Sommerfeld-Runge Law in Three and Four Dimensions

H. POEVERLEIN*
Air Force Cambridge Research Laboratories, Bedford, Massachusetts
(Received February 5, 1962)

The Sommerfeld-Runge law, a law of geometric optics expressing that the propagation vector is irrotational, is useful for general deductions and for computations of geometric-optical wave fields, for which it proves without supplementary notions sufficient under a wide range of conditions. Application to inhomogeneous anisotropic media, though not necessitating recourse to the concept of rays, yields the equation of ray paths. Absorption and lateral intensity variation of waves are accounted for by introducing complex propagation vectors.

A four-dimensional generalization of the law is applicable to modulated waves and time-varying media. The four-dimensional considerations, to which it leads, deal with group propagation (in time-constant and time-varying media), with isotropy in space-time (relating to the law $uv=c^2$), and with refraction and reflection at discontinuities in time and at moving boundaries. A brief discussion of focusing and diffraction in the scope of four-dimensional geometric optics is also given.

1. FUNDAMENTALS

1.1. The Sommerfeld-Runge Law

A WAVE field in geometric-optical approximation is supposed to be fully described by the propagation vector as a function of the coordinates. A solution of a propagation problem may in geometric optics be given in two forms: either as a representation of propagation vectors (or wave normals) or as a picture of ray paths. Wave normals and ray directions are not identical in anisotropic media. Wave normals have the obvious meaning of normals to the phase planes. Ray paths are paths of propagation in some physical sense (see Sec. 2.3). The intensity of waves, although not so much a matter of geometric optics, is derivable in some approximation by assuming energy flow along the ray paths. The limitations of this idea will not be investigated in this study which is based on geometric optics.

When some emission of waves occurs in a medium whose characteristics at all places are given, the wave normals and propagation vectors throughout the wave field are determined by theorems of geometric optics such as Fermat's principle or Snell's law. In case of anisotropic media both these theorems involve wave normals and ray directions. The proper application of Snell's law in this case will be discussed later (Sec. 2.2). The Sommerfeld-Runge law is another law of geometric optics determining the wave normals or propagation vectors. It is equivalent to Snell's law in application to boundaries and inhomogeneous isotropic media, as will be seen, yet it is quite generally usable in its simple formulation, not requiring reference to ray directions in anisotropic media.

The direction of the wave normal and the refractive index n in combination may be represented by a vector $\mathbf{n}$. The Sommerfeld-Runge law,[1–3] stated by Sommerfeld and Runge as early as 1911, says that $\mathbf{n}$ is irrotational,

$$\nabla\times\mathbf{n}=0. \tag{1}$$

Introduction of the propagation vector

$$\mathbf{k}=\mathbf{n}k_0 \tag{2}$$

(i.e., the product of $\mathbf{n}$ and the propagation constant in vacuum k_0) leads to the alternative formulation of the law:

$$\nabla\times\mathbf{k}=0. \tag{3}$$

The Sommerfeld-Runge law is an immediate consequence of the existence of a uniquely defined wave function,

$$u=C\exp\left[i\left(\omega t-\int\mathbf{k}\cdot d\mathbf{r}\right)\right], \tag{4}$$

in which the propagation vector $\mathbf{k}$ is a function of the coordinates ($\mathbf{r}$). This wave function is a first approximation of the WKB type.[4–6] The quantities u and C may be scalar or vectorial; C may vary slowly with coordinates, but no time dependence of the three quantities C, ω, and $\mathbf{k}$ is considered at present. In order that the wave function be uniquely defined, there must be[2,7]

$$\oint\mathbf{k}\cdot d\mathbf{r}=0, \tag{5}$$

over any closed path. This postulate yields the Sommerfeld-Runge law, Eq. (3).[7]

The wave function, Eq. (4), and the Sommerfeld-

* Present address: Elektrophysikalisches Institut der Technische Hochschule München, Munich 2, Germany.

[1] A. Sommerfeld and J. Runge, Ann. Physik (Leipzig) **35**, 277 (1911).
[2] A. Sommerfeld, *Optik* (Akademische Verlagsgesellschaft, Leipzig, 1959), 2nd ed., pp. 306–309.
[3] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, New York, 1959), pp. 123–124, 129–131, and 681–683.
[4] S. Flügge and H. Marschall, *Rechenmethoden der Quantentheorie* (Springer-Verlag, Berlin, 1952), 2nd ed., pp. 152–165.
[5] K. Suchy, Ann. Physik (Leipzig) **13**, 178 (1953).
[6] H. Bremmer, in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1958), Vol. 16, pp. 552–557.
[7] G. B. Whitham, J. Fluid Mechanics **9**, 347 (1960).

Runge law are usable under all conditions appropriate for geometric optics, including both sufficiently slowly varying media and discontinuity surfaces (boundaries between materials), at which the refracted wave is considered as a continuation of the incident wave. In transition from the incident wave to the refracted wave, the Sommerfeld-Runge law requires continuity of the $\mathbf{k}$ (or $\mathbf{n}$) components tangential to the discontinuity surface.[1] Discontinuity of a tangential $\mathbf{k}$ component would lead to an infinite $\nabla\times\mathbf{k}$ and prevent a continuation of the wave field described by Eq. (4).[8] The requirement of continuity of the tangential $\mathbf{k}$ components is equivalent with Snell's law in this case.[1] The amplitude factor C, of course, will, in general, be discontinuous at the boundary.

The reflected wave originating at a discontinuity surface is another continuation of the incident wave. The postulate of continuity of the tangential $\mathbf{k}$ component in transition from the incident to the reflected wave provides the necessary continuity between the wave functions of the incident and reflected waves (disregarding C).

As to the characteristics of the medium or the wave field, the only limitation is the assumption of slow or discontinuous variations, which is imposed by geometric optics. The medium may be anisotropic and inhomogeneous in a general manner and the propagation vector may be real or complex.

Sommerfeld and Runge dealt with isotropic, nondissipative media only. Suchy[9] in his studies of the transition between ray and wave optics used the Sommerfeld-Runge law as a law of refraction also for inhomogeneous anisotropic media.

No assumption on the nature of the waves will be made in the present paper. The field may be scalar or vectorial and the waves may be electromagnetic, hydrodynamic, matter waves, or any other type of waves.

1.2. Four-Dimensional Generalization

When the propagation vector $\mathbf{k}$ and the angular frequency ω vary slowly with coordinates and with time, the wave function may be written

$$u=C\exp\left[i\int(\omega dt-\mathbf{k}\cdot d\mathbf{r})\right]. \tag{6}$$

Also the amplitude factor C is now assumed to vary slowly with time and coordinates. The postulate of unique definition of the present wave function leads to

$$\oint(\omega dt-\mathbf{k}\cdot d\mathbf{r})=0. \tag{7}$$

[8] This argumentation implies that the wave function, Eq. (4), is usable at a very steep, in the limiting case, infinite gradient of the propagation characteristics and that $\mathbf{k}$ remains finite in the limiting case, thus permitting no discontinuity of the exponential term of the wave function.

[9] K. Suchy, Ann. Physik (Leipzig) **11**, 113 (1952).

(to be integrated over any closed $\mathbf{r}$, t path) or with introduction of the four-vectors

$$\begin{aligned}\boldsymbol{r}&=(\mathbf{r},ict),\\ \boldsymbol{k}&=(\mathbf{k},i\omega/c).\end{aligned} \tag{8}$$

to

$$\oint\boldsymbol{k}\cdot d\boldsymbol{r}=0. \tag{9}$$

Boldface italic is used to denote four-vectors. Equation (9) as four-dimensional analog of Eq. (5) yields the four-dimensional generalization of the Sommerfeld-Runge law,

$$\square\times\boldsymbol{k}=0. \tag{10}$$

The four-dimensional operator replacing ∇ is

$$\square=\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z},-\frac{i}{c}\frac{\partial}{\partial t}\right). \tag{11}$$

Equation (10) consists of six component equations, three representing the three-dimensional Sommerfeld-Runge law, Eq. (3), the other three, in three-dimensional vector notation, reading

$$\nabla\omega+\partial\mathbf{k}/\partial t=0. \tag{12}$$

A variation of the frequency, as referred to in Eq. (12), may result from an emission of variable frequency or from propagation in time-varying media.

The three-dimensional Sommerfeld-Runge law and Eq. (12), now appearing combined in the four-dimensional Sommerfeld-Runge law, have been used in conjunction by Whitham[7] in a treatise on group propagation. Alternative derivations of Eq. (12) are based on a consideration of the phase of the wave (MacDonald[10]) or on the postulate that in propagation of a wave train the number of waves is conserved (Whitham[7]). The one-dimensional version of Eq. (12) [Eq. (37)] is more often met with in the literature. A study of group propagation on basis of the four-dimensional Sommerfeld-Runge law will follow later (Secs. 4.1 and 4.2).

The four-dimensional Sommerfeld-Runge law applies to media varying in space-time, provided the variations are slow enough to render the wave function, Eq. (6), a good approximation. Analogy with the three-dimensional case shows that, in addition, the law is applicable to discontinuity surfaces in space-time, if discontinuity of C is allowed for (cf. Sec. 4.3).

1.3. Equivalence with Snell's Law in Isotropic Media

At boundaries between materials the three-dimensional Sommerfeld-Runge law was seen to lead to a continuity condition equivalent with Snell's law. In order to deduce Snell's law for *inhomogeneous isotropic media*, we multiply Eq. (3) vectorially by $\mathbf{k}$, thus

[10] G. J. F. MacDonald, J. Geophys. Res. **66**, 3639 (1961).

obtaining

$$\mathbf{k}\times(\nabla\times\mathbf{k})=0. \tag{13}$$

A simple transformation yields

$$(\mathbf{k}\cdot\nabla)\mathbf{k}=k\nabla k, \tag{14}$$

with k denoting the magnitude of $\mathbf{k}$.

Because in isotropic media the direction of $\mathbf{k}$ (i.e., the wave normal) is the ray direction, Eq. (14) may be interpreted as saying that in progression in the ray direction the $\mathbf{k}$ components parallel to the surfaces of constant k stay constant. This is Snell's law combined with the statement that the wave normal remains in the plane of incidence (now considered part of Snell's law).

Equation (13) relates to only two components of $\nabla\times\mathbf{k}$, the components normal to $\mathbf{k}$. Disappearance of the third component, as equally required by the Sommerfeld-Runge law, is expressed by

$$\mathbf{k}\cdot(\nabla\times\mathbf{k})=0. \tag{15}$$

This equation, in fact, is deducible from the theorem of Malus and Dupin,[3] which states that surfaces normal to the $\mathbf{k}$ lines (wave surfaces) do exist. The deduction given by Sommerfeld and Runge for homogeneous media [cf. their Eq. (4), which corresponds to our Eq. (15)] can readily be transferred to inhomogeneous media.

If the existence of wave surfaces and the resulting Eq. (15) are taken as self-evident, we may identify the Sommerfeld-Runge law with Eq. (14) or, in application to isotropic media, with Snell's law. The Sommerfeld-Runge law, however, has the advantage of appearing in a very simple vectorial formulation.

1.4. Eikonal and Phase

The Sommerfeld-Runge law, Eq. (1), indicates that $\mathbf{n}$ is the gradient of a scalar, the eikonal,[1–3,7]

$$\mathbf{n}=\nabla S. \tag{16}$$

We may alternately write

$$\mathbf{k}=k_0\nabla S. \tag{16a}$$

The four-dimensional Sommerfeld-Runge law suggests introduction of an eikonal Φ in space-time, corresponding to

$$k=\Box\Phi. \tag{17}$$

This eikonal is the analog of k_0S. It appears, multiplied by $-i$, in the exponent of the wave function, Eq. (6), and consequently represents the running phase of the wave, if it is real. A complex Φ includes phase and attenuation coefficient.

2. ANISOTROPIC MEDIA

2.1. Basic Facts

In an anisotropic medium, the propagation constant is dependent on the direction of propagation. The dependence on both the location and the direction is expressed by an equation[11–13]

$$F(\mathbf{r},\mathbf{k})=0. \tag{18}$$

The function F characterizes the medium throughout the space under consideration. No attention is paid now to variation with frequency; the frequency is assumed to be constant.

Equation (18) with assumption of a fixed $\mathbf{r}$ determines a surface in $\mathbf{k}$ space, which is the "refractive index surface"[13,14] scaled up by the vacuum propagation constant k_0. (Various names are found for this or an equivalent surface in the literature.) Double-valuedness of the surface with the possibility of coincidence of two branches at some place, indicating coupling of two waves, is not a matter of geometric optics and will not be considered here.

In a specific wave field, in which a definite wave normal is assigned to each location, Eq. (18) enables us to compute the magnitude of $\mathbf{k}$ for all locations. A picture of the wave normals or their connecting lines, the phase trajectories (Fig. 1), may therefore be considered an adequate description of a wave field in a medium of known characteristics. The phase trajectories in anisotropic media are, however, not ray paths; ray directions, in general, differ from wave normals, as is known.

2.2. Variation of k in Propagation

The Sommerfeld-Runge law and Eq. (18) together provide a sufficient number of equations to have $\mathbf{k}$ determined throughout a wave field, when it is given at some boundaries or in the vicinity of the sources of waves. Let us assume that $\mathbf{k}$ has been computed for all points on a surface in the wave field. Then, there are three conditions determining the increment of $\mathbf{k}$ in a differential departure from the surface to a neighboring point: The $\nabla\times\mathbf{k}$ components parallel to the surface must disappear and $dF=0$. As can be seen from Eq. (23) and the succeeding explanations, a necessary supposition is that there exists a ray path element by which the point is linked with the surface. If such a ray path element does not exist or, in other words, the surface is parallel to the ray direction (direction of ∇_kF), the three conditions are interdependent or incompatible.

Constancy of F in a wave field, as required, is expressed by

$$\nabla F=0. \tag{19}$$

In differentiation, it has to be noticed that F depends on

[11] R. Courant and D. Hilbert, *Mathematische Physik* (Springer-Verlag, Berlin, 1937), Vol. 2, pp. 82–92.

[12] J. Bazer and O. Fleischman, Research Report No. MH-10, Division of Electromagnetic Research, Institute of Mathematical Sciences, New York University, New York, 1959 (unpublished), pp. 14–19.

[13] K. G. Budden, *Radio Waves in the Ionosphere* (Cambridge University Press, New York, 1961), pp. 200, 252–254, and 276–278.

[14] H. Poeverlein, Z. Naturforsch. **5a**, 492 (1950).

$\mathbf{r}$ directly and indirectly through $\mathbf{k}$. By introduction of a gradient ∇_r referring to the direct dependence on $\mathbf{r}$ and a gradient ∇_k in $\mathbf{k}$ space, Eq. (19) becomes

$$\nabla_r F + \nabla(\mathbf{k}\cdot\nabla_k)F = 0. \tag{20}$$

The closing bracket in the last term might be placed before or after F, however, the operator ∇ is supposed to apply to $\mathbf{k}$ only. This suggests the transformation

$$\nabla(\mathbf{k}\cdot\nabla_k)F = \nabla_k F\times(\nabla\times\mathbf{k}) + (\nabla_k F\cdot\nabla)\mathbf{k}. \tag{21}$$

Now we introduce the Sommerfeld-Runge law, Eq. (3). It allows us to simplify Eq. (21) to

$$\nabla(\mathbf{k}\cdot\nabla_k)F = (\nabla_k F\cdot\nabla)\mathbf{k}, \tag{22}$$

and thus to transform Eq. (20) into

$$\nabla_r F + (\nabla_k F\cdot\nabla)\mathbf{k} = 0. \tag{23}$$

Equation (23) expresses the variation of $\mathbf{k}$ in progression in the direction of $\nabla_k F$ and will be seen to describe propagation of the waves. The equation is related to Hamilton's equations of geometric optics, which are an alternative formulation dealing with propagation. The Hamiltonian minus a constant may be taken as function F in the above representation. Constancy of the Hamiltonian corresponds to the equation determining the characteristics of the medium, Eq. (18). Derivations of Hamilton's equations frequently make use of the fact that $\mathbf{k}$ is representable as a gradient,[11–13] occasionally also of the Sommerfeld-Runge law.[7,13] Budden[13] arrives at a version of Eq. (23), specialized for one-dimensionally stratified media, [his Eq. (14.16)] in an intermediate stage of his derivation of Hamilton's equations.

The gradient $\nabla_r F$ relates to the stratification of the medium as it appears under the assumption of a constant direction of $\mathbf{k}$; in a surface element normal to $\nabla_r F$, the condition $F=0$ leads to a constant magnitude k for fixed direction of $\mathbf{k}$. Equation (23) thus is found to say that, for progression in the direction indicated, the $\mathbf{k}$ components parallel to the stratification (i.e., normal to $\nabla_r F$) do not vary. This is Snell's law as formulated, for example, by Haselgrove.[15] It refers to progression in the direction of $\nabla_k F$, which in the limiting case of an isotropic medium becomes the direction of $\mathbf{k}$. That this direction of progression is the "ray direction" will be understood in the next section. The reference to both the ray direction and the stratification makes Snell's law in the present form impractical for propagation problems of a general character.

The previous version of Snell's law, deduced from Eq. (14) for isotropic media, remains as an alternative form applicable to anisotropic media, though one must avoid identifying the direction of progression in this version, i.e., the direction of $\mathbf{k}$, with the ray direction. The gradient ∇k, appearing in this version, is obviously a measure of the variation of the magnitude k in space, thus seeming to describe some kind of stratification, but it is defined only for the wave field under consideration and is not representative of the medium.

2.3. Ray Paths

The concept of ray paths is based on the idea that generation of a wave field at or around some location entails propagation of a wave along a certain path, a ray path. The field on the ray path thus is supposed to arise from that generated at the initial point. Along the entire ray path, consequently, some quantity characterizing the wave field is determined by the field at the initial point.

Equation (23) shows that a path element parallel to $\nabla_k F$ is a ray path element in this sense and the propagation vector $\mathbf{k}$ is the quantity determined along a ray path or ray path element. The equation yields $d\mathbf{k}$ for every path element parallel to $\nabla_k F$. A path composed of such path elements and being described by

$$\nabla_k F\times d\mathbf{s} = 0, \tag{24}$$

hence, is seen to have the characteristic that $\mathbf{k}$ is determined in all its points when it is given in an initial point. Equation (24) is the equation of ray paths. It has been derived from Eq. (18), the equation describing the medium, without introducing anything but the Sommerfeld-Runge law and some axiomatic idea on ray paths.

Uniqueness of our ray path element follows from the fact that Eq. (18) and the Sommerfeld-Runge law together leave some freedom for choosing the variation of $\mathbf{k}$ along any line element that is not parallel to $\nabla_k F$.

The ray direction as the direction of $\nabla_k F$ is the normal to the surface $F=0$ in $\mathbf{k}$ space and to the refractive index surface, which is similar to the surface $F=0$ (cf. Sec. 2.1). This is a well-known law of geometric optics of anisotropic media.[13,14]

In conventional terms, a ray is a beam of waves showing an intensity structure with a maximum in the center. Such a beam can be represented by complex propagation vectors. We simply have to introduce in a plane normal to a given ray path a small imaginary part of $\mathbf{k}$, pointing toward the ray path, disappearing on the ray path, and increasing with departure from the ray path. Because $\mathbf{k}$ is left unaltered (real) on the ray path, it will remain real on the entire ray path and maximum intensity will be encountered along the path. From this representation it becomes evident that a beam of waves in fact follows the course of a ray path as determined by the above formulas [in particular Eq. (24)].

The model of a beam of waves suggests splitting an extensive wave field up into an assemblage of independent adjacent beams. The separate consideration of individual beams leads to the concept of energy propagation as taking place along ray paths.

[15] J. Haselgrove, *Report of the Physical Society Conference on the Physics of the Ionosphere* (Physical Society, London, 1955), pp. 355–364.

Curves along which some quantity is not open to choice appear as characteristic curves in the theory of partial differential equations. The ray paths, in fact, are characteristic curves of some partial differential equation: the eikonal equation (or "Hamilton-Jacobi equation"), which is obtained by substituting $k_0\nabla S$ for $\mathbf{k}$ in Eq. (18).[11,12]

2.4. Example

The above explanations are illustrated by Figs. 1 to 3. In the example chosen, waves are emitted from a point source and propagated in a medium varying with one coordinate, z. The medium is homogeneous and isotropic in the vicinity of the source (located at $z=0$) and starts varying, becoming at the same time anisotropic, at $z=z_0$. Only one plane, containing the source, is shown in the figures. The underlying propagation problem will be outlined later.

Phase trajectories are represented in Fig. 1, ray paths in Fig. 2. Phase trajectories and ray paths coincide in the isotropic part of space ($z<z_0$), but are separated and intersect each other in traversing the anisotropic area ($z>z_0$).

Figure 1 or Fig. 2 might be considered as the solution of the present propagation problem. The *wave field* in geometric-optical approximation is readily obtained from Fig. 1. The phase surfaces are the normal surfaces to the phase trajectories and the propagation constants corresponding to the given wave normal directions follow from the characteristics of the medium (cf. Sec. 2.1). The *ray paths* (Fig. 2), on the other hand, may be thought of as having more physical significance.

The Sommerfeld-Runge law, referring to propagation vectors only, yields as immediate solution of a propagation problem the propagation vectors and hence the representation of phase trajectories, Fig. 1. Most other geometric-optical theorems (e.g. Fermat's principle or Snell's law in the first version of Sec. 2.2) involve propagation vectors and ray directions, thus requiring more complicated computational procedures in determining the propagation vectors. Figure 3, showing phase trajectories and the intensity structure of a beam guided along a ray path, is the picture that might emerge from introduction of a complex $\mathbf{k}$ in the Sommerfeld-Runge law as explained in Sec. 2.3. This representation of a beam, based on geometric optics, fails in the case of focusing, i.e., in the present example, at the caustic separating the "light" from the "dark" area.

The waves in our example are supposed to be matter waves corresponding to monoenergetic electron beams. A nonvarying scalar potential and a vector potential disappearing below z_0 and increasing linearly above z_0, being pointed there in the x direction, are assumed. This corresponds to a constant magnetic field of the y direction (direction normal to the plane of drawing) in the region above z_0. The graphical constructions were begun with the ray paths (Fig. 2), because they are now

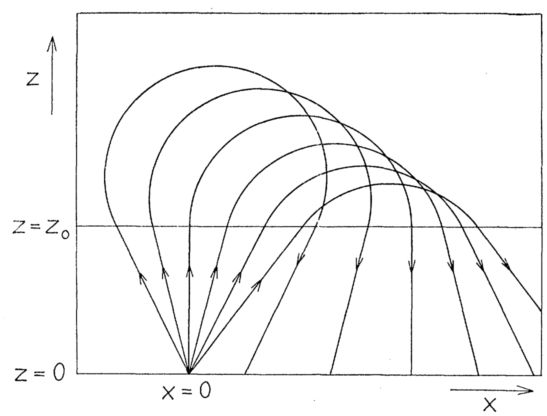

FIG. 2. Ray paths in the example of Fig. 1.

particle paths consisting of straight lines in the field-free space and parts of circles in the constant magnetic field. The propagation vector for given particle velocity $\mathbf{u}$ and vector potential $\mathbf{A}$ is derived from the Schrödinger equation as being

$$\mathbf{k}=(1/\hbar)(e\mathbf{A}+m\mathbf{u}). \tag{25}$$

The refractive index surfaces (obtained by varying the direction of $\mathbf{u}$) are eccentric spheres.[14,16]

3. ATTENUATED WAVES

The wave function, Eq. (4), with insertion of a complex propagation vector $\mathbf{k}$ describes attenuated waves. The Sommerfeld-Runge law, following from the unique definition of the wave function, applies to complex $\mathbf{k}$. Theorems based on the ray concept, however, become useless when complex ray directions (directions of $\nabla_k F$) lead to complex coordinates for which the medium is not defined.

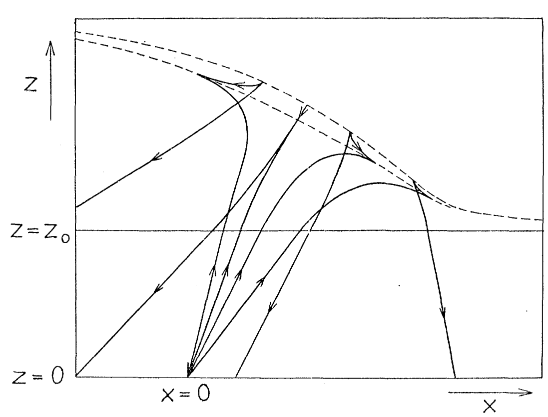

FIG. 1. Example of phase trajectories (normal trajectories to the phase planes). The wave source is located at $x=0$, $z=0$. In the region $z>z_0$, the medium is anisotropic and variable with z. Dashed lines are loci of cusps.

[16] M. Cotte, Ann. phys. (Paris) **10**, 333 (1938).

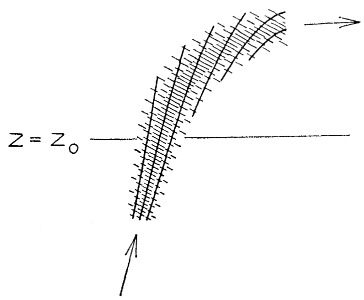

FIG. 3. Beam of waves. The strong lines are phase trajectories (corresponding to the example of Figs. 1 and 2). The density of hatching indicates the intensity structure of a beam. Slow intensity variation due to contraction or expansion of the beam is not shown.

The real and imaginary parts of a propagation vector

$$\mathbf{k}=\mathbf{p}+i\mathbf{q}, \tag{26}$$

in general, have different directions. The imaginary part expresses both attenuation due to dissipation of energy (absorption) and lateral intensity variation as considered in Sec. 2.3 and Fig. 3.

The two types of attenuation are easily separated in *isotropic media*, as will be shown now. The propagation constant k in an isotropic medium is a function of the coordinates only. The imaginary part of k (being fixed for a given location) indicates then energy dissipation. From

$$p^2-q^2+i2(\mathbf{p}\cdot\mathbf{q})=k^2, \tag{27}$$

we conclude in this case that a $\mathbf{q}$ component parallel to $\mathbf{p}$ relates to energy dissipation, whereas a $\mathbf{q}$ component normal to $\mathbf{p}$, expressing lateral attenuation, does not involve energy dissipation. Keller[17] gives examples of complex $\mathbf{k}$ in nondissipative isotropic media. He uses (in isotropic media) complex (or "imaginary") rays, which correspond to our complex propagation vectors.

A constant k for a given location in an isotropic medium is an approximation implied by geometric optics. From the refined formulation of the eikonal equation,[5]

$$k_0(k_0\nabla S+i\nabla)\cdot\nabla S=\kappa^2 \tag{28}$$

(with κ being a constant of the medium at a considered location), we obtain by means of Eq. (16a) the relationship

$$(\mathbf{k}+i\nabla)\cdot\mathbf{k}=\kappa^2, \tag{29}$$

or

$$p^2-q^2-\nabla\cdot\mathbf{q}+i[2(\mathbf{p}\cdot\mathbf{q})+\nabla\cdot\mathbf{p}]=\kappa^2. \tag{30}$$

Notice that in case of small $\mathbf{q}$ (corresponding to low attenuation) the term $\nabla\cdot\mathbf{q}$ may be essential for the deviation of p^2 from the real part of κ^2.

4. CONSIDERATIONS IN FOUR DIMENSIONS

4.1. General Statements

In space-time, a four-dimensional version of the Sommerfeld-Runge law has been seen to apply (Sec. 1.2). All deductions made from the three-dimensional Sommerfeld-Runge law may therefore be transcribed to four dimensions.

The inclination of the four-dimensional propagation vector $\boldsymbol{k}$ [Eqs. (8)] toward the time axis is

$$\tan\alpha=-ikc/\omega, \tag{31}$$

or

$$\tan\alpha=-in, \tag{32}$$

and is apparently $-i$ times the refractive index n. For electromagnetic waves in vacuum (characterized by $n=1$), the space component and time component of $\boldsymbol{k}$ are of equal magnitude.

A medium being isotropic in four dimensions is one in which the magnitude of $\boldsymbol{k}$ at a given point in space-time is independent of the direction of $\boldsymbol{k}$ in space-time. The square of this magnitude is

$$\boldsymbol{k}\cdot\boldsymbol{k}=(n^2-1)\omega^2/c^2. \tag{33}$$

Four-dimensional isotropy means invariance of this expression for varying direction of $\mathbf{k}$ in space and for varying angular frequency ω, which multiplied by i/c represents the fourth component of $\boldsymbol{k}$. Electromagnetic waves in vacuum are a trivial example of four-dimensional isotropy. The corresponding magnitude of $\boldsymbol{k}$ is zero. In the case of electromagnetic waves in a plasma, we have the refractive index formula

$$(n^2-1)\omega^2=-\omega_p{}^2. \tag{34}$$

Comparison with Eq. (33) shows that this is another case of four-dimensional isotropy. Despite the isotropy, there is, however, an asymmetry between space and time coordinates, resulting from the appearance of the imaginary unit in all time components. Four-dimensional isotropy (of matter waves) and its implications were treated by Synge.[18] For further remarks on four-dimensional isotropy it may be referred to the following Sec. 4.2 and to a paper in preparation.

As in three dimensions, a medium may be described by an equation

$$F(\boldsymbol{r},\boldsymbol{k})=0, \tag{35}$$

which now determines the characteristics of the medium for all points in space-time and for all frequencies. For a fixed point $(\boldsymbol{r})$ and a fixed direction of $\mathbf{k}$, Eq. (35) is the "dispersion equation."

In three dimensions, the condition $F=0$ together with the Sommerfeld-Runge law led to Eq. (23), an equation expressing the variation of $\mathbf{k}$ in progression along certain path elements, which were recognized as being differentials of ray paths. Equation (23) and its derivation are readily transcribed to four dimensions. (The four-dimensional transcription of the double vector product of Eq. (21) is a four-vector composed of $\Box_k F$ and $\Box\times\boldsymbol{k}$.) With retention of the three-dimensional symbols and introduction of the new function F,

[17] J. B. Keller, Proceedings of Symposia in Applied Mathematics (American Mathematical Society) **8**, 27 (1958).

[18] J. L. Synge, *Geometrical Mechanics and De Broglie Waves* (Cambridge University Press, New York, 1954), in particular pp. 6–59.

the transcribed Eq. (23) becomes in four dimensions the set of equations

$$\begin{aligned}\nabla_r F+[\nabla_k F\cdot\nabla-(\partial F/\partial\omega)(\partial/\partial t)]\mathbf{k}=0,\\ -\partial F/\partial t+[\nabla_k F\cdot\nabla-(\partial F/\partial\omega)(\partial/\partial t)]\omega=0.\end{aligned}\tag{36}$$

It may be emphasized that this set of equations follows from the four-dimensional Sommerfeld-Runge law in conjunction with Eq. (35), the characteristic equation of the medium.

Equations (36) may be used as basic equations of a four-dimensional geometric optics. Synge[18] developed an alternative four-dimensional geometric-optical theory by means of four-dimensional generalization of Hamilton's theory of geometric optics. Equations (36) and the four-dimensional Hamiltonian formulations are relevant to the propagation of wave groups. Traveling wave groups obviously are the four-dimensional analog of ray paths. An auxiliary variable, in Hamilton's equations needed to denote space-time points on the four-dimensional group paths, does not appear in the present formulations. Equation (35) is referred to by Synge as "slowness equation." The "slowness" in his terms is a four-vector proportional to $\boldsymbol{k}$. The following four-dimensional considerations are partly based on Eqs. (36), partly on the four-dimensional Sommerfeld-Runge law immediately. Some emphasis is placed on specialized situations permitting a simple interpretation.

4.2. Group Propagation

Group propagation in a *time-constant medium* may be studied, as Lighthill and Whitham[7,19] demonstrated, by following a section of a definite frequency in propagation of a variable-frequency wave train. Equation (12) and the three-dimensional Sommerfeld-Runge law, together making up the four-dimensional Sommerfeld-Runge law, prove to be immediately usable for this. Only the simple example of *spatially one-dimensional propagation* may now be depicted in Whitham's representation.[7] Equation (12) in the one-dimensional case reads

$$\partial\omega/\partial z+\partial k/\partial t=0,\tag{37}$$

or, because k (assumed in the z direction) is not explicitly dependent on t,

$$\partial\omega/\partial z+(\partial k/\partial\omega)(\partial\omega/\partial t)=0.\tag{38}$$

This formulation shows that in traveling with the velocity

$$u=1/(\partial k/\partial\omega),\tag{39}$$

which is the one-dimensional group velocity, one sees a constant frequency.

Intensity variation within a wave train, corresponding to the conventional notion of a wave group, is obtained by introducing a small varying imaginary part of ω. Zero imaginary part, relating to maximum intensity in the wave group, is found to travel with group velocity, just as any definite ω. We may recall the three-dimensional analog: complex $\mathbf{k}$, representing a beam of waves that is propagated along ray paths.

In the general case of a *medium varying with time and spatial coordinates*, the laws of group propagation follow from a four-dimensional transcription of the considerations on ray paths that emerged from Eq. (23). Equations (36), the four-dimensional analog of Eq. (23), show that a group path element in space-time is parallel to

$$\Box_k F=[\nabla_k F,\ -ic(\partial F/\partial\omega)].\tag{40}$$

These group path elements form a group path, on which $\boldsymbol{k}$, once given in an initial space-time point, is determined everywhere. The group velocity, being ic times the ratio between space and time components of Eq. (40), is

$$d\mathbf{s}/dt=-\nabla_k F/(\partial F/\partial\omega).\tag{41}$$

The velocity $d\mathbf{s}/dt$ has the direction of a ray (direction of $\nabla_k F$) and, as Eq. (42) will prove, a projection on the $\mathbf{k}$ direction equal to the differential quotient $d\omega/dk$ for given location and time. It thus is in agreement with conventional formulas for the group velocity,[13,20] though its applicability is now not limited to conditions of no time variation. Differentiation of $F=0$ for constant $\boldsymbol{r}$ and fixed direction of $\mathbf{k}$ yields the expression for $d\omega/dk$ corresponding to Eq. (41)

$$d\omega/dk=-(\mathbf{k}\cdot\nabla_k F)/k(\partial F/\partial\omega).\tag{42}$$

In a *time-constant medium*, the behavior of the waves in group propagation is found as it is supposed to be: According to the second equation of Eqs. (36) the frequency in a wave group remains constant. Maximum intensity, earlier (in the spatially one-dimensional case) seen to correspond to a disappearing imaginary part of ω, follows group propagation.

In *time-varying media*, the variation of $\boldsymbol{k}$ in a wave group is determined by Eqs. (36). Apparently, $d\boldsymbol{k}$ in a wave group is parallel to $\Box_r F$. If we take the z axis in the direction of the three-dimensional gradient $\nabla_r F$, this means that in a wave group there is

$$dk_x=dk_y=0;\tag{43}$$

(this is Snell's law) and

$$\frac{d\omega}{dk_z}=-\frac{\partial F}{\partial t}\bigg/\frac{\partial F}{\partial z}.\tag{44}$$

The expression on the right-hand side of this equation represents a virtual velocity of the medium; traveling with this velocity V in the z direction, one finds the medium invariable for fixed direction of $\mathbf{k}$ and fixed frequency. Equation (44) may be written

$$V dk_z-d\omega=0.\tag{45}$$

[19] M. J. Lighthill and G. B. Whitham, Proc. Roy. Soc., **A229**, 281 (1955).

[20] K. Suchy, Ann. Physik (Leipzig) **14**, 412 (1954).

A *four-dimensionally isotropic medium*, as noted in Sec. 4.1, exhibits a magnitude of $\boldsymbol{k}$ independent of the direction of $\mathbf{k}$ and of ω. From Eq. (33), we obtain in this case

$$k^2-\omega^2/c^2=\text{const}, \tag{46}$$

or

$$kdk-(1/c^2)\omega d\omega=0, \tag{47}$$

and hence the group velocity

$$u=d\omega/dk=c^2(k/\omega). \tag{48}$$

The phase velocity is $v=\omega/k$. Thus, we have[18]

$$uv=c^2. \tag{49}$$

This law is known to be valid for electromagnetic waves in plasmas, the case recognized as four-dimensionally isotropic in Sec. 4.1, and for matter waves in field-free space, provided their frequency is defined properly (in fact, proportional to the energy mc^2 of the moving particle). The law, Eq. (49), is characteristic of four-dimensional isotropy. Matter waves thus are seen to represent another case of four-dimensional isotropy. In all cases in which Eq. (49) does not hold, we deal obviously with four-dimensionally anisotropic media.

Isotropy with respect to one spatial coordinate and the time coordinate is found in propagation of electromagnetic waves in wave guides. Equation (49) applies also to this case.

4.3. Time-Varying Media

A time variation of the frequency and the propagation vector may result from a modulation of the emissions at the wave source or from a variation of the medium with time. This section deals with variable media, for the most part, under specializing assumptions. The variation may be continuous or discontinuous.

The locations in space-time at which a discontinuous variation of the medium is encountered may be given by a relationship between all four coordinates of space-time (or between some of them). This relationship describes in four-dimensional space a discontinuity surface (in fact, a three-dimensional structure) and in three-dimensional space, in general, a moving boundary. At the discontinuity surface, we have to postulate continuity of the tangential $\boldsymbol{k}$ components in order to keep $\square\times\boldsymbol{k}$ finite.

A *sudden, simultaneous variation* of the medium throughout space represents a discontinuity in time. The discontinuity surface in four-dimensional space, corresponding to $t=t_0$, is the three-dimensional space viewed in this instant. Consequently, the spatial propagation vector $\mathbf{k}$, comprehending all spatial components of $\boldsymbol{k}$, has to stay continuous. The frequency has to jump in accordance with the condition prescribing the characteristics of the medium before and after the discontinuous variation ($F=0$).

The condition $F=0$ with a given $\mathbf{k}$ will, in general, yield more than one solution for ω. Frequently, two solutions are obtained: a positive and a negative ω of the same amount. The propagation vector $\mathbf{k}$ is supposed to be the same for the two solutions (with respect to magnitude and direction). The wave corresponding to negative ω, however, travels in the direction opposite to $\mathbf{k}$. We term this wave "reflected" and the other, which travels in the direction of the "incident wave," "refracted," understanding by refraction and reflection the processes at a discontinuity in time. In case of emission from a localized source, the refracted wave is an expanding wave, the reflected wave a collapsing wave.[21]

An interesting special case of a varying medium to which attention was called by Morgenthaler[22] is a *medium varying continuously in time*, but being *homogeneous in space*. Equations (36) with the condition $\nabla_r F=0$ indicate in this case constancy of $\mathbf{k}$ in group propagation. The frequency in a wave group must therefore vary. The case apparently is inverse to that of a pure spatial variation of the medium, in which ω was seen to remain constant in a wave group. The time variation, of course, has to be slow enough to leave the four-dimensional Sommerfeld-Runge law valid.

Now a *discontinuity of the medium at a moving boundary*, a case treated in some generality by Synge,[18] will be considered. In a small space-time volume, the boundary is assumed to be plane and to travel with a constant velocity V in the z direction as described by

$$z-Vt=0. \tag{50}$$

The function F (characterizing the medium) can, around the considered part of the discontinuity surface, be thought of as containing space and time coordinates only in the combination $z-Vt$ and varying discontinuously as $z-Vt$ passes through zero.

The quantity $z-Vt$ is the scalar product of $\boldsymbol{r}$ with a vector $(0, 0, 1, iV/c)$. The discontinuity surface, consequently, is normal to this vector, and the three $\boldsymbol{k}$ components parallel to the discontinuity surface, which have to be continuous, are k_x, k_y, and $(i/c)(Vk_z-\omega)/(1-V^2/c^2)^{1/2}$. As in three-dimensional theory, continuity of these components is required in transition from the incident wave to both the refracted and the reflected waves departing from the discontinuity surface.

Formulation of the present continuity requirement for a differential step of a medium yields Eqs. (43) and (45), which describe the variation of $\mathbf{k}$ and ω in group propagation. This is analogous to the inference in the three-dimensional considerations that Snell's law, originally stated for boundaries, applies on a ray path to differential steps of a continuously varying medium.

The limiting cases $V=0$ and $V\to\infty$ refer to a stationary boundary and to a discontinuity in time. In slight deviation from the limiting cases, the reflected wave shows shifts of the frequency and of k proportional to V or $1/V$, respectively (provided, certain singular condi-

[21] In terms of *advanced potentials* the collapsing electromagnetic wave corresponds to a sink at the place of the actual source.

[22] F. R. Morgenthaler, IRE Trans. **MTT-6**, 167 (1958).

tions are excluded). The shift proportional to V is known as Doppler effect. The shift proportional to $1/V$, which is a consequence of nonsimultaneity of the variation in the entire space, may accordingly be called "inverse Doppler effect."

Under more general conditions, in particular relating to intermediate V, there may be a larger number of solutions for waves leaving a discontinuity surface. This may prevent a clear distinction between refracted and reflected waves. Rydbeck[23] discusses a similar phenomenon in emission of radiation from a moving source in a dispersive medium (in fact, a plasma). In the one-dimensional problem, occasionally more than two emitted waves (as would correspond to the two directions of propagation) are found.

As to the processes leading to a time-variation of the medium, various possibilities may be noted. In addition to motions and deformations of material bodies, there are shock fronts, in which the state of matter appears to be propagated, eventually without transportation of matter. Ionization or dissociation of a medium by ultraviolet light is a possible means for almost instantaneous or properly timed variation of a medium in a wider volume. In ionospheric radio wave propagation, frequency fluctuations have been observed,[24,25] which may be a Doppler effect due to displacement of the ionospheric reflection level or, more generally, the result of a modification of the ionosphere while the waves pass through it.

4.4. Focusing and Diffraction

All the considerations of Secs. 4.1 to 4.3, being subject to the limitations of geometric optics, are a part of a four-dimensional geometric optics. We may now touch on a topic of particular interest in geometric optics: *focusing*. Focusing of first degree occurs at a *caustic*, which in four-dimensional space is a three-dimensional structure, corresponding in $\mathbf{r}$ space to a moving (and perhaps varying) surface. The waves approach a caustic from one side and leave it toward the same side. In the case of a caustic at rest, this requires that the rays (or ray paths) are tangential to the caustic (the caustic is an envelope surface to the ray paths). At a moving caustic, the group velocities of the two waves coincide and their components normal to the caustic must correspond with the velocity of the caustic in this direction. Otherwise, the wave field as determined by the group propagation concept (Sec. 4.2) would extend beyond the caustic.

A somewhat unusual focusing effect is encountered in the collapsing wave field resulting from a reflection of an expanding wave field at a discontinuity in time (Sec. 4.3). With a small jump of the medium, the collapsing wave field resembles closely the emitted expanding field and is, consequently, well focused on the source of the emission. The focusing is spatial only, yet in three dimensions.

It may be noted that in four dimensions focusing up to the fourth degree, including time-focusing, might be thought of.

A geometric-optical treatment of *diffraction*, as suggested in three dimensions by Keller,[17] is possible in four dimensions. Diffraction in space-time is to be expected, in particular, at discontinuity lines (i.e., singly infinite assemblages of space-time points) and at two-dimensional discontinuity surfaces (i.e., doubly infinite assemblages of space-time points). The $\boldsymbol{k}$ components parallel to the lines or surfaces have to remain continuous in transition from incident to diffracted waves, just as the $\mathbf{k}$ components in three-dimensional diffraction theory.[17] The relationship between this concept of diffraction and the process of refraction (taking place at a triply infinite assemblage of space-time points) is elucidated by Synge's consideration of "refraction through a hole,"[18] which, in fact, represents diffraction.

4.5. Concluding Remarks

The above considerations, developing out of the four-dimensional Sommerfeld-Runge law, are not the only way to arrive at the various statements of the four-dimensional theory. In place of the four-dimensional Sommerfeld-Runge law, for instance, the postulate of a uniquely defined, continuous phase of the waves could be used to start with. The Sommerfeld-Runge law, on the other hand, proved to be very suggestive by calling attention to a number of processes in space-time, many of them analogs of well-known three-dimensional phenomena. The Sommerfeld-Runge law in three and four dimensions is not only a means for general deductions but, as the discussion of a three-dimensional example (in Sec. 2.4) indicated, also a tool for computation purposes.

The presented four-dimensional theory, as a geometric-optical approach, is applicable to discontinuities and to sufficiently shallow gradients in space-time. In three dimensions, the WKB method with inclusion of internal reflections and related procedures[6,26] deal with arbitrary profiles of a medium by using geometric optics at differential slabs of the medium. Transfer of such procedures to four dimensions leads to wave theories based on four-dimensional geometric optics.

ACKNOWLEDGMENTS

The author gratefully acknowledges stimulating discussions he had in an early stage of the work with Dr. K. Suchy. Thanks are due to Gerald M. Daniels for assistance in preparing the manuscript.

[23] O. E. H. Rydbeck, Research Report No. 10, Research Laboratory of Electronics, Chalmers University of Technology, Gothenburg, Sweden, 1960 (unpublished).

[24] T. Ogawa, Proc. IRE, **46**, 1934 (1958).

[25] K. L. Chan and O. G. Villard, Jr., J. Geophys. Res. **67**, 973 (1962).

[26] H. Poeverlein, J. Research Natl. Bur. Standards **65D**, 465 (1961), in particular Appendix 2, pp. 472–473.