETRIC OPTICS

a) An estimate of the width of the pericaustic zone in the case of a simple caustic. Indistinguishability of rays in the pericaustic zone

We can estimate the width of the caustic zone by considerations that rest on criterion II: we can naturally consider a point $\mathbf{r}$ to lie inside the caustic zone if each of the rays lies inside the Fresnel volume of an adjacent ray (Fig. 10). This directly implies the physical indistinguishability of the rays in the caustic zone; when the Fresnel volumes overlap substantially, one cannot distinguish the rays (i.e., determine their parameters separately) by using an opaque screen with an aperture. The reason is that one cannot obscure the path of one ray without touching the Fresnel volume of the adjacent ray.

Overlap of the Fresnel volumes corresponds to the condition

$$k_0 |\psi_1 - \psi_2| < \pi. \tag{7.1}$$

Simple calculations by perturbation theory show[14,15] that, near a simple (nonsingular) caustic, we have

$$|\psi_1 - \psi_2| \approx \frac{4}{3}\beta^{1/2} l_N^{3/2}, \quad \beta \equiv 2n_c^2 \nu_{rel}. \tag{7.2}$$

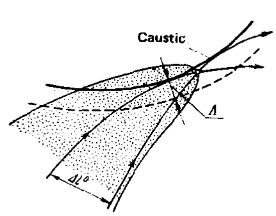

FIG. 10.

Here l_N is the distance along the normal to the caustic, n_c is the refractive index at the caustic, $v_{rel} = |v_{caust} - v_{ray}\cos\delta|$ is a quantity that determines the relative curvature of the ray and of the caustic, v_{caust} is the curvature of the normal cross-section of the caustic in the direction of the ray, v_{ray} is the curvature of the ray at its point of contact with the caustic, and δ is the angle between the normal to the caustic and the principal normal to the ray (the normal to the caustic is assumed to be directed into the light region). Upon substituting (7.2) into (7.1), we obtain the following estimate for the width of the caustic zone $|l_N| < l_c$:

$$l_c = \left(\frac{3\pi}{4k_0}\right)^{2/3}\beta^{-1/3} = 1.77\,(k_0^2\beta)^{-1/3} = 1.77\Lambda. \tag{7.3}$$

This estimate differs only by the numerical coefficient from the distance between the caustic and the first maximum of the Airy function describing the field near a simple caustic[14,17]:

$$l_1 = 1.02\Lambda, \quad \Lambda = (k_0^2\beta)^{-1/3}. \tag{7.4}$$

Equation (7.4) is a generalization of the corresponding expression given in Ref. 8, Sec. 59, for a homogeneous medium having $n=1$, when $\Lambda = (\rho/2k_0^2)^{1/3}$, where $\rho = v_{caust}^{-1}$ is the radius of the curvature of the caustic.

b) An estimate of the field near a caustic in the two-dimensional problem

If Δl^0 is the initial width of the ray tube, while $l_c \sim \Lambda$ is the width of the caustic zone (see Fig. 10), then, upon replacing ΔS^0 by Δl^0 and ΔS_c by Λ in Eq. (4.5), we get the following expression for the field near the caustic:

$$|u_{foc}| \sim |A^0|\left|\frac{n^0\Delta l^0}{n_k\Lambda}\right|. \tag{7.5}$$

The formulas of the Airy caustic asymptotes for the wave field[14,17] predict the same order of magnitude. A detailed analysis[1] has shown that the maximal value of the field (at the distance $l_1 = 1.02\Lambda$ from the caustic) exceeds the estimate of (7.5) by a factor of only 1.34.

Let us consider the uniform asymptotic representations of the field near caustics, in particular, the Airy asymptote[18]

$$u(r) = e^{ik_0\theta - i(\pi/4)}\left[(-\xi)^{-1/4}(iA_1 + A_2)v(\xi) + i(-\xi)^{-1/4}(iA_1 - A_2)v'(\xi)\right]. \tag{7.6}$$

Here $v(\xi)$ is the Airy function, and we have

$$\theta = \frac{1}{2}(\psi_1 + \psi_2), \quad \frac{2}{3}(-\xi)^{3/2} = \frac{1}{2}(\psi_1 - \psi_2).$$

We note that these representations employ the geometric-optical values of the eikonals $\psi_{1,2}$ and the amplitudes $A_{1,2}$. However, the rays that one uses to calculate the quantities $\psi_{1,2}$ and $A_{1,2}$ have lost any direct physical meaning. This means that near caustics the rays retain, in B.E. Kinber's graphic expression, the function of a *geometric skeleton that bears the flesh of the waves*.[19,20]

c) Estimates of the field at a focus in the presence of spherical aberration

Let the field in the initial plane $z=0$ have the quadratic phase: $u^0 = \exp(-ik_0\rho^2/2F)$. Then, when $z=F$, a caustic cusp is formed instead of a point focus. An elementary calculation by the formulas of Sec. 3 gives the following value of the Fresnel radius in the plane $z=0$:

$$a_f \approx (4F^3\lambda)^{1/4}. \tag{7.7}$$

We can estimate the transverse dimension $l_\perp$ of the spot by Eq. (5.5) by substituting therein the effective radius $b_{eff} \approx a_f$ for the radius b of the lens. This yields $l_\perp \sim (\lambda^3F/4)^{1/4}$. Upon assuming in (4.5) that $\Delta S^0 = \pi a_f^2$ and $\Delta S_{foc} = \pi l_\perp^2$, we get the following estimate for the field averaged over the spot:

$$|u_{foc}| \sim |A^0|\sqrt{\frac{4F}{\lambda}}. \tag{7.8}$$

This approximate value is smaller than the exact value of the field at the center of the spot by a factor of only $\pi/\sqrt{2} = 2.2$.

d) Focusing indices of the field at caustics

As we see from the estimates given above, the field at caustics and at foci shows a power-function dependence on the wavelength. If we employ the small parameter $\mu = 1/k_0Ln$, then the field at a common-type caustic is estimated to be

$$|u_{foc}| \sim |A^0|\,\mu^{-\sigma_f}. \tag{7.9}$$

Here σ_f is the so-called *focusing index*.

Evidently σ_f depends on the character of the focusing. Thus, for a simple caustic we have $\sigma_f = 1/6$, for a two-dimensional caustic cusp $\sigma_f = 1/4$, for a two-dimensional focus and for spherical aberration $\sigma_f = 1/4$, and for a caustic loop that constricts to a point $\sigma_f = 3/10$. The largest value $\sigma_f = 1$ is attained at an ideal focus, while the smallest value ($\sigma_f = 0$) corresponds to the absence of focusing. Values of σ_f calculated by diffraction theory for a number of typical caustics are given, e.g., in Refs. 1, 21, 22, and 35. These values are very convenient for especially rough estimates of the fields in focal zones, especially in nonlinear problems.

8. DIFFRACTION OF WAVES IN INHOMOGENEOUS MEDIA

a) The form of the Fresnel volume for a plane wave in a plane-layered medium

By employing the ray equations in a plane-layered medium having the refractive index $n = n(z)$, we can easily determine the cross-section a_f of the Fresnel volume in the horizontal planes $z = z'$:

$$a_f(z') = \sqrt{\lambda_0 \frac{\partial}{\partial s_0}\int_{z'}^{z}\frac{s_0\,dz}{\sqrt{n^2(z) - s_0^2}}}. \tag{8.1}$$

Here z is the coordinate of the point of observation, $s^0 = n^0\sin\theta^0$, θ^0 is the initial angle of incidence of the ray, and n^0 is the refractive index at the level $z = z^0$ [Fig. 11a]. With vertical incidence of the ray on the layer ($s_0 = 0$), we have

$$a_f = \sqrt{\lambda_0\int_{z'}^{z}\frac{dz}{n(z)}}. \tag{8.2}$$

According to (8.2), if $n(z)$ diminishes along the ray, the

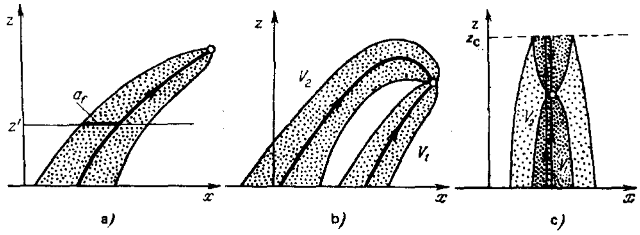

FIG. 11.

Fresnel volume is somewhat narrower than in the corresponding homogeneous medium having $n = \text{const} = n(z)$, whereas, if $n(z)$ increases, the Fresnel volume becomes broader: $a_f(z') > [\lambda_0(z-z')/n(z)]^{1/2}$.

b) Form of the Fresnel volume in a layered medium in the presence of a caustic

In this case the Fresnel volume is bounded by two envelopes that correspond to the two rays arriving at the observation point. With oblique incidence of the wave, the volumes V_1 and V_2 are separated in space [Fig. 11b], while with vertical incidence, the volume V_1 lies completely inside the volume V_2 [see Fig. 11c]. Here the volume V_2 is bounded by a surface that has the shape of a stocking partially turned inside out (for more details, see Ref. 1).

c) Estimates of the width of the Fresnel volume in grazing incidence of the rays on a layered medium: limits of applicability of the ray description of the propagation of radio waves in the ionosphere

In grazing propagation (angle of incidence θ close to $\pi/2$), the rays have a slightly curved trajectory. Over the entire course of the latter, the permittivity deviates insignificantly from unity, since even the minimal value $\varepsilon_{\min} = \sin^2\theta$, which is reached at the point of reflection, is close to unity when $\theta \sim \pi/2$. Therefore the maximal half-width a_f of the Fresnel volume can be estimated by Eq. (2.7) for a homogeneous medium having the permittivity $\varepsilon \sim 1$: $a_{f,\max} \sim (\lambda_0 L)^{1/2}/2$ (L is the distance along the horizontal).

According to criterion I for applicability of geometric optics, the decline in the permittivity $\varepsilon(z)$ over the diameter $2a_f \sim (\lambda_0 L)^{1/2}$ of the Fresnel volume must be small in comparison with unity:

$$\frac{\partial \varepsilon}{\partial z} 2a_f \ll 1. \tag{8.3}$$

The characteristic scale $H \sim \varepsilon / |\partial\varepsilon/\partial z|$ for the ionosphere amounts to no less than 50–200 km, so that, for a wavelength $\lambda_0 = 15$ m, we obtain from (8.3) the value $L \ll 10^5 - 10^6$ km. This distance is considerably greater than the circumference of the Earth. Hence, in the case of almost horizontal rays, we can assume that the geometric-optical method is suitable for describing at least single-skip propagation.

The problem can be solved similarly of the applicability of the geometric-optical approximation in the neighborhood of the critical Pedersen ray traveling along the axis of an anti-waveguide.[1] The estimates obtained here agree with the diffraction calculation performed in Ref. 23.

d) The near and far fields of an antenna in an inhomogeneous medium

As the point of observation moves away from the antenna in an inhomogeneous medium, the cross-section of the Fresnel volume a_f in the plane of the antenna first rises, and then strongly decreases on approaching a caustic. If the cross-section a_f becomes smaller than the dimension b of the antenna, then we can speak of a local coupling characteristic of the *near* field that occurs between the field at the aperture and the field at the remote point of observation. This feature of fields in inhomogeneous media has been found and described in Refs. 24 and 25.

9. LIMITS OF APPLICABILITY OF SPACE-TIME GEOMETRIC OPTICS

a) Necessary conditions of applicability

In problems of propagation of nonstationary waves (pulses, wave packets) in dispersive media, the approximation of space-time geometric optics plays the same important role as the ordinary approximation of geometric optics does in the propagation of monochromatic waves.[26-28]

If the properties of the medium and of the field vary slowly enough, then we can treat the nonstationary wave

$$u(\mathbf{r}, t) = A(\mathbf{r}, t)\, e^{i\varphi(\mathbf{r}, t)} \tag{9.1}$$

as locally plane and locally monochromatic at distances $|\Delta r| \sim \lambda$ and intervals $\Delta t \sim \tau$. Here λ is some mean wavelength in the medium, and τ is its mean period. The conditions of slowness needed for this have the form

$$\mu = \max\left\{\frac{\lambda}{L},\ \frac{\tau}{T},\ \frac{\tau_0}{T}\right\} \ll 1. \tag{9.2}$$

Here τ_0 is the characteristic scale of the frequency dispersion (the relationship between τ and τ_0 is arbitrary in the general case), and L and T are respectively the space and time scales of variation of the field and of the medium.

When the conditions of (9.2) are satisfied, one can easily calculate the amplitude A and the eikonal φ by employing space-time rays.[26-28]

b) The space-time Fresnel volume

Let the initial conditions for the field of the nonstationary wave $u(\mathbf{r}, t)$ have the form

$$u(\mathbf{r}^0, t^0) = A(\mathbf{r}^0, t^0)\, e^{i\varphi(\mathbf{r}^0, t^0)} \equiv A^0(\xi)\, e^{i\varphi^0(\xi)}. \tag{9.3}$$

Let them be assigned on an arbitrary hypersurface Σ^0 whose parametric equations are $\mathbf{r} = \mathbf{r}^0(\xi)$, $t = t^0(\xi)$, or $\mathbf{R} = \mathbf{R}^0(\xi)$, where $\mathbf{R} \equiv (\mathbf{r}, t)$, $\mathbf{R}^0 \equiv (\mathbf{r}^0, t^0)$, and $\xi = \{\xi_1, \xi_2, \xi_3\}$. Here the ξ_j are the curvilinear coordinates in Σ^0. Let us introduce the reference space-time ray $\mathbf{R}_1 \to \mathbf{R}_2$ having the phase advance $\varphi_{\text{ref}} = \varphi^0(\mathbf{R}_1) + \Phi(\mathbf{R}_1, \mathbf{R}_2)$ and the virtual ray $\mathbf{R}_1' \to \mathbf{R}' \to \mathbf{R}_2$ having the phase advance $\varphi_{\text{virt}} = \varphi^0(\mathbf{R}_1') + \Phi(\mathbf{R}_1', \mathbf{R}') + \Phi(\mathbf{R}', \mathbf{R}_2)$. Here $\mathbf{R}_1 = (\mathbf{r}_1, t_1)$ and $\mathbf{R}_1' = (\mathbf{r}_1', t_1')$ are the points of emergence of the rays from the hypersurface Σ^0, and $\mathbf{R}_2 = (\mathbf{r}_2, t_2)$ is the point of observation in four-dimensional space. Then the equation of the boundary of the Fresnel volume can be written in the form

$$F(\mathbf{R}') = |\phi_{\text{virt}} - \phi_{\text{ref}}| - \pi$$
$$= |[\varphi^0(\mathbf{R}_1') + \Phi(\mathbf{R}_1', \mathbf{R}') + \Phi(\mathbf{R}_1', \mathbf{R}_2)] - [\varphi^0(\mathbf{R}_1) + \Phi(\mathbf{R}_1, \mathbf{R}_2)]| - \pi = 0. \quad (9.4)$$

Here $\mathbf{R}' = (\mathbf{r}', t')$ is the running point of the boundary. We note that one can derive an approximate equation of the boundary from (9.4) by expansion in a Taylor series.

c) The form of the space-time Fresnel volume in a homogeneous medium

Let us concretize Eq. (9.4) as applied to the problem of propagation of a plane frequency-modulated pulse in a homogeneous dispersive medium having the refractive index $n(\omega)$. Let the phase of the pulse be $\varphi(0, t) = \varphi^0(t)$ in the initial plane $z = 0$. That is, the initial frequency of the pulse varies according to the law $\omega^0 = -\partial\varphi^0/\partial t$. We can then derive the following expression from (9.4) for the half-width of the Fresnel zone in the plane $z = 0$:

$$\tau_f^0 = \sqrt{2\pi \left| k_0'' \frac{zz_c}{z - z_c} \right|}. \quad (9.5)$$

Here we have $k_0'' = d^2k_0(\omega)/d\omega^2|_{\omega=\omega_0}$, and z_c is the coordinate of the point of contact of the reference ray with the caustic:

$$z_c = -\left[k_0'' \frac{d\omega^0(\eta)}{d\eta}\right]^{-1}. \quad (9.6)$$

Equation (9.5) is an analog of the expression (2.7) for a spherical monochromatic wave.

In the absence of frequency modulation, so that $\omega^0 = \text{const}$, according to (10.11) we have $|z_c| \to \infty$, and

$$\tau_f^0 = \sqrt{2\pi |k_0''| z}. \quad (9.7)$$

Figure 12 shows the general form of the Fresnel volume in the zt plane. The cross-section of this volume at the level $z = z'$ is

$$\tau_f' = \sqrt{2\pi |k_0''| (z - z')}. \quad (9.8)$$

We note that Eq. (9.7) agrees with the estimate of the characteristic time interval $\sim(|k_0''|z)^{1/2}$ derived in Refs. 29 and 30 by the method of the space-time parabolic equation.

d) Conditions for applicability of space-time geometric optics

By analogy with criterion I of Sec. 3, we require satisfaction of conditions of the type

$$\tau_f \left|\frac{\partial A}{\partial t}\right| \ll A, \quad \tau_f \left|\frac{\partial \omega}{\partial t}\right| \ll \omega. \quad (9.9)$$

By using these conditions, one can estimate the dimensions of the region of inapplicability in various situations, e.g., near a shadow boundary in space-time (near a pulse front), in the neighborhood of space-time caustics and foci, both in homogeneous and inhomogeneous dispersive media. The indicated estimates are taken by analogy with the spatial case. We shall consider only a single example here which pertains to compression of frequency-modulated signals in dispersive media (other examples are given in Ref. 1).

e) Estimation of the field at a space-time focus: pulse compression

Let us choose the frequency-modulation law $\omega^0(t)$ such that all the space-time rays converge at a single point (z_f, t_f) (Fig. 13). The dimension of the first Fresnel zone for $z = 0$ is determined by Eq. (9.5):

$$\tau_f^0 = \sqrt{2\pi |k_0''| \frac{zz_f}{z - z_f}}. \quad (9.10)$$

Owing to the inequalities of (9.9), the approximation of space-time geometric optics loses force when τ_f is comparable with the length T of the original pulse: $\tau_f \sim T$. We can find from this condition the spatial dimension of the focal spot (see Fig. 13):

$$|z - z_f| \sim 2\pi |k_0''| \frac{z_f^2}{T^2} \equiv l. \quad (9.11)$$

The distance l from the focus corresponds to the time interval

$$T_{\min} = T\frac{l}{z_f} = 2\pi |k_0''| \frac{z_f}{T}. \quad (9.12)$$

The latter amounts to the duration of the pulse at the focus.

We can obtain the field at the focus from the condition of energy balance

$$|A^0|^2 T = A_{\text{foc}}^2 T_{\min}.$$

This implies that the gain in amplitude of the compressed signal is

$$\frac{A_{\text{foc}}}{A^0} \approx \sqrt{\frac{T}{T_{\min}}} = \frac{T}{\sqrt{2\pi |k_0''| z_f}}. \quad (9.13)$$

The estimates that have been derived agree with the results of Refs. 29 and 30, which were obtained by the method of the space-time parabolic equation. They also admit verification by using the spectral approach, which has been described in detail by Vaĭnshteĭn[31] (see also Ref. 27).

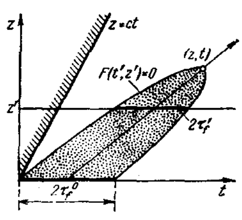

FIG. 12.

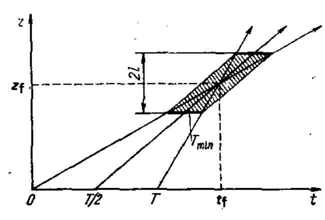

FIG. 13.

10. CONCLUSION

The treatment that we have carried out shows that ordinary geometric optics, when supplemented by the concept of the Fresnel volumes of the rays, enables one to carry out a global analysis of high-frequency wavefields, including not only the elucidation of the qualitative structure of the fields, but also the derivation of quantitative estimates of the field, even in zones of inapplicability of geometric optics. Such an analysis is composed of the following elements:

1. Determining the geometric-optical field, which presupposes finding the rays, phase fronts, and divergences of the ray tubes and other ray parameters.

2. Using the heuristic criteria to find the regions of applicability and inapplicability of geometric optics and the boundaries between them.

3. Estimating the error in the region of applicability of geometric optics.

4. Heuristic estimates of the amplitude of the field in regions of inapplicability of geometric optics.

When necessary, the analysis can also be supplemented by estimates of the exponentially small scattered fields in regions into which the rays do not penetrate.[9)]

The realization of this program depends substantially on the solvability of the equations of geometric optics, and primarily the ray equations. As we know, solutions exist in analytic form only for a limited number of special cases that fall far short of spanning the entire spectrum of actual wave problems of electrodynamics, acoustics, optics, and seismology. Under these conditions, it seems desirable to us, and in some applications even necessary, to develop a universal numerical program of analysis of high-frequency fields. This might be based on calculating the rays and their corresponding Fresnel volumes, and would envisage finding the boundaries of applicability of the ray method and roughly determining the field in regions of inapplicability of geometric optics.

The concepts that we have expressed above can also be transferred without substantial changes to the *quasi-classical wave functions* in quantum mechanics. Introduction of the Fresnel volume of the classical trajectories enables one to formulate the limiting admissible rate of change of the potential in space, to determine the width of caustic zones, to estimate the wave function in the neighborhood of caustics (e.g., in the rainbow zone in a problem of scattering by a potential), etc.

Moreover, we can generalize the criteria that we have formulated to vector fields by requiring in addition that the variation of the polarization of the wave over the transverse section of the Fresnel volume should be small.[1] Reference 1 also points out features of the form of the Fresnel volumes in anisotropic media and discusses problems of the reality of caustics and of possibilities of localizing complex rays.

In closing, the authors express their sincere appreciation to L. A. Vaĭnshteĭn for valuable advice and remarks.

9) A method for estimating these exponentially weak wave fields by using convergent series with the zero-order approximation of geometric optics as the leading term has been proposed in Ref. 1.

1. Yu. A. Kravtsov and Yu. I. Orlov, O granitsakh primenimosti metoda geometricheskoĭ optiki (On the Limits of Applicability of Geometric Optics), Lectures at the 5th All-Union School on Diffraction and Propagation of Waves, Chelyabinsk, 1979; in: Sovremennye problemy rasprostraneniya i rasseyania voln (Modern Problems of Propagation and Scattering of Waves), IRÉ AN SSSR, Moscow, 1979, p. 76.
2. S. M. Rytov, Dokl. Akad. Nauk SSSR **18**, 263 (1938).
3. S. M. Rytov, Tr. Fiz. Inst. Akad. Nauk SSSR **2**, 1 (1940).
4. Yu. A. Kravtsov, in: Analiticheskie metody v teorii difraktsii i rasprostraneniya voln (Analytical Methods in the Theory of Diffraction and Propagation of Waves), Sovet po akustike, Moscow, 1979, p. 257.
5. V. M. Babich and V. S. Buldyrev, Asimptoticheskie metody v zadachakh difraktsii korotkikh voln (Asymptotic Methods in Problems of Diffraction of Short Waves), Nauka, Moscow, 1972.
6. V. A. Borovikov and B. E. Kinber, Geometricheskaya teoriya difraktsii (Geometric Theory of Diffraction), Svyaz', Moscow, 1978.
7. R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals, McGraw-Hill, New York, 1965 (Russ. Transl., Mir, Moscow, 1968).
8. E. L. Feĭnberg, Rasprostranenie radiovoln vdol' zemnoĭ poverkhnosti (Propagation of Radio Waves Along the Earth's Surface), Izd-vo AN SSSR, Moscow, 1961.
9. F. B. Chernyĭ, Rasprostranenie radiovoln (Propagtion of Radio Waves), Sov. radio, Moscow, 1972.
10. H. L. Bertoni, L. B. Felsen, and A. Hessel, IEEE Trans. Antennas Propagation **AP-19**, 226 (1971).
11. Yu. A. Kravtsov and Z. I. Feĭzulin, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **12**, 886 (1969).
12. V. A. Fok, Problemy difraktsii i rasprostraneniya élektromagnitnykh voln (Problems of Diffraction and Propagation of Electromagnetic Waves), Sov. radio, Moscow, 1970 (Engl. Transl. of 1st ed., Electromagnetic Diffraction and Propagation Problems, Pergamon, Oxford, 1965).
13. L. M. Brekhovskikh, Volny v sloistykh sredakh (Waves in Layered Media), 2nd ed., Nauka Moscow, 1973 (Engl. Transl. of 2nd ed., Academic Press, New York, 1980).
14. Yu. L. Gazaryan, in: Voprosy dinamicheskoĭ teorii rasprostraneniya seĭsmicheskikh voln (Problems of the Dynamical Theory of Propagation of Seismic Waves), Izd-vo Leningr. un-ta, Moscow, 1961, No. 5, p. 73.
15. V. L. Groshev and Yu. A. Kravtsov, Izv. Vyssh, Uchebn. Zaved. Radiofiz. **11**, 1812 (1968).
16. Yu. A. Kravtsov, *ibid.* **7**, 664 (1964).
17. Yu. I. Orlov, Tr. MÉI, No. 194, 103 (1974).
18. L. D. Landau and E. M. Lifshits, Teoriya polya (Field Theory), Fizmatgiz, Moscow, 1960 (Engl. Transl., The Classical Theory of Fields, rev. 2nd ed., Addison-Wesley, Reading, Mass., 1962).
19. Yu. A. Kravtsov, Akust. Zh. **14**, 1 (1968) [Sov. Phys. Acoust. **14**, 1 (1968)].
20. M. V. Berry and K. E. Mount, Rep. Prog. Phys. **35**, 315 (1972).
21. V. I. Arnol'd, Matematicheskie metody klassicheskoĭ mekhaniki (Mathematical Methods of Classical Mechanics), Nauka, Moscow, 1974 (Engl. Transl., Springer-Verlag, New York, 1978).
22. M. V. Berry, Adv. Phys. **25**, 1 (1976).

[23]M. V. Tinin, in: Issledovaniya po geomagnetizmu, aéronomii i fizike Solntsa (Studies in Geomagnetism, Aeronomy, and Solar Physics), Nauka, Moscow, 1974, No. 29, p. 157.
[24]I. G. Kondrat'ev and G. V. Permitin, Izv. Vyssh. Uchebn. Zaved. Radiofiz. 13, 1795 (1970).
[25]G. V. Permitin, *ibid.* 16, 254 (1973).
[26]Yu. A. Kravtsov, L. A. Ostrovsky, and N. S. Stepanov, Proc. IEEE 62, 1492 (1974).
[27]Yu. I. Orlov, Prostranstvenno-vremennáya difraktsiya impul'sov (Space-Time Diffraction of Pulses), Lectures at the 5th All-Union School on Diffraction and Propagation of Waves, Chelyabinsk, 1979; in, Pryamye i obratnye zadachi teorii difraktsii (Direct and Reverse Problems of the Theory of Diffraction), IRÉ AN SSSR , Moscow 1979, pp. 5–114.
[28]V. M. Babich, V. S. Buldyrev, and I. A. Molotkov, Metod vozmushcheniĭ v teorii rasprostraneniya voln (The Perturbation Method in the Theory of Propagation of Waves), Lectures at the 5th All-Union School on Diffraction and Propagation of Waves, Chelyabinsk, 1979; in: Teoriya rasprostraneniya voln v neodnorodnykh i nelineĭnykh sredakh (Theory of Propagation of Waves in Inhomogeneous and Nonlinear Media), IRÉ AN SSSR, Moscow 1979, p. 28.
[29]P. V. Bliokh, Izv. Vyssh. Uchebn. Zaved. Radiofiz. 7, 460 (1964).
[30]A. G. Litvak and V. I. Talanov, *ibid.* 10, 539 (1967).
[31]L. A. Vaĭnshteĭn, Usp. Fiz. Nauk 118, 339 (1976) [Sov. Phys. Usp. 19, 189 (1976)].
[32]E. S. Fradkin, in: Problemy teoreticheskoĭ fiziki (Problems of Theoretical Physics), Nauka, Moscow, 1969, p. 386.
[33]Feynman Path Integrals, in: Proc. of Intern. Colloquium, Marseille, May 1978, Springer-Verlag, Berlin, Heidelberg, New York, 1979 (Lecture Notes in Physics, Vol. 106).
[34]V. L. Ginzburg, Rasprostranenie élektromagnitnykh voln v plazme (Propagation of Electromagnetic Waves in Plasma), Nauka, Moscow, 1967 (Engl. Transl., Gordon and Breach, New York, 1962).
[35]Yu. A. Kravtsov and Yu. I. Orlov, Geometricheskaya optika neodnorodnykh sred (Geometric Optics of Inhomogeneous Media), Nauka, Moscow, 1980.

Translated by M. V. King
Edited by R. T. Beyer