JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 9, NUMBER 2 FEBRUARY 1968

Dirac Algebra and the Six-Dimensional Lorentz Group*

A. TEN KATE
Institute for Theoretical Physics, University of Nijmegen, The Netherlands

(Received 15 June 1967)

We establish the relationship between the Dirac algebra and the six-dimensional Lorentz group. By considering in an appropriate way the fifteen excentrical basis elements of the Dirac algebra as the components of an antisymmetrical tensor in six dimensions, the commutation as well as the anti-commutation rules of the algebra can be written in a six-covariant way. The group of automorphisms of the real Dirac algebra turns out to be isomorphic to the proper six-dimensional Lorentz group. However, this result is very sensitive to the specific choice of Lorentz metric.

INTRODUCTION

The Dirac algebra has been the object of many investigations. These investigations have two important physical results. First, a unified description of boson and fermion fields can be given in terms of this algebra,[1] and second, traces of unitary symmetry are found in its ideal structure.[2] On the other hand the higher-dimensional Lorentz groups grow more and more important in group-theoretical physics.

Therefore it may be important to point out the relationship between the Dirac algebra and the six-dimensional Lorentz group, that is, the group of linear homogeneous transformations of a six-dimensional space which preserves the invariant diagonal metric $(+1, -1, -1, -1, -1, -1)$. An extensive review of the mathematical properties of the algebra has been given by Rashevskij.[3] The present paper adds some new mathematical properties to our knowledge of the algebra.

In Sec. I, we give the definition of the algebra and a short review of its most important features. Section II is devoted to the six-dimensional Lorentz group. A definition is given, and its local and global structure is specified. Then, in Sec. III, we put forward the six-dimensional aspects of the algebra, while, in Sec. IV, the automorphism group of the real algebra is proved to be isomorphic to the proper six-dimensional Lorentz group. The full four-dimensional group is a subgroup of this proper six-dimensional group. In the first four sections we use a Lorentz metric $g^{\mu\nu}$ with diagonal elements $(1, -1, -1, -1)$. In Sec. V we state similar results for the Dirac algebra belonging to the metric $g'^{\mu\nu}$ with $(-1, +1, +1, +1)$.

I. DIRAC ALGEBRA

Let us start with a four-dimensional space with coordinates x^μ and a Lorentz metric $g^{\mu\nu}$ ($\mu, \nu = 0, 1, 2, 3$) with diagonal elements $(1, -1, -1, -1)$, the other elements vanishing. The Clifford algebra corresponding to this metric is called the Dirac algebra. It has sixteen basis elements, namely a scalar element I, four vector elements γ^μ, six tensor elements $\gamma^{\mu\nu} = \frac{1}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)$, four pseudovector elements $\gamma^{\mu 5} = \gamma^\mu\gamma^5$, and a pseudoscalar element $\gamma^5 = \gamma^0\gamma^1\gamma^2\gamma^3$. To

* This work was supported by the Foundation for Fundamental Research on Matter (FOM).
[1] S. Teitler, Nuovo Cimento, Suppl. 3, 1 (1965).
[2] S. Teitler, J. Math. Phys., 7, 1739 (1966).
[3] P. K. Rashevskij, Am. Math. Soc. 6, 1 (1957).

avoid confusion with the usual convention $\gamma^4 = i\gamma^0$ we omit the index 4. The four vector elements satisfy the relation

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}. \tag{1}$$

All the other elements can be obtained from the four vector elements by multiplication, as indicated above. As a consequence, relation (1) fixes the multiplication of the algebra. All linear combinations of these basis elements with complex coefficients define the complex algebra. However, we restrict ourselves in the present paper to the real algebra, that is, to linear combinations with real coefficients. Only in Sec. IV will we need the complex extension of the algebra. There we use the important fact that the complex Dirac algebra is isomorphic to the algebra of complex 4×4 matrices. In that isomorphism all the basis elements of the algebra except the scalar one correspond to traceless matrices.[4]

Relation (1) is covariant under Lorentz transformations. If $x'^\mu = \omega^\mu{}_\nu x^\nu$ defines a Lorentz transformation (that is, $\omega^\kappa{}_\mu\omega^\lambda{}_\nu g^{\mu\nu} = g^{\kappa\lambda}$), then $\gamma'^\mu = \omega^\mu{}_\nu\gamma^\nu$ also satisfies (1). Now we can construct an automorphism of the real algebra by means of the correspondence

$$\begin{aligned} &I \leftrightarrow I,\\ &\gamma^\mu \leftrightarrow \gamma'^\mu = \omega^\mu{}_\nu\gamma^\nu,\\ &\gamma^{\mu\nu} \leftrightarrow \gamma'^{\mu\nu} = \omega^\mu{}_\kappa\omega^\nu{}_\lambda\gamma^{\kappa\lambda},\\ &\quad\text{etc.}\end{aligned} \tag{2}$$

It is easily verified that this correspondence preserves all the algebraic operations. Thus it is an automorphism. Now every Lorentz transformation ω implies such an automorphism $A(\omega)$. On the other hand we can ask ourselves the question: Can every automorphism of the algebra be generated by an ω or is the correspondence between automorphisms and Lorentz transformations one-to-one? The answer must be no. The group of automorphisms is much larger than the four-dimensional Lorentz group. This is shown in Secs. III and IV.

II. SIX-DIMENSIONAL LORENTZ GROUP

Now we consider a six-dimensional space with coordinates X^k and a metric G^{KL} ($K, L = 0, 1, 2, 3, 5, 6$) with diagonal elements $(1, -1, -1, -1, -1, -1)$ and vanishing off-diagonal elements. A six-dimensional coordinate transformation $X'^K = \Omega^K{}_L X_L$ is called a Lorentz transformation if

$$\Omega^K{}_M\Omega^L{}_N G^{MN} = G^{KL}. \tag{3}$$

[4] A. Messiah, *Mécanique quantique* (Dunod et Cie., Paris, 1960), p. 774.

These transformations form a Lie group (we call it the six group). The local structure of such a group is characterized by the commutation rules of its generators. The six group has fifteen generators, which may be considered to be the components of an antisymmetrical tensor M^{KL}. They satisfy the commutation rules

$$[M^{KL}, M^{MN}]_- = G^{KN}M^{LM} - G^{KM}M^{LN} + G^{LM}M^{KN} - G^{LN}M^{KM}. \tag{4}$$

The global structure is specified by the connectivity of the parameter space. This manifold consists of four disjoint parts. Two of them have $\det\Omega = +1$ and the other two $\det\Omega = -1$. One of the parts with $\det\Omega = +1$ contains the identity transformation. It has $\Omega_{00} > 0$ and will be called the proper six group.

III. SIX-DIMENSIONAL ASPECTS OF THE ALGEBRA

To clarify the six-dimensional aspects of the algebra, we write the fifteen excentrical basis elements, i.e., all the basis elements except the scalar one, as the elements of an antisymmetrical 6×6 matrix. Let us define this matrix Γ^{KL} to be

$$\begin{aligned}\Gamma^{KL} &= \tfrac{1}{4}(\gamma^K\gamma^L - \gamma^L\gamma^K) \quad \text{for} \quad (K, L = 0, 1, 2, 3, 5),\\ \Gamma^{K6} &= -\Gamma^{6K} = \tfrac{1}{2}\gamma^K \quad \text{for} \quad (K = 0, 1, 2, 3, 5).\end{aligned} \tag{5}$$

The matrix elements of Γ are not numbers, but they are elements of the real Dirac algebra. The arrangement (5) will prove to be very convenient. For if we compute the commutation rules of the elements Γ^{KL}, we find

$$[\Gamma^{KL}, \Gamma^{MN}]_- = G^{KN}\Gamma^{LM} - G^{KM}\Gamma^{LN} + G^{LM}\Gamma^{KN} - G^{LN}\Gamma^{KM}. \tag{6}$$

Equation (6) may be verified in a straightforward manner by using the definition of Γ^{KL} (5) and the multiplication rules for the basis elements (1). Equation (6) is identical to Eq. (4), so that the real Dirac algebra can be used to represent the six group. In addition, if we let the matrix Γ^{KL} transform as a tensor of rank two under six-dimensional Lorentz transformations, then relation (6) is "six covariant"; i.e., if Γ^{KL} satisfies (6), then $\Gamma'^{KL} = \Omega^K{}_M\Omega^L{}_N\Gamma^{MN}$ does also.

In the same way, the anticommutation rules of the Γ^{KL}'s may be calculated and found to be

$$[\Gamma^{KL}, \Gamma^{MN}]_+ = \tfrac{1}{2}(G^{KN}G^{LM} - G^{KM}G^{LN}) + \tfrac{1}{2}\epsilon^{KLMN}{}_{ST}\,\Gamma^{ST}, \tag{7}$$

where ϵ^{KLMNST} is a completely antisymmetrical tensor with $\epsilon^{012356} = +1$. Because of the odd number of negative signs in the metric, this implies that $\epsilon_{012356} = -1$. Note that (1) may be obtained from (7) by setting $L = N = 6$. Equation (7) is only covariant under six-dimensional Lorentz transformations with $\det \Omega = +1$ since the ϵ tensor changes sign if $\det \Omega = -1$. Addition of Eqs. (6) and (7) for the multiplication of two Γ's gives

$$\Gamma^{KL}\Gamma^{MN} = (G^{KN}\Gamma^{LM} - G^{KM}\Gamma^{LN} + G^{LM}\Gamma^{KN} - G^{LN}\Gamma^{KM}) + \tfrac{1}{4}(G^{KN}G^{LM} - G^{KM}G^{LN}) + \tfrac{1}{4}\epsilon^{KLMN}{}_{ST}\Gamma^{ST}. \tag{8a}$$

The multiplication in the algebra is completely specified if we add one more relation to (8a):

$$I\Gamma^{KL} = \Gamma^{KL}I = \Gamma^{KL}. \tag{8b}$$

Considering I as a six-scalar, Eqs. (8a) and (8b) turn out to be covariant under all Lorentz transformations $\Omega^K{}_L$ with $\det \Omega = +1$.

As a consequence of the covariance of Eqs. (8a) and (8b), we can construct with every $\Omega(\det + 1)$ an automorphism $A(\Omega)$ of the real algebra:

$$\begin{aligned} I &\leftrightarrow I, \\ \Gamma^{KL} &\leftrightarrow \Gamma'^{KL} = \Omega^K{}_M\Omega^L{}_N\Gamma^{MN}. \end{aligned} \tag{9}$$

However, we can imagine that two different Ω's generate one and the same automorphism.

Theorem 1: If Ω and Ω' generate the same automorphism, then $\Omega' = \pm\Omega$.

Proof: If $A(\Omega) \equiv A(\Omega')$, then

$$\Omega^K{}_M\Omega^L{}_N - \Omega^K{}_N\Omega^L{}_M = \Omega'^K{}_M\Omega'^L{}_N - \Omega'^K{}_N\Omega'^L{}_M. \tag{10}$$

Now multiplying Eq. (10) by $(\Omega^{-1})^S{}_K(\Omega^{-1})^T{}_L$ and summing over K and L, we find that the matrix elements of $\bar{\Omega} = \Omega^{-1}\Omega'$ must satisfy

$$\bar{\Omega}^S{}_M\bar{\Omega}^T{}_N - \bar{\Omega}^S{}_N\bar{\Omega}^T{}_M = G^S{}_MG^T{}_N - G^S{}_NG^T{}_M. \tag{11}$$

This condition is nontrivial only for $S \neq T$ and $M \neq N$. In that case the left-hand side of Eq. (11) is just the determinant of the 2×2 submatrix obtained from the matrix $\bar{\Omega}$ by crossing the rows S and T with the columns M and N. Equation (11) tells us that all these 2×2 submatrices are singular, except those which contain two diagonal elements. In the last case the so-called "principal minors" equal one.

Now, we consider the matrix $\bar{\Omega}^{-1}$. This matrix can be constructed from $\bar{\Omega}$ in the usual way. $\bar{\Omega}^{-1K}{}_L = \pm$ minor of $\bar{\Omega}^K{}_L$ because $\det \bar{\Omega} = 1$. If $K \neq L$, the minor of $\bar{\Omega}^K{}_L$ is zero. This can be seen by expanding this minor in the subminors of two rows, one of which must be the Lth. Thus the off-diagonal elements of $\bar{\Omega}^{-1}$ are all zero, and the same is valid for $\bar{\Omega}$ itself. Because the principal minors equal 1, the diagonal elements of $\bar{\Omega}$ must be all $+1$ or all -1. But that means that $\Omega' = \pm\Omega$. Q.E.D. Now if $\det \Omega$ is $+1$, then $\det(-\Omega)$ is also $+1$. Thus both transformations $+\Omega$ and $-\Omega$ belong to the subgroup with $\det +1$. This subgroup consists of two disjoint parts, one of which is the proper six group. If Ω belongs to one of the parts, then $-\Omega$ belongs to the other. That means that one and only one of the two transformations $+\Omega$ and $-\Omega$ belongs to the proper six group. Consequently, all automorphisms of the type (9) are generated by one and only one element of the proper six group.

The automorphisms (2) are also of type (9). Hence the full four-dimensional Lorentz group is a subgroup of the proper six group. The proper four transformations correspond to those transformations of the proper six group which leave the 5 and 6 axes unchanged. In our arrangement (5), space and time inversion correspond, respectively, to inversion of the 1, 2, 3, 5 and 1, 2, 3, 6 axes.

IV. AUTOMORPHISM GROUP

To complete the picture we need to prove that every automorphism A is generated by an Ω.

Theorem 2: Every automorphism can be written in the form (9). The proof of this theorem is divided in five steps.

Step 1. Every automorphism $A(\Gamma \leftrightarrow \Gamma')$ of the real algebra implies an automorphism of the complex algebra and every automorphism of the complex algebra is inner; i.e., there exists an S in the complex algebra so that $\Gamma' = S\Gamma S^{-1}$.[5] The automorphism fixes the S up to a complex factor $\neq 0$.[6]

Step 2. Now we use the isomorphism of the complex Dirac algebra on the algebra of complex 4×4 matrices. Stating the result of step 1 in terms of matrices, we may say that every automorphism is generated by a nonsingular matrix S. Since S is fixed up to a complex factor, we may choose an S with

[5] Reference 3, p. 33ff.
[6] Reference 3, p. 25.

$\det S = +1$ without loss of generality. Now every nonsingular matrix can be written as the exponent of another matrix[7]:

$$S = \exp T. \tag{12}$$

The matrix T corresponds to an element of the complex algebra. Thus it can always be written as

$$T = aI - a_{KL}\Gamma^{KL} \tag{13}$$

with complex a and a_{KL}, the latter being antisymmetrical in K and L. However, since $\det S = \exp(\operatorname{trace} T)$ and all excentrical basis elements correspond to traceless matrices, we may choose $a = 0$. Finally, every automorphism $A(\Gamma \leftrightarrow \Gamma')$ may be written as

$$\Gamma'^{M}{}_{N} = e^{-a_{KL}\Gamma^{KL}}\Gamma^{M}{}_{N}\, e^{+a_{ST}\Gamma^{ST}}. \tag{14}$$

with complex a_{KL}.

Step 3. We expand (14) in a series by means of the general matrix identity

$$(\exp A)B(\exp -A) = \sum_{n=0}^{\infty} \frac{1}{n!} A^{(n)}\{B\}, \tag{15}$$

where

$$A^{(0)}\{B\} = B$$

and

$$A^{(n)}\{B\} = [A, [A, \cdots, [A, B]_{-} \cdots]_{-}]_{-} \quad (n \text{ commutators}).$$

In our case $A = -a_{KL}\Gamma^{KL}$ and $B = \Gamma^{M}{}_{N}$. Using the commutation rules for the Γ's (6), we find

$$A^{(1)}\{B\} = [A, B] = 2(a^{M}{}_{K}\Gamma^{K}{}_{N} - \Gamma^{M}{}_{K}a^{K}{}_{N}).$$

Considering $a^{K}{}_{L}$ and $\Gamma^{K}{}_{L}$ as 6×6 matrices and suppressing the matrix indices, we get

$$A^{(1)}\{\Gamma\} = 2(a\Gamma - \Gamma a). \tag{16a}$$

And, in general,

$$A^{(n)}\{\Gamma\} = 2(aA^{(n-1)}\{\Gamma\} - A^{(n-1)}\{\Gamma\}a). \tag{16b}$$

Step 4. Now we define a complex six-dimensional Lorentz transformation Ω by

$$\Omega = \exp(a_{KL}D^{KL}), \tag{17}$$

where D^{KL} are the fifteen 6×6 matrices which represent the generators of the six group in the defining representation

$$(D^{KL})^{S}{}_{T} = G^{KS}G^{L}{}_{T} - G^{LS}G^{K}{}_{T}. \tag{18}$$

[7] F. R. Gantmacher, *Matrix Theory* (Chelsea Publishing Company, New York, 1959), Vol. 1, p. 239.

Using the antisymmetry of the $(D^{KL})^{ST} = -(D^{KL})^{ST}$, we can see immediately that Ω is a Lorentz transformation:

$$\begin{aligned}\Omega^{S}{}_{T} &= [\exp(a_{KL}D^{KL})]^{S}{}_{T} \\ &= G^{SV}G_{TW}[\exp(-a_{KL}D^{KL})]^{W}{}_{V} \\ &= G^{SV}G_{TW}\Omega^{-1W}{}_{V},\end{aligned} \tag{19}$$

and from this we get

$$\Omega^{S}{}_{T}\Omega^{U}{}_{V}G^{TV} = G^{SU}.$$

Thus Ω is a complex Lorentz transformation and $\det \Omega = +1$ because of its definition and the tracelessness of the D^{KL}'s. Now we consider the tensor transformation

$$\Gamma''^{MN} = \Omega^{M}{}_{S}\Omega^{N}{}_{T}\Gamma^{ST}. \tag{20}$$

We prove that this transformation is identical to transformation (14). To do this we use (19):

$$\Gamma''^{M}{}_{N} = (\exp a_{KL}D^{KL})^{M}{}_{S}\Gamma^{S}{}_{W}\exp(-a_{ST}D^{ST})^{W}{}_{N}, \tag{21}$$

and, suppressing the index summation, we write in terms of 6×6 matrices

$$\Gamma'' = (\exp a_{KL}D^{KL})\Gamma(\exp -a_{ST}D^{ST}). \tag{21'}$$

This is an expression of the type (15), and it can be reduced along similar lines. Now, $A = a_{KL}D^{KL}$ and $B = \Gamma$. Using (18), we find the commutation rules between A and B:

$$A^{(1)}\{B\} = [A, B] = 2(a\Gamma - \Gamma a) \tag{22a}$$

and, in general,

$$A^{(n)}\{B\} = 2(aA^{(n-1)}\{B\} - A^{(n-1)}\{B\}a). \tag{22b}$$

Comparing this result with (16), we conclude that every term of the series (21) is equal to the corresponding term of (14). Hence $\Gamma''^{MN} = \Gamma'^{MN}$. But that means that we have found a complex Lorentz transformation which generates our automorphism A of the real algebra. We still prove that such a real automorphism can only be generated by real Ω's. This will complete the proof of Theorem 2.

Step 5. Although Ω may be complex, we know that $\Omega^{S}{}_{T}\Omega^{U}{}_{V} - \Omega^{S}{}_{V}\Omega^{T}{}_{U}$ is real because A is an automorphism of the real algebra. Now if Ω is a complex Lorentz transformation, then the complex conjugate Ω^* is also. But then Ω and Ω^* generate the same automorphism, and we are left with two possibilities, $\Omega^* = +\Omega$ and $\Omega^* = -\Omega$, since Theorem 1 holds also for complex Ω. In the first case $\Omega^* = +\Omega$, the

Lorentz transformation is real and that is just what we want. In the second case $\Omega^* = -\Omega$, the Lorentz transformation would be purely imaginary, and we show that purely imaginary six-dimensional Lorentz transformations do not exist. Suppose Ω is a purely imaginary Lorentz transformation,

$$\Omega^K{}_M \Omega^L{}_N G^{MN} = G^{KL}.$$

Because Ω is imaginary, the matrix $Z^S{}_T = i\Omega^S{}_T$ is real. This real matrix satisfies the equality

$$Z^K{}_M Z^L{}_N G^{MN} = -G^{KL}. \tag{23}$$

Now let us consider the rows of this matrix to be vectors. Then (23) asserts that the first row is a spacelike vector, while the other five rows are timelike. Moreover, all the vectors are orthogonal to each other. But five orthogonal, timelike, real vectors do not exist. Thus a real matrix Z satisfying (23) does not exist; so that the case $\Omega^* = -\Omega$ may be excluded. But then the proof is given because $\Omega = \Omega^*$ is real. Consequently, the group of automorphisms of the real Dirac algebra is isomorphic to the proper six group.

V. INVERSE METRIC

If we start with a metric $g'^{\mu\nu}$ $(-1, +1, +1, +1)$ instead of $g^{\mu\nu}$ $(+1, -1, -1, -1)$, we do not find the same results. Although the complex Clifford algebras belonging to the two metrics $g^{\mu\nu}$ and $g'^{\mu\nu}$ are isomorphic to each other, the real algebras are not. Nevertheless, all the preceding results will still hold if we replace G^{KL} $(+1, -1, -1, -1, -1, -1)$ by G'^{KL} $(-1, +1, +1, +1, -1, -1)$. All formulas and statements remain valid if we substitute $g'^{\mu\nu}$ for $g^{\mu\nu}$ and G'^{KL} for G^{KL}. However, we must make one important exception. Because of the change of signature in the six-dimensional Lorentz group, the last statement of Step 5 in the proof of Theorem 2 is not valid. Indeed, the new six group, which leaves G'^{KL} invariant, contains purely imaginary elements in its complex extension. Consider, for example, the matrix

$$E = \begin{pmatrix} 0 & i & 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 & 0 & i \\ 0 & 0 & i & 0 & 0 & 0 \\ 0 & 0 & 0 & i & 0 & 0 \end{pmatrix}.$$

This is a purely imaginary element of the new six group. As a consequence, Theorem 2 breaks down. It is readily verified that E generates an automorphism of the real algebra, and this automorphism is not of type (9) because Ω is not real. However, if we modify Theorem 2 in such a way that purely imaginary Ω's are also admitted, then Theorem 2 remains valid.

Theorem 2': All automorphisms of the real Dirac algebra belonging to $g'^{\mu\nu}$ can be written in the form

$$\Gamma'^{KL} = \Omega^K{}_M \Omega^L{}_N \Gamma^{MN}, \tag{24}$$

where Ω is a real or purely imaginary element of the new six group. Now every purely imaginary Ω can be written as the product of a real Ω; the E, and the product of two imaginary Ω's, is a real Ω. Hence, the parameter space of the group of automorphisms of the new algebra splits up into two disjoint parts. One of these parts contains the identity and is an invariant subgroup of index 2. This subgroup is isomorphic to the proper six group which leaves G'^{KL} invariant.

It is remarkable how sensitive these considerations are to the specific choice of the metric.

ACKNOWLEDGMENTS

The author wishes to thank Professor A. G. M. Janner, H. O. Singh Varma, and C. Dullemond for some discussions on this subject.

Representations and Classes in Groups of Finite Order

A. GAMBA
Istituto di Fisica dell'Università, Genova, Italy

(Received 27 April 1967)

There are two main results in this paper. First, it is shown that we can develop a theory of classes in close analogy to the usual theory of representations. We can introduce concepts, such as reducible and irreducible classes, sum and product of classes, reduction of a class when going from a group to a subgroup, etc. Second, it is shown that it is possible to associate a "magic square" to each group. It is related to the numbers of pairs of commuting elements between classes and it can be used immediately to find the structure of the "tensor operators" of the group.

1. INTRODUCTION

The whole physical literature about the theory of representations of groups of finite order amounts essentially to a few standard procedures. One sets up the table of characters of the irreducible representations—a square table with representations as column entries and classes as row entries. Then, using the orthonormal properties of characters, one can easily perform such mathematical operations as the reduction of a representation into irreducible ones, the reduction of a representation from a group to a subgroup, etc. This amounts to using only one-half of the orthonormal properties of characters. It is well known that orthonormal properties similar to those obtained by multiplying columns of the table of characters also hold when multiplying rows.[1] In Sec. 3 we study the row equivalent of the usual column theory. It will therefore be a theory of classes, in close analogy to the usual theory of representations.

The second part of this paper (Sec. 4) shows that, to each group of finite order, one can associate a "magic square" related to the numbers of pairs of commuting elements between classes. This magic square can be immediately translated into a table of characters that shows clearly the structure of the "tensor operators" of the group.

2. NOTATION

In order to simplify the reading of this paper, each step will be illustrated with examples taken from the two groups π_4 and π_3 (the symmetric permutation group on four and three variables, respectively). The tables of characters of these groups are given in Tables I and II. Whenever we speak of a group, the reader can always refer to Table I for an explanation of symbols and notation. Whenever we speak of a subgroup, Table II is to be consulted. This procedure eliminates the need of a lengthy and cumbersome list of symbols and notation. By simply looking at the two tables, for example, the reader can easily see that columns refer to representations and rows to classes, that an irreducible representation of a group is indicated by R_k and an irreducible representation of a subgroup by ρ_k, etc. In short, the notation used in this paper becomes self-explanatory.

TABLE I. Table of characters of the group π_4 of order $N = 24$.

No. of group elements in each class	Irreducible classes \ Irreducible representations	R_1	R_2	R_3	R_4	R_5
$N = 1$	C_1	$n_1 = 1$	$n_2 = 3$	$n_3 = 2$	$n_4 = 3$	$n_5 = 1$
$N_2 = 6$	C_2	1	1	0	−1	−1
$N_3 = 8$	C_3	1	0	−1	0	1
$N_4 = 6$	C_4	1	−1	0	1	−1
$N_5 = 3$	C_5	1	−1	2	−1	1

[1] Only the basic concepts of group theory (found in almost any book on the subject) are necessary to read the present paper. Therefore, we give only two general references, one for mathematicians and one for physicists: W. Burnside, *Theory of Groups of Finite Order* (Dover Publications, Inc., New York, 1955), 2nd ed.; J. S. Lomont, *Applications of Finite Groups* (Academic Press Inc., New York, 1959).

TABLE II. Table of characters of the group π_3 of order $m = 6$.

No. of group elements in each class	Irreducible classes \ Irreducible representations	ρ_1	ρ_2	ρ_3
$m_1 = 1$	γ_1	$\nu_1 = 1$	$\nu_2 = 2$	$\nu_3 = 1$
$m_2 = 3$	γ_2	1	0	-1
$m_3 = 2$	γ_3	1	-1	1

3. DUALITY BETWEEN REPRESENTATIONS AND CLASSES

Let us call "irreducible class" a set of group elements, ordinarily called a class of conjugate elements. For example, in π_4, there are five irreducible classes (C_1, C_2, C_3, C_4, C_5 of Table I). We can then establish the following "dual" properties: at left we have written the well-known properties of representations (sometimes only indicating them in a sketchy form) for the purpose of comparison with the dual properties of classes written at the right.

To every irreducible representation R_k, we can associate a *column* of the table of characters. For example, to R_2 in π_4, we can associate

$$X_{R_2}(C_k) = \begin{pmatrix} 3 \\ 1 \\ 0 \\ -1 \\ -1 \end{pmatrix} \quad C_k \text{ variable.} \tag{1r}$$

$X_{R_2}(C_k)$ is the character of the irreducible representation R_2. The characters of irreducible representations are orthonormal, i.e.,

$$\frac{1}{N}\sum_k N_k X^*_{R_m}(C_k) X_{R_n}(C_k) = \delta_{mn}. \tag{2r}$$

The *sum* $R_k + R_l$ of two irreducible representations is $\cdots$.

Through the operation of the sum, we can define (reducible) representations. For example,

$$\mathbf{R} = \sum_k \alpha_k R_k. \tag{3r}$$

The character of the representation $\mathbf{R}$ is defined as

$$X_{\mathbf{R}}(C_l) = \sum_k \alpha_k X_{R_k}(C_l). \tag{4r}$$

For example, the representation $\mathbf{R} = R_2 + 3R_5$ has the character

$$X_{\mathbf{R}}(C_k) = \begin{pmatrix} 6 \\ -2 \\ 3 \\ -4 \\ 2 \end{pmatrix}. \tag{5r}$$

To every irreducible class C_k, we can associate a *row* of the table of characters. For example to C_2 in π_4, we can associate

$$X_{R_k}(C_2) = (1, 1, 0, -1, -1) \quad R_k \text{ variable.} \tag{1c}$$

$X_{R_k}(C_2)$ is the character of the irreducible class C_2. The characters of irreducible classes are orthonormal, i.e.,

$$\frac{(N_m N_n)^{\frac{1}{2}}}{N}\sum_k X^*_{R_k}(C_m) X_{R_k}(C_n) = \delta_{mn}. \tag{2c}$$

The *sum* $C_k + C_l$ of two irreducible classes is the set of group elements belonging to both C_k and C_l.

Through the operation of the sum, we can define (reducible) classes. For example,

$$\mathbf{C} = \sum_k \beta_k C_k. \tag{3c}$$

The character of the class $\mathbf{C}$ is defined as

$$X_{R_l}(\mathbf{C}) = \frac{1}{N_{\mathbf{C}}}\sum_k \beta_k N_k X_{R_l}(C_k), \tag{4c}$$

where $N_{\mathbf{C}}$ is the number of group elements in class $\mathbf{C}$ ($N_{\mathbf{C}} = \sum_k \beta_k N_k$).

For example, the class $\mathbf{C} = C_2 + 3C_5$ has the character

$$X_{R_k}(\mathbf{C}) = (1, -\tfrac{1}{5}, \tfrac{6}{5}, -1, \tfrac{1}{5}) \quad (N_{\mathbf{C}} = 15). \tag{5c}$$

Using (2r), one can easily compute how many times (α_k) the irreducible representation R_k is contained in the representation $\mathbf{R}$:

$$\alpha_k = \frac{1}{N} \sum_l N_l \mathrm{X}^*_{R_k}(C_l) \mathrm{X}_{\mathbf{R}}(C_l). \tag{6r}$$

For example, with (5r), one gets

$$\alpha_4 = \tfrac{1}{24}[3 \cdot 6 + 6(-1)(-2) + 8 \cdot 0 \cdot 3 + 6 \cdot 1 \cdot (-4) + 3(-1) \cdot 2] = 0,$$
$$\alpha_5 = \tfrac{1}{24}[1 \cdot 6 + 6(-1)(-2) + 8 \cdot 1 \cdot 3 + 6(-1)(-4) + 3 \cdot 1 \cdot 2] = 3, \text{ etc.} \tag{7r}$$

The *product* $\mathbf{R} \times \mathbf{R}'$ of two representations is $\cdots$.

The product representation has for character the product of its characters:

$$\mathrm{X}_{\mathbf{R}\times\mathbf{R}'}(C_k) = \mathrm{X}_{\mathbf{R}}(C_k) \cdot \mathrm{X}_{\mathbf{R}'}(C_k) \tag{8r}$$

For example, the product representation $\mathbf{R} \times R_3$ [$\mathbf{R}$ being given by (5r)] has for its character

$$\mathrm{X}_{\mathbf{R}\times\mathbf{R}_3}(C_k) = \begin{pmatrix} 12 \\ 0 \\ -3 \\ 0 \\ 4 \end{pmatrix}. \tag{9r}$$

It is, of course, easy to decompose the product representation into irreducible representations using formula (6r). In our example (9r), one obtains

$$\mathbf{R} \times R_3 = R_2 + 3R_3 + R_4: \tag{10r}$$

Reduction of a representation $\mathbf{R}$ into irreducible representations of a subgroup.

Knowing which classes of the group contain irreducible classes of the subgroup

$$C_s \supset \gamma_t \quad (\supset \text{ means "contains"}), \tag{11r}$$

one can write for the character in the subgroup

$$\mathrm{X}_{\mathbf{R}}(\gamma_t) = \mathrm{X}_{\mathbf{R}}(C_s), \tag{12r}$$

and then proceed to the reduction using (6r).

The order of the subgroup is the number of matrices of the representation $\mathbf{R}$ which are carried over into the subgroup.

Using (2c), one can easily compute how many times (β_k) the irreducible class C_k is contained in the class $\mathbf{C}$:

$$\beta_k = \frac{N_{\mathbf{C}}}{N} \sum_l \mathrm{X}^*_{R_l}(C_k) \mathrm{X}_{R_l}(\mathbf{C}). \tag{6c}$$

For example, with (5c), one gets

$$\beta_4 = \tfrac{15}{24}[1 \cdot 1 + (-1)(-\tfrac{1}{5}) + 0 \cdot \tfrac{6}{5} + 1 \cdot (-1) + (-1) \cdot \tfrac{1}{5}] = 0,$$
$$\beta_5 = \tfrac{15}{24}[1 \cdot 1 + (-1)(-\tfrac{1}{5}) + 2 \cdot \tfrac{6}{5} + (-1)(-1) + 1 \cdot \tfrac{1}{5}] = 3, \quad \text{etc.} \tag{7c}$$

The *product*, $\mathbf{C} \times \mathbf{C}'$ of two classes is the class obtained by taking all the group elements obtained through multiplication of one element of class $\mathbf{C}$ by one element of class $\mathbf{C}'$. (It is easily seen that $\mathbf{C} \times \mathbf{C}' = \mathbf{C}' \times \mathbf{C}$.)

The product class has for its character

$$\mathrm{X}_{R_k}(\mathbf{C} \times \mathbf{C}') = \frac{1}{n_k} \mathrm{X}_{R_k}(\mathbf{C}) \cdot \mathrm{X}_{R_k}(\mathbf{C}'). \tag{8c}$$

For example, the product class $\mathbf{C} \times C_3$ [$\mathbf{C}$ being given by (5c)] has for its character

$$\mathrm{X}_{R_k}(\mathbf{C} \times C_3) = (1, 0, -\tfrac{3}{5}, 0, \tfrac{1}{5}). \tag{9c}$$

It is, of course, easy to decompose the product class into irreducible classes using formula (6c). In our example (9c), one obtains

$$\mathbf{C} \times C_3 = 4C_2 + 9C_3 + 4C_4: \tag{10c}$$

Reduction of a class $\mathbf{C}$ into irreducible classes of a subgroup.

Knowing the decomposition of the representations R_k into irreducible representations ρ_t of the subgroup

$$R_s = \sum_t \omega_{st} \rho_t, \tag{11c}$$

one can write for the character in the subgroup

$$\mathrm{X}_{\rho_t}(\mathbf{C}) = \sum_s \omega_{st} \mathrm{X}_{R_s}(\mathbf{C}) \Big/ \sum_s \omega_{s1} \mathrm{X}_{R_s}(\mathbf{C}), \tag{12c}$$

and then proceed to the reduction using (6c).

The number of elements of class $\mathbf{C}$ which are carried over into the subgroup is given by

$$m_c = \frac{m}{N} N_{\mathbf{C}} \sum_s \omega_{s1} \mathrm{X}_{R_s}(\mathbf{C}). \tag{12c'}$$

For example, in order to reduce the representation of π_4 [given in (9r)] with respect to the subgroup π_3, one first observes that

$$\begin{cases} C_1 \supset \gamma_1, \\ C_2 \supset \gamma_2, \\ C_3 \supset \gamma_3, \\ C_4 \supset \text{no classes of } \pi_3, \\ C_5 \supset \text{no classes of } \pi_3; \end{cases} \tag{13r}$$

then one writes for the character in π_3

$$\mathrm{X}_{\mathbf{R}\times R_3}(\gamma_i) = \begin{pmatrix} 12 \\ 0 \\ -3 \end{pmatrix}. \tag{14r}$$

Finally, using (6r), one obtains

$$\mathbf{R} \times R_3 = \rho_1 + 5\rho_2 + \rho_3. \tag{15r}$$

Algebra of representations. In a group with n irreducible representations, we can associate to every representation R_k an $n \times n$ matrix, such that the sum and the product of representations become the sum and the product of the corresponding matrices. Since the product of representations is commutative, all the corresponding matrices commute with one another. One association is as follows: The matrix element $(R_k)_{il}$ on the ith row and lth column of the matrix R_k is given by

$$(R_k)_{il} = \frac{1}{N} \sum_s N_s \mathrm{X}_{R_k}(C_s)\mathrm{X}^*_{R_i}(C_s)\mathrm{X}_{R_l}(C_s). \tag{16r}$$

For example, with π_4 we obtain (all missing elements are equal to zero)

$$R_1 = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & \end{pmatrix},$$

$$R_2 = \begin{pmatrix} & 1 & & & \\ 1 & 1 & 1 & 1 & \\ & 1 & & 1 & \\ & 1 & 1 & 1 & 1 \\ & & & 1 & \end{pmatrix},$$

$$R_3 = \begin{pmatrix} & & 1 & & \\ & 1 & & 1 & \\ 1 & & 1 & & 1 \\ & 1 & & 1 & \\ & & 1 & & \end{pmatrix}, \tag{17r}$$

For example, in order to reduce the class of π_4 given in (9c) with respect to the subgroup π_3, one first observes that

$$\begin{cases} R_1 = \rho_1, \\ R_2 = \rho_1 + \rho_2, \\ R_3 = \rho_2, \\ R_4 = \rho_2 + \rho_3, \\ R_5 = \rho_3; \end{cases} \tag{13c}$$

then one writes for the character in π_3

$$\mathrm{X}_{\rho_i}(\mathbf{C} \times C_3) = (1, -\tfrac{3}{5}, \tfrac{1}{5}). \tag{14c}$$

Finally, using (6c), one obtains

$$\mathbf{C} \times C_3 \supset 4\gamma_2 + 9\gamma_3. \tag{15c}$$

Algebra of classes. In a group with n irreducible classes, we can associate to every class C_k an $n \times n$ matrix, such that the sum and the product of classes become the sum and the product of the corresponding matrices.

Since the product of classes is commutative, all the corresponding matrices commute with one another. One association is as follows: The matrix element $(C_k)_{il}$ on the ith row and lth column of the matrix C_k is given by

$$(C_k)_{il} = \frac{N_k N_l}{N} \sum_s \frac{\mathrm{X}_{R_s}(C_k)\mathrm{X}^*_{R_s}(C_i)\mathrm{X}_{R_s}(C_l)}{n_s}. \tag{16c}$$

For example, with π_4 we obtain (all missing elements are equal to zero)

$$C_1 = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix},$$

$$C_2 = \begin{pmatrix} & 6 & & & \\ 1 & & 4 & & 1 \\ & 3 & & 3 & \\ & & 4 & & 2 \\ & 2 & & 4 & \end{pmatrix},$$

$$C_3 = \begin{pmatrix} & & 8 & & \\ & 4 & & 4 & \\ 1 & & 4 & & 3 \\ & 4 & & 4 & \\ & & 8 & & \end{pmatrix}, \tag{17c}$$

(Eq. 17r *cont.*)

$$R_4 = \begin{pmatrix} & & & 1 & \\ & 1 & 1 & 1 & 1 \\ & 1 & & 1 & \\ 1 & 1 & 1 & 1 & \\ & 1 & & & \end{pmatrix},$$

$$R_5 = \begin{pmatrix} & & & & 1 \\ & & & 1 & \\ & & 1 & & \\ & 1 & & & \\ 1 & & & & \end{pmatrix}. \tag{17r}$$

The set of matrices (16r) is obtained as follows. Write each irreducible representation as an n-dimensional unit vector:

$$R_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \quad R_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix},$$

$$R_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \quad \text{etc.} \tag{18r}$$

Then, knowing the result of the products $R_k \times R_1$, $R_k \times R_2$, $R_k \times R_3, \cdots$, from the table of characters, one can immediately write down the matrix elements of R_k. Since all the matrices (17r) commute with one another, they have the same set of n eigenvectors in common. It is easily seen that the eigenvectors of the matrices (17r) are the following ones:

$$c_1 = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 3 \\ 1 \end{pmatrix}, \quad c_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \\ -1 \end{pmatrix}, \quad c_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix},$$

$$c_4 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad c_5 = \begin{pmatrix} 1 \\ -1 \\ 2 \\ -1 \\ 1 \end{pmatrix}, \tag{19r}$$

(Eq. 17c *cont.*)

$$C_4 = \begin{pmatrix} & & & 6 & \\ & & 4 & & 2 \\ & 3 & & 3 & \\ 1 & & 4 & & 1 \\ & 4 & & 2 & \end{pmatrix},$$

$$C_5 = \begin{pmatrix} & & & & 3 \\ & 1 & & 2 & \\ & & 3 & & \\ & 2 & & 1 & \\ 1 & & & & 2 \end{pmatrix}. \tag{17c}$$

The set of matrices (16c) is obtained as follows. Write each irreducible class as an n-dimensional unit vector:

$$C_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \quad C_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix},$$

$$C_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \quad \text{etc.} \tag{18c}$$

Then, knowing the result of the products $C_k \times C_1$, $C_k \times C_2$, $C_k \times C_3, \cdots$, from the table of characters, one can immediately write down the matrix elements of C_k. Since all the matrices (17c) commute with one another, they have the same set of n eigenvectors in common. It is easily seen that the eigenvectors of the matrices (17c) are the following ones:

$$r_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad r_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \\ -1 \\ -1 \end{pmatrix}, \quad r_3 = \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \\ 2 \end{pmatrix},$$

$$r_4 = \begin{pmatrix} 3 \\ -1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad r_5 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \tag{19c}$$

i.e., vectors having the class characters as components. The eigenvalue of the matrix R_k, corresponding to the eigenvector c_l, is $X_{R_k}(C_l)$. Thus, for example,

$$R_3c_4 = 0, \quad R_3c_5 = 2c_5, \quad \text{etc.} \tag{20r}$$

This result holds in general; i.e.,

Theorem: The eigenvectors of R_k are the vectors c_l:

$$c_l = \begin{pmatrix} X^*_{R_1}(C_l) \\ X^*_{R_2}(C_l) \\ X^*_{R_3}(C_l) \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \tag{21r}$$

and

$$R_kc_l = X_{R_k}(C_l)c_l. \tag{22r}$$

The proof is easily obtained by writing Eq. (22r), explicitly using (16r) and (21r), and simplifying the resulting expression with formula (2c).

i.e., vectors having the representation characters as components. The eigenvalue of the matrix C_k, corresponding to the eigenvector r_l, is $(N_k/n_l)X_{Rl}(C_k)$. Thus, for example,

$$C_3r_4 = 0, \quad C_3r_5 = 8r_5, \quad \text{etc.} \tag{20c}$$

This result holds in general, i.e.,

Theorem: The eigenvectors of C_k are the vectors r_l:

$$r_l = \begin{pmatrix} X^*_{R_l}(C_1) \\ X^*_{R_l}(C_2) \\ X^*_{R_l}(C_3) \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \tag{21c}$$

and

$$C_kr_l = \frac{N_k}{n_l} X_{R_l}(C_k)r_l. \tag{22c}$$

The proof is easily obtained by writing Eq. (22c), explicitly using (16c) and (21c), and simplifying the resulting expression with formula (2r).

4. A "MAGIC SQUARE" ASSOCIATED TO ANY FINITE GROUP

Consider the N_k group elements belonging to an irreducible class C_k. Call them $C_k^{(1)}, C_k^{(2)}, \cdots, C_k^{(s)}, \cdots, C_{(k)}^{(N_k)}$. Let U be an arbitrary element of the group. From the definition of class, we see that

$$UC_k^{(s)}U^{-1} = C_k^{(t)}, \tag{23}$$

i.e., the elements of a class are transformed into themselves. Expressing each element $C_k^{(s)}$ as a unit vector in an N_k-dimensional space, we have

$$C_k^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \quad C_k^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \quad C_k^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \quad \text{etc.} \tag{24}$$

We can write Eq. (23) in the form

$$M(U)C_k^{(s)} = C_k^{(t)}, \tag{25}$$

where $M(U)$ is a matrix. We have thus obtained an N_k-dimensional representation of the group. We call it a C_k-class representation of the group R_{C_k} and we proceed to reduce it into irreducible representations. It is easily seen from Eq. (23) that the value of the character of the representation R_{C_k} corresponding to the element U is *the number of elements in class C_k that commute with U*, i.e., $X_{R_{C_k}}(C_s)$ is equal to the number of elements of class C_k that commute with a fixed element of class C_s.

For example, the characters for the class representations of π_4 are given in Table III. If we multiply each row of the table of characters of class representations by the corresponding number of group elements in that class, we obtain a "magic square" where the sum of each row and of each column is always equal to the order of the group. See Table IV for an example.

The number at the intersection of the kth row with the lth column in the magic square is the number of pairs of commuting elements, one from class C_k and one from class C_l. The close relation between "class commutators" and characters of class representations is thus made transparent.

Having the characters, we can reduce the class representations in the usual fashion. For example,

TABLE III. Characters of the class representations of π_4.

No. of group elements in each class	Irreducible classes \ Class representations	R_{C_1}	R_{C_2}	R_{C_3}	R_{C_4}	R_{C_5}
1	C_1	1	6	8	6	3
6	C_2	1	2	0	0	1
8	C_3	1	0	2	0	0
6	C_4	1	0	0	2	1
3	C_5	1	2	0	2	3

with π_4 we obtain

$$\begin{cases} R_{C_1} = R_1, \\ R_{C_2} = R_1 + R_2 + R_3, \\ R_{C_3} = R_1 + R_2 + R_4 + R_5, \\ R_{C_4} = R_1 + R_3 + R_4, \\ R_{C_5} = R_1 + R_3. \end{cases} \tag{26}$$

Let us note in passing that the number of times (α_k) each irreducible representation R_k appears in the whole set of (irreducible) class representations is simply

$$\alpha_k = \sum_l \chi_{R_k}(C_l). \tag{27}$$

For example, in π_4

$$\begin{aligned} \alpha_1 &= 1 + 1 + 1 + 1 + 1 = 5, \\ \alpha_2 &= 3 + 1 + 0 - 1 - 1 = 2, \text{ etc.} \end{aligned} \tag{28}$$

What is the meaning of the class representation and its reduction? It is easily seen that our definition (23)–(25) is equivalent to the definition of a "tensor operator" in the physical literature.[2]

The tensor belonging to R_1 (one and only one for each class representation) is the sum of all group elements in a class; it is the "class operator." The entire set of class operators constitutes "the complete set of commuting observable in the group" in the language of quantum mechanics. The formulas of this section show how many (and which) tensor operators are contributed by the elements of each class. An explicit construction can be obtained by the usual "projection operator" technique: The part of the operator $C_k^{(s)}$ that belongs to the irreducible representation R_l is given by

$$\frac{n_l}{N} \sum_{\mu,\nu} \chi^*_{R_l}(C_\mu^{(\nu)}) \cdot C_\mu^{(\nu)} C_k^{(s)} (C_\mu^{(\nu)})^{-1}. \tag{29}$$

Since an example in π_4 would require too much space, we give an example in π_3. Writing for simplicity

$$\begin{aligned} 1 = P_{(1)(2)(2)}, \quad A = P_{(12)(3)}, \quad B = P_{(13)(2)}, \\ C = P_{(1)(23)}, \quad D = P_{(123)}, \quad E = P_{(132)}, \end{aligned} \tag{30}$$

we obtain for A, using formula (29),

$$\begin{cases} \frac{1}{3}(A + B + C), & \text{belongs to } \rho_1, \\ \frac{1}{3}(2A - B - C), & \text{belongs to } \rho_2. \end{cases} \tag{31}$$

For D we obtain

$$\begin{cases} \frac{1}{2}(D + E), & \text{belongs to } \rho_1, \\ \frac{1}{2}(D - E), & \text{belongs to } \rho_3, \end{cases} \tag{32}$$

and so on. The tensor operators of the group π_3 are the following:

$$\begin{aligned} &1, A + B + C, \ D + E, && \text{belong to } \rho_1, \\ &A - B, \ A - C, && \text{belong to } \rho_2, \\ &D - E, && \text{belongs to } \rho_3. \end{aligned} \tag{33}$$

We have seen in this section that a magic square is associated with each finite group, one which gives the number of pairs of commuting elements between two irreducible classes. This magic square can be immediately interpreted as a table of characters for the class representations, which, in turn, gives the structure of the tensor operators of the group.

It is easy to surmise that it should be possible to write down the "magic square" immediately—or very simply—from the table of characters of the irreducible representations. However, so far I have not succeeded in finding such a connection.

TABLE IV. Class commutators in the group π_4 (Magic square).

	C_1	C_2	C_3	C_4	C_5
C_1	1	6	8	6	3
C_2	6	12	0	0	6
C_3	8	0	16	0	0
C_4	6	0	0	12	6
C_5	3	6	0	6	9

[2] See J. S. Lomont, Ref. 1, p. 85.

Structure of the Poincaré Generators

GORDON N. FLEMING
Department of Physics, The Pennsylvania State University, University Park, Pennsylvania

(Received 29 June 1967)

The Poincaré generators for an open system augmented by the interaction parts of the full Poincaré generators are shown to satisfy a closed set of coupled differential equations having a form which is independent of the nature of the interaction parts. The differential equations are formulated in the hyperplane formalism, the differentiation being with respect to the hyperplane parameters. The general solutions of the equations are studied, yielding relations among the augmented generators that must be preserved in the limit of zero interaction, i.e., for a closed system. Introducing a hyperplane-dependent Hamiltonian density in a manner not implying local field theory, the obtained relations are shown to yield expressions for *all* the generators of a closed system in terms of the Hamiltonian density and its derivatives alone.

1. INTRODUCTION

The principle of relativity along with the Lorentz transformation rules demand that if one inertial observer can measure a particular observable at an instant then any other inertial observer can measure the "same" observable on an arbitrary spacelike hyperplane. In particular, the concept of a time-dependent *observable* must always be generalizable to a concept of a hyperplane-dependent observable.[1]

These statements are almost academic when one is concerned only with observables like the Poincaré generators of a closed system which are time, and hence hyperplane, independent. However, consideration of physical systems which are not closed, on the one hand, or of the structure of the Poincaré generators in terms of dynamical variables which may be regarded as more fundamental, on the other hand, does require the general point of view of the *hyperplane formalism* if all the facets of a Poincaré invariant theory are to be explored.[2]

In this article the problem of the relations between the Poincaré generators of open systems and the structure of the Poincaré generators for closed systems in terms of a hyperplane-dependent Hamiltonian density are studied in the framework of the hyperplane formalism. In particular, it is shown that the latter problem has a complete solution, without any appeal to quantum field theory of the Lagrangian form or otherwise, via the consideration of the former problem.

In spirit this investigation stands somewhere between those going under the heading of axiomatic or asymptotic field theory[3] and those commonly referred to as Lagrangian field theory.[4] The concern for some details of the internal structure of the Poincaré generators, as well as the modest rigor of some of the subsequent manipulations, reflects sympathy for the more conventional forms of relativistic quantum theory. At the same time, the attempt to extract the information from a small set of seemingly very general principles, rather than assuming a very powerful but possibly internally inconsistent starting point such as an all-encompassing action principle, reflects admiration for the caution of the more modern school.

In Sec. 2 the Poincaré generators for a composite system are considered and are assumed to be separable into contributions from the constituent subsystems and the interaction between them. Arguments are presented for the circumstances under which the interaction parts of the generators acquire an especially simple form, viz., the choice of a fundamental set of dynamical variables with *interaction-independent equal hyperplane commutators.* The simple form of the interaction parts provokes the introduction of the *hyperplane generators.*[1] In Sec. 3 it is shown that the generators of any *one* of the subsystems, augmented by the interaction parts, satisfy a closed set of coupled differential equations among themselves. Furthermore, the form of these differential equations is completely independent of the nature or strength of the interaction between the subsystems or of the nature of the *other* subsystem. Consequently the differential equations can be used to study properties of the generators of a system which must hold regardless of whether the

[1] G. N. Fleming, J. Math. Phys. 7, 1959 (1967).
[2] Examples of the use of the hyperplane formalism in the analysis of particular physical problems are given in G. N. Fleming, Phys. Rev. **137**, B188 (1965); **154**, 1475 (1967). A restricted use of the notation and ideas of the hyperplane formalism in conventional quantum field theory occurs in the classic book by J. M. Jauch and F. Rohrlich, *Theory of Photons and Electrons* (Addison-Wesley Publishing Company, Inc., Reading, Mass., 1955).
[3] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (W. A. Benjamin, Inc., New York 1964); R. Jost, *The General Theory of Quantized Fields* (American Mathematical Society, 1966).
[4] For a modern rendition of the general theory see B. S. DeWitt, *Dynamical Theory of Groups and Fields* (Gordon and Breach Science Publishers, New York, 1965).

system is open or closed. In Sec. 4 the general solutions of the coupled differential equations are sought and, to a considerable extent, found without making simplifying assumptions concerning the details of the interaction. The limiting case of zero interaction is then considered to obtain further results. The results of this section are explicit expressions for the hyperplane generators of reorientations of the hyperplane and rotations within the hyperplane, i.e., the $N_\mu(\eta, \tau)$ and $J_\mu(\eta, \tau)$, for all values of the hyperplane parameters (η_μ, τ) in terms of the hyperplane Hamiltonian $H(\eta, \tau)$ and the generator of reorientation at a particular τ value, $N_\mu(\eta, \tau_0)$.[1] The final discussion of the generator $K_\mu(\eta, \tau)$ is already implied in the author's earlier paper on the hyperplane formalism. Finally, in Sec. 5, after introducing the *hyperplane Hamiltonian density* via a familiar trick[5] that is independent of field theory in the usual sense, the expressions obtained in Sec. 4 are applied to a closed system to yield expressions for *all* the generators in terms of the hyperplane Hamiltonian density and its hyperplane derivatives. In these last manipulations a principle of *maximal causality* is invoked which demands that the generators be expressible as functionals of dynamical variables defined on only one hyperplane.[6] A summary in Sec. 6 briefly recapitulates the final results and compares them to the corresponding equations that follow from a Lagrangian theory of local quantized fields.

2. PARTITIONING THE GENERATORS

Let P_μ and $M_{\mu\nu}$ be the Hermitian generators of the Poincaré group defined in the quantum-mechanical state space of a closed physical system. The generators then have no time dependence. If the physical system can be regarded as composed of two interacting subsystems, then one may expect to be able to decompose the generators for the entire system into contributions from each subsystem and from the interaction. A "contribution from a subsystem" means a part depending only on dynamical variables referring or "belonging" to the subsystem of interest. The contribution from the interaction is a part in which dynamical variables from both subsystems are combined in a nonadditive way. These characterizations are a long way from uniquely determining the individual contributions, and it is, in fact, not clear that such a decomposition can be effected in a meaningful way in all interesting cases. Nevertheless, one's physical intuition indicates that the partitioning makes sense for "weak" coupling between the subsystems and such coupling will suffice for present purposes. The more precise delineation of the contributions to the full generators is taken up after considering the time dependence of the contributions.

The various parts of the full generators do depend on the time, providing there is some coupling between the subsystems, and one may write

$$P_\mu = P_\mu^{(1)}(t) + P_\mu^{(2)}(t) + V_\mu(t), \tag{2.1a}$$

$$M_{\mu\nu} = M_{\mu\nu}^{(1)}(t) + M_{\mu\nu}^{(2)}(t) + U_{\mu\nu}(t), \tag{2.1b}$$

where the superscripts denote the parts referring to the individual subsystems and V_μ and $U_{\mu\nu}$ are the interaction parts. This notation, however, already destroys the manifest covariance of the formalism by introducing the time variable of a particular inertial frame. Manifest covariance is highly desirable in the type of investigation intended here, where one wants as many as possible of the demands of the principle of Poincaré invariance to be satisfied automatically. This difficulty is easily bypassed by realizing that Poincaré invariance itself demands that if the decomposition of the full generators can be carried out at *any definite time in any inertial frame*, then it can also be carried out on *any spacelike hyperplane in any definite inertial frame*.[1] Otherwise one could not translate the decomposition, carried out in one frame, into the language of any other frame.[7] Thus the parts of the full generators become operators defined on arbitrary spacelike hyperplanes (η_μ, τ), and one writes

$$P_\mu = P_\mu^{(1)}(\eta, \tau) + P_\mu^{(2)}(\eta, \tau) + V_\mu(\eta, \tau), \tag{2.2a}$$

$$M_{\mu\nu} = M_{\mu\nu}^{(1)}(\eta, \tau) + M_{\mu\nu}^{(2)}(\eta, \tau) + U_{\mu\nu}(\eta, \tau). \tag{2.2b}$$

Notice that in the preceding discussion no reference, and therefore no commitment, to quantum field theory has been made.

Now suppose that the basic dynamical variables, in terms of which the parts of the generators and all other dynamical variables are expressed, are chosen to have *fixed equal-hyperplane commutation relations*

[5] In general, if $[A, B] = 0$ and

$$\langle B'|A|B''\rangle \equiv 2\pi\delta(B' - B'')\langle B'|a|B'\rangle,$$

then

$$A = \int_{-\infty}^{\infty} d\lambda a(\lambda),$$

where

$$a(\lambda) \equiv e^{iB\lambda}ae^{-iB\lambda}.$$

[6] In Lagrangian field theory this assumption is a consequence of the stronger assertion that the basic field operators defined on a *single* spacelike hypersurface are sufficient to construct a complete set of commuting observables. Interesting observations on the feasability of relaxing this requirement are provided by W. C. Davidon and H. Ekstein, J. Math. Phys. 5, 1588 (1964); R. Haag and D. Kastler, J. Math. Phys. 5, 848 (1964).

[7] R. M. F. Houtappel, H. Van Dam, and E. P. Wigner, Rev. Mod. Phys. 37, 595 (1965). See also E. P. Wigner, Nuovo Cimento 3, 517 (1956).

among themselves. The phrase "equal-hyperplane" refers to the fact that the basic dynamical variables can also be defined on arbitrary spacelike hyperplanes and that any dynamical variable $A(\eta, \tau)$ is to be expressed in terms of the basic dynamical variables on the (η, τ) hyperplane. The choice of fixed commutators corresponds to the use of generalized coordinates and canonical momenta, rather than generalized coordinates and their time derivatives, as basic variables in classical mechanics or nonrelativistic quantum mechanics. The virtue of the choice is that many important dynamical variables acquire a form which is independent of the presence or nature of interactions.[8]

For example, in the quantum theory of local fields, if the Poincaré generators are expressed as spatial integrals of functions of the fields and their derivatives, then, in the presence of derivative coupling, *all* the generators acquire interaction parts.[4] On the other hand, if the generators are expressed in terms of the fields, their *spatial* derivatives, and their canonically conjugate fields, then the generators of spatial translations and rotations do not acquire interaction parts—even when the canonically conjugate fields appear in the interaction part of the Hamiltonian. The crucial feature of the canonically conjugate fields is that they have fixed equal-time commutators with the original fields, while the time derivatives of the original fields have equal-time commutators depending on the presence or absence of derivative coupling.

On the basis of this choice of basic dynamical variables then, one can conclude that the interaction parts in (2.2) have the form

$$V_\mu(\eta, \tau) = \eta_\mu V(\eta, \tau), \tag{2.3a}$$

$$U_{\mu\nu}(\eta, \tau) = \eta_\nu U_\mu(\eta, \tau) - \eta_\mu U_\nu(\eta, \tau), \tag{2.3b}$$

where

$$\eta^\mu U_\mu(\eta, \tau) = 0. \tag{2.3c}$$

The reason is that the projections of P_μ and $M_{\mu\nu}$ orthogonal to η_λ, i.e.,

$$P_\mu - \eta_\mu \eta P \equiv K_\mu(\eta) \tag{2.4a}$$

and

$$-\tfrac{1}{2}\epsilon_{\mu\alpha\beta\gamma} M^{\alpha\beta}\eta^\gamma \equiv J_\mu(\eta), \tag{2.4b}$$

are generators of translations and "rotations" *within* the hyperplane.[1] As such their commutators with arbitrary dynamical variables do not involve the *dynamical* dependence of the variables on the hyperplane parameters (η, τ), but only the *kinematical* transformation properties of the dynamical variables under the Poincaré group. The commutation relations, then, must not depend on the presence or nature of interactions.[9] Therefore, if the dynamical variables and the generators are expressed in terms of basic variables with fixed commutators, the projections of the generators orthogonal to η_λ cannot be dependent.

The preceding comments justify (2.3a,b). Equation (2.3c) then simply removes an ambiguity in the definition of $U_\mu(\eta, \tau)$ with no loss in generality.

Upon introducing the projections of the Poincaré generators parallel to η_λ, i.e.,

$$\eta P \equiv H(\eta) \tag{2.4c}$$

and

$$M_{\mu\nu}\eta^\nu \equiv N_\mu(\eta), \tag{2.4d}$$

to obtain the complete set of *hyperplane generators* one can rewrite (2.2) in the form

$$K_\mu(\eta) = K_\mu^{(1)}(\eta, \tau) + K_\mu^{(2)}(\eta, \tau), \tag{2.5a}$$

$$J_\mu(\eta) = J_\mu^{(1)}(\eta, \tau) + J_\mu^{(2)}(\eta, \tau), \tag{2.5b}$$

$$H(\eta) = H^{(1)}(\eta, \tau) + H^{(2)}(\eta, \tau) + V(\eta, \tau), \tag{2.5c}$$

$$N_\mu(\eta) = N_\mu^{(1)}(\eta, \tau) + N_\mu^{(2)}(\eta, \tau) + U_\mu(\eta, \tau), \tag{2.5d}$$

explicitly displaying the appearance of the interaction terms in the generators which induce changes in the hyperplane parameters, (η, τ).

From the familiar commutation relations of the full Poincaré generators among themselves one can deduce the commutation relations among the hyperplane generators. They are used later, and so are put down here[1]:

$$[K_\mu, K_\nu] = [K_\nu, H] = [H, J_\mu] = 0, \tag{2.6a}$$

$$[J_\mu, K_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}K^\alpha\eta^\beta, \tag{2.6b}$$

$$[J_\mu, J_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}J^\alpha\eta^\beta, \tag{2.6c}$$

$$[J_\mu, N_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}N^\alpha\eta^\beta, \tag{2.6d}$$

$$[N_\mu, N_\nu] = -i\hbar\epsilon_{\mu\nu\alpha\beta}J^\alpha\eta^\beta, \tag{2.6e}$$

$$[K_\mu, N_\nu] = i\hbar(g_{\mu\nu} - \eta_\mu\eta_\nu)H, \tag{2.6f}$$

$$[N_\mu, H] = i\hbar K_\mu. \tag{2.6g}$$

In the absence of interaction between the subsystems, the parts $K_\mu^{(i)}$, $H^{(i)}$, $J_\mu^{(i)}$, $N_\mu^{(i)}$, $(i = 1, 2)$ become the hyperplane generators for closed systems. Since the equal-hyperplane commutation relations of these

[8] H. Ekstein, Phys. Rev. **153**, 1397 (1967). This very interesting paper on the nature of, and relations among, observables for open systems adopts a pessimistic view towards the application of its main ideas to Poincaré invariant theories. The pessimism is not shared by the present author, who believes that the consideration of the generalization of Ekstein's group G_t to the group $G(\eta, \tau)$, which leaves the arbitrary (η, τ) hyperplane invariant, would yield all the desired results.

[9] P. A. M. Dirac, Rev. Mod. Phys. **34**, 592 (1962). See also *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, Yeshiva University, 1964), Chap. 4.

parts among themselves cannot be interaction-dependent, it follows that any two parts with different superscripts commute, while the parts with a fixed superscript satisfy (2.6) among themselves.

3. FUNDAMENTAL DIFFERENTIAL EQUATIONS

If the transformation properties of an arbitrary dynamical variable under the Poincaré group are known, then the commutation relations of the dynamical variable with the hyperplane generators can immediately be written down. The transformation equations for the full Poincaré generators are well known:

$$P'_\mu = \Lambda^\nu_\mu P_\nu \tag{3.1a}$$

and

$$M'_{\mu\nu} = \Lambda^\lambda_\mu \Lambda^\rho_\nu M_{\lambda\rho} + a_\mu \Lambda^\lambda_\nu P_\lambda - a_\nu \Lambda^\lambda_\mu P_\lambda, \tag{3.1b}$$

where the transformation is

$$x'_\mu = \Lambda^\nu_\mu x_\nu + a_\mu . \tag{3.1c}$$

Consequently the full hyperplane generators transform according to

$$K'_\mu(\eta') = \Lambda^\nu_\mu K_\nu(\eta), \tag{3.2a}$$
$$H'(\eta') = H(\eta), \tag{3.2b}$$
$$J'_\mu(\eta') = \Lambda^\nu_\mu J_\nu(\eta) - \tfrac{1}{2}\epsilon_{\mu\alpha\beta\gamma} a^\alpha K'^\beta(\eta')\eta^{\gamma\prime}, \tag{3.2c}$$
$$N'_\mu(\eta') = \Lambda^\nu_\mu N_\nu(\eta) + (a_\mu - \eta'_\mu \eta' a)H'(\eta') - (\eta' a)K'_\mu(\eta'), \tag{3.2d}$$

where

$$\eta'_\mu = \Lambda^\nu_\mu \eta_\nu . \tag{3.2e}$$

Since the transformation properties of the subsystem hyperplane generators must be interaction-independent, they must have the same form as (3.2), with the proviso that the prime refers to the parameter τ as well as η_μ, where

$$\tau' = \tau + a_\mu \eta^{\mu\prime}. \tag{3.2f}$$

The transformation properties of the interaction parts follow from the substitution of (2.5) into (3.2) and the application of the transformation rules for the subsystem generators. The results are

$$V'(\eta', \tau') = V(\eta, \tau) \tag{3.3a}$$

and

$$U'_\mu(\eta', \tau') = \Lambda^\nu_\mu U_\nu(\eta, \tau) + (a_\mu - \eta'_\mu \eta' a)V'(\eta', \tau'). \tag{3.3b}$$

From the preceding transformation rules the commutation relations of the subsystem generators and the interaction parts with the full generators can be obtained as special cases of the general commutation relations appearing as Eqs. (6.19–22) in the author's earlier article setting forth the hyperplane formalism.[1] Some of the equations are the hyperplane analogs of the Heisenberg equations of motion for the subsystem generators and interaction parts. The equations are:

$$[K_\mu, H^{(i)}] = [K_\mu K^{(i)}_\nu] = [K_\mu, V] = 0, \tag{3.4a}$$
$$[K_\mu, J^{(i)}_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}K^{(i)\alpha}\eta^\beta, \tag{3.4b}$$
$$[K_\mu, N^{(i)}_\nu] = i\hbar(g_{\mu\nu} - \eta_\mu\eta_\nu)H^{(i)}, \tag{3.4c}$$
$$[K_\mu, U_\nu] = i\hbar(g_{\mu\nu} - \eta_\mu\eta_\nu)V; \tag{3.4d}$$

$$[H, H^{(i)}] = -i\hbar\partial H^{(i)}/\partial\tau, \tag{3.5a}$$
$$[H, K^{(i)}_\mu] = -i\hbar\partial K^{(i)}_\mu/\partial\tau, \tag{3.5b}$$
$$[H, J^{(i)}_\mu] = -i\hbar J^{(i)}_\mu/\partial\tau, \tag{3.5c}$$
$$[H, N^{(i)}_\mu] = -i\hbar(K^{(i)}_\mu + \partial N^{(i)}_\mu/\partial\tau), \tag{3.5d}$$
$$[H, V] = -i\hbar\partial V/\partial\tau, \tag{3.5e}$$
$$[H, U_\mu] = -i\hbar\partial U_\mu/\partial\tau; \tag{3.5f}$$

$$[J_\mu, H^{(i)}] = [J_\mu, V] = 0, \tag{3.6a}$$
$$[J_\mu, K^{(i)}_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}K^{(i)\alpha}\eta^\beta, \tag{3.6b}$$
$$[J_\mu, J^{(i)}_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}J^{(i)\alpha}\eta^\beta, \tag{3.6c}$$
$$[J_\mu, N^{(i)}_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}N^{(i)\alpha}\eta^\beta, \tag{3.6d}$$
$$[J_\mu, U_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}U^\alpha\eta^\beta; \tag{3.6e}$$

$$[N_\mu, H^{(i)}] = i\hbar\partial H^{(i)}/\partial\eta^\mu, \tag{3.7a}$$
$$[N_\mu, K^{(i)}_\nu] = i\hbar(\eta_\nu K^{(i)}_\mu + \partial K^{(i)}_\nu/\partial\eta^\mu), \tag{3.7b}$$
$$[N_\mu, J^{(i)}_\nu] = i\hbar(\eta_\nu J^{(i)}_\nu + \partial J^{(i)}_\nu/\partial\eta^\mu), \tag{3.7c}$$
$$[N_\mu, N^{(i)}_\nu] = i\hbar(\eta_\nu N^{(i)}_\nu + \partial N^{(i)}_\nu/\partial\eta^\mu), \tag{3.7d}$$
$$[N_\mu, V] = i\hbar\partial V/\partial\eta^\mu, \tag{3.7e}$$
$$[N_\mu, U_\nu] = i\hbar(\eta_\nu U_\mu + \partial U_\nu/\partial\eta^\mu). \tag{3.7f}$$

In using Eqs. (3.7) one must handle the partial derivative with respect to η^μ with care. Being a unit vector, the four components of η^μ are not independent. The consistent application of the rule

$$\frac{\partial\eta^\mu}{\partial\eta^\nu} \equiv g^\mu_\nu - \eta^\mu\eta_\nu \tag{3.8}$$

guarantees the correct result.

Now suppose that the physical system of interest is the subsystem ($i = 2$), and that subsystem ($i = 1$) is to be varied, as well as the coupling between the subsystems, in order to probe the structure of subsystem ($i = 2$). At the outset of such a study it is natural to seek relations between quantities of interest, i.e., quantities referring to the system of interest, which have a form independent of the nature of, or the coupling to, the probing system. It would make very little sense to tamper with such relations in any approximation method applied to the study of the structure of the system of interest. Furthermore, such relations would themselves illuminate the structure of all systems possessing the quantities appearing in the relations. In this spirit one may now proceed to demonstrate a set of such relations between the parts of the full generators that involve dynamical variables associated with subsystem ($i = 2$), and, in a certain sense, the set of relations is *complete.*

The parts of the generators concerned are

$$K'_\mu \equiv K_\mu^{(2)}, \tag{3.9a}$$

$$H' \equiv H^{(2)} + V, \tag{3.9b}$$

$$J'_\mu \equiv J_\mu^{(2)}, \tag{3.9c}$$

and

$$N'_\mu \equiv N_\mu^{(2)} + U_\mu. \tag{3.9d}$$

The obvious and trivial coupling-independent relations among these quantities are

$$[K'_\mu, K'_\nu] = 0, \tag{3.10a}$$

$$[J'_\mu, K'_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}K'^\alpha\eta^\beta, \tag{3.10b}$$

and

$$[J'_\mu, J'_\nu] = i\hbar\epsilon_{\mu\nu\alpha\beta}J'^\alpha\eta^\beta. \tag{3.10c}$$

Equally obvious, but not quite so trivial, are the equations of motion for K'_μ and J'_μ:

$$[H', K'_\mu] = -i\hbar\partial K'_\mu/\partial\tau, \tag{3.11a}$$

$$[H', J'_\mu] = -i\hbar\partial J'_\mu/\partial\tau, \tag{3.11b}$$

$$[N'_\mu, K'_\nu] = i\hbar(\eta_\nu K'_\mu + \partial K'_\nu/\partial\eta^\mu), \tag{3.11c}$$

and

$$[N'_\mu, J'_\nu] = i\hbar(\eta_\nu J'_\mu + \partial J'_\nu/\partial\eta^\mu). \tag{3.11d}$$

Finally, the truly surprising relations (proofs below) are

$$[N'_\mu, H'] = i\hbar(\partial N'_\mu/\partial\tau + \partial H'/\partial\eta^\mu) \tag{3.12a}$$

and

$$[N'_\mu, N'_\nu] = i\hbar\{\partial N'_\nu/\partial\eta^\mu - \partial N'_\mu/\partial\eta^\nu + \eta_\nu N'_\mu - \eta_\mu N'_\nu + \epsilon_{\mu\nu\alpha\beta}J'^\alpha\eta^\beta\}. \tag{3.12b}$$

Thus, regardless of the nature of subsystem ($i = 1$) or of the coupling between the two subsystems, the commutators of the primed generators can be expressed *linearly* in terms of the primed generators and their hyperplane derivatives *alone.* For reasons to be described below, the equations (3.12) regarded as *differential equations* are referred to as the *fundamental differential equations.*

The proofs of (3.12a, b) are as follows:

$$\begin{aligned} i\hbar K_\mu &= [N_\mu, H] = [N_\mu, H'] + [N_\mu, H^{(1)}] \\ &= [N_\mu, H'] + [N_\mu^{(1)}, H^{(1)}] + [N'_\mu, H^{(1)}] \\ &= [N_\mu, H'] + i\hbar K_\mu^{(1)} + [N'_\mu, H] - [N'_\mu, H']. \end{aligned}$$

Therefore,

$$\begin{aligned} i\hbar K'_\mu &= [N_\mu, H'] + [N'_\mu, H] - [N'_\mu, H'] \\ &= i\hbar(\partial H'/\partial\eta^\mu + K'_\mu + \partial N'_\mu/\partial\tau) - [N'_\mu, H'], \end{aligned} \tag{3.13}$$

and (3.12a) is obtained. The last steps in (3.13) employed (3.7a, e) and (3.5d, f) For (3.12b) consider

$$\begin{aligned} -i\hbar\epsilon_{\mu\nu\alpha\beta}J^\alpha\eta^\beta &= [N_\mu, N_\nu] = [N_\mu, N'_\nu] + [N_\mu, N_\nu^{(1)}] \\ &= [N_\mu, N'_\nu] + [N_\mu^{(1)}, N_\nu^{(1)}] + [N'_\mu, N_\nu^{(1)}] \\ &= [N_\mu, N'_\nu] - i\hbar\epsilon_{\mu\nu\alpha\beta}J^{(1)\alpha}\eta^\beta + [N'_\mu, N_\nu] - [N'_\mu, N'_\nu]. \end{aligned}$$

Therefore,

$$\begin{aligned} -i\hbar\epsilon_{\mu\nu\alpha\beta}J'^\alpha\eta^\beta &= i\hbar(\eta_\nu N'_\mu + \partial N'_\nu/\partial\eta^\mu - \eta_\mu N'_\nu - \partial N'_\mu/\partial\eta^\nu) \\ &\quad - [N'_\mu, N'_\nu], \end{aligned} \tag{3.14}$$

and (3.12b) is obtained, where (3.7d, f) were used.

4. SOLUTIONS OF THE FUNDAMENTAL EQUATIONS

For the remainder of this paper the operator $H'(\eta, \tau)$, which becomes the hyperplane Hamiltonian of the system of interest in the limit of no interaction, is regarded as given for all (η, τ). The problem, then, is to deduce the remaining generators N'_μ, J'_μ, and K'_μ from H' using (3.10–12). Such an approach places the hyperplane Hamiltonian in a preferred role among the generators, and it is indeed the intention of the author to investigate how far one may go in extracting physically relevant information from the hyperplane Hamiltonian alone. Conceivably one may be able to bypass the customary Lagrangian origins of the canonical approach to relativistic quantum theory, which are characterized by ill-defined functional derivatives of operators with respect to operators.[10] Regardless of one's attitude towards such a program, however, the display of all the generators as functionals of the hyperplane Hamiltonian remains interesting. (To keep the notation simple, the primes are hereafter dropped from the generators of interest.)

The procedure will be as follows: Firstly, the general solution of (3.12a) is obtained yielding the possible N_μ; secondly, (3.12b) is used to define a J_μ for each N_μ; thirdly, the equations of motion for J_μ, (3.11b, d), is considered for obtaining constraints on the general expressions for N_μ and J_μ; fourthly, it is demonstrated that $\partial H/\partial\eta^\mu$, while not equal to K_μ in

[10] In some of the modern approaches to quantum field theory these concepts have been sharpened by restriction to functional differentiation with respect to free asymptotic fields. See F. Rohrlich, J. Math. Phys. **5**, 324 (1964).

the presence of interaction, behaves exactly as K_μ under commutation in the limit of zero interaction.[11]

A. General Solution of (3.12a)

First consider the solution of the equation

$$\frac{\partial}{\partial\tau}G(\tau;\eta;\tau_0)=\frac{i}{\hbar}H(\eta,\tau)G(\tau;\eta;\tau_0), \tag{4.1}$$

subject to the boundary condition

$$G(\tau_0;\eta;\tau_0)=I. \tag{4.2}$$

The formal expression of the solution is well known[12]:

$$G(\tau;\eta;\tau_0)=\mathcal{T}\exp\left\{\frac{i}{\hbar}\int_{\tau_0}^{\tau}d\tau' H(\eta,\tau')\right\}, \tag{4.3}$$

where $\mathcal{T}$ denotes *τ-ordered products* in the power series expansion of the exponential for $\tau>\tau_0$ and *anti-τ-ordered products* for $\tau<\tau_0$. More precisely, the solution may be defined by

$$G(\tau;\eta;\tau_0)=\lim_{N\to\infty}\prod_{n=0}^{N}\left\{I+\frac{i}{\hbar}H\left(\eta,\tau-\frac{n}{N}(\tau-\tau_0)\right)\frac{\tau-\tau_0}{N}\right\}, \tag{4.4}$$

with the understanding that the nth factor stands in the nth position *to the right* of the $n=0$ factor.

From (4.4) one may easily infer the familiar results

$$G(\tau;\eta;\tau_0)^\dagger=G(\tau_0;\eta;\tau) \tag{4.5}$$

and

$$G(\tau;\eta;\tau')G(\tau';\eta;\tau'')=G(\tau';\eta;\tau''), \tag{4.6}$$

for $\tau\geq\tau'\geq\tau''$ or $\tau\leq\tau'\leq\tau''$. More difficult to infer, but equally valid, is

$$G(\tau;\eta;\tau_0)^\dagger=G(\tau;\eta;\tau_0)^{-1}, \tag{4.7}$$

thereby justifying (4.6) for any finite τ, τ', and τ''. Finally, the relation

$$\frac{\partial G(\tau;\eta;\tau_0)}{\partial\eta^\mu}=\frac{i}{\hbar}\int_{\tau_0}^{\tau}d\tau' G(\tau;\eta;\tau')\frac{\partial H(\eta,\tau')}{\partial\eta^\mu}G(\tau';\eta;\tau_0), \tag{4.8}$$

which also follows directly from (4.4), will be useful.

A particular solution of (3.12a) is given by

$$N_\mu(\eta,\tau;\tau_0)=i\hbar\frac{\partial G(\tau;\eta;\tau_0)}{\partial\eta^\mu}G(\tau_0;\eta;\tau), \tag{4.9}$$

where

$$N_\mu(\eta,\tau_0;\tau_0)=0. \tag{4.10}$$

If the previous statement is valid, then the general solution of (3.12a) is

$$N_\mu(\eta,\tau)=i\hbar\frac{\partial G(\tau;\eta;\tau_0)}{\partial\eta^\mu}G(\tau_0;\eta;\tau)+G(\tau;\eta;\tau_0)N_\mu(\eta,\tau_0)G(\tau_0;\eta;\tau). \tag{4.11}$$

where $N_\mu(\eta,\tau_0)$ is an arbitrary Hermitian operator. This follows because the second term on the right is the general solution of the homogeneous equation corresponding to (3.12a). To verify that (4.9) is indeed a particular solution of (3.12a) it is convenient to use (4.8) to obtain

$$N_\mu(\eta,\tau;\tau_0)=-\int_{\tau_0}^{\tau}d\tau' G(\tau;\eta;\tau')\frac{\partial H(\eta,\tau')}{\partial\eta^\mu}G(\tau';\eta;\tau). \tag{4.12}$$

Then

$$\begin{aligned}\frac{\partial N_\mu(\eta,\tau;\tau_0)}{\partial\tau}&=-\frac{\partial H(\eta,\tau)}{\partial\eta^\mu}-\int_{\tau_0}^{\tau}d\tau'\frac{\partial}{\partial\tau}\times\left\{G(\tau;\eta;\tau')\frac{\partial H(\eta,\tau')}{\partial\eta^\mu}\times G(\tau';\eta;\tau)\right\},\\&=-\frac{\partial H(\eta,\tau)}{\partial\eta^\mu}+\frac{i}{\hbar}\left[H(\eta,\tau),-\int_{\tau_0}^{\tau}d\tau' G(\tau;\eta;\tau_0)\frac{\partial H(\eta,\tau')}{\partial\eta^\mu}\times G(\tau';\eta;\tau)\right],\\&=-\frac{\partial H(\eta,\tau)}{\partial\eta^\mu}+\frac{i}{\hbar}[H(\eta,\tau),N_\mu(\eta,\tau;\tau_0)],\end{aligned} \tag{4.13}$$

by direct differentiation.

One immediately objectionable feature of the most general solution (4.11) is that it appears to express $N_\mu(\eta,\tau)$ in terms of dynamical variables on an entire family of hyperplanes parallel to (η,τ). Much of the discussion in Sec. 2 was based on the assumption that the generators on a hyperplane can be expressed in terms of basic dynamical variables defined on that hyperplane alone. This physical requirement restricts the general solution somewhat, and the restriction will be incorporated in Sec. 5.

B. Definition of J_μ

Consider the η_μ derivative of (4.11). It is (in abbreviated notation)

$$\frac{\partial N_\mu(1)}{\partial\eta^\nu}=i\hbar\frac{\partial^2 G_{10}}{\partial\eta^\nu\partial\eta^\mu}G_{01}+i\hbar\frac{\partial G_{10}}{\partial\eta^\mu}\frac{\partial G_{01}}{\partial\eta^\nu}+\frac{\partial G_{10}}{\partial\eta^\nu}N_\mu(0)G_{01}+G_{10}\frac{\partial N_\mu(0)}{\partial\eta^\nu}G_{01}+G_{10}N_\mu(0)\frac{\partial G_{01}}{\partial\eta^\nu}. \tag{4.14}$$

[11] This last result was already obtained in Ref. 1 by direct differentiation of $H(\eta)$, and the discussion presented here may be regarded as a demonstration of internal consistency.

[12] Equations (4.3) and (4.4) are very reminiscent of equations commonly derived in the interaction picture which, by Haag's theorem, may not exist. See, however, M. Guenin, Commun. Math. Phys. 3, 120 (1966) for an interesting reevaluation of Haag's theorem.

But $G_{01} = G_{10}^{-1}$. Hence

$$\frac{\partial G_{01}}{\partial \eta^\nu} = \frac{\partial G_{10}^{-1}}{\partial \eta^\nu} = -G_{10}^{-1}\frac{\partial G_{10}}{\partial \eta^\nu}G_{10}^{-1} = -G_{01}\frac{\partial G_{10}}{\partial \eta^\nu}G_{01}. \tag{4.15}$$

Therefore (4.14) becomes

$$\frac{\partial N_\mu}{\partial \eta^\nu} = i\hbar\frac{\partial^2 G_{10}}{\partial \eta^\nu \partial \eta^\mu}G_{01} - i\hbar\frac{\partial G_{10}}{\partial \eta^\mu}G_{01}\frac{\partial G_{10}}{\partial \eta^\nu}G_{01} + \left[\frac{\partial G_{10}}{\partial \eta^\nu}G_{01}, G_{10}N_\mu(0)G_{01}\right] + G_{10}\frac{\partial N_\mu(0)}{\partial \eta^\nu}G_{01}, \tag{4.16}$$

and, consequently,

$$\begin{aligned}\frac{\partial N_\mu(1)}{\partial \eta^\nu} - \frac{\partial N_\nu(1)}{\partial \eta^\mu} = {} & i\hbar\left\{\frac{\partial^2 G_{10}}{\partial \eta^\nu \partial \eta^\mu} - \frac{\partial^2 G_{10}}{\partial \eta^\mu \partial \eta^\nu}\right\}G_{01} \\ & + \frac{i}{\hbar}\left[i\hbar\frac{\partial G_{10}}{\partial \eta^\mu}G_{01}, i\hbar\frac{\partial G_{10}}{\partial \eta^\nu}G_{01}\right] \\ & + \frac{i}{\hbar}\left[G_{10}N_\mu(0)G_{01}, i\hbar\frac{\partial G_{10}}{\partial \eta^\nu}G_{01}\right] \\ & + \frac{i}{\hbar}\left[i\hbar\frac{\partial G_{10}}{\partial \eta^\mu}G_{01}, G_{10}N_\nu(0)G_{01}\right] \\ & + G_{10}\left\{\frac{\partial N_\mu(0)}{\partial \eta^\nu} - \frac{\partial N_\nu(0)}{\partial \eta^\mu}\right\}G_{01}.\end{aligned} \tag{4.17}$$

From (3.12b), however, the left-hand side of (4.17) must be

$$\begin{aligned}&\eta_\nu N_\mu(1) - \eta_\mu N_\nu(1) + \frac{i}{\hbar}[N_\mu(1), N_\nu(1)] + \epsilon_{\mu\nu\alpha\beta}J^\alpha(1)\eta^\beta \\ &= i\hbar\left\{\eta_\nu\frac{\partial G_{10}}{\partial \eta^\mu}G_{01} - \eta_\mu\frac{\partial G_{10}}{\partial \eta^\nu}G_{01}\right\} \\ &\quad + G_{10}\{\eta_\nu N_\mu(0) - \eta_\mu N_\nu(0)\}G_{01} \\ &\quad + \frac{i}{\hbar}\left[i\hbar\frac{\partial G_{10}}{\partial \eta^\mu}G_{01} + G_{10}N_\mu(0)G_{01}, \right. \\ &\qquad\qquad \left. i\hbar\frac{\partial G_{10}}{\partial \eta^\nu}G_{01} + G_{10}N_\nu(0)G_{01}\right] \\ &\quad + \epsilon_{\mu\nu\alpha\beta}J^\alpha(1)\eta^\beta.\end{aligned} \tag{4.18}$$

The equating of the right-hand sides of (4.17) and (4.18) yields

$$\begin{aligned}&i\hbar\left\{\frac{\partial^2 G_{10}}{\partial \eta^\nu \partial \eta^\mu} - \frac{\partial^2 G_{10}}{\partial \eta^\mu \partial \eta^\nu} - \eta_\nu\frac{\partial G_{10}}{\partial \eta^\mu} + \eta_\mu\frac{\partial G_{10}}{\partial \eta^\nu}\right\}G_{01} \\ &= -G_{10}\left\{\frac{\partial N_\mu(0)}{\partial \eta^\nu} - \frac{\partial N_\nu(0)}{\partial \eta^\mu} + \eta_\mu N_\nu(0) - \eta_\nu N_\mu(0)\right. \\ &\qquad \left. - \frac{i}{\hbar}[N_\mu(0), N_\nu(0)]\right\}G_{01} + \epsilon_{\mu\nu\alpha\beta}J^\alpha(1)\eta^\beta.\end{aligned} \tag{4.19}$$

The left-hand side of (4.19) can be shown to vanish upon using the prescription

$$\frac{\partial}{\partial \eta^\mu} \equiv (g_\mu^\lambda - \eta_\mu\eta^\lambda)\left[\frac{\partial}{\partial \eta^\lambda}\right], \tag{4.20}$$

where $[\partial/\partial\eta^\lambda]$ indicates differentiating as though the η^λ were independent variables. Thus

$$\begin{aligned}\frac{\partial^2 G_{10}}{\partial \eta^\nu \partial \eta^\mu} &= \frac{\partial}{\partial \eta^\nu}\left\{(g_\mu^\lambda - \eta_\mu\eta^\lambda)\left[\frac{\partial G_{10}}{\partial \eta^\lambda}\right]\right\} \\ &= -(g_{\mu\nu} - \eta_\mu\eta_\nu)\eta^\lambda\left[\frac{\partial G_{10}}{\partial \eta^\lambda}\right] - \eta_\mu(g_\nu^\lambda - \eta_\nu\eta^\lambda)\left[\frac{\partial G_{10}}{\partial \eta^\lambda}\right] \\ &\quad + (g_\nu^\rho - \eta_\nu\eta^\rho)(g_\mu^\lambda - \eta_\mu\eta^\lambda)\left[\frac{\partial^2 G_{10}}{\partial \eta^\rho \partial \eta^\lambda}\right] \\ &= -\eta_\mu\frac{\partial G_{10}}{\partial \eta^\nu} + F_{\mu\nu},\end{aligned} \tag{4.21}$$

where $F_{\mu\nu} = F_{\nu\mu}$, provided that

$$\left[\frac{\partial^2 G_{10}}{\partial \eta^\rho \partial \eta^\lambda}\right] = \left[\frac{\partial^2 G_{10}}{\partial \eta^\lambda \partial \eta^\rho}\right], \tag{4.22}$$

which shall be assumed. Substituting (4.21) into the left-hand side of (4.19) for both second derivatives yields, finally,

$$\begin{aligned}&G_{10}\left\{\frac{\partial N_\mu(0)}{\partial \eta^\nu} - \frac{\partial N_\nu(0)}{\partial \eta^\mu} + \eta_\mu N_\nu(0) - \eta_\nu N_\mu(0)\right. \\ &\qquad \left. - \frac{i}{\hbar}[N_\mu(0), N_\nu(0)]\right\}G_{01} = \epsilon_{\mu\nu\alpha\beta}J^\alpha(1)\eta^\beta.\end{aligned} \tag{4.23}$$

This last equation yields $J_\alpha(1)$ explicitly in terms of the G_{10} operator and the arbitrary $N_\mu(0)$.

C. Equations of Motion for J_μ

A comparison of (3.21b) for $\tau = \tau_0$ with (4.23) yields

$$G_{10}J_\mu(0)G_{01} = J_\mu(1), \tag{4.24}$$

and this is precisely the requirement that $J_\mu(1)$ must satisfy in order for (3.11b) to hold. Therefore no constraint is placed on $N_\mu(0)$ by (3.11b).

To investigate (3.11d), first solve (3.12b) for J_μ in the form

$$J_\mu = -\epsilon_{\mu\alpha\beta\gamma}\left\{\frac{\partial N^\alpha}{\partial \eta_\beta} - \frac{i}{\hbar}N^\alpha N^\beta\right\}\eta^\gamma. \tag{4.25}$$

Upon substitution into (3.11d) this yields

$$\begin{aligned}&\epsilon_{\nu\alpha\beta\gamma}\left[N_\mu, \frac{\partial N^\alpha}{\partial \eta_\beta}\right]\eta^\gamma - \frac{i}{\hbar}\epsilon_{\nu\alpha\beta\gamma}[N_\mu, N^\alpha N^\beta]\eta^\gamma \\ &= i\hbar\eta_\nu\epsilon_{\mu\alpha\beta\gamma}\left\{\frac{\partial N^\alpha}{\partial \eta_\beta} - \frac{i}{\hbar}N^\alpha N^\beta\right\}\eta^\gamma \\ &\quad + i\hbar\epsilon_{\nu\alpha\beta\gamma}\left\{\frac{\partial^2 N^\alpha}{\partial \eta_\mu \partial \eta_\beta} - \frac{i}{\hbar}\left(\frac{\partial N^\alpha}{\partial \eta^\mu}N^\beta + N^\alpha\frac{\partial N^\beta}{\partial \eta^\mu}\right)\right\}\eta^\beta.\end{aligned} \tag{4.26}$$

I have not been able to determine the detailed nature of this constraint on N_μ. It is possible to show, however, that it does not constrain H or the relation between H and the (hitherto) arbitrary $N_\mu(0)$ in any way. To see this, substitute (4.11) and (4.24) into (3.11d) to obtain

$$\left[i\hbar\frac{\partial G_{10}}{\partial\eta^\mu}G_{01}+G_{10}N_\mu(0)G_{01},\,G_{10}J_\nu(0)G_{01}\right]$$
$$=i\hbar\left\{\eta_\nu G_{10}J_\mu(0)G_{01}+\frac{\partial}{\partial\eta^\mu}(G_{10}J_\nu(0)G_{01})\right\},$$
$$=i\hbar\left\{\eta_\nu G_{10}J_\mu(0)G_{01}+\frac{\partial G_{10}}{\partial\eta^\mu}G_{01}G_{10}J_\nu(0)G_{01}\right.$$
$$\left.+\,G_{10}\frac{\partial J_\nu(0)}{\partial\eta^\mu}G_{01}-G_{10}J_\nu(0)G_{01}\frac{\partial G_{10}}{\partial\eta^\mu}G_{01}\right\}. \quad (4.27)$$

Hence (3.11d) is identical to

$$G_{10}[N_\mu(0),J_\nu(0)]G_{01}=G_{10}i\hbar\left\{\eta_\nu J_\mu(0)+\frac{\partial J_\nu(0)}{\partial\eta^\mu}\right\}G_{01},$$

and, consequently, equivalent to

$$[N_\mu(0),J_\nu(0)]=i\hbar\left\{\eta_\nu J_\mu(0)+\frac{\partial J_\nu(0)}{\partial\eta^\mu}\right\}, \quad (4.28)$$

where

$$J_\nu(0)=-\epsilon_{\nu\alpha\beta\gamma}\left\{\frac{\partial N^\alpha(0)}{\partial\eta_\beta}-\frac{i}{\hbar}N^\alpha(0)N^\beta(0)\right\}. \quad (4.29)$$

Thus (4.26) is equivalent to the corresponding equation obtained by replacing N by $N(0)$ everywhere.

D. K_μ and the Role of $\partial H/\partial\eta^\mu$

Equations (3.10a, b) and (3.11a, c) do not enable one to get much of a hold on K_μ. All that is determined by these equations is that K_μ is a translationally invariant (3.10a), hyperplane four-vector (3.10b, 3.11c), not explicitly dependent on τ (3.11a). I do not, at present, know whether or not it is possible to express K_μ in terms of H and $N_\mu(0)$ alone in the presence of interactions. In the limit of zero interaction, however, the matter is easily resolved.

Let $A_\alpha^{(2)}(\eta,\tau)$ be an arbitrary dynamical variable belonging to the subsystem (2) of Sec. 2, the physical system of interest. Let the transformation properties of this dynamical variable under arbitrary infinitesimal translations δa_μ be described by[1]

$$A_\alpha'^{(2)}(\eta,\tau')=A_\alpha^{(2)}(\eta,\tau)+T_\alpha^{(2)}(\eta,\tau)\delta a, \quad (4.30)$$

where the primes *here* refer to transformed variables. This relation determines the commutator of $A_\alpha^{(2)}$ with H_T, the total hyperplane Hamiltonian, and hence with the partial hyperplane Hamiltonian $H\equiv H^{(2)}+V$ as

$$[H,A_\alpha^{(2)}]=i\hbar\{T_\alpha^{(2)\mu}\eta_\mu-\partial A_\alpha^{(2)}/\partial\tau\}. \quad (4.31)$$

Taking the derivative of both sides with respect to η^μ yields

$$[\partial H/\partial\eta^\mu,A_\alpha^{(2)}]+[H,\partial A_\alpha^{(2)}/\partial\eta^\mu]$$
$$=i\hbar(g_{\mu\nu}-\eta_\mu\eta_\nu)T_\alpha^{(2)\nu}+i\hbar\left\{\frac{\partial T_\alpha^{(2)\nu}}{\partial\eta^\mu}\eta_\nu-\frac{\partial^2A_\alpha^{(2)}}{\partial\eta^\mu\partial\tau}\right\}. \quad (4.32)$$

But from (4.30)

$$\frac{\partial A_\alpha'^{(2)}(\eta,\tau')}{\partial\eta^\mu}=\frac{\partial A_\alpha^{(2)}(\eta,\tau)}{\partial\eta^\mu}+\frac{\partial T_\alpha^{(2)\nu}(\eta,\tau)}{\partial\eta^\mu}\delta a_\nu, \quad (4.33)$$

so that

$$\left[H_T,\frac{\partial A_\alpha^{(2)}}{\partial\eta^\mu}\right]=i\hbar\left\{\frac{\partial T_\alpha^{(2)\nu}}{\partial\eta^\mu}\eta_\nu-\frac{\partial^2A_\alpha^{(2)}}{\partial\tau\partial\eta^\mu}\right\}. \quad (4.34)$$

Hence, if the order of differentiation is immaterial for the second-order derivatives on the right-hand sides of (4.32) and (4.34), (and in a *closed* system this is demanded for consistency) then (4.32) becomes

$$[\partial H/\partial\eta^\mu,A_\alpha^{(2)}]$$
$$=i\hbar(g_{\mu\nu}-\eta_\mu\eta_\nu)T_\alpha^{(2)\nu}-[H^{(1)},\partial A_\alpha^{(2)}/\partial\eta^\mu]. \quad (4.35)$$

The commutator involving $H^{(1)}$ cannot be set equal to zero so long as there is coupling between the subsystems. The hyperplane derivatives of quantities expressed entirely in terms of the basic dynamical variables of one subsystem acquire contributions from the other subsystems interacting with the first. Nevertheless, when the interaction does vanish, the commutator with $H^{(1)}$ disappears from (4.35) and

$$[\partial H/\partial\eta^\mu,A_\alpha^{(2)}]=i\hbar(g_{\mu\nu}-\eta_\mu\eta_\nu)T_\alpha^{(2)\nu} \quad (4.36)$$

results. But, with or without interaction, K_μ must satisfy

$$[K_\mu,A_\alpha^{(2)}]=i\hbar(g_{\mu\nu}-\eta_\mu\eta_\nu)T_\alpha^{(2)\nu}, \quad (4.37)$$

and hence

$$\partial H/\partial\eta^\mu\to K_\mu \quad (4.38)$$

as the interaction between the systems is turned off.[13]

5. FIELDLIKE EXPRESSIONS FOR THE GENERATORS OF CLOSED SYSTEMS

Let $\{|k_\mu,\alpha(\eta,\tau)\rangle\}$ be a complete orthonormal set of basis vectors such that

$$K_\mu(\eta)_T\,|k_\mu,\alpha(\eta,\tau)\rangle=k_\mu\,|k_\mu;\alpha(\eta,\tau)\rangle, \quad (5.1)$$

where $K_\mu(\eta)_T$ is the total hyperplane momentum operator, and where α denotes the eigenvalues of a set of other observables which, along with $K_\mu(\eta)_T$, form a complete commuting set on the hyperplane

[13] It may occur to the reader that $\partial H/\partial\eta^\mu+\partial N_\mu/\partial\tau=K_\mu$ is the appropriate generalization of the zero-interaction result. Alas, no!

(η, τ). Then

$$\begin{aligned}\langle k_\mu'; \alpha'(\eta,\tau)|\, H(\eta,\tau)\, |k_\mu; \alpha(\eta,\tau)\rangle \\ \equiv (2\pi\hbar)^3\delta^3_\eta(k'-k)\langle k_\mu; \alpha'(\eta,\tau)|\, h(\eta,\tau)\, |k_\mu; \alpha(\eta,\tau)\rangle\end{aligned} \tag{5.2}$$

defines the operator $h(\eta,\tau)$ on the hyperplane momentum shell, where $\delta^3_\eta(k'-k)$ satisfies

$$\delta(\eta k' - \eta k)\delta^3_\eta(k'-k) = \delta^4(k'-k). \tag{5.3}$$

Defining $\mathcal{H}(x;\eta)$ by[14]

$$\mathcal{H}(x;\eta) \equiv e^{-i/\hbar K_\mu(\eta)_T x^\mu} h(\eta, \eta x) e^{i/\hbar K_\mu(\eta)_T x^\mu} \tag{5.4}$$

yields

$$\begin{aligned}\langle k_\mu'; \alpha'(\eta,\tau)|\, H(\eta,\tau)\, |k_\mu; \alpha(\eta,\tau)\rangle \\ \equiv \langle k_\mu'; \alpha'(\eta,\tau)| \int d^4x\,\delta(\eta x - \tau)\mathcal{H}(x;\eta)\, |k_\mu; \alpha(\eta,\tau)\rangle\end{aligned} \tag{5.5}$$

or

$$H(\eta,\tau) = \int d^4x\,\delta(\eta x - \tau)\mathcal{H}(x;\eta), \tag{5.6}$$

since the eigenvectors form a complete basis, and

$$\int d^4x\,\delta(\eta x - \tau) e^{(i/\hbar)(k_\mu - k_\mu')x^\mu} = (2\pi\hbar)^3\delta^3_\eta(k'-k), \tag{5.7}$$

for $\eta k = \eta k' = 0$.

The foregoing construction of the fieldlike expression (5.6) is, in perhaps slightly modified form, familiar to many.[5] It has been reproduced here to emphasize the *independence* of expressions like (5.6) and the assumptions that are generally regarded as comprising a bona fide field theory of relativistic quantum physics. For the remainder of this section, expressions similar to (5.6) are sought for the remaining hyperplane generators, and $\mathcal{H}(x;\eta)$ is seen to play a fundamental role in these expressions. Nevertheless, a field theory of quantum phenomena is not implied by these considerations (at least not the usual kind of field theory),[15] although the results certainly sit well in a field-theoretic context.

A. Generator $N_\mu(\eta,\tau)$

Substituting (5.6) into (4.12) yields

$$\begin{aligned}N_\mu(\eta,\tau;\tau_0) &= -\int_{\tau_0}^{\tau} d\tau' G(\tau;\eta;\tau')\int d^4x \\ &\quad\times \Big\{\delta'(\eta x - \tau')(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta) \\ &\quad + \delta(\eta x - \tau')\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu}\Big\} G(\tau';\eta;\tau) \\ &= \int_{\tau_0}^{\tau} d\tau' G(\tau;\eta;\tau')\int d^4x \\ &\quad\times \Big\{\frac{\partial}{\partial\tau'}\delta(\eta x - \tau')(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta)\Big\} \\ &\quad\times G(\tau';\eta;\tau) \\ &\quad - \int_{\tau_0}^{\tau} d\tau' G(\tau;\eta;\tau')\int d^4x\,\delta(\eta x - \tau')\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu} \\ &\quad\times G(\tau';\eta;\tau).\end{aligned} \tag{5.8}$$

But

$$\begin{aligned}&\int_{\tau_0}^{\tau} G(\tau;\eta;\tau')\int d^4x\Big\{\frac{\partial}{\partial\tau'}\delta(\eta x - \tau')(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta)\Big\} \\ &\quad\times G(\tau';\eta;\tau) \\ &= \int_{\tau_0}^{\tau} d\tau'\frac{\partial}{\partial\tau'}\Big\{G(\tau;\eta;\tau')\int d^4x\,\delta(\eta x - \tau')(x_\mu - \eta_\mu\eta x) \\ &\quad\times \mathcal{H}(x;\eta)G(\tau';\eta;\tau)\Big\} \\ &\quad - \frac{i}{\hbar}\int_{\tau_0}^{\tau} d\tau' G(\tau;\eta;\tau')\int d^4x\,\delta(\eta x - \tau')(x_\mu - \eta_\mu\eta x) \\ &\quad\times [\mathcal{H}(x;\eta), H(\eta,\tau')]G(\tau';\eta;\tau), \\ &= \int d^4x\,\delta(\eta x - \tau)(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta) - G(\tau;\eta;\tau_0) \\ &\quad\times \int d^4x\,\delta(\eta x - \tau_0)(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta)G(\tau_0;\eta;\tau) \\ &\quad - \frac{i}{\hbar}\int_{\tau_0}^{\tau} d\tau' G(\tau;\eta;\tau')\int d^4x\,\delta(\eta x - \tau')(x_\mu - \eta_\mu\eta x \\ &\quad\times [\mathcal{H}(x;\eta), H(\eta,\tau')] \times G(\tau';\eta;\tau).\end{aligned} \tag{5.9}$$

At this point in the calculation it is judicious to turn *off* the coupling between the physical system of interest and the subsystem $H^{(1)}$. Since the relevant expressions of the previous sections hold for arbitrary nonvanishing coupling, it will be *assumed* that they hold in the limit of no coupling. The interaction-augmented generators then become the generators of a closed system, and since $\mathcal{H}(x;\eta)$ is invariant under translations (passive), it follows that

$$[\mathcal{H}(x;\eta), H(\eta,\tau')] = [\mathcal{H}(x;\eta), H(\eta)] = i\hbar\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta). \tag{5.10}$$

Hence

$$G(\tau;\eta;\tau')\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta)G(\tau';\eta;\tau) = \eta\frac{\partial}{\partial x}\mathcal{H}(x + \eta(\tau - \tau');\eta) \tag{5.11}$$

---

[14] $h(\eta,\tau)$ is deliberately left unspecified *off* the hyperplane momentum shell, since none of the subsequent results depend on such a specification.

[15] The present author is, at the time of writing, uncertain as to the exact relationship between the formalism recently proposed by J. Schwinger, Phys. Rev. **152**, 1219 (1966) and the results obtained here.

and

$$G(\tau;\eta;\tau')\frac{\partial}{\partial\eta^\mu}\mathcal{H}(x;\eta)G(\tau';\eta;\tau)$$
$$=\frac{\partial}{\partial\eta^\mu}\mathcal{H}(x+\eta(\tau-\tau');\eta)-(\tau-\tau')$$
$$\times\left\{\frac{\partial}{\partial x^\mu}-\eta_\mu\eta\frac{\partial}{\partial x}\right\}\mathcal{H}(x+\eta(\tau-\tau');\eta). \quad (5.12)$$

We substitute (5.10), and then (5.11), into (5.9), and (5.12) into the last term of (5.8) and note that, for any $\mathcal{F}$,

$$\int d^4x\delta(\eta x-\tau')(x_\mu-\eta_\mu\eta x)\mathcal{F}(x+\eta(\tau-\tau');\eta)$$
$$\equiv\int d^4x\delta(\eta x-\tau)(x_\mu-\eta_\mu\eta x)\mathcal{F}(x;\eta). \quad (5.13)$$

This yields, finally, upon discarding surface terms,

$$N_\mu(\eta,\tau;\tau_0)=\int d^4x\delta(\eta x-\tau)\Big\{(x_\mu-\eta_\mu\eta x)\mathcal{H}(x;\eta)$$
$$+(\tau-\tau_0)\left((x_\mu-\eta_\mu\eta x)\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta)-\frac{\partial}{\partial\eta^\mu}\mathcal{H}(x;\eta)\right)\Big\}$$
$$-G(\tau;\eta;\tau_0)\int d^4x\delta(\eta x-\tau_0)(x_\mu-\eta_\mu\eta x)$$
$$\times\mathcal{H}(x;\eta)G(\tau_0;\eta;\tau). \quad (5.14)$$

The general solution for $N_\mu(\eta)$ now appears as[16]

$$N_\mu(\eta)=N_\mu(\eta,\tau;\tau_0)+G(\tau;\eta;\tau_0)N_\mu(\eta,\tau_0)G(\tau_0;\eta;\tau),$$
$$=\int d^4x\delta(\eta x-\tau)\Big\{(x_\mu-\eta_\mu\eta x)\mathcal{H}(x;\eta)+(\tau-\tau_0)$$
$$\times\left((x_\mu-\eta_\mu\eta x)\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta)-\frac{\partial}{\partial\eta^\mu}\mathcal{H}(x;\eta)\right)\Big\}$$
$$+G(\tau;\eta;\tau_0)\Big\{N_\mu(\eta,\tau_0)$$
$$-\int d^4x\delta(\eta x-\tau_0)(x_\mu-\eta_\mu\eta x)\mathcal{H}(x;\eta)\Big\}$$
$$\times G(\tau_0;\eta;\tau). \quad (5.15)$$

With this expression we can answer the objection raised just after Eq. (4.13) concerning the expression of $N_\mu'(\eta,\tau)$ in terms of dynamical variables defined on the hyperplane (η,τ) alone. Clearly such a form is possible for $N_\mu(\eta\ )$ if and only if

$$N_\mu(\eta,\tau_0)=\int d^4x\delta(\eta x-\tau_0)(x_\mu-\eta_\mu\eta x)\mathcal{H}(x;\eta). \quad (5.16)$$

This latter expression must then be tested for consistency in (4.26). One notes that since the hyperplane parameter τ does not appear explicitly anywhere in (4.26), (5.16) may not satisfy (4.26) for any $\tau_0\neq 0$ (see note added in proof). This possibility will be anticipated in the final form for $N_\mu(\eta)$:

$$N_\mu(\eta)=\int d^4x\delta(\eta x-\tau)\Big\{(x_\mu-\eta_\mu\eta x)\mathcal{H}(x;\eta)$$
$$+\tau\left((x_\mu-\eta_\mu\eta x)\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta)-\frac{\partial}{\partial\eta^\mu}\mathcal{H}(x;\eta)\right)\Big\}. \quad (5.17)$$

The demand that the generators be expressible in terms of dynamical variables on *one* hyperplane may, in the fashion of the times, be called the *principle of maximal causality*.

B. Generator $J_\mu(\eta)$

The expression (4.25) is nonlinear in N_μ. It can easily be replaced by a linear expression for the special case of a closed system. The point is that for a closed system the analog of (2.6e)

$$[N_\mu,N_\nu]=-i\hbar\epsilon_{\mu\nu\alpha\beta}J^\alpha\eta^\beta \quad (5.18)$$

must hold, and, upon substitution into (3.12b), yields

$$-\epsilon_{\mu\nu\alpha\beta}J^\alpha\eta^\beta=\tfrac12\{\partial N_\nu/\partial\eta^\mu-\partial N_\mu/\partial\eta^\nu+\eta_\nu N_\mu-\eta_\mu N_\nu\} \quad (5.19)$$

or

$$J_\mu=\tfrac12\epsilon_{\mu\alpha\beta\gamma}\frac{\partial N^\beta}{\partial\eta_\alpha}\eta^\gamma. \quad (5.20)$$

Direct substitution of (5.17) into (5.20) yields

$$J_\mu=\tfrac12\epsilon_{\mu\alpha\beta\gamma}\int d^4x\Big\{\delta(\eta x-\tau)x^\beta\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\alpha}$$
$$+\tau\Big[\delta(\eta x-\tau)x^\beta\eta\frac{\partial}{\partial x}\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\alpha}$$
$$-\delta'(\eta x-\tau)x^\alpha\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\beta}\Big]\Big\}\eta^\gamma$$
$$=\tfrac12\epsilon_{\mu\alpha\beta\gamma}\int d^4x\delta(\eta x-\tau)x^\beta\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\alpha}\eta^\gamma$$
$$+\frac{\tau}{2}\epsilon_{\mu\alpha\beta\gamma}\int d^4x\eta\frac{\partial}{\partial x}\Big\{\delta(\eta x-\tau)x^\beta\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\alpha}\Big\}\eta^\gamma$$
$$=\int d^4x\delta(\eta x-\tau)\tfrac12\epsilon_{\mu\alpha\beta\gamma}x^\beta\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\alpha}\eta^\gamma, \quad (5.21)$$

where, in obtaining the *first* integral expression, the relations

$$\epsilon_{\mu\alpha\beta\gamma}\int d^4x\delta(\eta x-\tau)x^\beta\frac{\partial}{\partial x_\alpha}\mathcal{H}(x;\eta)\eta^\gamma$$
$$=\epsilon_{\mu\alpha\beta\gamma}\int d^4x\delta(\eta x-\tau)\frac{\partial}{\partial x_\alpha}\{x^\beta\mathcal{H}(x;\eta)\}\eta^\gamma$$
$$=0 \quad (5.22)$$

[16] The τ dependence of $N_\mu(\eta)$ is absent because the interaction has been turned off. The τ_0 dependence of $N_\mu(\eta,\tau_0)$ is retained since $N_\mu(\eta,\tau_0)$ can be chosen arbitrarily in this approach. In particular, the "causal" choice (5.16) *is* τ_0 dependent.

and

$$\epsilon_{\mu\alpha\beta\gamma}\frac{\partial^2\mathcal{H}(x;\eta)}{\partial\eta_\alpha\partial\eta_\beta}\eta^\gamma$$
$$= \tfrac{1}{2}\epsilon_{\mu\alpha\beta\gamma}\left\{\frac{\partial^2\mathcal{H}(x;\eta)}{\partial\eta_\alpha\partial\eta_\beta} - \frac{\partial^2\mathcal{H}(x;\eta)}{\partial\eta_\beta\partial\eta_\alpha}\right\}\eta^\gamma$$
$$= \tfrac{1}{2}\epsilon_{\mu\alpha\beta\gamma}\left\{\eta^\alpha\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\beta} - \eta^\beta\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\alpha}\right\}\eta^\gamma = 0 \quad (5.23)$$

have been used.

C. Generator $K_\mu(\eta)$

The relation (4.38) yields immediately

$$K_\mu(\eta) = \int d^4x\left\{\delta'(\eta x - \tau)(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta) + \delta(\eta x - \tau)\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu}\right\}$$
$$= \int d^4x\delta(\eta x - \tau) \times \left\{\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu} - (x_\mu - \eta_\mu\eta x)\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta)\right\}. \quad (5.24)$$

Comparing (5.24) with (5.17), one obtains the familiar-looking result

$$N_\mu(\eta) = \int d^4x\delta(\eta x - \tau)(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta) - \tau K_\mu(\eta). \quad (5.25)$$

Similarly, if one defines

$$j(x;\eta) \equiv \eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta) + \left\{\frac{\partial}{\partial x^\mu} - \eta_\mu\eta\frac{\partial}{\partial x}\right\}\frac{1}{2}\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\mu} \quad (5.26)$$

and

$$\mathcal{K}_\mu(x;\eta) \equiv \frac{1}{2}\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu} - (x_\mu - \eta_\mu\eta x)j(x;\eta), \quad (5.27)$$

then, after integration by parts,

$$K_\mu(\eta) = \int d^4x\delta(\eta x - \tau)\mathcal{K}_\mu(x;\eta) \quad (5.28)$$

and

$$J_\mu(\eta) = -\int d^4x\delta(\eta x - \tau)\epsilon_{\mu\alpha\beta\gamma}x^\alpha\mathcal{K}^\beta(x;\eta)\eta^\gamma. \quad (5.29)$$

Again, these are familiar-looking results.

D. Constraint on N_μ

The relation (5.16) must be tested for consistency with (4.26). For a closed system, however, (4.26) is a direct consequence of the much simpler relation

$$[N_\mu, N_\nu] = i\hbar\left(\eta_\nu N_\mu + \frac{\partial N_\nu}{\partial\eta^\mu}\right), \quad (5.30)$$

which must also hold. This, in turn, must be consistent with the *hyperplane equation of motion* (see note added in proof)

$$[N_\mu(\eta), \mathcal{H}(x;\eta)] = i\hbar\left\{\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu} - (x_\mu - \eta_\mu\eta x)\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta)\right\}, \quad (5.31)$$

which is a direct consequence of the scalar transformation properties of the hyperplane Hamiltonian density. From (5.16) applied to (5.31), we get

$$[N_\mu(\eta), N_\nu(\eta)] = \left[N_\mu(\eta), \int d^4x\delta(\eta x - \tau_0)(x_\nu - \eta_\nu\eta x)\mathcal{H}(x;\eta)\right]$$
$$= i\hbar\int d^4x\delta(\eta x - \tau_0)\left\{\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu}(x_\nu - \eta_\nu\eta x) - (x_\nu - \eta_\nu\eta x)(x_\mu - \eta_\mu\eta x)\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta)\right\}. \quad (5.32)$$

Applying (5.16) to (5.30) yields

$$[N_\mu(\eta), N_\nu(\eta)] = i\hbar\int d^4x\Big\{\delta(\eta x - \tau_0)\eta_\nu(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta) + \delta'(\eta x - \tau_0)(x_\mu - \eta_\mu\eta x)(x_\nu - \eta_\nu\eta x)\mathcal{H}(x;\eta)$$
$$- \delta(\eta x - \tau_0)(g_{\mu\nu} - \eta_\mu\eta_\nu)\eta x\mathcal{H}(x;\eta) - \delta(\eta x - \tau_0)\eta_\nu(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta) + \delta(\eta x - \tau_0)(x_\nu - \eta_\nu\eta x)\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu}\Big\}$$
$$= i\hbar\int d^4x\delta(\eta x - \tau_0)\left\{(x_\nu - \eta_\nu\eta x)\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu} - (x_\mu - \eta_\mu\eta x)(x_\nu - \eta_\nu\eta x)\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta)\right\}$$
$$- \tau_0(g_{\mu\nu} - \eta_\mu\eta_\nu)H(\eta). \quad (5.33)$$

Clearly (5.32) and (5.33) are consistent if and only if $\tau_0 = 0$, as was suspected earlier.

6. SUMMARY

The principal results may be conveniently and briefly summarized as follows. For a closed system the hyperplane generators K_μ and J_μ are related to H and N_μ, respectively, via the equations

$$K_\mu = \partial H/\partial\eta^\mu \quad (6.1)$$

and (5.20). If the system is allowed to interact in an arbitrary way with some other arbitrary system, then (3.12a) holds for the system generators augmented by the interaction terms and has (4.11) as its most general solution. In the limit of zero interaction between the systems, the general solution can be written as (5.15) in terms of a hyperplane Hamiltonian density $\mathcal{H}(x;\eta)$, and the particular solution (5.17) is the only one satisfying maximal causality and consistent with (5.30, 31), both of which are demanded for a closed system. With the expressions (5.6) and (5.17) then, for H and N_μ in terms of the hyperplane Hamiltonian density, one can use (6.1) and (5.20) to express K_μ and J_μ in terms of the hyperplane Hamiltonian density. Thus, independently of a Lagrangian or field-theoretic approach, the hyperplane Hamiltonian density achieves a fundamental status, at least with respect to the Poincaré generators. The final expressions of the generators are given here for convenience:

$$H(\eta) = \int d^4x\,\delta(\eta x - \tau)\mathcal{H}(x;\eta), \tag{6.2a}$$

$$K_\mu(\eta) = \int d^4x\,\delta(\eta x - \tau) \times \left\{\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu} - (x_\mu - \eta_\mu\eta x)\eta\frac{\partial}{\partial x}\mathcal{H}(x;\eta)\right\}, \tag{6.2b}$$

$$N_\mu(\eta) = \int d^4x\,\delta(\eta x - \tau)\{(x_\mu - \eta_\mu\eta x)\mathcal{H}(x;\eta)\} - \tau K_\mu(\eta), \tag{6.2c}$$

$$J_\mu(\eta) = \tfrac{1}{2}\,\epsilon_{\mu\alpha\beta\gamma}\int d^4x\,\delta(\eta x - \tau)x^\beta\,\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta_\alpha}\,\eta^\gamma. \tag{6.2d}$$

It should be noticed that these latter expressions are *not* in terms of dynamical variables with fixed equal-hyperplane commutators—the single exception being $H(\eta)$ if $\mathcal{H}(x;\eta)$ is expressed in terms of such variables. Thus the hyperplane derivatives of $\mathcal{H}(x;\eta)$ appearing in N_μ, K_μ, and J_μ involve the corresponding derivatives of the basic dynamical variables, and these derivatives possess interaction-dependent equal-hyperplane commutators among themselves and with undifferentiated dynamical variables. Of course, once the basic dynamical variables have been chosen and their transformation properties determined, the specification of $\mathcal{H}(x;\eta)$ enables one to replace the hyperplane derivatives by functions of the basic dynamical variables. Only after this is done, however, do K_μ and J_μ assume a form independent of the presence or nature of interactions within the system.

It is instructive to compare the results (6.2) with the corresponding relations that follow from Lagrangian field theory. Denoting the conventional symmetric stress–energy–momentum tensor density by $\theta_{\mu\nu}(x)$, one recalls

$$P_\mu = \int d^4x\,\delta(\eta x - \tau)\theta_{\mu\nu}(x)\eta^\nu. \tag{6.3}$$

Hence, since

$$H(\eta) = \eta^\mu P_\mu = \int d^4x\,\delta(\eta x - \tau)\eta^\mu\theta_{\mu\nu}(x)\eta^\nu, \tag{6.4}$$

it follows that one can choose

$$\mathcal{H}(x;\eta) = \eta^\mu\theta_{\mu\nu}(x)\eta^\nu. \tag{6.5}$$

Now with this choice we have

$$\begin{aligned}\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu} &= \theta_{\mu\nu}(x)\eta^\nu - \eta_\mu\eta^\lambda\theta_{\lambda\nu}(x)\eta^\nu \\ &\quad + \eta^\lambda\theta_{\lambda\mu}(x) - \eta^\lambda\theta_{\lambda\nu}(x)\eta^\nu\eta_\mu \\ &= 2\{g^\nu_\mu - \eta_\mu\eta^\lambda\}\theta_{\lambda\nu}(x)\eta^\nu\end{aligned} \tag{6.6}$$

and

$$\begin{aligned}K_\mu(\eta) &= P_\mu - \eta_\mu H(\eta) \\ &= \int d^4x\,\delta(\eta x - \tau)\{\theta_{\mu\nu}(x)\eta^\nu - \eta_\mu\eta^\lambda\theta_{\lambda\nu}(x)\eta^\nu\} \\ &= \int d^4x\,\delta(\eta x - \tau)\,\frac{1}{2}\,\frac{\partial\mathcal{H}(x;\eta)}{\partial\eta^\mu}.\end{aligned} \tag{6.7}$$

This last result does not look much like (6.2b). But if (5.28), which is equivalent to (6.2b), is considered, it is seen that (6.7) would follow from

$$j(x;\eta) = 0. \tag{6.8}$$

But

$$\begin{aligned}j(x;\eta) &\equiv \eta\frac{\partial}{\partial x}\{\eta^\mu\theta_{\mu\nu}(x)\eta^\nu\} \\ &\quad + \left\{\frac{\partial}{\partial x^\mu} - \eta_\mu\eta\frac{\partial}{\partial x}\right\}(g^{\mu\lambda} - \eta^\mu\eta^\lambda)\theta_{\lambda\nu}(x)\eta^\nu \\ &= \frac{\partial}{\partial x_\mu}\,\theta_{\mu\nu}(x)\eta^\nu = 0\end{aligned} \tag{6.9}$$

for a closed system. In a similar manner the expressions (6.2c, d), which should now be almost obvious, can be derived.

Note Added in Proof: Equation (5.31) is in error. One should add to the right side the term,

$$i\hbar\eta x[\partial/\partial x^\mu - \eta_\mu(\eta\partial/\partial x)]H(x;\eta).$$

This would yield an extra term on the right side of (5.32) making (5.32) consistent with (5.33) for *all* τ_0. Thus the discussion between (5.16) and (5.17) is misleading and (5.17) is slightly restrictive.

Note on the WKB Method

R. K. NESBET
IBM Research Laboratory, San Jose, California

(Received 17 May 1967)

In the phase integral WKB method, a solution $u_1(z)$ of a second-order linear differential equation is represented in terms of its logarithmic derivative $iy(z)$ which satisfies a simple nonlinear first-order equation. This representation does not in general lead directly to an independent second solution of the original equation. However, if $y(z)$ is expressed in the form $q(z) + iq'(z)/2q(z)$, where $q(z)$ satisfies a nonlinear second-order equation, then $q(z)$ can be used to determine a second solution to the original equation. These two solutions remain linearly independent throughout their domain of definition. It is shown that $q(z)$ is given by the sum of alternate terms in the well-known asymptotic expansion of $y(z)$. Any two linearly independent solutions of the original equation, normalized so that the Wronskian is $-2i$, give $q(z)$ in the form $(u_1u_2)^{-1}$.

The phase integral WKB method[1] is concerned with differential equations of the form

$$\frac{d^2u}{dz^2} + Q^2u = 0, \tag{1}$$

where Q^2 is a function of z. Zeros of Q^2 are referred to as transition points, corresponding to classical turning points in a one-dimensional Schrödinger equation.

A solution $u_1(z)$ of Eq. (1) can be expressed in terms of its logarithmic derivative $iy(z)$ in the form[1]

$$u_1(z) = \exp\left[i\int^z y(\zeta)\, d\zeta\right], \tag{2}$$

where $y(z)$ satisfies the equation

$$iy' = y^2 - Q^2. \tag{3}$$

In order to match given boundary conditions, it is necessary to find a second solution to Eq. (1) that is linearly independent of $u_1(z)$. An obvious choice, if Q^2 is real on the real axis, is the complex conjugate of u_1, but this is a solution of Eq. (1) only on the real axis and is not always independent of u_1. Reversing the sign of i in Eq. (2) does not in general give a solution of Eq. (1).

Another approach to Eq. (1) is to find a function $q(z)$ such that the transformation of variables[1]

$$u(z) = q^{-\frac{1}{2}}\phi, \tag{4a}$$

$$w(z) = \int^z q(\zeta)\, d\zeta, \tag{4b}$$

converts Eq. (1) into the form

$$\frac{d^2\phi}{dw^2} + \phi = 0. \tag{5}$$

[1] N. Fröman and P. O. Fröman, *JWKB Approximation, Contributions to the Theory* (North-Holland Publishing Co., Amsterdam, 1965); J. Heading, *An Introduction to Phase-Integral Methods* (John Wiley & Sons, Inc., New York, 1962). The extensive literature on this subject is reviewed in these two references.

Two linearly independent solutions of Eq. (1) are then

$$u_1(z) = q(z)^{-\frac{1}{2}} \exp i\int^z q(\zeta)\, d\zeta, \tag{6a}$$

$$u_2(z) = q(z)^{-\frac{1}{2}} \exp -i\int^z q(\zeta)\, d\zeta. \tag{6b}$$

The Wronskian for these functions is

$$u_1u_2' - u_2u_1' = -2i, \tag{7}$$

so that they are linearly independent throughout their domain of definition. The function $q(z)$ must satisfy the equation

$$q^{\frac{1}{2}} \frac{d^2}{dz^2} q^{-\frac{1}{2}} = q^2 - Q^2. \tag{8}$$

If Eqs. (2) and (6a) are compared, it is clear that the functions y and q must be related by

$$y(z) = q(z) + iq'(z)/2q(z), \tag{9}$$

or

$$iq' = 2q(y - q). \tag{10}$$

The two first-order equations, Eqs. (3) and (10), together imply Eq. (8).

It follows from Eqs. (6) and (9) that

$$u_2(z) = \exp i\int^z (y(\zeta) - 2q(\zeta))\, d\zeta \tag{11}$$

is a solution of Eq. (1), linearly independent of $u_1(z)$, given by Eq. (2), throughout the domain of definition of $q(z)$.

If Q^2 in Eq. (1) is replaced by λ^2Q^2, where λ is a large real parameter, then Eq. (3) leads to a well-known asymptotic expansion of y in decreasing powers of λ,

$$y(z) = \sum_{n=-1}^{\infty} y_n\lambda^{-n}, \tag{12}$$

where

$$y_{-1} = \pm Q, \quad (13a)$$

$$iy'_n = \sum_{\nu=-1}^{n+1} y_\nu y_{n-\nu}, \quad n = -1, 0, 1, \cdots. \quad (13b)$$

It follows immediately from Eq. (10) that $q(z)$ is given by the terms of odd index in Eq. (12),

$$q(z) = \sum_{k=0}^{\infty} y_{2k-1}\lambda^{-2k+1}. \quad (14)$$

This can be seen by writing Eq. (12) as separate sums of odd and even terms,

$$y = y_{\text{odd}} + y_{\text{even}}. \quad (15)$$

Then since $\lambda^2 Q^2$ is even, the even part of Eq. (3) is

$$iy'_{\text{even}} = y^2_{\text{even}} + y^2_{\text{odd}} - \lambda^2 Q^2, \quad (16)$$

and the odd part is

$$i\, y'_{\text{odd}} = 2\, y_{\text{odd}}\, y_{\text{even}}. \quad (17)$$

If $q(z)$ is set equal to y_{odd} this equation is identical to Eq. (10). Equation (14) gives an asymptotic expansion of $q(z)$ about transition points of Eq. (1).

If u_1 and u_2 are any two independent solutions of Eq. (1), normalized so that the Wronskian is $-2i$, as in Eq. (7), then Eqs. (6) imply that

$$q(z) = (u_1 u_2)^{-1}. \quad (18)$$

This result can be checked by substitution in Eq. (8), using Eqs. (1) and (7). Since the Wronskian is invariant under unimodular linear transformations of the functions u_1 and u_2, a more general solution of Eq. (8) is given by

$$q(z) = [(au_1 + bu_2)(cu_1 + du_2)]^{-1}, \quad (19)$$

where the constant coefficients a, b, c, d satisfy

$$ad - bc = 1. \quad (20)$$

Let z_0 be a transition point of Eq. (1). Because of the so-called Stokes phenomenon,[1] the asymptotic series of Eq. (14), for different values of arg $(z - z_0)$, can represent functions $q(z)$ that correspond to different unimodular transformations in Eq. (19).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Unitary Representations of the Affine Group

ERIK W. ASLAKSEN
Department of Physics, Lehigh University, Bethlehem, Pennsylvania
and
Bell Telephone Laboratories Inc., Allentown, Pennsylvania
AND
JOHN R. KLAUDER
*Bell Telephone Laboratories Inc., Murray Hill, New Jersey**

(Received 26 July 1967)

The unitary representations of the affine group, or the group of linear transformations without reflections on the real line, have been found previously by Gel'fand and Naimark. The present paper gives an alternate proof, and presents several properties of the representations which will be used in a later application of this group to continuous representations of Hilbert space. The development follows closely that used by von Neumann to prove the uniqueness of the Schrödinger operators.

1. INTRODUCTION

In several previous papers a theory of continuous representations of Hilbert space has been developed.[1] Central to this theory is the concept of an overcomplete family of states (OFS). For mechanical systems whose dynamics can be described by two canonical variables p, q, the OFS takes on the form

$$\{\Phi[p, q]\} = \{U[p, q]\Phi_0\},$$

where Φ_0 is a suitable unit vector in Hilbert space, and $U[p, q]$ is a two-parameter family of unitary operators. This family and its parametrization are chosen so as to be consonant with the dynamics of the system under consideration; e.g., the parameters p, q are interpreted as the classical dynamical variables.

* Present Address: Department of Physics, Syracuse University, Syracuse, N.Y.

[1] J. R. Klauder, J. Math. Phys. 4, 1055 (1963); 5, 177 (1964).

where

$$y_{-1} = \pm Q, \tag{13a}$$

$$iy'_n = \sum_{\nu=-1}^{n+1} y_\nu y_{n-\nu}, \quad n = -1, 0, 1, \cdots. \tag{13b}$$

It follows immediately from Eq. (10) that $q(z)$ is given by the terms of odd index in Eq. (12),

$$q(z) = \sum_{k=0}^{\infty} y_{2k-1}\lambda^{-2k+1}. \tag{14}$$

This can be seen by writing Eq. (12) as separate sums of odd and even terms,

$$y = y_{\text{odd}} + y_{\text{even}}. \tag{15}$$

Then since $\lambda^2 Q^2$ is even, the even part of Eq. (3) is

$$iy'_{\text{even}} = y^2_{\text{even}} + y^2_{\text{odd}} - \lambda^2 Q^2, \tag{16}$$

and the odd part is

$$i\, y'_{\text{odd}} = 2\, y_{\text{odd}}\, y_{\text{even}}. \tag{17}$$

If $q(z)$ is set equal to y_{odd} this equation is identical to Eq. (10). Equation (14) gives an asymptotic expansion of $q(z)$ about transition points of Eq. (1).

If u_1 and u_2 are any two independent solutions of Eq. (1), normalized so that the Wronskian is $-2i$, as in Eq. (7), then Eqs. (6) imply that

$$q(z) = (u_1 u_2)^{-1}. \tag{18}$$

This result can be checked by substitution in Eq. (8), using Eqs. (1) and (7). Since the Wronskian is invariant under unimodular linear transformations of the functions u_1 and u_2, a more general solution of Eq. (8) is given by

$$q(z) = [(au_1 + bu_2)(cu_1 + du_2)]^{-1}, \tag{19}$$

where the constant coefficients a, b, c, d satisfy

$$ad - bc = 1. \tag{20}$$

Let z_0 be a transition point of Eq. (1). Because of the so-called Stokes phenomenon,[1] the asymptotic series of Eq. (14), for different values of arg $(z - z_0)$, can represent functions $q(z)$ that correspond to different unimodular transformations in Eq. (19).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Unitary Representations of the Affine Group

ERIK W. ASLAKSEN
Department of Physics, Lehigh University, Bethlehem, Pennsylvania
and
Bell Telephone Laboratories Inc., Allentown, Pennsylvania
AND
JOHN R. KLAUDER
*Bell Telephone Laboratories Inc., Murray Hill, New Jersey**

(Received 26 July 1967)

The unitary representations of the affine group, or the group of linear transformations without reflections on the real line, have been found previously by Gel'fand and Naimark. The present paper gives an alternate proof, and presents several properties of the representations which will be used in a later application of this group to continuous representations of Hilbert space. The development follows closely that used by von Neumann to prove the uniqueness of the Schrödinger operators.

1. INTRODUCTION

In several previous papers a theory of continuous representations of Hilbert space has been developed.[1] Central to this theory is the concept of an overcomplete family of states (OFS). For mechanical systems whose dynamics can be described by two canonical variables p, q, the OFS takes on the form

$$\{\Phi[p, q]\} = \{U[p, q]\Phi_0\},$$

where Φ_0 is a suitable unit vector in Hilbert space, and $U[p, q]$ is a two-parameter family of unitary operators. This family and its parametrization are chosen so as to be consonant with the dynamics of the system under consideration; e.g., the parameters p, q are interpreted as the classical dynamical variables.

* Present Address: Department of Physics, Syracuse University, Syracuse, N.Y.

[1] J. R. Klauder, J. Math. Phys. 4, 1055 (1963); 5, 177 (1964).

It is our aim to utilize the continuous-representation theory for unusual classical systems, as for instance systems whose dynamical variables are constrained in one or another way so as, e.g., to lie on the surface of a sphere, or in a half-plane, say $p > 0$.

The gravitational theory is known to be of this particular kind since the metric-tensor variables must of necessity satisfy certain positivity requirements. This suggests the study of systems in which the parameters or dynamical variables are constrained to be positive. Consequently we look for a suitable group to define our OFS. Of particular interest is what may be called the affine group, the group of linear transformations without reflections on the real line: $x \to p^{-1}x - q$, a two-parameter, non-Abelian group which we contend embodies many of the features sought for in an ultimate quantization of the gravitational field.

The unitary representations of the affine group have been previously found by Gel'fand and Naimark.[2] In the present paper we give a different treatment of the problem, one which follows closely the classical treatment given by von Neumann[3] for the canonical group, and which exposes the properties of the group in the same mathematical language as we use in applying the group to dynamical problems.

2. PROPERTIES OF THE AFFINE GROUP

In our particular parametrization the elements of the two-parameter Lie group are given by

$$U[p, q] = e^{-iqP}e^{i \ln pB}, \tag{1}$$

where $-\infty < q < \infty$, $0 < p < \infty$, and P, B are self-adjoint operators on the abstract Hilbert space $\mathcal{H}$. Their commutation relation, which determines the Lie algebra, is given by

$$[P, B] = -iP. \tag{2}$$

The elements of the group can conveniently be written as

$$U[p, q] = V[q]W[p],$$

where

$$V[q] = e^{-iqP}; \quad W[p] = e^{i \ln pB}.$$

From the commutation relation (2) we can deduce

$$e^{isB}Pe^{-isB} = e^{-s}P.$$

This, together with the operator identity

$$e^{B}e^{P}e^{-B} = \exp(e^{B}Pe^{-B}),$$

allows us to write down the group multiplication law

$$U[r, s]U[p, q] = U[rp, s + r^{-1}q]. \tag{3}$$

The commutation relation (2) can also be re-expressed in terms of the operators $V[q]$ and $W[p]$ as

$$V[q]W[p] = W[p]V[pq]. \tag{4}$$

We now want to find the representations of this group, that is, a correspondence between the operators $U[p, q]$ and a two-parameter group of linear, unitary transformations on a separable Hilbert space. This space $\mathfrak{R}$ may be realized as the space of all functions $f(k)$ such that

$$\int_{-\infty}^{\infty} |f(k)|^2\, dk < \infty$$

or in other words the linear space L^2 on the whole real line. The operators on this space are again denoted by $U[p, q]$.

Following von Neumann,[3] we seek an integral operator of the form

$$E = \iint b(p, q)U[p, q]\, dp\, dq, \tag{5}$$

with the following properties:

$$E^{\dagger} = E; \quad E^2 = E, \tag{6}$$

i.e., E is a projection operator. The domain of integration in Eq. (5) is $-\infty < q < \infty$ and $0 < p < \infty$. We say that $b(p, q)$ is the *kernel* of the operator E, and we assume that $b(p, q)$ is square integrable, which is sufficient for E to be defined on all of $\mathfrak{R}$.

Note here the close connection with the work of Peter and Weyl[4]; $U[p, q]$ is the Weyl canonical form, and E is analogous to the "group numbers" used by them.

3. PROPERTIES OF THE PROJECTION OPERATOR E

We now calculate several properties of the operator E which are consequences of the group property.

(a) Let

$$E = \iint b(p, q)U[p, q]\, dp\, dq.$$

Then

$$U[r, s]E = \iint b(p, q)U[r, s]U[p, q]\, dp\, dq$$
$$= \iint b(p, q)U[rp, s + r^{-1}q]\, dp\, dq.$$

With a change of variables we find that

$$U[r, s]E = \iint b(pr^{-1}, r(q - s))U[p, q]\, dp\, dq, \tag{7}$$

so the kernel of $U[r, s]E$ is $b(pr^{-1}, r(q - s))$.

[2] I. M. Gel'fand and M. A. Naimark, Dokl. Akad. Nauk SSSR, **55**, 570 (1947).

[3] J. von Neumann, Math. Ann. **104**, 570 (1931).

[4] F. Peter and H. Weyl, Math. Ann. **96**, 737 (1926).

In the same way we find

$$EU[r, s] = \iint r^{-1}b(pr^{-1}, q - sp^{-1})U[p, q]\,dp\,dq. \quad (8)$$

(b) Let E_1 have the kernel $b_1(p, q)$ and E_2 the kernel $b_2(p, q)$. Then the kernel of E_1E_2 is, from (7),

$$\iint b_1(r, s)b_2(pr^{-1}, r(q - s))\,dr\,ds. \quad (9)$$

(c) We now consider

$$E^\dagger = \iint b^*(p, q)U^\dagger[p, q]\,dp\,dq,$$

where

$$\begin{aligned} U^\dagger[p, q] &= e^{-i\ln pB}e^{iqP} = W[p^{-1}]V[-q], \\ &= V[-pq]W[p^{-1}], \\ &= U[p^{-1}, -qp], \end{aligned} \quad (10)$$

as follows from Eq. (4).

After relabeling, we find

$$E^\dagger = \iint p^{-1}b^*(p^{-1}, -pq)U[p, q]\,dp\,dq,$$

and as a general requirement on $b(p, q)$ for E to be Hermitian we obtain

$$p^{-1}b^*(p^{-1}, -pq) = b(p, q). \quad (11)$$

(d) The kernel of EE follows from (9), and to satisfy $EE = E$ we require

$$\iint b(r, s)b(pr^{-1}, r(g - s))\,dr\,ds = b(p, q). \quad (12)$$

We want to consider the integration over s in Eq. (12), and for simplicity we assume that $b(r, s)$ is a rational function. Let us extend s to a complex variable z, and introduce the notation $b(r, s) = b(r, z) = b(z)$, with r as an implicit parameter, $b(pr^{-1}, r(q - s)) = b(pr^{-1}, r(q - z)) = b'(z)$ with r, p, q as implicit parameters.

We further assume that $b(z)b'(z) \to 0$ as $|z| \to \infty$ such that the integral along a semicircle in the upper or lower complex z plane goes to zero as the radius of the circle goes to infinity. There is no essential loss of generality in these assumptions since, as is quite evident, there exists a set of such functions which is dense in L^2.

The fact that Eq. (12) holds assures us that the integral exists for all values of r, p, q; in particular the poles in the z plane never lie on the real axis. Let $z = z_0(r)$ be a pole for $b(z)$, then $z = q - r^{-1}z_0(pr^{-1})$ is a pole for $b'(z)$; thus, there is a one-to-one correspondence between the poles of $b(z)$ and $b'(z)$. Since $p, r > 0$, such "corresponding" poles lie on opposite sides of the real axis.

Furthermore, we assert that the residues at two corresponding poles are of opposite sign and equal magnitude. Assume that $b(z)$ has n poles, and that the assertion is true for $n - 1$ of the distinct poles, in the sense of a partial-fraction expansion. The value of the integral

$$\int_C b(z)b'(z)\,dz$$

is the same whether we close the integration path C in the upper or lower half-plane, thus the sum of the residues in the upper half-plane, $\sum_{i=1}^n R_i^u$, is equal to the negative sum of the residues in the lower half plane, $-\sum_{i=1}^n R_i^l$. But by assumption

$$\sum_{i=1}^{n-1} R_i^u + R_n^u = -\sum_{i=1}^{n-1} R_i^l - R_n^l$$

implies $R_n^u = -R_n^l$. The assertion is obviously true for $n = 1$, thus for all n.

Let E_+ be an arbitrary operator with kernel $b_+(p, q)$. Then there exists a unique associated operator E_- with kernel $b_-(p, q) \equiv b_+(p, -q)$, such that $E_+E_- = 0$, i.e., the two operators are orthogonal. This follows from the preceding conclusion about the residues, since the kernel of E_+E_- will have its corresponding poles in the same half-plane, so that the integration over s in (9) already gives zero. In particular this is true when E_+ and E_- are projection operators, so that the corresponding subspaces are orthogonal.

(e) We now calculate the kernel of the operator $E_\pm U[l, m]E_\pm$. Since

$$U[l, m]E_\pm = \iint b_\pm(pl^{-1}, l(q - m))U[p, q]\,dp\,dq,$$

$$E_\pm U[l, m]E_\pm = \iiiint b_\pm(r, s)U[r, s] \times b_\pm(pl^{-1}, l(q - m))U[p, q]\,dr\,ds\,dp\,dq.$$

After using Eq. (7) and making a change of variables, we get for the kernel of $E_\pm U[l, m]E_\pm$

$$\iint b_\pm(r, s)b_\pm\left(\frac{p}{lr}, l(r(q - s) - m)\right)dr\,ds. \quad (13)$$

Adopting the notation of (d), we can write the integral over s as

$$\int b_\pm(z)b_\pm'(z)\,dz.$$

If $z = z_0(r)$ is a pole for $b_\pm(z)$, then

$$z = q - mr^{-1} - (rl)^{-1}z_0\left(\frac{p}{rl}\right)$$

is the corresponding pole for $b_\pm'(z)$, and again we see that, since p, r, $l > 0$, the corresponding poles lie on opposite sides of the real axis. Thus the discussion given in (d) holds for this case also, and we immediately get the additional result that

$$E_\pm U[l, m]E_\mp = 0. \tag{14}$$

4. A SPECIFIC KERNEL

In order to proceed further, and also as an illustration of the foregoing arguments, we introduce a specific function for $b_\pm(p, q)$. From the general theory of continuous representations, which will be detailed in a subsequent paper, we find that a projection operator will have the form

$$|\psi\rangle\langle\psi| = \iint (\psi, PU_0^\dagger[p, q]\psi)U[p, q]\frac{dp\,dq}{2\pi}.$$

Choosing $\varphi = 2k^{\frac{1}{2}}e^{-k}$ for $k > 0$ and $\varphi = 0$ for $k < 0$, where φ is the element of $\mathfrak{R}$ corresponding to ψ, and using the representation of Gel'fand and Naimark[2] for U_0, we get for the kernel

$$\begin{aligned} b_\pm(p, q) &= 4\int_0^\infty ke^{-k}e^{\pm iqk}p^{-2}ke^{-kp^{-1}}\,dk \\ &= 8p^{-2}(1 + p^{-1} \mp iq)^{-3}, \end{aligned} \tag{15}$$

which is of the rational type assumed. It should be noted that using one particular representation as a guide to finding a suitable kernel does not impair the generality of our argument, since the representation of $U[p, q]$ in $E_\pm$ is left completely open.

(a) Using Eq. (9), $E_\pm E_\pm$ has the kernel

$$\begin{aligned} 64\iint & r^{-2}(1 + r^{-1} \mp is)^{-3} \\ &\times p^{-2}(1 + rp^{-1} \mp ir(q - s))^{-3}\frac{dr\,ds}{2\pi} \\ &= \frac{32}{\pi}p^{-2}\int_{-\infty}^\infty r^{-3}\,dr\int_{-\infty}^\infty (s \pm a)^{-3}(s - c^\pm)^{-3}\,ds, \end{aligned}$$

where

$$a \equiv i(1 + r^{-1}), \quad c^\pm \equiv q \pm i(r^{-1} + p^{-1}).$$

After the integration over s the kernel is

$$584p^{-2}\int_0^\infty r^{-3}(r^{-1} + d^\pm)^{-5}\,dr,$$

where

$$d^\pm \equiv \tfrac{1}{2}(1 + p^{-1} \pm iq),$$

and integrating over r gives $8p^{-2}(1 + p^{-1} \mp iq)^{-3}$. In summary, we note that

$$E_\pm = 8\iint p^{-2}(1 + p^{-1} \mp)q)^{-3}U[p, q]\frac{dp\,dq}{2\pi}.$$

Furthermore, E_+ and E_- are Hermitian, and $E_\pm E_\mp = 0$, since the integration over s gives zero, as expected.

(b) Using Eq. (13), the kernel of $E_\pm U[l, m]E_\pm$ is given by

$$\frac{32}{\pi}p^{-2}l^{-1}\int_0^\infty r^{-3}\,dr\int_{-\infty}^\infty (s \pm u)^{-3}(s - v^\pm)^{-3}\,ds,$$

where

$$\begin{aligned} u &= i(1 + r^{-1}), \\ v^\pm &= q - mr^{-1} \pm i((lr)^{-1} + p^{-1}). \end{aligned}$$

The result of the integration over s is

$$384ip^{-2}l^{-1}(m \mp i(1 + l^{-1}))^{-5}\int_0^\infty r^{-3}(r^{-1} + h^\pm)^{-5}\,dr,$$

where

$$h^\pm = \frac{q \pm i(1 + p^{-1})}{-m \pm i(1 + l^{-1})}.$$

The integrand has a pole at $r = -(h^\pm)^{-1}$, but this pole never lies on the positive real axis, as is seen from the following calculation:

$$(h^\pm)^{-1} = \frac{[-m \pm i(1 + l^{-1})][q \mp i(1 + p^{-1})]}{q^2 + (1 + p^{-1})^2}.$$

For the pole to lie on the real axis, $\mathrm{Im}\,(h^\pm)^{-1} = 0$, or

$$q(1 + l^{-1}) = -m(1 + p^{-1}),$$

$$q = -m\frac{1 + p^{-1}}{1 + l^{-1}}.$$

But then

$$\mathrm{Re}\,(h^\pm)^{-1} = \frac{m^2(1 + p^{-1})(1 + l^{-1})^{-1} + (1 + l^{-1})(1 + p^{-1})}{q^2 + (1 + p^{-1})^2} > 0.$$

Thus, the pole can never lie on the positive real axis, and the integral exists for all p, q, l, and m.

The result of the integration over r is

$$-32p^{-2}l^{-1}[m \mp i(1 + l^{-1})]^{-2}(1 + p^{-1} \mp iq)^{-3}.$$

With $E_\pm U[l, m]E_\pm = C_\pm(l, m)E_\pm$, we get

$$C_\pm(l, m) = -4l^{-1}[m \mp i(1 + l^{-1})]^{-2}. \tag{16}$$

Since $U[1, 0] = I$, we must obtain $C_\pm(1, 0) = 1$, which is evidently satisfied by Eq. (16).

5. THE INVARIANT SUBSPACES OF $\mathfrak{R}$

Consider the solutions to the equation $E_\pm f = f$, $f \in \mathfrak{R}$. Since $E_\pm$ is bounded as a projection operator, the solutions form a closed linear subspace of $\mathfrak{R}$, denoted by $\mathcal{M}_\pm$. The subspace orthogonal to $\mathcal{M}_\pm$ is $\mathcal{N}_\pm$, defined by $E_\pm g = 0$ for each $g \in \mathcal{N}_\pm$.

If $f, g \in \mathcal{M}_\pm$, then

$$\begin{aligned}(U[\alpha,\beta]f, U[\gamma,\delta]g) &= (U[\alpha,\beta]E_\pm f, U[\gamma,\delta]E_\pm g)\\ &= (E_\pm f, U^\dagger[\alpha,\beta]U[\gamma,\delta]E_\pm g)\\ &= (f, E_\pm U[\gamma\alpha^{-1}, -\alpha\beta+\alpha\delta]E_\pm g)\\ &= C_\pm(\gamma\alpha^{-1}, -\alpha\beta+\alpha\delta)(f,g), \quad (17)\end{aligned}$$

where we have used Eqs. (9) and (16). From Eq. (14) we see that, if $f \in \mathcal{M}_+$ and $g \in \mathcal{M}_-$,

$$(U[\alpha,\beta]f, U[\gamma,\delta]g) = (f, E_+U[\gamma\alpha^{-1}, -\alpha\beta+\alpha\delta]E_-g) = 0.$$

The space $\mathfrak{R}$ is separable, so we can introduce an orthonormal basis $\{\varphi_\pm^i, i = 1, 2, \cdots\}$ for $\mathcal{M}_\pm$. It then follows that

$$(U[\alpha,\beta]\varphi_\pm^m, U[\gamma,\delta]\varphi_\pm^n) = C_\pm(\gamma\alpha^{-1}, -\alpha\beta+\alpha\delta)\delta_{n,m}.$$

Let $\mathfrak{T}_\pm^n = [\overline{U[p,q]\varphi_\pm^n}]$, the closed subspace spanned by vectors of the form $U[p,q]\varphi_\pm^n$ for all p, q such that $0 < p < \infty$, $-\infty < q < \infty$. The subspaces $\mathfrak{T}_\pm^n$ and $\mathfrak{T}_\pm^m$ are orthogonal for $m \neq n$, and $\mathfrak{T}_\pm^n$ and $\mathfrak{T}_\mp^m$ are orthogonal for all m, n. We can therefore form the following direct sum:

$$\sum_n \oplus \mathfrak{T}_\pm^n = \mathfrak{S}_\pm,$$

and denote by $\mathfrak{J}$ the complement of $\mathfrak{S}_+ \oplus \mathfrak{S}_-$ in $\mathfrak{R}$.

The transformation $U[p,q]$ generates an automorphism on $\mathfrak{T}_\pm^n$, and since $U^{-1}[p,q] = U^\dagger[p,q] = U[p^{-1}, -qp]$, $\mathfrak{T}_\pm^n$ is invariant under $\{U[p,q]\}$. $\mathfrak{S}_\pm$ and $\mathfrak{J}$ are then also invariant.

That $\mathfrak{S}_\pm = \sum_n \oplus \mathfrak{T}_\pm^n$ implies that $\varphi_\pm^n \in \mathfrak{S}_\pm$ for each n, which further implies that $\mathfrak{S}_\pm \supset \mathcal{M}_\pm$, since $\{\varphi_\pm^i\}$ is a basis for $\mathcal{M}_\pm$. Thus $\mathfrak{J} \subset \mathcal{N}_+ \cap \mathcal{N}_-$, and $E_+f = E_-f = 0$ for each $f \in \mathfrak{J}$. The invariance of $\mathfrak{J}$ under $U[p,q]$ can then be written as

$$f \in \mathfrak{J} \text{ implies } E_+U[r,s]f = E_-U[r,s]f = 0, \quad \text{for all } r, s. \quad (18)$$

Consider the expression $E_\pm U[r,s]f$. Using Eq. (8), we write it as

$$\iint r^{-1}b_\pm(pr^{-1}, q - sp^{-1})U[p,q]f\,dp\,dq,$$

and if we apply Stone's theorem[5] to the commutative subgroup $V[q]$, we find

$$\iint r^{-1}b_\pm(pr^{-1}, q - sp^{-1})e^{-iqk}W[p]f(k)\,dp\,dq = \int r^{-1}\tilde{b}_\pm(pr^{-1}, k)e^{iksp^{-1}}W[q]f(k)\,dp, \quad (19)$$

with

$$\tilde{b}_\pm(pr^{-1}, k)e^{iksp^{-1}} \equiv \int b_\pm(pr^{-1}, q - sp^{-1})e^{-iqk}\,dp.$$

[5] F. Riesz and B. Sz-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1965).

The integral over q is just the Fourier transform of $b_\pm$ on the second argument, and the transformed function $\tilde{b}_\pm$ is therefore identically zero only if $b_\pm \equiv 0$.

However, we see that if $b_+(p,q)$ has poles (as discussed in Sec. 3) in the lower half-plane only, then $\tilde{b}_+$ actually vanishes for $k \leq 0$, hence $E_+f = 0$ if $f(k) = 0$ for $k > 0$. Likewise, $E_-f = 0$ if $f(k) = 0$ for $k < 0$. Consequently, after the integration over q we may restrict k to have either only positive values or only negative values. The kernel given by Eq. (15) is of the above mentioned type, since it has one pole only, and henceforth we shall interpret the + sign in $b_+(p,q)$ to mean that $b_+(p,q)$ has poles in the lower half-plane only. In the case of a more general kernel $b(p,q)$ with poles in both half-planes we may decompose it into two such simple kernels:

$$b(p,q) = b_+^{(1)}(p,q) + b_-^{(2)}(p,q),$$
$$b(p,-q) = b_-^{(1)}(p,q) + b_+^{(2)}(p,q).$$

For completeness we also remark that, if the kernel of a projection operator is found by using an arbitrary representation of U_0, this kernel will have poles in the one half-plane if and only if the representation is irreducible.

To investigate the integration over p, we note that r and s are two independent, arbitrary parameters, so we may write rs instead of s, thus Eq. (9) can be written as

$$\int_0^\infty r^{-1}\tilde{b}_\pm(pr^{-1}, k)e^{iksrp^{-1}}W[p]f(k)\,dp.$$

With $p = e^\xi$ and $r = e^\eta$, the integral is

$$\int_{-\infty}^\infty e^{-\eta}\hat{b}_\pm(\xi - \eta, k)\hat{F}(\xi, k)\,d\xi, \quad (20)$$

where

$$\hat{b}_\pm(\xi - \eta, k) \equiv \tilde{b}_\pm(e^{\xi-\eta}, k)e^{ikse^{\eta-\xi}},$$

and

$$\hat{F}(\xi, k) \equiv e^\xi W[e^\xi]f(k).$$

But Eq. (20) is the convolution of $\hat{b}_\pm(\eta, k)$ with $\hat{F}(\eta, k)$, which is identically zero if and only if the product of the Fourier transforms of $\hat{b}_\pm(\eta, k)$ and $\hat{F}(\eta, k)$, say $\beta_\pm(\alpha, k)$ and $\varphi(\alpha, k)$, respectively, is identically zero. Since $b_\pm(r,s)$ is assumed to be a rational function, $\beta_\pm(\alpha, k)$ is an analytic function of α for each fixed k, as is seen by expanding $b_\pm(r,s)$ in a partial fractions expansion and carrying out the two transforms. Thus, for all fixed k, except for possibly a set of measure zero, $\beta_\pm(\alpha, k)$ does not vanish on any interval $a < \alpha < b$. Therefore, $\beta_\pm \cdot \varphi = 0$ implies $\varphi = 0$, which further implies $\hat{F} = 0$ which finally implies $f(k) = 0$ for all k except possibly a set of measure zero.

We can now strengthen Eq. (18) and state the

Lemma: $f \in \mathfrak{Z}$ implies $E_+U[r,s]f = E_-U[r,s]f = 0$ for all r, s, which further implies $f = 0$.

Define

$$\mathfrak{R}_+ = \{f(k): f(k) = 0 \quad \text{for} \quad k \leq 0\},$$
$$\mathfrak{R}_- = \{f(k): f(k) = 0 \quad \text{for} \quad k \geq 0\},$$

then we have

$$\mathfrak{R} = \mathfrak{R}_+ \oplus \mathfrak{R}_-,$$

and

$$\mathfrak{R}_+ = \sum_n \oplus \mathfrak{T}_+^n, \quad \mathfrak{R}_- = \sum_n \oplus \mathfrak{T}_-^n.$$

From the definition of the subspaces $\mathfrak{T}_\pm^n$ we see that there is a one-to-one correspondence between any two of them, and by Eq. (17) this correspondence is an isometry, thus representations on these subspaces are unitarily equivalent.

We can summarize our results as follows:

Theorem: The affine group admits, for each realization of its elements as linear, unitary transformations on a separable Hilbert space, two faithful, inequivalent representations. Each of these reduces to finitely or countably many equivalent, irreducible representations.

The number of equivalent representations will of course depend upon the particular realization of the elements $U[p,q]$ as linear transformations.

6. A PARTICULAR REPRESENTATION

We now turn our attention to the particular realization of $U[p,q]$ given by Gel'fand and Naimark[2]:

$$U[p,q]f(k) = e^{-iqk}p^{-\frac{1}{2}}f(p^{-1}k).$$

We seek the solutions to $E_\pm f = f$; we carry through the calculation in $\mathfrak{R}_+$ only, since it is identical in $\mathfrak{R}_-$.

$$\begin{aligned} E_+f(k) &= 8\iint p^{-2}(1 + p^{-1} - iq)^{-3} \\ &\quad \times e^{-iqk}p^{-\frac{1}{2}}f(p^{-1}k)\,\frac{dp\,dq}{2\pi} \\ &= \frac{4}{\pi}\int_0^\infty dp\, p^{-\frac{5}{2}}f(p^{-1}k) \\ &\quad \times \int_{-\infty}^{\infty}(1 + p^{-1} - iq)^{-3}e^{-iqk}\,dq \\ &= 4k^2e^{-k}\int_0^\infty p^{-\frac{5}{2}}e^{-p^{-1}}f(p^{-1}k)\,dp \\ &= 4k^{\frac{1}{2}}e^{-k}\int_0^\infty p^{\frac{1}{2}}e^{-p}f(p)\,dp. \end{aligned}$$

The equation $E_+f = f$ has one and only one solution up to a scale factor c, namely $f(k) = ck^{\frac{1}{2}}e^{-k}$. This is then the one and only normalized basis vector φ, and the corresponding subspace $P = \overline{[U[p,q]\varphi]} = \mathfrak{R}_+$.

7. EXTENSION TO N INDEPENDENT DEGREES OF FREEDOM

Consider a system with N degrees of freedom, described by $2N$ variables. The operators P_α and B_α, where $\alpha = 1, 2, \cdots N$, are defined by

$$[P_\alpha, B_\beta] = -i\delta_{\alpha\beta}P_\alpha,$$
$$[P_\alpha, P_\beta] = [B_\alpha, B_\beta] = 0.$$

We define $2N$ one-parameter families of unitary operators by

$$V_\alpha[q_\alpha] = e^{-iq_\alpha P_\alpha}, \quad W_\alpha[p_\alpha] = e^{i\ln p_\alpha B_\alpha}.$$

In terms of these operators the commutation relations are

$$V_\alpha[q_\alpha]W_\alpha[p_\alpha] = V_\alpha[q_\alpha(1-p_\alpha)]W_\alpha[p_\alpha]V_\alpha[q_\alpha],$$
$$[V_\alpha[q_\alpha], W_\beta[p_\beta]] = 0 \quad \text{for} \quad \alpha \neq \beta,$$
$$[V_\alpha[q_\alpha], V_\beta[q_\beta]] = [W_\alpha[p_\alpha], W_\beta[p_\beta]] = 0.$$

We further define

$$\mathfrak{V}[\mathbf{q}] = \prod_{\alpha=1}^N V_\alpha[q_\alpha], \quad \mathfrak{W}[\mathbf{p}] = \prod_{\alpha=1}^N W_\alpha[p_\alpha],$$

where

$$\mathbf{q} = \{q_\alpha\} \quad \text{and} \quad \mathbf{p} = \{p_\alpha\}.$$

The elements of the $2N$-parameter group are then

$$U[\mathbf{p},\mathbf{q}] = \mathfrak{V}[\mathbf{q}]\mathfrak{W}[\mathbf{p}].$$

As a representation space, consider the function space $\mathfrak{R}$ with elements $f(\mathbf{k})$, where $\mathbf{k} = \{k_1, k_2, \cdots, k_N\}$ and $\int |f(\mathbf{k})|^2\,d\mathbf{k} < \infty$. Now, exactly as in the case of the two-parameter group, we can find 2^N operators

$$E_\omega = \prod_{\alpha=1}^N E_\pm^\alpha,$$

with

$$E_\pm^\alpha = \iint b_\pm^\alpha(p_\alpha, q_\alpha)V_\alpha[q_\alpha]W_\alpha[p_\alpha]\,\frac{dp_\alpha\,dq_\alpha}{2\pi},$$

where

$$E_\pm^\alpha E_\pm^\alpha = E_\pm^\alpha, \quad E_\pm^\alpha E_\mp^\alpha = 0,$$
$$(E_\pm^\alpha)^\dagger = E_\pm^\alpha,$$

and ω is an integer between one and 2^N.

The whole proof proceeds just as before, and if the operators $V_\alpha[q_\alpha]$ and $W_\alpha[p_\alpha]$ are given by

$$V_\alpha[q_\alpha]f(\mathbf{k}) = e^{-iq_\alpha k_\alpha}f(\mathbf{k}),$$
$$W_\alpha[p_\alpha]f(\mathbf{k}) = p_\alpha^{-\frac{1}{2}}f(k_1, \cdots, k_{\alpha-1}, p_\alpha^{-1}k_\alpha, k_{\alpha+1}, \cdots, k_N),$$

we find that there are 2^N inequivalent, irreducible, faithful, unitary representations.

These representations can be numbered as follows: Consider the sequence $b_1, b_2, \cdots, b_N$. Let $b_n = \theta(k_n)$ such that $\theta(k_n) = 1$ if $k > 0$, $\theta(k_n) = 0$ if $k < 0$. Then we have an ordered sequence of zeros and ones which may be read as a binary number, the identification number for this particular representation.

Clebsch-Gordan Coefficients for Space Groups*

D. B. LITVIN AND J. ZAK
Department of Physics, Technion—Israel Institute of Technology, Haifa, Israel

(Received 6 June 1967)

It is shown that to find Clebsch-Gordan coefficients of space groups (both single and double), the representations of the groups of **k** alone are required. This is another example demonstrating the well-accepted fact that in applications of space groups it is sufficient to know the representations of the groups of **k**. Final formulas are derived that enable the calculation of the Clebsch-Gordan coefficients from the representations of the groups of **k**. As an example the spin-orbit coupling in solids is considered.

I. INTRODUCTION

In applications of group theory in physics the problem very often arises of decomposing a direct product of two irreducible representations into a sum of irreducible parts. A classical example is the addition of angular momentum in quantum mechanics. In the theory of solid-state physics such a decomposition is required in defining selection rules in scattering processes in crystals.[1] Sometimes more detailed information is needed, as when one has to express a product of two wavefunctions $\psi_i^{(\alpha)}\psi_j^{(\beta)}$, which are specified according to irreducible representations of some symmetry group, α being the index of the representation and i of the row, by means of functions $\psi_k^{(\gamma)}$ that undergo transformations according to irreducible representations of the same group. The elements of the matrix that gives the connection between $\Psi_i^{(\alpha)}\psi_j^{(\beta)}$ and $\psi_k^{(\gamma)}$ are called the Clebsch-Gordan coefficients. These coefficients are of general interest for each specific symmetry group. For example, they are of very great use for the three-dimensional rotation group and different kinds of SU groups.

In solids the symmetry groups are space groups, and this paper deals with the question of finding the Clebsch-Gordan coefficients for them. The method used is one developed by Koster[2] for finite groups. It is shown that the finding of Clebsch-Gordan coefficients for space groups can be reduced to formulas containing only the representations of groups of the vector **k**. A similar result was obtained before, for the decomposition of direct products of representations of space groups when one is interested in selection rules in crystals.[3]

As an example, it is shown how the Clebsch-Gordan coefficients are obtained for the spin-orbit coupling in solids.

II. GENERAL FORMALISM

Any space group G can be decomposed for a specific **k**, a vector in the first Brillouin zone, into q left cosets

$$G = (\alpha_0 \mid \mathbf{A}_0)\mathcal{K} + (\alpha_i \mid \mathbf{A}_1)\mathcal{K} + \cdots + (\alpha_{q-1} \mid \mathbf{A}_{q-1})\mathcal{K},$$

where $(\alpha_0 \mid \mathbf{A}_0) = (\epsilon \mid 0)$ is the unit element and $\mathcal{K}$, the little group of the vector **k**, is the set of elements $\{(\beta \mid \mathbf{B})\}$ with the property that $\beta\mathbf{k} = \mathbf{k}$ or $\beta\mathbf{k} = \mathbf{k} + \mathbf{K}$ where **K** is a lattice vector of k space. We will denote these two relations by $\beta\mathbf{k} \doteq \mathbf{k}$. The set of elements $\{(\alpha_i \mid \mathbf{A}_i)\}$, the representing elements, has the property that $\alpha_i\mathbf{k} = \mathbf{k}_i$ where $\mathbf{k}_i \not\doteq \mathbf{k}$. The q vectors $\mathbf{k}_i$, i.e., $\mathbf{k}_0 = \mathbf{k}, \mathbf{k}_1, \cdots, \mathbf{k}_{q-1}$, form the star of the vector **k** denoted by S_k.

The elements of G can be written as

$$(\alpha \mid \mathbf{A}) = (\alpha \mid \mathbf{v}(\alpha) + \mathbf{a}) = (\epsilon \mid \mathbf{a})(\alpha \mid \mathbf{v}(\alpha)),$$

where **a** is a primitive translation and $\mathbf{v}(\alpha)$ is either zero or a nonprimitive translation associated with the operator α. We note that $\mathbf{v}(\epsilon) = 0$.

An irreducible representation of the space group G is characterized by the vector **k** and its star, and the irreducible representation of the group of the vector **k**. We denote an irreducible representation of the space group G by $D_{k^*}^r$, where k^* denotes the specific vector **k** in the Brillouin zone and its star, and r denotes the irreducible representation of the group of the vector **k**. The irreducible representation $D_{k^*}^r$ is a $n = dq$ dimensional irreducible representation, q is the number of vectors in the star of **k**, and d the dimension of the rth irreducible representation of the little group of the vector **k**.

We take the irreducible representation of G in the standard form, i.e., the representation of the elements of the invariant subgroup of translations is of the

* Based on a thesis submitted by D. B. Litvin to the Graduate School of the Technion—Israel Institute of Technology, in partial fulfilment of the requirements for the degree of Master of Science, January 1967.
[1] E. J. Elliott and R. Loudon, J. Phys. Chem. Solids 15, 146 (1960).
[2] G. F. Koster, Phys. Rev. 109, 227 (1958).
[3] J. Zak, J. Math. Phys. 3, 1278 (1962).

following form[4]:

$$D^{r}_{k^*}[\epsilon \mid \mathbf{a}] = \begin{pmatrix} e^{-i\mathbf{k}\cdot\mathbf{a}}I & & & & \\ & e^{-i\mathbf{k}_1\cdot\mathbf{a}}I & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & e^{-i\mathbf{k}_{q-1}\cdot\mathbf{a}}I \end{pmatrix}$$

The unit matrix I is of dimension d. For a general element $D^{r}_{k^*}(G)$ is divided into blocks of dimension d. There will be q rows and columns of blocks. We say then that $D^{r}_{k^*}(G)$ has $n = dq$ rows (columns) and q block rows (block columns).

The $\theta\mu$th block of $D^{r}_{k^*}(G)$, denoted by $D^{r}_{k^*\theta\mu}(G)$, is nonzero, when for all $\mathbf{a}$

$$e^{i\alpha\mathbf{k}_\mu\cdot\mathbf{a}} = e^{i\alpha_\theta\mathbf{k}\cdot\mathbf{a}}.$$

We have for this nonzero block of dimension d:

$$D^{r}_{k^*\theta\mu}[\alpha \mid \mathbf{v}(\alpha) + \mathbf{a}] = D^{r}_{k}[\beta_{\theta\mu} \mid \mathbf{v}(\beta_{\theta\mu}) + \mathbf{b}_{\theta\mu}], \quad (1)$$

where $D^{r}_{k}[\beta_{\theta\mu} \mid \mathbf{v}(\beta_{\theta\mu}) + \mathbf{b}_{\theta\mu}]$ is the d-dimensional irreducible representation of the little group of the vector $\mathbf{k}$ and $(\beta_{\theta\mu} \mid \mathbf{v}(\beta_{\theta\mu}) + \mathbf{b}_{\theta\mu})$ is found from the relation

$$\begin{aligned}(\alpha \mid \mathbf{v}(\alpha) + \mathbf{a})(\alpha_\mu \mid \mathbf{v}(\alpha_\mu) + \mathbf{a}_\mu) \\ = (\alpha_\theta \mid \mathbf{v}(\alpha_\theta) + \mathbf{a}_\theta)(\beta_{\theta\mu} \mid \mathbf{v}(\beta_{\theta\mu}) + \mathbf{b}_{\theta\mu}). \quad (2)\end{aligned}$$

Hence $(\alpha_\mu \mid \mathbf{v}(\alpha_\mu) + \mathbf{a}_\mu)$ and $(\alpha_\theta \mid \mathbf{v}(\alpha_\theta) + \mathbf{a}_\mu)$ are the representing elements, such that $\alpha_\mu\mathbf{k} = \mathbf{k}_\mu$ and $\alpha_\theta\mathbf{k} = \mathbf{k}_\theta$; we assume that $\mathbf{a}_\mu = \mathbf{a}_\theta = 0$.

Defining the nonprimitive translation associated with $\beta_{\theta\mu}$ as

$$\mathbf{v}(\beta_{\theta\mu}) = \alpha_\theta^{-1}(\alpha\mathbf{v}(\alpha_\mu) + \mathbf{v}(d) - \mathbf{v}(d_\theta)),$$

we derive from Eq. (2) that

$$\alpha\alpha_\mu = \alpha_\theta\beta_{\theta\mu} \quad (3)$$

and

$$\mathbf{b}_{\theta\mu} = \alpha_\theta^{-1}\mathbf{a}. \quad (4)$$

Rewriting Eq. (1) and using Eq. (4), we have for the nonzero blocks of $D^{r}_{k^*}$

$$\begin{aligned}D^{r}_{k^*\mu\theta}[\alpha \mid \mathbf{v}(\alpha) + \mathbf{a}] &= D^{r}_{k}[\beta_{\theta\mu} \mid \mathbf{v}(\beta_{\theta\mu}) + \alpha_\theta^{-1}\mathbf{a}] \\ &= e^{-i\mathbf{k}_\theta\cdot\mathbf{a}}D^{r}_{k}[\beta_{\theta\mu} \mid \mathbf{v}(\beta_{\theta\mu})], \quad (5)\end{aligned}$$

where $\beta_{\theta\mu}$ and $\mathbf{v}(\beta_{\theta\mu})$ are given by Eqs. (3) and (4), respectively.

If the direct product of two irreducible representations of the space group G is reducible, we have

$$D^{r}_{k^*} \times D^{r'}_{k'^*} = \sum_{mn} c_{mn} D^{r''_n}_{k''_m{}^*},$$

where c_{mn} is the number of times the irreducible representation appears in the reduced form.

We calculate c_{mn} from

$$c_{mn} = \frac{1}{gh}\sum_{G} \mathrm{X}^{r}_{k^*}(G)\mathrm{X}^{r'}_{k'^*}(G)\mathrm{X}^{r''_n}_{k''_m{}^*}(G)^*,$$

where $\mathrm{X}^{r}_{k^*}(G)$ is the character of $D^{r}_{k^*}(G)$, and gh is the order of the space group G; h is the order of the invariant subgroup of translations T, and g the order of the factor group G/T.

The direct product is put into reduced form by a similarity transformation using a unitary matrix U:

$$U^{-1}[D^{r}_{k^*} \times D^{r'}_{k'^*}]U = \begin{pmatrix}\text{reduced form of the} \\ \text{direct product}\end{pmatrix}.$$

We assume that the irreducible representation $D^{r''_1}_{k''_1{}^*}$, a $n'' = q''d''$ dimensional representation, appears c times in the reduced form. (In general the dimension of $D^{r''_n}_{k''_m{}^*}$ is denoted as $n''_{mn} = q''_m d''_n$ and its multiplicity in the reduced form as c_{mn}; when referring to $D^{r''_1}_{k''_1{}^*}$, for typographical reasons only, we will drop the indexes m and n from the dimensionality and multiplicity.) If these are the first irreducible representations in the reduced form, i.e., the first irreducible representations along the diagonal of the reduced form, then the first $cq''d''$ columns of U are calculated from the equation[2]

$$\begin{aligned}\frac{d''q''}{gh}\sum_{G}[D^{r}_{k^*}(G) \times D^{r'}_{k'^*}(G)]_{mn}[D^{r''_1}_{k''_1{}^*}(G)]^*_{i''j''} \\ = U_{mi''}U^*_{nj''} + U_{m(d''q''+i'')}U^*_{n(d''q''+j'')} \\ + U_{m(2d''q''+i'')}U^*_{n(2d''q''+j'')} \\ + U_{m([c-1]d''q''+i'')}U_{n([c-1]d''q''+j'')}. \quad (6)\end{aligned}$$

This gives the elements of the first $cq''d''$ columns of U in terms of the known quantities

$$\frac{d''q''}{gh}\sum_{G}[D^{r}_{k^*}(G) \times D^{r'}_{k'^*}(G)]_{mn}[D^{r''_1}_{k''_1{}^*}(G)]^*_{i''j''}.$$

We also know that

$$[D^{r}_{k^*}(G) \times D^{r'}_{k'^*}(G)]_{mn} \equiv D^{r}_{k^*}(G)_{ij}D^{r'}_{k'^*}(G)_{i'j'},$$

where m and n are used as an abbreviation for the double indices (i, i') and (j, j'), respectively. The possible values of the indices are as follows:

$$\begin{aligned} m, n &= 1, 2, \cdots, qdq'd', \\ i, j &= 1, 2, \cdots, qd, \\ i', j' &= 1, 2, \cdots, q'd', \\ i'', j'' &= 1, 2, \cdots, q''d''. \end{aligned}$$

qd is the dimension of the irreducible representation

[4] G. F. Koster, *Space Groups and their Representations* (Academic Press Inc., New York, 1957).

D^{r}_{k*}, $q'd'$ of $D^{r'}_{k'*}$, $q''d''$ of $D^{r_1''}_{k_1''*}$, and $qdq'd'$ of the direct product $D^{r}_{k*} \times D^{r'}_{k'*}$.

Let us introduce a new indexation which will show clearly the division of the irreducible representations into blocks. D^{r}_{k*} is a qd-dimensional irreducible representation which is divided into $q \times q$ blocks of dimension $d \times d$. Let θ be the index of the block rows, and μ be the index of the block columns. A specific block of D^{r}_{k*} is denoted by (θ, μ), the intersection of the θth block row and the μth block column.

The ith row of D^{r}_{k*} can be denoted as the ηth row of the θth block row, and the jth column denoted as the πth column of the μth block column, that is,

$$\begin{aligned} i &= \theta d + \eta, \quad \theta, \mu = 0, 1, \cdots, q-1, \\ j &= \mu d + \pi, \quad \eta, \pi = 1, 2, \cdots, d. \end{aligned} \tag{7}$$

Using this indexation, the (i,j)th element of D^{r}_{k*} is denoted as the $(\theta d + \eta)(\mu d + \pi)$th element. It is the $(\eta\pi)$th element of the $(\theta\mu)$th block of D^{r}_{k*}. The rows and columns of $D^{r'}_{k'*}$ and $D^{r_1''}_{k_1''*}$ are denoted in a similar manner.

Recalling that $G = \{(\alpha \mid \mathbf{v}(\alpha) + \mathbf{a})\}$ and using Eqs. (1) and (5), we can write

$$\begin{aligned} D^{r}_{k^*}(G)_{(\theta d+\eta)(\mu d+\pi)} \\ = e^{-i\mathbf{k}_\theta\cdot\mathbf{a}} D^{r}_{k}[\beta_{\theta\mu} \mid \mathbf{v}(\beta_{\theta\mu})]\delta(\alpha\alpha_\mu - \alpha_\theta\beta_{\theta\mu}). \end{aligned} \tag{8}$$

The $(\theta\mu)$th block of $D^{r}_{k*}(G)$ is nonzero only if α fulfills the condition written in the delta function. Two similar expressions can be written for elements of the irreducible representations $D^{r'}_{k'*}$ and $D^{r_1''}_{k_1''*}$.

Using the new indexation, and writing the sum on the elements of the space group G as

$$\sum_G = \sum_{(\epsilon|\mathbf{a})} \sum_{(\alpha|\mathbf{v}(\alpha))}$$

we then rewrite Eq. (6) in the following form:

$$\begin{aligned} &\frac{1}{h}\sum_{(\epsilon|\mathbf{a})} e^{-i(\mathbf{k}_\theta+\mathbf{k}'_{\theta'}-\mathbf{k}_1''_{\theta''})\cdot\mathbf{a}} \frac{d''q''}{g}\sum_{(\alpha|\mathbf{v}(\alpha))} D^{r}_{k}[\beta_{\theta\mu} \mid \mathbf{v}(\beta_{\theta\mu})]_{\eta\pi} \\ &\times D^{r'}_{k'}[\beta'_{\theta'\mu'} \mid \mathbf{v}(\beta'_{\theta'\mu'})]_{\eta'\pi'} D^{r_1''}_{k_1''}[\beta''_{\theta''\mu''} \mid \mathbf{v}(\beta''_{\theta''\mu''})]^*_{\eta''\pi''} \\ &\qquad \times \delta(\alpha\alpha_\mu - \alpha_\theta\beta_{\theta\mu}) \\ &\qquad \times \delta(\alpha\alpha'_{\mu'} - \alpha'_{\theta'}\beta'_{\theta'\mu'})\delta(\alpha\alpha''_{\mu''} - \alpha''_{\theta''}\beta''_{\theta''\mu''}) \\ &= U_{(\theta d+\eta;\theta'd'+\eta')(\theta''d''+\eta'')} U^*_{(\mu d+\pi;\mu'd'+\pi')(\mu''d''+\pi'')} \\ &\qquad + \cdots + U_{(\theta d+\eta;\theta'd'+\eta')([c-1]d''q''+\theta''d''+\eta'')} \\ &\qquad \times U^*_{(\mu d+\pi;\mu'd'+\pi')([c-1]d''q''+\mu''d''+\pi'')}. \end{aligned} \tag{9}$$

The basis functions of the irreducible representation D^{r}_{k*} are $\psi^{\mathbf{k}_\theta}_\eta$, and transform under elements of the space group as follows:

$$(\alpha \mid \mathbf{A})\psi^{\mathbf{k}_\theta}_\eta = \sum_{\mu,\pi} D^{r}_{k^*}[\alpha \mid \mathbf{A}]_{(\mu d+\pi)(\theta d+\eta)}\psi^{\mathbf{k}_\theta}_\pi. \tag{10}$$

The functions $\psi^{\mathbf{k}}_\eta$, $\theta = 0$, form the basis functions of the irreducible representation D^{r}_{k} of the little group $\mathcal{K}$. The functions $\psi^{\mathbf{k}_\theta}_\eta$ are related to the functions $\psi^{\mathbf{k}}_\eta$ by the following relation[4]:

$$\psi^{\mathbf{k}_\theta}_\eta = (\alpha_\theta \mid \nu(\alpha_\theta))\psi^{\mathbf{k}}_\eta, \tag{11}$$

where $(|\alpha_\theta\ \mathbf{v}(\alpha_\theta))$ are the representing elements of the group G. Similar relations hold for the basis functions $\psi^{\mathbf{k}'_{\theta'}}_{\eta'}$ and $\psi^{\mathbf{k}_1''\theta''}_{\eta''}$ of the irreducible representations $D^{r'}_{k'*}$ and $D^{r''}_{k_1''*}$, respectively.

The basis functions of the direct product are $\psi^{\mathbf{k}_\theta}_\eta \psi^{\mathbf{k}'_{\theta'}}_{\eta'}$. The elements of U give the coefficients of the linear combinations of these functions, which form the basis functions of the c irreducible representations $D^{r_1''}_{k_1''*}$ appearing in the reduced form, that is,

$$\psi^{\mathbf{k}_1''\theta''}_{\eta''} = \sum_{\substack{\theta\theta' \\ \eta\eta'}} U_{(\theta d+\eta;\theta'd'+\eta')(ld''q''+\theta''d''+\eta'')}\psi^{\mathbf{k}_\theta}_\eta \psi^{\mathbf{k}'_{\theta'}}_{\eta'}, \tag{12}$$

where $l = 0, 1, \cdots, c-1$.

The columns of U calculated in Eq. (9) can be shown to be divided into blocks. The $cq''d''$ columns are divided into sections, and each section divided into $qq' \times q''$ blocks of dimension $dd' \times d''$. A specific block is denoted by $(\theta\theta')(lq'' + \theta'')$, the $(\theta\theta')(\theta'')$th block of the lth section.

Equation (9) facilitates the calculation of the set of c blocks $(\theta\theta')(lq'' + \theta'')$, $l = 0, 1, \cdots, c-1$, for each trio of values of θ, θ', and θ''. Once we have chosen a specific trio we may perform the first sum in the equation, for

$$\frac{1}{h}\sum_{(\epsilon|\mathbf{a})} e^{-i(\mathbf{k}_\theta+\mathbf{k}'_{\theta'}-\mathbf{k}_1''_{\theta''})\cdot\mathbf{a}} = \begin{cases} 0 & \text{if} \quad \mathbf{k}_\theta + \mathbf{k}'_{\theta'} - \mathbf{k}''_{1\theta''} \not\doteq 0 \\ 1 & \text{if} \quad \mathbf{k}_\theta + \mathbf{k}'_{\theta'} - \mathbf{k}''_{1\theta''} \doteq 0. \end{cases} \tag{13}$$

By choosing in Eq. (9) the indices $\mu\mu'\mu''$ and $\pi\pi'\pi''$, equal to $\theta\theta'\theta''$ and $\eta\eta'\eta''$, respectively, and using Eq. (13), one finds that sums of the squares of elements of the $(\theta\theta')(lq'' + \theta'')$th blocks are equal to zero if $\mathbf{k}_\theta + \mathbf{k}'_{\theta'} - \mathbf{k}''_{1\theta''} \not\doteq 0$. Consequently, the elements of the blocks for which $\mathbf{k}_\theta + \mathbf{k}'_{\theta'} - \mathbf{k}''_{1\theta''} \not\doteq 0$ are zero. In the following we assume that $\mathbf{k}_\theta + \mathbf{k}'_{\theta'} - \mathbf{k}''_{1\theta''} \doteq 0$. We also note that in the second sum of Eq. (9) we do not have to sum over all α, but only over those that simultaneously fulfill the three delta conditions.

Let us begin by calculating the first block column of each section, that is, the $(\theta\theta')(lq'')$th blocks, where we have taken $\theta'' = 0$. We first choose $\theta = \theta' = \theta'' = 0$ and calculate the $(00)(lq'')$th blocks, using Eq. (9.)

We have assumed that $\mathbf{k} + \mathbf{k}' - \mathbf{k}_1'' \doteq 0$, where $\mathbf{k}$ is the first vector of the star of $\mathbf{k}$, $\mathbf{k}'$ the first vector of the star of $\mathbf{k}'$, and $\mathbf{k}_1''$, the first vector of the star of $\mathbf{k}_1''$.

In our notation we have

$$\mathbf{k}_0 = \mathbf{k}, \quad \alpha_0 = \epsilon;$$
$$\mathbf{k}_0' = \mathbf{k}', \quad \alpha_0' = \epsilon;$$
$$\mathbf{k}_{1_0}'' = \mathbf{k}_1'', \quad \alpha_0'' = \epsilon.$$

We choose $\mu = \mu' = \mu'' = 0$ and with this choice, the three delta conditions in Eq. (9) are

$$\alpha = \beta, \quad \alpha = \beta', \quad \alpha = \beta''.$$

Let us denote the elements of G which simultaneously fulfill the three delta conditions by $\hat{\beta}$. We have then:

$$\{\hat{\beta}\} = \{\beta\} \cap \{\beta'\} \cap \{\beta''\}. \tag{14}$$

The set of elements $\{\hat{\beta}\}$ is the intersection of the point groups associated with the three little groups $\mathcal{K}$, $\mathcal{K}'$, and $\mathcal{K}''$.

To find the elements $\hat{\beta}$ the following remark is useful. One can define the direct product of two stars, $S_\mathbf{k} \times S_{\mathbf{k}'}$, the star $S_\mathbf{k}$ of the vector $\mathbf{k}$ times the star $S_{\mathbf{k}'}$ of the vector $\mathbf{k}'$ as the aggregate of all vectors formed by adding vectorally one vector of the star of $\mathbf{k}$ and one vector of the star of $\mathbf{k}'$.[5] One may write

$$S_\mathbf{k} \times S_{\mathbf{k}'} = \sum_m \epsilon_m S_{\mathbf{k}_m''},$$

where ϵ_m is the number of times the star $S_{\mathbf{k}_m''}$ appears in the direct product $S_\mathbf{k} \times S_{\mathbf{k}'}$. Two types of stars appear in the reduced form of the direct product of the two stars. If $\epsilon_m = 1$ we speak of $S_{\mathbf{k}_m''}$ as a star of the first kind, and if $\epsilon_m > 1$, as a star of the second kind.

It is clear that the elements of G which simultaneously leave the vectors $\mathbf{k}$ and $\mathbf{k}'$ invariant also leave $\mathbf{k}_1''$ invariant. Now, if $S_{\mathbf{k}_1''}$ is a star of the first kind, then the point group $\{\beta''\}$ contains only these elements, and relation (11) reduces to

$$\{\hat{\beta}\} = \{\beta''\}. \tag{15}$$

However, if $S_{\mathbf{k}_1''}$ is a star of the second kind, then $\{\beta''\}$ contains additional elements that, while leaving $\mathbf{k}_1''$ invariant, do not leave $\mathbf{k}$ and $\mathbf{k}'$ invariant. Relation (11) reduces in this case to (see Appendix A):

$$\{\hat{\beta}\} = \{\beta\} \cap \{\beta''\}. \tag{16}$$

We are now in a position to use Eq. (9) to calculate the (00)(lq'') blocks; this equation becomes

$$\frac{d''q''}{g} \sum_{(\hat{\beta}|\mathbf{v}(\hat{\beta}))} D_k^r[\hat{\beta} \mid \mathbf{v}(\hat{\beta})]_{\eta\pi} D_{k'}^{r'}[\hat{\beta} \mid \mathbf{v}(\hat{\beta})]_{\eta'\pi'} D_{k_1''}^{r_1''}[\hat{\beta} \mid \mathbf{v}(\hat{\beta})]^*_{\eta''\pi}$$
$$= U_{(\eta\eta')(\eta'')} U^*_{(\pi\pi')(\pi'')} + U_{(\eta\eta')(d''q''+\eta'')} U^*_{(\pi\pi')(d''q''+\pi'')}$$
$$+ \cdots + U_{(\eta\eta')([c-1]d''q''+\eta'')} U^*_{(\pi\pi')([c-1]d''q''+\eta'')}, \tag{17}$$

where the elements $\{\hat{\beta}\}$ are given by Eq. (12) or (13).

[5] J. L. Birman, Phys. Rev. **127**, 1093 (1962).

If the irreducible representation $D_{\mathbf{k}_1''*}^{r_1''}$ appears only once in the reduced form of the direct product, i.e., $c = 1$ then to calculate the (00)(0)th block of the $q''d''$ columns of U associated with $D_{\mathbf{k}_1''*}^{r_1''}$ we use Eq. (17) for the case $c = 1$. The right-hand side of (17) is then

$$U_{(\eta\eta')(\eta'')} U^*_{(\pi\pi')(\pi'')}. \tag{18}$$

According to Koster[1] one finds specific values of $\pi\pi'\pi''$ and $\eta\eta'\eta''$ such that the left-hand side of Eq. (17) is nonzero, and then holds $\pi\pi'\pi''$ to these specific values and lets $\eta\eta'\eta''$ run over their possible values. For each trio of values of $\eta\eta'\eta''$ one uses Eq. (17) with Eq. (18) to calculate $U_{(\eta\eta')(\eta'')} U^*_{(\pi\pi')(\pi'')}$, and thereby one derives $dd'd''$ equations for the $dd'd''$ elements of the (00)(0)th block of the columns of the matrix U. The general case for an arbitrary c is obtained in a similar manner.[2]

Now we find the other blocks in the first block columns of each section, the $(\theta\theta')(lq'')$th blocks, for which $\mathbf{k}_\theta + \mathbf{k}_{\theta'}' - \mathbf{k}_1'' \doteq 0$.[6]

At the beginning of this section, the space group G was divided for a chosen vector $\mathbf{k}$ into q left cosets. The vector $\mathbf{k}$ defined the aggregate of vectors $\alpha_\theta \mathbf{k} = \mathbf{k}_\theta$, the star of the vector $\mathbf{k}$. If we were to choose any other of the vectors $\mathbf{k}_\theta$ we could again divide the space group G into q left cosets and define the star of the vector $\mathbf{k}_\theta$. The stars of the vectors $\mathbf{k}$ and $\mathbf{k}_\theta$ are identical, the star is defined by giving any one of its vectors; but the division of the space group G is in general different.

We redefine the first star by the vector $\mathbf{k}_\theta$ instead of $\mathbf{k}$, and define the little group of the vector $\mathbf{k}_\theta$. The functions $\psi_\eta^{\mathbf{k}_\theta}$, for the specific θ, form the basis functions of $D_{\mathbf{k}_\theta}^r$, the irreducible representation of the little group of the vector $\mathbf{k}_\theta$. The second star is redefined by $\mathbf{k}_{\theta'}'$ instead of $\mathbf{k}'$, and the star of the vector $\mathbf{k}_1''$ remains defined by the vector $\mathbf{k}_1''$.

The irreducible representation $D_{\mathbf{k}_1''*}^{r_1''}$ again appears c times in the reduced form of the direct product of the irreducible representations in the redefinition. To solve for the $(\theta\theta')(lq'')$th blocks of U is equivalent to finding the (00)(lq'')th blocks of the matrix $\bar{U}$, which reduces the direct product of the irreducible representations in the redefinition (see Appendix B).

$$\bar{U}_{(\eta\eta')(lq''d''+\eta'')} = U_{(\theta d+\eta;\theta' d'+\eta')(lq''d''+\eta'')} \tag{19}$$

holds for all values of the indexes η, η', η'', and l.

[6] If $S_{\mathbf{k}_m''}$ is defined by $\mathbf{k} + \mathbf{k}' - \mathbf{k}_m'' \doteq 0$ and is a star of the first kind, then the (00)(lq'')th blocks are the only nonzero blocks of the first block columns. If $S_{\mathbf{k}_m''}$ is a star of the second kind, then additional nonzero blocks, $(\theta\theta')(lq'')$ for which $\mathbf{k}_\theta + \mathbf{k}_{\theta'}' - \mathbf{k}_m'' \doteq 0$, are in general such that $\theta = \theta'$.

Following steps (1) to (17) we have in our redefinition

$$\begin{aligned}\frac{d''g''}{g} & \sum_{(\hat{\beta}|\mathbf{v}(\hat{\beta}))} D^{r}_{k_\theta}[\hat{\beta} \mid \mathbf{v}(\hat{\beta})]_{\eta\pi} \\ & \times D^{r'}_{k'\theta'}[\hat{\beta} \mid \mathbf{v}(\hat{\beta})]_{\eta'\pi'} D^{r_1''}_{k_1''}[\hat{\beta} \mid \mathbf{v}(\hat{\beta})]_{\eta''\pi''} \\ = {} & \bar{U}_{(\eta\eta')(\eta'')} \bar{U}^*_{(\pi\pi')(\pi'')} + \bar{U}_{(\eta\eta')(d''q''+\eta'')} \bar{U}^*_{(\pi\pi')(d''q''+\pi'')} \\ & + \cdots + \bar{U}_{(\eta\eta')([c-1]d''q''+\eta'')} \bar{U}^*_{(\pi\pi')([c-1]d''q''+\pi'')}. \end{aligned} \tag{20}$$

If the star of the vector $\mathbf{k}_1''$ is of the first kind, then

$$\{\hat{\beta}\} = \{\beta''\},$$

and if of the second kind,

$$\{\hat{\beta}\} = \{\alpha_\theta \beta \alpha_\theta^{-1}\} \cap \{\alpha'_{\theta'} \beta' \alpha'^{-1}_{\theta'}\}. \tag{21}$$

Again, by finding specific values of $\pi\pi'\pi''$ such that the left-hand side of Eq. (20) is nonzero, the number of trios of $\pi\pi'\pi''$ depending on the value of c, we find the elements of the $(00)(lq'')$ blocks of $\bar{U}$, and therefore by Eq. (19) the $(\theta\theta')(lq'')$ blocks of U.

We have shown a method to calculate the blocks in the first block column of each section, i.e., the $(\theta\theta')(lq'')$th blocks, $\theta'' = 0$, for which $\mathbf{k}_\theta + \mathbf{k}'_{\theta'} - \mathbf{k}_1'' \doteq 0$. In both Eqs. (17) and (20) it is necessary to know only the irreducible representations of the factor group $\mathcal{K}/T$, $\mathcal{K}'/T$, and $\mathcal{K}''/T$, where T is the invariant subgroup of translations.

Instead of using Eq. (9) to calculate the remaining nonzero blocks of elements, it is advantageous to review the structure of the basis functions of irreducible representations of space groups. Using properties of this structure we derive an alternative method to calculate the remaining blocks.

For the sake of simplicity we consider only the first of the c sections which we are calculating. The results are, of course, applicable to every section.

The basic functions of the direct product are $\psi^{\mathbf{k}\theta}_\eta \psi^{\mathbf{k}'\theta'}_{\eta'}$. The elements of U give the coefficients of the linear combinations of these functions which form the basis functions of the irreducible representation $D^{r_1''}_{\mathbf{k}_1''*}$; we rewrite Eq. (12) by

$$\psi^{\mathbf{k}_1''\theta''}_{\eta''} = \sum_{\substack{\theta\theta' \\ \eta\eta'}} U_{(\theta d+\eta;\theta' d'+\eta')(\theta'' d''+\eta'')} \psi^{\mathbf{k}\theta}_\eta \psi^{\mathbf{k}'\theta'}_{\eta'}. \tag{22}$$

The sum is not on all possible values of θ and θ', but only on those values which fulfill the condition $\mathbf{k}_\theta + \mathbf{k}'_{\theta'} - \mathbf{k}''_{1\theta''} \doteq 0$. To denote this, we replace the sign of summation $\sum_{\theta\theta'}$ with $\sum^{\mathbf{k}''_{1}\theta''}_{\theta\theta'}$, where $\mathbf{k}'_{1\theta''}$ denotes that θ and θ' take only those values for which the condition is fulfilled.

In particular, for $\theta'' = 0$, we have

$$\psi^{\mathbf{k}_1''}_{\eta''} = \sum^{\mathbf{k}_1''}_{\substack{\theta\theta' \\ \eta\eta'}} U_{(\theta d+\eta;\theta' d'+\eta')(\eta'')} \psi^{\mathbf{k}\theta}_\eta \psi^{\mathbf{k}'\theta'}_{\eta'}. \tag{23}$$

The basis function of $D^{r_1''}_{\mathbf{k}_1''*}$ must fulfill relation (11), that is

$$\psi^{\mathbf{k}_1''\theta''}_{\eta''} = (\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''}))\psi^{\mathbf{k}_1''}_{\eta''}. \tag{24}$$

For a specific θ'' we have, using Eq. (22), that

$$\psi^{\mathbf{k}_1''\theta''}_{\eta''} = \sum^{\mathbf{k}_1''\theta''}_{\substack{\mu\mu' \\ \pi\pi'}} U_{(\mu d+\pi;\mu' d'+\pi')(\theta'' d''+\eta'')} \psi^{\mathbf{k}\mu}_\pi \psi^{\mathbf{k}'\mu'}_{\pi'}, \tag{25}$$

where the values of μ and μ' are constrained by the condition $\mathbf{k}_\mu + \mathbf{k}'_{\mu'} - \mathbf{k}''_{1\theta''} \doteq 0$. On the other hand, substituting (23) into (24), we have

$$\psi^{\mathbf{k}_1''\theta''}_{\eta''} = \sum^{\mathbf{k}_1''}_{\substack{\theta\theta' \\ \eta\eta'}} U_{(\theta d+\eta;\theta' d'+\eta')(\eta'')} (\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''}))\psi^{\mathbf{k}\theta}_\eta \psi^{\mathbf{k}'\theta'}_{\eta'}. \tag{26}$$

For Eqs. (25) and (26) to be consistent, we must have

$$\begin{aligned} & (\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''}))\psi^{\mathbf{k}\theta}_\eta \psi^{\mathbf{k}'\theta}_{\eta'} \\ & = \sum^{\mathbf{k}_1''\theta''}_{\substack{\mu\mu' \\ \pi\pi'}} D^{r}_{k*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})]_{(\mu d+\pi)(\theta d+\eta)} \\ & \quad \times D^{r'}_{k'*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})]_{(\mu' d'+\pi')(\theta' d'+\pi')} \psi^{\mathbf{k}\mu}_\pi \psi^{\mathbf{k}'\mu'}_{\pi'} \\ & = \sum^{\mathbf{k}_1''\theta''}_{\substack{\mu\mu' \\ \pi\pi'}} (D^{r}_{k*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})] \\ & \quad \times D^{r'}_{k'*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})])_{(\mu d+\pi,\mu' d'+\pi')(\theta d+\eta;\theta' d'+\eta')} \psi^{\mathbf{k}\mu}_\pi \psi^{\mathbf{k}'\mu'}_{\pi'}. \end{aligned} \tag{27}$$

Substituting this into (26), we have

$$\begin{aligned} \psi^{\mathbf{k}_1''\theta''}_{\eta''} = & \sum^{\mathbf{k}_1''\theta''}_{\substack{\mu\mu' \\ \pi\pi'}} \sum^{\mathbf{k}_1''}_{\substack{\theta\theta' \\ \eta\eta'}} (D^{r}_{k*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})] \\ & \times D^{r'}_{k'*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})])_{(\mu d+\pi;\mu' d'+\pi')(\theta d+\eta;\theta' d'+\eta')} \\ & \times U_{(\theta d+\eta;\theta' d'+\eta')(\eta'')} \psi^{\mathbf{k}\mu}_\pi \psi^{\mathbf{k}'\mu'}_{\pi'}, \end{aligned}$$

and comparing this to Eq. (25), we see that

$$\begin{aligned} & U_{(\mu d+\pi;\mu' d'+\pi')(\theta'' d''+\eta'')} \\ & = \sum^{\mathbf{k}_1''}_{\substack{\theta\theta' \\ \eta\eta'}} (D^{r}_{k*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})] \\ & \quad \times D^{r'}_{k'*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})])_{(\mu d+\pi;\mu' d'+\pi')(\theta d+\eta)(\theta' d'+\eta')} \\ & \quad \times U_{(\theta d+\eta;\theta' d'+\eta')(\eta'')}. \end{aligned} \tag{28}$$

By this important relation the $(\theta'' d'' + \eta'')$th column of elements is related to the elements of the (η'')th column. Once the elements of the first block column are known, the remaining elements of the section are calculated using Eq. (28).

In the general case when $D^{r_1''}_{k_1''*}$ appears c times in the reduced form of the direct product, we can generalize Eq. (28) as

$$\begin{aligned} & U_{(\mu d+\pi;\mu' d'+\pi')(ld''q''+\theta'' d''+\eta'')} \\ & = \sum^{\mathbf{k}_1''}_{\substack{\theta\theta' \\ \eta\eta'}} (D^{r}_{k*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})] \\ & \quad \times D^{r'}_{k'*}[\alpha''_{\theta''} \mid \mathbf{v}(\alpha''_{\theta''})])_{(\mu d+\pi;\mu' d'+\pi')(\theta d+\pi;\theta' d'+\pi')} \\ & \quad \times U_{(\theta d+\eta;\theta' d'+\eta')(ld''q''+\eta'')}. \end{aligned} \tag{29}$$

Once the first block column of each section has been calculated, the remaining elements of each section are derived, using Eq. (29), and thus the first $cq''d''$ columns of U are calculated.

If the c irreducible representations $D^{r_1''}_{k_1''*}$ are not the only irreducible representations appearing in the reduced form of the direct product $D^{r}_{k*} \times D^{r'}_{k'*}$, then there is a second one, $D^{r_n''}_{k_m''*}$, that appears c_{mn} times. We assume these are the c_{mn} irreducible representations following the c irreducible representations $D^{r_1''}_{k_1''*}$. The dimension of $D^{r_n''}_{k_m''*}$ is $n''_{mn} \doteq q''_m d''_n$. We divide the $C_{mn}q''_m d''_n$ columns of U following the $cq''d''$ columns previously calculated into blocks of dimension $dd' \times d''_n$ and c_{mn} sections. The elements of the blocks of these c_{mn} sections are calculated in the same manner as the blocks of the c sections of the first $cd''q''$ column of U.

If there are additional irreducible representations appearing in the reduced form, we repeat the above procedure until we have exhausted all the irreducible representations that appear in the reduced form of the direct product $D^{r}_{k*} \times D^{r'}_{k'*}$.

Thus we have obtained a method to calculate the elements of the matrix U, the Clebsch–Gordan coefficients: with each irreducible representation $D^{r_n''}_{k_m''*}$, that appears c_{mn} times in the reduced form of the direct product $D^{r}_{k*} \times D^{r'}_{k'*}$, we associate $c_{mn}d''_n q''_m$ columns of U. These columns are divided into blocks of dimension $dd' \times d''_n$ and into c_{mn} sections. The nonzero blocks $(\theta\theta')(lq''_m + \theta'')$ must fulfill the condition $\mathbf{k}_\theta + \mathbf{k}'_{\theta'} - \mathbf{k}''_{m\theta''} \doteq 0$. The nonzero blocks in the first block column of each section are calculated, using Eq. (17) or (20); and finally, the elements of the remaining nonzero blocks are calculated, using Eq. (29). In the calculation of the Clebsch–Gordan coefficients it is not necessary to know the irreducible representations of the space group G. In both Eqs. (17) and (20) only the irreducible representations of the factor groups $\mathcal{K}/T$, $\mathcal{K}'/T$, and $\mathcal{K}''/T$ enter into the calculations.

So far the formalism is quite general, applicable also for nonsymmorphic space groups on the boundary of the Brillouin zone. There are, however, simplifications for symmorphic groups and nonsymmorphic groups in the interior of the Brillouin zone.

Equation (5) for nonsymmorphic groups in the interior of the Brillouin zone may be written as

$$D^{r}_{k^*\theta\mu}[\alpha \mid \mathbf{v}(\alpha) + \mathbf{a}] = e^{-i\mathbf{k}\cdot\mathbf{v}(\beta_{\theta\mu})}e^{-i\mathbf{k}_\theta\cdot\mathbf{a}}\Gamma^{r}_{k}(\beta_{\theta\mu}), \tag{30}$$

where we have used

$$D^{r}_{k}[\beta_{\theta\mu} \mid \mathbf{v}(\beta_{\theta\mu})] = e^{-i\mathbf{k}\cdot\mathbf{v}(\beta_{\theta\mu})}\Gamma^{r}_{k}(\beta_{\theta\mu}).$$

Γ^{r}_{k} is an irreducible representation of the point group associated with the factor group $\mathcal{K}/T$.

In subsequent calculations, for each irreducible representation whose vector $\mathbf{k}$ is in the interior of the Brillouin zone, instead of

$$D^{r}_{k}[\hat\beta \mid \mathbf{v}(\hat\beta)]_{\eta\pi},$$

we write

$$e^{-i\mathbf{k}\cdot\mathbf{v}(\hat\beta)}\Gamma^{r}_{k}(\hat\beta)_{\eta\pi}.$$

In the case where all three vectors $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{k}''$ are in the interior of the Brillouin zone, the left-hand side of Eq. (17) will read

$$\frac{d''q''}{g}\sum_{(\beta\mid\mathbf{v}(\beta))} e^{-i(\mathbf{k}+\mathbf{k}'-\mathbf{k}_m'')\cdot\mathbf{v}(\beta)}\Gamma^{r}_{k}(\hat\beta)_{\eta\pi}\Gamma^{r'}_{k'}(\hat\beta)_{\eta'\pi'}\Gamma^{r_n''}_{k_m''}(\hat\beta)^{*}_{\eta''\pi''}. \tag{31}$$

This can further be simplified by noting that

$$e^{-i(\mathbf{k}+\mathbf{k}'-\mathbf{k}_m'')\cdot\mathbf{v}(\hat\beta)} = \begin{cases} 1 & \text{if } \mathbf{k}+\mathbf{k}'-\mathbf{k}''_m = 0, \\ e^{-i\mathbf{K}\cdot\mathbf{v}(\hat\beta)} & \text{if } \mathbf{k}+\mathbf{k}'-\mathbf{k}''_m = \mathbf{K}, \end{cases}$$

where $\mathbf{K}$ is a reciprocal lattice vector.

Therefore Eq. (31) becomes

$$\frac{d''q''}{g}\sum_{(\hat\beta)} \Gamma^{r}_{k}(\hat\beta)_{\eta\pi}\Gamma^{r'}_{k'}(\hat\beta)_{\eta'\pi'}\Gamma^{r_n''}_{k_m''}(\hat\beta)^{*}_{\eta''\pi''}, \tag{32}$$

if $\mathbf{k}+\mathbf{k}'-\mathbf{k}''_m = 0$, or

$$\frac{d''q''}{g}\sum_{(\beta\mid\mathbf{v}(\beta))} e^{-i\mathbf{K}\cdot\mathbf{v}(\hat\beta)}\Gamma^{r}_{k}(\hat\beta)_{\eta\pi}\Gamma^{r'}_{k'}(\hat\beta)_{\eta'\pi'}\Gamma^{r_n''}_{k_m''}(\hat\beta)^{*}_{\eta''\pi''}, \tag{33}$$

if $\mathbf{k}+\mathbf{k}'-\mathbf{k}''_m = \mathbf{K}$.

In the same manner, the left-hand side of (20) becomes

$$\frac{d''q''}{g}\sum_{(\hat\beta)} \Gamma^{r}_{k_\theta}(\hat\beta)_{\eta\pi}\Gamma^{r'}_{k'_{\theta'}}(\hat\beta)_{\eta'\pi'}\Gamma^{r_n''}_{k_m''}(\hat\beta)^{*}_{\eta''\pi''}, \tag{34}$$

if $\mathbf{k}+\mathbf{k}'-\mathbf{k}''_m = 0$, or

$$\frac{d''q''}{g}\sum_{(\beta\mid\mathbf{v}(\beta))} e^{-i\mathbf{K}\cdot\mathbf{v}(\hat\beta)}\Gamma^{r}_{k_\theta}(\hat\beta)_{\eta\pi}\Gamma^{r'}_{k'_{\theta'}}(\hat\beta)_{\eta'\pi'}\Gamma^{r_n'}_{k_m''}(\hat\beta)^{*}_{\eta''\pi''}, \tag{35}$$

if $\mathbf{k}+\mathbf{k}'-\mathbf{k}''_m = \mathbf{K}$.

For symmorphic groups we know that $\mathbf{v}(\alpha) = 0$ for all α. Equation (30) then reads

$$D^{r}_{k^*\theta\mu}[\alpha \mid \mathbf{a}] = e^{-i\mathbf{k}_\theta\cdot\mathbf{a}}\Gamma^{r}_{k}(\beta_{\theta\mu}),$$

and denoting $(\hat\beta \mid 0)$ simply as $(\hat\beta)$, the left-hand side of (17) reads

$$\frac{d''q''}{g}\sum_{(\hat\beta)} \Gamma^{r}_{k}(\hat\beta)_{\eta\pi}\Gamma^{r'}_{k'}(\hat\beta)_{\eta'\pi'}\Gamma^{r_n''}_{k_m''}(\hat\beta)^{*}_{\eta''\pi''}. \tag{36}$$

The left-hand side of Eq. (20) for symmorphic groups becomes

$$\frac{d''q''}{g}\sum_{(\hat\beta)} \Gamma^{r}_{k_\theta}(\hat\beta)_{\eta\pi}\Gamma^{r'}_{k'_{\theta'}}(\hat\beta)_{\eta'\pi'}\Gamma^{r_n''}_{k_m''}(\hat\beta)^{*}_{\eta''\pi''}. \tag{37}$$

We see, therefore, that for symmorphic groups and nonsymmorphic groups in the interior of the Brillouin zone, the Clebsch–Gordan coefficients of space groups can be obtained using only the irreducible representations of point groups.

In conclusion of this section let us note that the formalism developed here is applicable to both single and double space groups.

III. SPIN-ORBIT COUPLING

As an example of the application of the general theory developed in the previous section, let us treat the spin-orbit coupling in solids. Although the group-theoretical aspects of this problem have been considered before,[7] it is worthwhile to give an approach to it from the point of view of Clebsch–Gordan coefficients of the whole space group, which is the subject of this paper.

The Schrödinger equation of an electron moving in a crystal with a periodic potential V is

$$\left[\frac{\mathbf{P}^2}{2m} + V\right]\psi = \epsilon\psi. \tag{38}$$

The potential V has both the point and translation symmetry of the lattice, and the symmetry of the space group G associated with the lattice.

The Schrödinger equation when spin-orbit interaction is taken into account is[7]

$$\left[\frac{\mathbf{P}^2}{2m} + V + \frac{\hbar}{4m^2c^2}(\nabla V \times \mathbf{P})\cdot\boldsymbol{\sigma}\right]\Phi = E\Phi, \tag{39}$$

where $\boldsymbol{\sigma}$ are the Pauli-spin matrices. With the inclusion of spin, the symmetry group of the Hamiltonian is $\bar{G}$, the double group[7] of G.

Consider now the problem that arises when one has to find the eigenfunctions of Eq. (39) in the lowest order of perturbation theory. Let us denote by $\psi_\eta^{\mathbf{k}\theta}$ the orbital eigenfunctions of Eq. (38), and by $\psi_{\eta'}^{0}$ the spin function of the electron. The superscript $\mathbf{k}' = 0$ of $\psi_{\eta'}^{0}$ is a consequence of the fact that spin functions are invariant under translations. The functions $\psi_\eta^{\mathbf{k}\theta}$ and $\psi_{\eta'}^{0}$ can be looked upon as belonging to bases of irreducible representations of the double space group $\bar{G}$. Assume that the orbital function transforms according to a representation $D_{k^*}^{r}(\bar{G})$. The spin function $\psi_{\eta'}^{0}$ undergoes a transformation according to $D_0^{\frac{1}{2}}(\bar{G})$. In the lowest order of perturbation theory the eigenfunctions Φ of Eq. (39), say $\psi_{\eta''}^{\mathbf{k}\theta''}$, are linear combinations of the products $\psi_\eta^{\mathbf{k}\theta}\psi_{\eta'}^{0}$. The correct eigenfunctions $\psi_{\eta''}^{\mathbf{k}\theta''}$ of Eq. (39) in the lowest order of perturbation theory are those linear combinations of $\psi_\eta^{\mathbf{k}\theta}\psi_{\eta'}^{0}$ that transform according to irreducible representations of $\bar{G}$. The problem of finding the correct $\psi_{\eta''}^{\mathbf{k}\theta''}$ is therefore the reduction of direct products which was worked out in the previous section of this paper.

The functions $\psi_\eta^{\mathbf{k}\theta}\psi_{\eta'}^{0}$ form the basis of the direct product representation $D_{k^*}^{r}(\bar{G}) \times D_0^{\frac{1}{2}}(\bar{G})$. $D_{k^*}^{r}$ being an irreducible representation of the group G is also an irreducible representation of the double group $\bar{G}$. If the direct product is irreducible, the spin-orbit interaction causes no splitting, and eigenfunctions $\psi_{\eta''}^{\mathbf{k}\theta''}$ are equal to the product functions $\psi_\eta^{\mathbf{k}\theta}\psi_{\eta'}^{0}$. If the direct product is reducible, then

$$D_{k^*}^{r}(\bar{G}) \times D_0^{\frac{1}{2}}(\bar{G}) = \sum_n c_n D_{k^*}^{r_n''}(\bar{G}). \tag{40}$$

Since $\mathbf{k}' = 0$, the only star appearing in the reduced form is the star of the vector $\mathbf{k}$.

c_n is calculated from

$$c_n = \frac{1}{2gh}\sum X_{k^*}^{r}(\bar{G})X_0^{\frac{1}{2}}(\bar{G})X_{k^*}^{r_n''}(\bar{G})^*,$$

where gh is the order of the single group G, h being the order of the invariant subgroup of translations T, and g the order of the factor group G/T. Following Zak[3,8], this reduces to

$$c_n = \frac{q}{2g}\sum_{(\delta|\mathbf{v}(\delta))} \xi_k^r[\delta \mid \mathbf{v}(\delta)]X_0^{\frac{1}{2}}(\delta)\xi_k^{r_n''}[\delta \mid \mathbf{v}(\delta)]^*, \tag{41}$$

where $X_0^{\frac{1}{2}}(\delta)$ is the character of $D_0^{\frac{1}{2}}(\delta)$, $\xi_k^r[\delta \mid \mathbf{v}(\delta)]$ is the character of $D_k^r(\delta \mid \mathbf{v}(\delta))$, and $\xi_k^{r_n''}[\delta \mid \mathbf{v}(\delta)]$ is the character of $D_k^{r_n''}[\delta \mid \mathbf{v}(\delta)]$; q is the number of vectors in $S_\mathbf{k}$, the star of the vector $\mathbf{k}$. All the representations in Eqs. (40) and (41) are now representations of double space groups.

It is known[7] that

$$D_0^{\frac{1}{2}}(\bar{\beta}) = -D_0^{\frac{1}{2}}(\beta)$$

and

$$X_0^{\frac{1}{2}}(\bar{\beta}) = -X_0^{\frac{1}{2}}(\beta), \tag{42}$$

where $\bar{\beta}$ is the "barred" element of the double group. Since $D_{k^*}^{r}(G)$ is an irreducible representation of the single group as well, one has

$$D_k^r[\bar{\beta} \mid \mathbf{v}(\beta)] = D_k^r[\beta \mid \mathbf{v}(\beta)]$$

and

$$\xi_k^r[\bar{\beta} \mid \mathbf{v}(\beta)] = \xi_k^r[\beta \mid \mathbf{v}(\beta)]. \tag{43}$$

This being the case, to obtain a nonzero c_n it is necessary that the irreducible representation $D_{k^*}^{r_n''}(\bar{G})$ be such that

$$D_k^{r_n''}[\bar{\beta} \mid \mathbf{v}(\beta)] = -D_k^{r_n''}[\beta \mid \mathbf{v}(\beta)],$$

and

$$\xi_k^{r_n''}[\bar{\beta} \mid \mathbf{v}(\beta)] = -\xi_k^{r_n''}[\beta \mid \mathbf{v}(\beta)]. \tag{44}$$

[7] R. J. Elliott, Phys. Rev. 96, 280 (1954).

[8] J. Zak, Phys. Rev. 151, 464 (1966).

By dividing the sum in Eq. (41) into two, the sum on the elements $[\beta \mid \mathbf{v}(\beta)]$ and the sum on the elements $[\bar{\beta} \mid \mathbf{v}(\beta)]$ [both of which are equal by Eqs. (42), (43), and (44)], Eq. (41) reduces to

$$c_n = \frac{q}{g} \sum_{(\beta \mid \mathbf{v}(\beta))} \xi_k^r[\beta \mid \mathbf{v}(\beta)] X_0^{\frac{1}{2}}(\beta) \xi_k^{r_n''}[\beta \mid \mathbf{v}(\beta)]. \quad (45)$$

Now let us use the results obtained in the preceding sections: The blocks of U are denoted by $(\theta)(lq + \theta'')$; since $D_0^{\frac{1}{2}}(\bar{G})$ is not divided into blocks we have dropped the index θ'. The only nonzero blocks are those for which $\theta = \theta''$. To calculate the $(0)(lq)$th blocks we use Eq. (17), the left-hand side of this equation being in our case

$$\frac{d''q}{2g} \sum_{(\delta \mid \mathbf{v}(\delta))} D_k^r[\delta \mid \mathbf{v}(\delta)]_{\eta\pi} D_0^{\frac{1}{2}}(\delta)_{\eta'\pi'} D_k^{r_n''}[\delta \mid \mathbf{v}(\delta)]_{\eta''\pi''}.$$

Using Eqs. (42), (43), and (44), this becomes

$$\frac{d''q}{g} \sum_{(\beta \mid \mathbf{v}(\beta))} D_k^r[\beta \mid \mathbf{v}(\beta)]_{\eta\pi} D_0^{\frac{1}{2}}(\beta)_{\eta'\pi'} D_k^{r_n''}[\beta \mid \mathbf{v}(\beta)]_{\eta''\pi''}. \quad (46)$$

To calculate the remaining nonzero blocks we use Eqs. (10), (11), (27), and (28):

$$U_{(\theta d+\eta;\eta')(\theta d''+\eta'')} = D_0^{\frac{1}{2}}(\alpha_\theta)_{\eta'1} U_{(\eta'1)(\eta'')} + D_0^{\frac{1}{2}}(\alpha_\theta)_{\eta'2} U_{(\eta 2)(\eta'')}. \quad (47)$$

For an irreducible representation $D_{k*}^{r_n''}$ appearing c_n times in the reduced form of the direct produce, Eq. (47) is used to calculate the $(0)(lq)$th blocks, and the remaining nonzero blocks, the $(\theta)(lq + \theta)$th blocks, are calculated, using Eq. (48). This process is repeated for all irreducible representations appearing in the reduced form, and thus we calculate the matrix U, which reduces the direct product $D_{k*}^r(\bar{G}) \times D_0^{\frac{1}{2}}(\bar{G})$.

As mentioned before, the eigenfunctions Φ of Eq. (39), in the lowest order of perturbation theory, are the functions which form the basis functions of one of the c_n irreducible representations $D_{k*}^{r_n''}$ which appear in the reduced form of the direct product. The qd_n'' columns of U corresponding to this irreducible representation give the linear combinations of the product functions $\psi_\eta^{\mathbf{k}\theta}\psi_{\eta'}^0$ which form these functions, that is,

$$\psi_{\eta''}^{\mathbf{k}\theta''} = \sum_{\eta\eta'} U_{(\theta d+\eta;\eta')(lqd_n''+\theta''d_n''+\eta'')} \psi_\eta^{\mathbf{k}\theta}\psi_{\eta'}^0,$$

where $l = 0, 1, \cdots, c_n - 1$.

The method as given above is quite general, applicable for symmorphic and nonsymmorphic space groups with $\mathbf{k}$ on the boundary of or in the interior of the Brillouin zone. But for symmorphic groups and for nonsymmorphic groups within the interior of the Brillouin zone Eqs. (45) and (46) can be appreciatively simplified, using the results obtained in the previous section.

For symmorphic groups and nonsymmorphic groups with $\mathbf{k}$ in the interior of the Brillouin zone, Eq. (45) reduces to

$$c_n = \frac{q}{g} \sum_{(\beta)} \xi_k^r(\beta) X_0^{\frac{1}{2}}(\beta) \xi_k^{r_n''}(\beta), \quad (48)$$

where $\xi_k^r(\beta)$ is the character of $\Gamma_k^r(\beta)$, the irreducible representation of the point group formed by the set of elements $\{(\beta \mid 0)\}$.

It can be verified by inspection of the character tables of the point groups,[9] that for any point group excepting C_1 and C_i, c_n is either one or zero. For the point groups C_1 and C_i, c_n is either two or zero. Equation (17) will thus become [we have used Eq. (46) and dropped d'']

$$\frac{q}{g} \sum_{(\beta)} \Gamma_k^r(\beta)_{\eta\pi} D_0^{\frac{1}{2}}(\beta)_{\eta'\pi'} \Gamma_k^{r_n''}(\beta)^* = U_{(\eta\eta')(\eta'')} U^*_{(\pi\pi')(\pi'')} \quad (49)$$

for all point groups excepting C_1 and C_i. For these two latter point groups the right-hand side is replaced by

$$U_{(\eta\eta')(\eta'')} U^*_{(\pi\pi')(\pi'')} + U_{(\eta\eta')(d''q+\eta'')} U^*_{(\pi\pi')(d''q+\pi'')}.$$

In either case the sum is on the point group $\{(\beta \mid 0)\}$, the point group associated with the group of the vector $\mathbf{k}$, since the nonprimitive translations take no part in the calculation.

We see then that only for nonsymmorphic space groups with $\mathbf{k}$ on the boundary of the Brillouin zone do the nonprimitive translations enter into the calculations. For symmorphic groups and nonsymmorphic groups in the interior of the Brillouin zone only the properties of the irreducible representations of the point group enter into the calculation.

Finally, let us point our that the matrix U is completely determined by a matrix that reduces the direct product of representations of point groups. Indeed, we use Eq. (48) to calculate the number of times the irreducible representation $D_{k*}^{r_n''}(\bar{G})$ appears in the reduced form of the direct product $D_{k*}^r(\bar{G}) \times D_0^{\frac{1}{2}}(\bar{G})$. Equation (48) is also the equation necessary to calculate the number of times the irreducible representation $\Gamma_{k*}^{r_n''}(\beta)$ appears in the reduced form of the direct product $\Gamma_k^r(\beta) \times D_0^{\frac{1}{2}}(\beta)$. In addition, since g/q is the order of the point group

[9] G. F. Koster, J. O. Dimmock, R. G. Wheeler, and H. Statz, *Properties of the Thirty-two Point Groups* (Massachusetts Institute of Technology Press, Cambridge, Mass., 1963).

$\{(\beta \mid 0)\}$, Eq. (49) is also the equation for the elements of the section associated with the irreducible representation $\Gamma_k^{r_{n}''}(\beta)$ of the unitary matrix V, which reduces the direct product $\Gamma_k^r(\beta) \times D_0^{\frac{1}{2}}(\beta)$. Therefore, each section of V associated with the irreducible representation $\Gamma_k^{r_{n}''}(\beta)$ of the point group is the (0)(0)th block of the section of U association with $D_{k^*}^{r_{n}''}(\bar{G})$. The calculation of the unitary matrix V is a strictly point-group problem.

If $D_0^{\frac{1}{2}}(\beta)$ is an irreducible representation of the point group $\{(\beta \mid 0)\}$, $\Gamma_k^r(\beta) \times D_0^{\frac{1}{2}}(\beta)$ is then the direct product of two irreducible representations of the point group, and the matrix is known.[9] When $D_0^{\frac{1}{2}}(\beta)$ is irreducible,[10] by inspection of the character tables of the thirty-two point groups one finds that $\Gamma_k^r(\beta)$ is a one-dimensional representation, and the matrix V is then a two-dimensional unit matrix. In both cases then the (0)(0)th block of each section of U is known. The remaining nonzero blocks of each section of U are calculated using Eq. (47). Once U has been calculated, the eigenfunctions $\psi_\eta^{k\theta''}$ of Eq. (39) can be calculated.

APPENDIX A

Here we show that when $S_{k_m''}$ is a star of the first or second kind, Eq. (14) reduces to Eq. (15) or (16), respectively.

If λ is an element such that

$$\lambda \subset \{\beta\} \cap \{\beta'\},$$

then from $\mathbf{k} + \mathbf{k}' = \mathbf{k}_m''$ we have by multiplying from the left by λ

$$\lambda(\mathbf{k} + \mathbf{k}') = \lambda\mathbf{k}_m'',$$
$$\lambda\mathbf{k} + \lambda\mathbf{k}' = \lambda\mathbf{k}_m'',$$
$$\mathbf{k} + \mathbf{k}' = \lambda\mathbf{k}_m'',$$
$$\mathbf{k}_m'' = \lambda\mathbf{k}_m''.$$

Therefore

$$\lambda \subset \{\beta''\}$$

and

$$\{\beta\} \cap \{\beta'\} \subset \{\beta''\}. \tag{A1}$$

If λ is an element such that

$$\lambda \subset \{\beta''\},$$

then from $\mathbf{k} + \mathbf{k}' = \mathbf{k}_m''$ again multiplying from the left by λ,

$$\lambda\mathbf{k} + \lambda\mathbf{k}' \doteq \lambda\mathbf{k}_m'',$$
$$\lambda \ + \lambda\mathbf{k}' \doteq \mathbf{k}_m''. \tag{A2}$$

[10] $D_0^{\frac{1}{2}}(\beta)$ is a reducible representation for the point groups C_1, C_i, C_2, C_s, C_{2h}, C_4, S_4, C_{4h}, C_3, C_{3i}, C_6, C_{3h}, and C_{6h}; it is an irreducible representation for the remainder of the point groups.

We have either

$$\lambda\mathbf{k} \not\doteq \mathbf{k},$$
$$\lambda\mathbf{k}' \not\doteq \mathbf{k}',$$

or

$$\lambda\mathbf{k} \doteq \mathbf{k},$$
$$\lambda\mathbf{k}' \doteq \mathbf{k}'. \tag{A3}$$

[We cannot have

$$\lambda\mathbf{k} \doteq \mathbf{k}$$
$$\lambda\mathbf{k}' \not\doteq \mathbf{k}'$$

for inserting this into (A2) we would have

$$\mathbf{k} + \lambda\mathbf{k}' \doteq \mathbf{k}_m'',$$
$$\mathbf{k} + \lambda\mathbf{k}' \doteq \mathbf{k} + \mathbf{k}',$$
$$\lambda\mathbf{k}' \doteq \mathbf{k}',$$

which is a contradiction.]

We will then have, either

$$\mathbf{k}_\theta + \mathbf{k}_{\theta'}' \doteq \mathbf{k}_m''$$

or

$$\mathbf{k} + \mathbf{k}' \doteq \lambda\mathbf{k}_m'',$$

where $\lambda\mathbf{k} = \mathbf{k}_\theta$ and $\lambda\mathbf{k}' = \mathbf{k}_{\theta'}'$.

For $\mathbf{k}''$, which belongs to a star of the first kind, the first possibility in (A3) is forbidden by definition of a star of the first kind; the star $S_{k_m''}$ and consequently the vector $\mathbf{k}_m''$ defined by $\mathbf{k} + \mathbf{k}' = \mathbf{k}_m''$ appears only once in the direct product $S_\mathbf{k} \times S_{\mathbf{k}'}$. Therefore we have for stars of the first kind

$$\mathbf{k} + \mathbf{k}' \doteq \mathbf{k}_m''.$$

This means that

$$\lambda \subset \{\beta\} \cap \{\beta'\}$$

and therefore

$$\{\beta''\} \subset \{\beta\} \cap \{\beta'\}. \tag{A4}$$

From Eqs. (A1) and (A4) we have for stars of the first kind

$$\{\beta''\} = \{\beta\} \cap \{\beta'\}$$

and therefore Eq. (14) reduces to

$$\{\hat{\beta}\} = \{\beta''\}. \tag{A5}$$

For stars of the second kind, (A4) is not applicable, and using only (A2), Eq. (14) becomes

$$\{\hat{\beta}\} = \{\beta\} \cap \{\beta'\}. \tag{A6}$$

APPENDIX B

Here we prove Eq. (19), that is, by redefining the stars $S_\mathbf{k}$ and $S_{\mathbf{k}'}$ by the vectors $\mathbf{k}_\theta$ and $\mathbf{k}_{\theta'}'$, respectively, the $(00)(lq'')$th block of $\bar{U}$, the unitary matrix that reduces the direct $D_{k_\theta^*}^r(G) \times D_{k_{\theta'}'^*}^{r'}(G)$, is equal to the $(\theta\theta')(lq'')$th block of the unitary matrix U, which reduces the direct product $D_{k^*}^r(G) \times D_{k'^*}^r(G)$.

Let V be the unitary transformation that reorders the basis functions of $D^{r}_{k^*}(G)$ to form the basis functions of $D^{r}_{k_\theta{}^*}(G)$. We have

$$V \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \psi^{k\theta}_{\eta} \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} \psi^{k\theta}_{\eta} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \tag{B1}$$

where the new basis is some permutation of the old basis with the functions $\psi^{k\theta}_{\eta}$ in the top position.

Equation (B1) implies that the structure of V is of the following form:

$$\begin{pmatrix} 0 & \begin{matrix} 1 & & \\ & 1 & \\ & & \ddots \\ & & & 1 \end{matrix} & 0 \\ & 0 & \end{pmatrix}, \tag{B2}$$

where the large center zero denotes the θth block column. The irreducible representation $D^{r}_{k^*}(G)$ in the new basis is

$$VD^{r}_{k^*}(G)V^{-1} = D^{r}_{k_\theta{}^*}(G). \tag{B3}$$

In the same manner we define a unitary matrix W, which transforms the irreducible representation $D^{r'}_{k'^*}(G)$ into $D^{r'}_{k'_{\theta'}{}^*}(G)$.

$\bar{U}$ is defined as the unitary matrix which reduces the direct product $D^{r}_{k_\theta{}^*}(G) \times D^{r'}_{k'_{\theta'}{}^*}(G)$, that is,

$$\bar{U}^{-1}[D^{r}_{k_\theta{}^*}(G) \times D^{r}_{k'_{\theta'}{}^*}(G)]\bar{U} = (\text{reduced form}). \tag{B4}$$

Using Eq. (B3) and the comparable relation for $D^{r'}_{k'_{\theta'}{}^*}(G)$, we see that the left-hand side of the (B4) can be written as

$$\begin{aligned}\bar{U}^{-1}[VD^{r}_{k^*}(G)V^{-1} &\times WD^{r'}_{k'^*}(G)W^{-1}]\bar{U} \\ &= \bar{U}^{-1}(V \times W)[D^{r}_{k^*}(G)](V \times W)^{-1}\bar{U}.\end{aligned}$$

Using this and the definition of (B4) becomes

$$\bar{U}^{-1}(V \times W)U \text{ (reduced form) } U^{-1}(V \times W)^{-1}U = (\text{reduced form}).$$

Therefore

$$\bar{U}^{-1}(V \times W)U = I$$

and

$$\bar{U} = (V \times W)U. \tag{B5}$$

We now use (B5) to calculate an element of the $(00)(lq'')$th block of the first $cq''d''$ column of $\bar{U}$:

$$\bar{U}_{(\eta\eta')(lq''d''+\eta'')} = \sum_{\substack{aa'\\bb'}} V_{(\eta)(ad+b)}W_{(\eta')(a'd'+b')} \times U_{(ad+b;a'd'+b')(lq''d''+\eta'')}. \tag{B6}$$

From the structure of V and W we have

$$V_{(\eta)(ad+b)} = \delta(\theta - a)\delta(\eta - b),$$
$$W_{(\eta')(a'd'+b')} = \delta(\theta' - a')\delta(\eta' - b'),$$

and inserting this into (B6), we have

$$\bar{U}_{(\eta\eta')(lq''d''+\eta'')} = \sum_{\substack{aa'\\bb'}} U_{(ad+b;a'd'+b')(lq''d''+\eta'')} \times \delta(\theta - a)\delta(\eta - b)\delta(\theta' - a')\delta(\eta' - b')$$

and

$$\bar{U}_{(\eta\eta')(lq''d''+\eta'')} = U_{(\theta d+\eta;\theta'd'+\eta')(lq''d''+\eta'')}.$$

This last relation is Eq. (19).

Remarks on the Nature of Relativistic Particle Orbits

EDWARD H. KERNER
Sharp Physics Laboratory, University of Delaware, Newark, Delaware

(Received 30 January 1967)

Instantaneous action-at-a-distance relativistic particle dynamics, embraced in Newtonian-like equations of motion $\ddot{\chi}_i = F_i(\chi_i, \dot{\chi}_i)$ with suitable F's, is examined in once-integrated or "kinematical" form $\dot{\chi}_i = f_i(\chi_i, V_i)$ with V_i a set of first integrals transforming as velocities. The Lorentz covariance requirements on f_i are worked out and illustrative examples are given, including a family of many-valued ones. A general meaning for integrals of $\ddot{\chi}_i = F_i$ being in involution is adduced, and general counterparts to some well-known theorems in Hamiltonian dynamics are obtained accordingly. A novel elementary proof of the zero-interaction theorem is appended.

I. INTRODUCTION

The study of the motions of interacting relativistic particles has been built up since Maxwell's time from field theoretical ideas entailing simultaneously the description of the behavior of the mediating fields and their sources. The mathematical complications and physical inconsistencies of this scheme are well known. Effectively, no problem has been satisfactorily solved within it and no clear idea as to how relativistic orbits are made up has come forth from it.

In recent times the matter has been broached from the still older Newtonian tradition of action-at-a-distance. This somewhat atavistic turn had been thought to be ruled out, because of the noncovariance of simultaneity and because of the finiteness of propagation speeds of fields that can possibly couple particles to one another. But already in 1949 Wheeler and Feynman,[1] harking back to Schwartzschild and Fokker, showed how electrodynamics could be placed on a direct (though not instantaneous) interparticle-interaction basis, and Dirac, Thomas, and others[2] considered how to formulate a Hamiltonian theory of directly interacting relativistic particles. In the latter, however, world-line invariance had to be given up with a consequent major doubt of the meaning of the term "particle." Now, Van Dam and Wigner[3] on the one hand have discussed manifestly covariant but integro-differential equations of motion of interacting particles. On the other hand, Kerner[4] in a rudimentary way for electrodynamics, and Currie[5] and Hill,[6] generally and definitively, have shown how an ordinary Newtonian format of equations of motion (second-order equations of motion for each particle in the single coordinate time t) can be entirely compatible with the demands of special relativity; in particular, Hill[6] has shown this explicitly for interactions which in a conventional sense are said to go along light cones. Here the notion of world-line invariance is not only preserved but made central, and the redundancies due to the use of a proper time for each particle are eliminated. The cost, if it is that, is certain difficulties in putting together a Hamilton's principle for the dynamics, as foreshadowed in the zero-interaction theorem of Currie, Jordan, and Sudarshan.[7] But a way around[4] is to use the Lie–Königs theorem, giving up physical positions as canonical while keeping Lorentz transformations canonical; an essentially unique route to the Hamiltonian statement of relativistic particle dynamics in this way has now been established.[8]

In this paper, we further pursue the discussion of relativistic Newtonian equations of motion. We momentarily confine ourselves for simplicity of conception and exposition, and of results, to two particles moving in one dimension. It will be shown afterwards how to generalize the fundamental equations. The main question is, what restrictions on the particle orbits are compelled by translational invariance and Lorentz covariance requirements alone? In a way, the question is more geometrical than dynamical. It asks how to draw systems of world lines representing particle orbits in a space–time diagram such that the same rules applied in another space–time frame, inertial as was the first, give exactly the same systems of world lines. It is only by designation of what family of parameters is used to characterize the system of world lines that they are reckoned as *dynamical* orbits. So that, for instance, if

[1] J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **17**, 157 (1945); **21**, 425 (1949).
[2] P. A. M. Dirac, Rev. Mod. Phys. **21**, 392 (1949); L. H. Thomas Phys. Rev. **85**, 868 (1952); B. Bakamjian and L. H. Thomas, Phys. Rev. **92**, 1300 (1953); L. L. Foldy, Phys. Rev. **122**, 275 (1961).
[3] H. Van Dam and E. P. Wigner, Phys. Rev. **138**, B1576 (1965); **142**, 838 (1966).
[4] E. H. Kerner, J. Math. Phys. **6**, 1218 (1965).
[5] D. G. Currie, Phys. Rev. **142**, 817 (1966).
[6] R. N. Hill, Bull. Am. Phys. Soc. **11**, 96 (1966); J. Math. Phys. **8**, 201 (1967).
[7] D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963).
[8] R. N. Hill and E. H. Kerner, Phys. Rev. Letters **17**, 1156 (1966).

initial values of coordinates and of slopes of world lines (initial, that is, for some inertial observer) are chosen to mark out the system of lines, they can be considered solutions of second-order differential equations. Thus we obtain a 4-parameter family of lines for two particles in one dimension, which solves structurally the same equations of motion for all frames. There is no apparent reason of principle *a priori* why initial-position plus initial-velocity parametrization of orbits should be preferred over a higher-order and richer parametrization; physical experience seems to dictate only that the order of parametrization should be no lower.

First, we indicate how relativistic orbits can be discovered more directly and simply than by constructing a relativistic dynamics and then solving it for the orbits. A family of orbits in the one-dimensional two-particle case is developed, including specific illustrative examples. Second, we show how to extend these considerations to many particles in three dimensions. The main idea is to consider a set of first integrals of equations of motion rather than these equations themselves. In connection with this, several classical results of Hamiltonian theory on the relationships between symmetry transformations and integrals of motion and between first integrals and other integrals of motion, will be generalized for an arbitrary dynamics. Last, in the Appendix, a new and very simple proof of the zero-interaction theorem is sketched, as this theorem is important enough to all considerations of relativistic particle dynamics that it seems worthwhile seeing its workings from all possible sides.

II. DYNAMICS, "KINEMATICS," "GEOMETRY"—ONE DIMENSIONAL

In order that the dynamics $\dot{v}_i = F_i$ be structurally invariant under Lorentz transformations, the necessary and sufficient conditions on the "forces" $F_i(\chi, v_1, v_2)$, $\chi \equiv \chi_1 - \chi_2 =$ relative coordinate of two particles, have been developed by Currie[5] and by Hill[6]:

$$\chi v_2 \frac{\partial F_1}{\partial \chi} - (1 - v_1^2)\frac{\partial F_1}{\partial v_1} - (1 - v_2^2)\frac{\partial F_1}{\partial v_2} - \chi F_2 \frac{\partial F_1}{\partial v_2} = 3 v_1 F_1,$$

$$\chi v_1 \frac{\partial F_2}{\partial \chi} - (1 - v_1^2)\frac{\partial F_2}{\partial v_1} - (1 - v_2^2)\frac{\partial F_2}{\partial v_2} + \chi F_1 \frac{\partial F_2}{\partial v_1} = 3 v_2 F_2. \quad (1)$$

The velocity of light c is taken as unity. These nonlinear partial differential equations are not easy to solve. Suppose, though, that solutions have been constructed, so that the orbits can be formally represented from the once-integrated dynamics as

$$v_i(t) = v_{i0} + F_i(\chi_0, v_{10}, v_{20})t + \tfrac{1}{2}\left[(v_{10} - v_{20})\left(\frac{\partial F_i}{\partial \chi}\right)_0 + \left(F_1 \frac{\partial F_i}{\partial v_1}\right)_0 + \left(F_2 \frac{\partial F_i}{\partial v_2}\right)_0\right]t^2 + \cdots,$$

where $(\ )_0$ means that the enclosed quantity is evaluated using initial data $\chi_0 = \chi(0)$, $v_{i0} = v_i(0)$. This means that the motions can be thought to stem from the first-order equations[9]

$$v_i = \varphi_i(\chi, \alpha_1, \alpha_2), \quad (2)$$

with α_1 and α_2 being integration constants fixed by

$$v_{i0} = \varphi_i(\chi_0, \alpha_1, \alpha_2)$$

in terms of initial values. As only χ_0, v_{i0} enter α_k, we could just as well write

$$v_i = f_i(\chi; \chi_0, v_{10}, v_{20}), \quad (3)$$

with the restriction on f_i that $f_i(\chi_0; \chi_0, v_{10}, v_{20})$ be identically equal to v_{i0}.

The once-integrated equations (2) are what we can convene to call "kinematics" in view of their being of first order, though they carry the full dynamical contents of the original equations of motion $\dot{v}_i = F_i$ through their load of integration constants.

We can easily obtain, in a calculation paralleling that of Eqs. (1), a set of partial differential conditions on the f_i of Eq. (3) that guarantees the form invariance $v_i' = f_i(\chi'; \chi_0', v_{10}', v_{20}')$ of the kinematics under infinitesimal Lorentz transformations. But it is simpler to admit for the motion a conserved energy E and conserved momentum P that make up a four-vector, in view of the admitted time- and space-translational invariance of the dynamics. This involves no real loss of generality (see below) and is a physically interesting way to state first integrals of motion. We therefore replace Eqs. (2) by

$$v_i = \varphi_i(\chi, E, P), \quad (4)$$

where now E, P have simple well-defined transformation properties which the general α_1, α_2 of Eq. (2) lacked.

In the same way we can descend [as with Eq. (2)] down to a "geometrical" construction for the orbits

[9] E. P. Wigner, Progr. Theoret. Phys. (Kyoto) **11**, 437 (1954), remarks on the inappositeness of certain types of "kinematical" equations for relativistic particle motions.

by going to the twice-integrated equations of motion

$$\chi_i = \psi_i(t, \beta_1, \beta_2, \beta_3, \beta_4),$$

or as in Eq. (3),

$$\chi_i = \chi_{i0} + X_i(t, \chi_0, v_{10}, v_{20}).$$

Again, partial differential conditions on X_i may be built out of the requirement of covariance of these equations for the orbits.

The kinematics (4), sitting half-way between the dynamics and geometry, seem to have elements of special simplicity. Let us conduct an infinitesimal Lorentz transformation,[5,6] calling P on orbit 1 simultaneous with Q or orbit 2 in one frame, and P simultaneous with a neighboring point Q' on orbit 2 in a neighboring frame:

$$v_1'(P) = v_1(P) - v(1 - v_1^2(P)),$$

$$\begin{aligned}\chi_1'(P) - \chi_2'(Q') &\\ = \chi_1(P) - \chi_2(Q) &+ vv_2(P)(\chi_1(P) - \chi_2(Q)),\\ E' &= E - vP,\\ P' &= P - vE,\end{aligned}$$

which takes $v_1'(P) = \varphi_1(\chi_1'(P) - \chi_2'(Q'), E', P')$ to

$$v_1 - (1 - v_1^2)v = \varphi_1(\chi, E, P) + \left(\chi v_2 \frac{\partial\varphi_1}{\partial\chi} - P\frac{\partial\varphi_1}{\partial E} - E\frac{\partial\varphi_1}{\partial P}\right)v,$$

(correct to first order in the infinitesimal relative v of primed and unprimed frames) together with a similar equation in φ_2. Now covariance of the "kinematics" $v_i = \varphi_i(\chi, E, P)$ instructs that the φ_i necessarily and sufficiently satisfy

$$\varphi_1^2 - 1 = \chi\varphi_2\frac{\partial\varphi_1}{\partial\chi} - P\frac{\partial\varphi_1}{\partial E} - E\frac{\partial\varphi_1}{\partial P},$$

$$\varphi_2^2 - 1 = \chi\varphi_1\frac{\partial\varphi_2}{\partial\chi} - P\frac{\partial\varphi_2}{\partial E} - E\frac{\partial\varphi_2}{\partial P}. \tag{5}$$

These are the simple "kinematical" counterparts to Eq. (1). When solved we can if we wish regress back to dynamical equations $\dot{v}_i = F_i$ through differentiations and then eliminations of E and P, or else go forward by integration to obtain orbits. By inverting $E = E(\chi, v_1, v_2)$ and $P = P(\chi, v_1, v_2)$, we get these basic integrals of motion directly in terms of primitive variables. The nonlinear system (5) comprises a statement on all possible differentiable laws of two-particle one-dimensional relativistic "kinematics."

Introducing the new variables

$$\xi = \log x, \quad \eta = \tfrac{1}{2}\log\frac{P - E}{P + E}, \quad \zeta = \tfrac{1}{2}\log(P^2 - E^2),$$

they may be written

$$\varphi_1^2 - 1 = \varphi_2\frac{\partial\varphi_1}{\partial\xi} + \frac{\partial\varphi_1}{\partial\eta},$$

$$\varphi_2^2 - 1 = \varphi_1\frac{\partial\varphi_2}{\partial\xi} + \frac{\partial\varphi_2}{\partial\eta}. \tag{6}$$

The variable ζ, being a scalar invariant, has naturally fallen away.

One of the simplest solutions is that for which φ_i are independent of η. Then

$$\frac{d\varphi_1}{\varphi_1(1 - \varphi_1^2)} = \frac{d\varphi_2}{\varphi_2(1 - \varphi_2^2)}$$

and the integration proceeds directly to

$$v_1^2 = \varphi_1^2 = 1 - \frac{4\alpha\chi^2\chi_0^2}{(\chi^2 + \chi_0^2)^2},$$

$$v_2^2 = \varphi_2^2 = 1 + \frac{4\alpha'\chi^2\chi_0^2}{(\chi^2 - \chi_0^2)^2}.$$

χ_0 and α being integration constants and $\alpha + \alpha' = 1$. Differentiation and elimination of α, χ_0 here gives the Lorentz covariant dynamics

$$\dot{v}_1 = \frac{(1 - v_1^2)(v_2 - v_1)}{\chi v_2},$$

$$\dot{v}_2 = \frac{(1 - v_2^2)(v_2 - v_1)}{\chi v_1}.$$

These are further discussed by Hill.[10] This is a rather special case in that the φ_i are independent of η, therefore of P/E, and P, E here are not independent integrals of motion.

Another elementary solution is that for which $\varphi_1 = \varphi_2 = \varphi$ (then $v_1 - v_2 = 0$ and $\chi = \text{const}$) giving the integral

$$X(\eta + \tanh^{-1}\varphi, \xi - \tfrac{1}{2}\log(1 - \varphi^2)) = 0$$

of the one linear partial differential equation for φ (X = arbitrary function).

A third is $\varphi_1 = \pm 1$, and for φ_2,

$$\varphi_2^2 - 1 = \pm\frac{\partial\varphi_2}{\partial\xi} + \frac{\partial\varphi_2}{\partial\eta},$$

or

$$\varphi_2 = \tanh(\theta(\xi - \eta) \mp \xi)$$

with θ arbitrary. In what sense, if any, the particle 1 moving along a light cone can be thought of as the "kinematical" resurrection of a massless field, cannot be seen until the many particle case is examined in detail.

A fourth is that coming from the assumption that both φ_1 and φ_2 are functions of a linear function

[10] R. N. Hill, J. Math. Phys. 8, 1756 (1967).

$\xi + \gamma\eta$ of the independent variables. Then

$$\frac{1-\varphi_2^2}{1-\varphi_1^2} = \frac{d\varphi_2}{d\varphi_1}\frac{\varphi_1+\gamma}{\varphi_2+\gamma}, \tag{7}$$

or

$$\frac{\gamma+\varphi_2}{1-\varphi_2}\left(\frac{1+\varphi_2}{\gamma+\varphi_2}\right)^{\sigma} = \lambda\frac{\gamma+\varphi_1}{1-\varphi_1}\left(\frac{1+\varphi_1}{\gamma+\varphi_1}\right)^{\sigma} \quad \left(\sigma \equiv \frac{\gamma-1}{\gamma+1}\right), \tag{8}$$

λ being a constant of integration. More generally, if we take φ_2 to be a function $R(\varphi_1)$ of φ_1, then Eqs. (6) read:

$$\varphi_1^2 - 1 = R\varphi_{1\xi} + \varphi_{1\eta},$$
$$\frac{R^2-1}{R'} = \varphi_1\varphi_{1\xi} + \varphi_{1\eta},$$

or solving for $\varphi_{1\xi}$, $\varphi_{1\eta}$,

$$\varphi_{1\xi} = \frac{R'(\varphi_1^2-1)-(R^2-1)}{R'(R-\varphi_1)},$$
$$\varphi_{1\eta} = \frac{(R^2-1)R - R'\varphi_1(\varphi_1^2-1)}{R(R-\varphi_1)}.$$

The compatibility condition $\varphi_{1\xi\eta} = \varphi_{1\eta\xi}$ now produces the ordinary equation for R,

$$R'(R-\varphi_1)^2(R^2-1)(\varphi_1^2-1) \times \left\{R'' - (R')^2\frac{\varphi_1+3R}{R^2-1} + R'\frac{R+3\varphi_1}{\varphi_1^2-1}\right\} = 0,$$

whose first integral is just Eq. (7) and whose complete solution is then Eq. (8). That is, the restriction that φ_1 and φ_2 have common linear arguments can be dropped in favor of their being functions of one another.

The integral (8) does not hold for $\gamma = \pm 1$. In these cases, it is replaced by

$$\left(\frac{1+\varphi_2}{1-\varphi_2}\cdot\frac{1-\varphi_1}{1+\varphi_1}\right)\exp\frac{2(\varphi_2-\varphi_1)}{(1\pm\varphi_1)(1\pm\varphi_2)} = \text{const.}$$

To see Eq. (8) more simply, define $(1+\varphi_i)/(\gamma+\varphi_i) \equiv A_i z_i$, so that

$$\frac{A_2^\sigma z_2^\sigma}{\sigma - 1 - \sigma A_2 z_2} = \lambda\frac{A_1 z_1^\sigma}{\sigma - 1 - \sigma A_1 z_1} \equiv B(z_1),$$

or

$$z_2^\sigma + z_2 = (\sigma-1)(\sigma B)^{-\sigma/(\sigma-1)}B,$$

providing the scale factor A_2 is chosen according to $A_2^\sigma = \sigma A_2 B$.

The occurrence of branches may not be without physical interest nor may it be atypical, so that there is a hint that the notion of "orbit" itself may have to be generally understood rather differently and more broadly relativistically from what it is classically. The many valuedness of forces in relativistic dynamics has already been otherwise noticed[11] in connection with particular solutions to Eq. (1). Through simple particular choices of λ, γ, a number of explicit examples of "kinematics" can be worked out in terms of known functions.

III. GENERALIZATION

Now consider the problem of many particles in three dimensions. In order to follow the previous line of development, we need a set of first integrals of the starting dynamics with known transformation properties.

Let the dynamical equations $\ddot{\boldsymbol{\chi}}_i = \mathbf{F}_i$ $(i = 1, \cdots, n)$ be written as $\dot{\boldsymbol{\chi}}_i = \mathbf{v}_i$, $\dot{\mathbf{v}}_i = \mathbf{F}_i$, or collectively, $\dot{y}_i = f_i(y)$ with $y_i, \cdots, y_{3n}$ being the Cartesian components of $\boldsymbol{\chi}_1, \cdots, \boldsymbol{\chi}_n$ and $y_{3n+1}, \cdots, y_{6n}$ being those of $\mathbf{v}_i, \cdots, \mathbf{v}_n$, as in previous work.[8,10] Also writing $y_0 = t$, the Lorentz covariance of the dynamics can be stated by writing an infinitesimal pure Lorentz transformation as

$$y_k \to y_k + \epsilon g_k(y),$$
$$y_0 \to y_0 + \epsilon g_0(y)$$

with appropriate g's. Then,

$$\dot{y}_k \to \dot{y}_k + \epsilon\left(\dot{y}_\beta\frac{\partial g_k}{\partial y_\beta} - \dot{y}_k\dot{y}_\beta\frac{\partial g_0}{\partial y_\beta}\right),$$
$$f_k(y) \to f_k(y) + \epsilon g_\beta\frac{\partial f_k}{\partial y_\beta},$$

(sum on β from 0 to $6n$) correct to first order in ϵ, and calling $\mathcal{L}$ the operator $g_\beta\partial/\partial y_\beta$ and D the operator $f_\beta\partial/\partial y_\beta$ (substantial time differentiation under the motion $\dot{y}_\beta = f_\beta$) the covariance of the equations of motion is guaranteed when the f's satisfy

$$\mathcal{L}f_\beta = Dg_\beta - f_\beta Dg_0. \tag{9}$$

These conditions generalize in a convenient form those like Eq. (1) relating to the special case of two particles in one dimension. Invariance of the equations of motion to time and space translations and spatial rotations can be reckoned manifest when the forces $\mathbf{F}_i$ are taken to be time-independent three-vector functions of relative positions $\boldsymbol{\chi}_k - \boldsymbol{\chi}_l$ and velocities $\mathbf{v}_k$.

The existence of a set of Lorentz scalars that are time-independent constants of motion can be seen readily. If θ is such a scalar, explicitly translationally invariant through a dependence only on components

[11] E. H. Kerner, Phys. Rev. Letters **16**, 667 (1966).

of $\chi_k - \chi_l \equiv \chi_{kl}$ and $\mathbf{v}_k$, we must have

$$\begin{aligned} \mathfrak{L}_i\theta &= 0, \\ R_i\theta &= 0, \\ D\theta &= 0, \end{aligned} \tag{10}$$

where $\mathfrak{L}_i$, R_i are generators for infinitesimal pure Lorentz transformations and spatial rotations about three Cartesian axes. The commutator $(D, \mathfrak{L}_i) = D\mathfrak{L}_i - \mathfrak{L}_i D$ is

$$(D, \mathfrak{L}_i) = (Dg_\beta^{(i)})\frac{\partial}{\partial y_\beta} - (\mathfrak{L}_i f_\beta)\frac{\partial}{\partial y_\beta},$$

or using Eq. (9),

$$(D, \mathfrak{L}_i) = f_\beta(Dg_0^{(i)})\frac{\partial}{\partial y_\beta} = (Dg_0^{(i)})D. \tag{11}$$

In similar fashion,

$$\begin{aligned} (D, R_i) &= 0, & (R_i, R_j) &= \epsilon_{ijk}R_k, \\ (\mathfrak{L}_i, \mathfrak{L}_j) &= -\epsilon_{ijk}R_k, & (R_i, \mathfrak{L}_j) &= \epsilon_{ijk}\mathfrak{L}_k. \end{aligned}$$

Owing to these, the system of seven partial-differential equations (10) for θ is complete in that no new equation can be produced from $(A_m, A_n)\theta = 0$, where the A's are any of $\mathfrak{L}_i$, R_i, D. From the classical theory[12] of such systems, we know there will be $(6n - 3) - 7 = 6n - 10$ functionally independent integrals depending on the $6n - 3$ independent variables $(\chi_k - \chi_l)_i$, $(\mathbf{v}_k)_i$ (only $n - 1$ of the $\chi_k - \chi_l$ being independent); these can be referred to as $z_1, \cdots, z_{6n-3}$.

Travelling a similar path, the existence of four-vector integrals of motion can also be shown, but it will be simpler and more useful for our purposes to look for three-vector integrals $\mathbf{V}$ transforming as velocities. Accordingly, we ask for solutions to

$$\begin{aligned} \mathfrak{L}\mathbf{V} &= \mathbf{V}\mathbf{V} - I, \\ \mathbf{R}\mathbf{V} &= -I \times \mathbf{V}, \\ D\mathbf{V} &= 0, \end{aligned} \tag{12}$$

again presuming $\mathbf{V}$ to be translationally invariant by construction. This set, written out as $A_i V_l = J_{il}(V)$ $(i = 1, \cdots, 7;\ l = 1, 2, 3)$, is in the same sense as before, complete. Under a linear reorganization, $\alpha_{ki}(y)A_i V_l = \alpha_{ki}J_{il}$, or say $B_k V_l = K_{kl}$, it is easy to see that the set remains complete. By a suitable nonsingular choice of α, it can therefore be arranged that B_k is of the form $\partial/\partial z_k +$ terms independent of $\partial/\partial z_k$, so that the system can be written as $\partial V_l/\partial z_k = *$, where $*$ contains only the independent variables, the dependent variables and derivatives (parametric derivatives) of the latter other than $\partial V_l/\partial z_k$. The system now is in normal form, being in fact of König's type.[12] Its completeness signifies its complete integrability and we are therefore assured of the existence of a family of first integrals transforming as velocities.

As will be described below, based on the fact that in a free-particle limit, the particle velocities themselves are constants of motion, we may assume that for the n-particle problem there are at least n velocity-like first integrals that are time- and space-translationally invariant. The statement of "kinematics" can now be framed quite generally as

$$\mathbf{v}_i = \mathbf{\Phi}_i(\chi_{il}, \mathbf{V}_l) \tag{13}$$

[replacing Eq. (4)] with $\mathbf{\Phi}$'s disposed to make the "kinematics" Lorentz covariant; namely, under infinitesimal Lorentz transformation at velocity $\mathbf{v}$ for which

$$\begin{aligned} \mathbf{v}_i &\to \mathbf{v}_i + (\mathbf{v}_i\mathbf{v}_i - I)\cdot\mathbf{v}, \\ \chi_{il} &\to \chi_{il} + \mathbf{v}_l\chi_{il}\cdot\mathbf{v}, \end{aligned}$$

and

$$\mathbf{V}_l \to \mathbf{V}_l + (\mathbf{V}_l\mathbf{V}_l - I)\cdot\mathbf{v},$$

Eqs. (13) are to be maintained up to first order in $\mathbf{v}$, so that

$$\mathbf{\Phi}_i\mathbf{\Phi}_i - I = (\mathbf{V}_l\mathbf{V}_l - I)\cdot\frac{\partial\mathbf{\Phi}_i}{\partial\mathbf{V}_l} + \chi_{il}\mathbf{\Phi}_l\cdot\frac{\partial\mathbf{\Phi}_i}{\partial\chi_{il}}$$

[replacing Eq. (5)]. These are the general conditions on $\mathbf{\Phi}_i$ giving all covariant "kinematics." [To recover the equivalent of the previous one-dimensional result, write

$$\phi_1^2 - 1 = (V_1^2 - 1)\frac{\partial\phi_1}{\partial V_1} + (V_2^2 - 1)\frac{\partial\phi_1}{\partial V_2} + \chi\phi_2\frac{\partial\phi_1}{\partial\chi},$$

$$\phi_2^2 - 1 = (V_1^2 - 1)\frac{\partial\phi_2}{\partial V_1} + (V_2^2 - 1)\frac{\partial\phi_2}{\partial V_2} + \chi\phi_1\frac{\partial\phi_2}{\partial\chi},$$

or with $V_1 = -\tanh w_1$, $V_2 = -\tanh w_2$, and then $\eta' = \frac{1}{2}(w_1 + w_2)$, $\zeta' = \frac{1}{2}(w_1 - w_2)$, $\xi = \log\chi$,

$$\phi_1^2 - 1 = \phi_2\frac{\partial\phi_1}{\partial\xi} + \frac{\partial\phi_1}{\partial\eta'},$$

$$\phi_2^2 - 1 = \phi_1\frac{\partial\phi_2}{\partial\xi} + \frac{\partial\phi_2}{\partial\eta'},$$

as in Eq. (6).]

We now consider some properties of the velocity-like integrals that have been used to formulate "kinematics."

The thought behind their introduction, and a physical reason why they could have been expected to exist in the first place, is that in the important case that a particle dynamics $\ddot{\chi}_i = \mathbf{F}_i$ admits an asymptotic region where, under wide separation of all particles, the forces fall off fast enough to give free

[12] A. R. Forsyth, *Theory of Differential Equations* (Dover Publications, Inc., New York, 1959), Vol. 5.

particle motion, $\ddot{\boldsymbol{\chi}}_i \to 0$, then here the velocities $\mathbf{v}_i$ themselves are first integrals, so that it is natural to see whether there are general integrals $\mathbf{V}_i$ asymptotically agreeing with $\mathbf{v}_i$, and maintaining the transformation character of velocity in the interaction region as well. We can see directly that

$$D\mathbf{V}_i \equiv \left(\mathbf{v}_k \cdot \frac{\partial}{\partial \boldsymbol{\chi}_k} + \mathbf{F}_k \cdot \frac{\partial}{\partial \mathbf{v}_k}\right)\mathbf{V}_i = 0$$

is a Kowalewskian system for which, if $\mathbf{F}_k$ is regular and vanishes at $\boldsymbol{\chi}_{kl} = \infty$, there are solutions in a finite vicinity of the hypersurface $\boldsymbol{\chi}_{kl} = \infty$ in the combined position and velocity space, for which $\mathbf{V}_i \to \mathbf{v}_i$. Then from Eq. (12) we have [also using Eq. (11)]

$$D\{\mathfrak{L}\mathbf{V} - (\mathbf{V}\mathbf{V} - I)\} = \mathfrak{L}D\mathbf{V} + (D\mathbf{g}_0)(D\mathbf{V}) - (D\mathbf{V})\mathbf{V} - \mathbf{V}(D\mathbf{V}),$$

$$D\{\mathbf{R}\mathbf{V} + I \times \mathbf{V}\} = \mathbf{R}D\mathbf{V} + I \times D\mathbf{V},$$

so that as $D\mathbf{V} = 0$, both $\mathfrak{L}\mathbf{V} - (\mathbf{V}\mathbf{V} - I)$ and $\mathbf{R}\mathbf{V} + I \times \mathbf{V}$ will continue to vanish in the neighborhood of the infinite hypersurface as they do upon the hypersurface where $\mathbf{V}_i = \mathbf{v}_i$. This means that the asymptotic integrals $\mathbf{V}_i = \mathbf{v}_i$ do indeed propagate into the interaction region with their transformation character intact.[13]

Next we show that the integrals $\mathbf{V}_i$ which $\to \mathbf{v}_i$ are in involution, and that the remaining integrals of motion of $\dot{\mathbf{v}}_i = \mathbf{F}_i$ can be obtained by quadrature. This amplifies a well-known result in Hamiltonian theory.[14]

Again writing the second-order dynamics $\dot{\mathbf{v}}_i = \mathbf{F}_i$ as the system of first-order equations $\dot{y}_i = f_i(y)$, we always have for the latter the possibility[4] of representing them as Euler–Lagrange equations:

$$\left(\frac{\partial U_m}{\partial y_i} - \frac{\partial U_i}{\partial y_m}\right)\dot{y}_i = -\frac{\partial H}{\partial y_m},$$

or

$$\Gamma_{mi}\dot{y}_i = -\frac{\partial H}{\partial y_m}, \qquad (14)$$

or

$$\dot{y}_i = -\gamma_{im}\frac{\partial H}{\partial y_m}, \qquad (\Gamma\gamma = 1);$$

of the variational principle

$$\delta\int(U_i\dot{y}_i - H)\,dt = 0. \qquad (15)$$

[13] One of course cannot get from this appeal to the *local* existence theorem of Cauchy–Kowalewski a global result on the existence everywhere of integrals $\mathbf{V}_r$ that go over into $\mathbf{v}_r$ asymptotically.

[14] E. T. Whittaker, *Treatise on the Analytical Dynamics of Particles and Rigid Bodies* (Cambridge University Press, New York, 1960).

In principle $U_i\,dy_i$ $(i = 1, \cdots, 6n)$ can be reduced to $P_\alpha(y)\,dQ_\alpha(y)$ $(\alpha = 1, \cdots, 3n)$ (Pfaff's problem), whereupon the dynamics gets cast into Hamiltonian form (Lie–Königs theorem). But the Hamiltonian apparatus is really available without this explicit and in practice often prohibitively difficult reduction. For, if the reduction is made we have the general Poisson bracket

$$(A, B) = \frac{\partial A}{\partial Q_\alpha}\frac{\partial B}{\partial P_\alpha} - \frac{\partial A}{\partial P_\alpha}\frac{\partial B}{\partial Q_\alpha},$$

which is readily translated back to the primitive y variables,

$$\begin{aligned}(A, B) &= \left(\frac{\partial A}{\partial y_k}\frac{\partial B}{\partial y_l} - \frac{\partial A}{\partial y_l}\frac{\partial B}{\partial y_k}\right)\frac{\partial y_l}{\partial P_\alpha}\frac{\partial y_k}{\partial Q_\alpha} \\ &= \frac{\partial A}{\partial y_k}\frac{\partial B}{\partial y_l}\left(\frac{\partial y_k}{\partial Q_\alpha}\frac{\partial y_l}{\partial P_\alpha} - \frac{\partial y_k}{\partial P_\alpha}\frac{\partial y_l}{\partial Q_\alpha}\right) \\ &= \frac{\partial A}{\partial y_k}(y_k, y_l)\frac{\partial B}{\partial y_l}.\end{aligned}$$

But

$$\begin{aligned}\dot{y}_k = (y_k, H) &= \frac{\partial y_k}{\partial Q_\alpha}\frac{\partial H}{\partial P_\alpha} - \frac{\partial y_k}{\partial P_\alpha}\frac{\partial H}{\partial Q_\alpha} \\ &= (y_k, y_l)\frac{\partial H}{\partial y_l},\end{aligned}$$

and a comparison with Eq. (14) shows $(y_k, y_l) = -\gamma_{kl}$, so that in terms of primitive variables alone,[15]

$$(A, B) = -\frac{\partial A}{\partial y_k}\gamma_{kl}\frac{\partial B}{\partial y_l}. \qquad (16)$$

The framework of the variational principle, Eq. (15), is therefore entirely sufficient for "Hamiltonian" discussion, canonical variables as such being dispensable.

Now we can give primitive meaning to integrals of motion being in involution: If our velocity-like integrals $V_m = S_m(y)$ $(m = 1, \cdots, 3n)$ are in involution, it is to say in the sense of Eq. (16) that $(S_m, S_n) = 0$. Now the integrals which asymptotically go over into the particle velocities themselves are asymptotically in involution. By Jacobi's identity we have

$$(H, (S_n, S_m)) + (S_n, (S_m, H)) + (S_m, (H, S_n)) = 0$$

and since $(H, S) = DS = 0$ for all S,

$$D(S_n, S_m) = 0.$$

Thus the involution character of the velocity integrals is preserved at least a finite way into the interaction region.

These integrals, being half the total, and in involution with each other, may be used to get the remaining

[15] The possibility of this type of Poisson bracket was mentioned by J. L. Martin, Proc. Roy. Soc. (London) **A251**, 536 (1959).

integrals by one quadrature and thus can give a complete integration of the dynamics, as follows.

Let us call y_s or y_t the last half of the collection of all y variables, that is, the components of particle velocities, and y_a or y_b the first half, comprising the particle coordinates. Considering $V_m = S_m(y_a, y_s)$ to be inverted to $y_s = \psi_s(y_a, V)$, we have on the one hand using Eq. (16)

$$\dot{y}_s = (y_s, H) = -\gamma_{sl} H_l \quad (H_l \equiv \partial H/\partial y_l);$$

and on the other hand

$$\dot{y}_s = \frac{\partial \psi_s}{\partial y_a} \dot{y}_a = \frac{\partial \psi_s}{\partial y_a} (y_a, H) = -\psi_{sa}\gamma_{al} H_l,$$

where ψ_{sa} is written for $\partial\psi_s/\partial y_a$ so that

$$(\gamma_{sl} - \psi_{sa}\gamma_{al})H_l = 0,$$

or splitting the sum on l,

$$(\gamma_{st} - \psi_{sa}\gamma_{at})H_t = -(\gamma_{sb} - \psi_{sa}\gamma_{ab})H_b. \quad (17)$$

The involution character of the integrals is expressed by

$$\begin{aligned} 0 &= (y_s - \psi_s, y_t - \psi_t) \\ &= (y_s, y_t) + (\psi_s, \psi_t) - (\psi_s, y_t) - (y_s, \psi_t) \\ &= -\gamma_{st} - \psi_{sa}\gamma_{ab}\psi_{tb} + \psi_{sa}\gamma_{at} + \gamma_{sb}\psi_{tb}, \end{aligned} \quad (18)$$

or

$$\gamma_{st} - \psi_{sa}\gamma_{at} = (\gamma_{sb} - \psi_{sa}\gamma_{ab})\psi_{tb}.$$

Multiplying the last by H_t and summing on t yields

$$(\gamma_{st} - \psi_{sa}\gamma_{at})H_t = (\gamma_{sb} - \psi_{sa}\gamma_{ab})\psi_{tb}H_t,$$

whence by Eq. (17),

$$(\gamma_{sb} - \psi_{sa}\gamma_{ab})(H_b + \psi_{tb}H_t) = 0.$$

As will be seen in a moment, $\gamma_{sb} - \psi_{sa}\gamma_{ab}$ is nonsingular, so

$$\begin{aligned} H_b + \psi_{tb}H_t &\equiv \frac{\partial H}{\partial y_b} + \frac{\partial \psi_t}{\partial y_b}\frac{\partial H}{\partial y_t} \\ &= \frac{\partial H(y_a, \psi_s(y_a))}{\partial y_b} = 0. \end{aligned} \quad (19)$$

That is, $H(y_a, y_s)$ written in terms of y_a alone as $H' = H(y_a, \psi_s(y_a))$, by using the basic integrals $y_s = \psi_s(y_a)$, is independent of all y_a.

Corresponding to the partition of the y's into y_a, y_s, we may partition the matrices Γ and γ into

$$\Gamma = \begin{pmatrix} A & K \\ -\tilde{K} & B \end{pmatrix}, \quad \gamma = \begin{pmatrix} a & k \\ -\tilde{k} & b \end{pmatrix}$$

with all of A, B, a, b antisymmetric and the reciprocity $\Gamma\gamma = 1$ between the two is then embraced in

$$Aa - K\tilde{k} = 1, \quad (20a)$$

$$Ak + Kb = 0, \quad (20b)$$

$$\tilde{K}a + B\tilde{k} = 0, \quad (20c)$$

$$Bb - \tilde{K}k = 1, \quad (20d)$$

while the involution statement Eq. (18) is

$$b - \psi k + \tilde{k}\tilde{\psi} + \psi a\tilde{\psi} = 0, \quad (21)$$

where ψ stands for the matrix $(\psi)_{sa} \equiv \psi_{sa} = \partial\psi_s/\partial y_a$. Multiplying Eq. (21) on the left by K and replacing $K\tilde{k}$ by $Aa - 1$ [from Eq. (20a)] and Kb by $-Ak$ [from Eq. (20b)] and grouping terms together, gives

$$-\tilde{\psi} = (A + K\psi)(k - a\tilde{\psi}). \quad (22)$$

Note that ψ is nonsingular; for if $|\psi| = |\partial\psi_s/\partial y_a|$ were to vanish, then the y_s, the components of particle velocities, would not be independent. Hence, $A + K\psi$ and $k - a\tilde{\psi}$ are nonsingular; the transpose of the latter is $\tilde{k} + \psi a$ or $-(-\tilde{k}) + \psi a$, whose nonsingularity was just used in deriving Eq. (19). Multiplying Eq. (22) through by $\tilde{K}(A + K\psi)^{-1}$ on the left and putting in $\tilde{K}a$ as $-B\tilde{k}$ and $\tilde{K}k$ as $Bb - 1$ from Eq. (20c) and (20d), where they occur in the right-hand side, produces

$$-\tilde{K}(A + K\psi)^{-1}\tilde{\psi} = B(b + \tilde{k}\tilde{\psi}) - 1;$$

or using $b + \tilde{k}\tilde{\psi}$ from Eq. (21),

$$-\tilde{K}(A + K\psi)^{-1}\tilde{\psi} = -B\psi(A + K\psi)^{-1}\tilde{\psi} - 1,$$

and finally rearranging the latter,

$$\tilde{\psi}\tilde{K} = \tilde{\psi}B\psi + A + K\psi. \quad (23)$$

In other words, Eq. (23) is a purely formal algebraic consequence of Eq. (21). It has the following significance.

In the variational principle Eq. (15), we may write

$$U_i\, dy_i = U_a\, dy_a + U_s\, dy_s,$$

and may introduce $y_s = \psi_s(y_a)$ giving

$$U_i\, dy_i = \left(U_a + U_s \frac{\partial \psi_s}{\partial y_a}\right) dy_a \equiv W_a\, dy_a.$$

The differential form $W_a\, dy_a$ will be exact only when the integrability conditions

$$\frac{\partial W_a}{\partial y_b} = \frac{\partial W_b}{\partial y_a}$$

are met. Written out in full, these are

$$\Gamma_{ab} + \Gamma_{at}\psi_{tb} + \Gamma_{tb}\psi_{ta} + \Gamma_{st}\psi_{sa}\psi_{tb} = 0,$$

or in the language of the partitioned Γ matrix

$$A + K\psi - \tilde{\psi}\tilde{K} + \tilde{\psi}B\psi = 0,$$

which is just Eq. (23). This is to say: integrability and involution conditions are translations of one another.

Thus we have shown altogether that in consequence of the basic integrals being in involution, the action integrand

$$U_i\, dy_i - H\, dt$$

is an exact differential

$$\left[U_a(y_a, \psi_s(y_a)) + U_s(y_a, \psi_s)\frac{\partial\psi_s}{\partial y_a}\right] dy_a - H(y_a, \psi_s)\, dt$$
$$= d[M(y_a) - H't] \equiv dR(y_a, V, t),$$

when, employing these integrals, y_s is written in terms of y_a throughout.

Now we compute

$$\frac{d}{dt}\frac{\partial R}{\partial V_n} = \dot{y}_a \frac{\partial}{\partial y_a}\frac{\partial M}{\partial V_n} - \frac{\partial H'}{\partial V_n}$$
$$= \dot{y}_a \frac{\partial W_a}{\partial V_n} - \frac{\partial H}{\partial \psi_s}\frac{\partial \psi_s}{\partial V_n},$$

and in a subcomputation

$$\frac{\partial W_a}{\partial V_n} = \frac{\partial U_a}{\partial \psi_s}\frac{\partial \psi_s}{\partial V_n} + \frac{\partial U_t}{\partial \psi_s}\frac{\partial \psi_s}{\partial V_n}\frac{\partial \psi_t}{\partial y_a} + U_s \frac{\partial^2 \psi_s}{\partial V_n \partial y_a}.$$

Since $\dot{y}_a \partial\psi_t/\partial y_a$ is $\dot{y}_t$, the first two terms in $\dot{y}_a \partial W_a/\partial V_n$ amount to

$$\dot{y}_m\left(\frac{\partial V_m}{\partial \psi_s}\right)\left(\frac{\partial \psi_s}{\partial V_n}\right)$$

(m summed from 1 to $6n$), so placing $\dot{y}_m = -\gamma_{ml}H_l$ (l similar to m), we find

$$\frac{d}{dt}\frac{\partial R}{\partial V_n} = \frac{\partial \psi_k}{\partial V_n}\left\{-\gamma_{ml}H_l \frac{\partial U_m}{\partial \psi_s} - \frac{\partial H}{\partial \psi_s}\right\} + \dot{y}_a U_s \frac{\partial^2 \psi_s}{\partial V_n \partial y_a}.$$

Then using

$$\frac{\partial U_m}{\partial \psi_s} = -\Gamma_{sm} + \frac{\partial U_s}{\partial y_m}$$

and $\Gamma_{sm}\gamma_{ml} = \delta_{sl}$, the { } is

$$-\gamma_{ml}\frac{\partial U_s}{\partial y_m} H_l = \dot{y}_m \frac{\partial U_s}{\partial y_m} = \dot{U}_s,$$

while the last term is

$$\dot{y}_a \frac{\partial}{\partial y_a}\left(U_s \frac{\partial \psi_s}{\partial V_n}\right) - \dot{y}_a \frac{\partial U_s}{\partial y_a}\frac{\partial \psi_s}{\partial V_n}$$
$$= \frac{d}{dt}\left[U_s(y_a)\frac{\partial \psi_s}{\partial V_n}\right] - \dot{U}_s \frac{\partial \psi_s}{\partial V_n}.$$

This brings us to

$$\frac{d}{dt}\frac{\partial R}{\partial V_n} = \frac{d}{dt}\left(U_s \frac{\partial \psi_s}{\partial V_n}\right)$$

from which

$$\frac{\partial R}{\partial V_n} - U_s \frac{\partial \psi_s}{\partial V_n} = \text{constant of motion.}$$

These clearly are the remaining independent integrals of motion, all of them linear in t and all stemming out of the single quadrature for R,

$$R = \int W_a(y_a, V)\, dy_a - H't.$$

According to the foregoing development, the position of "kinematics" relative to dynamics and "geometry" can be summarized in the following schematic fashion:

"Kinematics" $\to$ Dynamics with half the integrals known and in involution $\to$ "Geometry" by construction of the remaining integrals.

IV. CONCLUSION: INTEGRALS AND SYMMETRIES

We conclude by showing in the language of the primitive variables y_α what the connections are between integrals and symmetry transformations. This recovers in general form some known results of Hamiltonian theory, which also have been discussed by Hill[10] in another context.

In the action integrand, Eq. (15), for uniformity write

$$(U_i \dot{y}_i - H)\, dt = U_i\, dy_i - H\, dt = U_\alpha\, dy_\alpha$$

with summation on α from 0–$6n$, dt being dy_0, and $-H$ being U_0. An infinitesimal symmetry transformation $y_\gamma \to y_\gamma + \epsilon g_\gamma(y)$ will be a canonical transformation just when

$$U_\alpha(y_\gamma + \epsilon g_\gamma)\, d(y_\alpha + \epsilon g_\alpha) - U_\alpha\, dy_\alpha$$
$$= \text{exact differential}, \quad dT(y),$$

viz., when

$$g_\gamma \frac{\partial U_\alpha}{\partial y_\gamma} + U_\alpha \frac{\partial g_\gamma}{\partial y_\alpha} = \frac{\partial T}{\partial y_\alpha}. \tag{24}$$

If this condition is satisfied, rearrange it through $U_\gamma \partial g_\gamma/\partial y_\alpha = \partial(U_\gamma g_\gamma)/\partial y_\alpha - g_\gamma \partial U_\gamma/\partial y_\alpha$, into

$$\left(\frac{\partial U_\alpha}{\partial y_\gamma} - \frac{\partial U_\gamma}{\partial y_\alpha}\right) g_\gamma \equiv -\Gamma_{\gamma\alpha} g_\gamma = \frac{\partial}{\partial y_\alpha}(T - U_\gamma g_\gamma). \tag{25}$$

Multiplying by $\dot{y}_\alpha$ and summing on α gives

$$-g_\gamma \Gamma_{\gamma\alpha}\dot{y}_\alpha = \dot{y}_\alpha \frac{\partial}{\partial y_\alpha}(T - U_\gamma g_\gamma) = D(T - U_\gamma g_\gamma),$$

but

$$\Gamma_{\gamma\alpha}\dot{y}_\alpha = \Gamma_{\gamma j}\dot{y}_j + \frac{\partial U_\gamma}{\partial y_0} - \frac{\partial U_0}{\partial y_\gamma}$$
$$\left(j = 1, \cdots, 6n; \frac{\partial U_\alpha}{\partial y_0} = 0\right)$$

vanishes because of the Euler–Lagrange equations of motion, Eq. (14). Therefore,

$$T - U_\gamma g_\gamma = \text{integral of motion going with symmetry transformation } y_\gamma \to y_\gamma + \epsilon g_\gamma. \quad (26)$$

This is Noether's theorem[16] for the dynamical scheme embraced in the variational principle Eq. (15). It should be emphasized that the use of the theorem requires two steps: first, the test, [Eq. (24)] whether or not a supposed symmetry transformation is canonically represented, and if it is, the evaluation of T; second, having T, the calculation of the integral [Eq. (26)].

A kind of converse to Noether's theorem is the construction of symmetry transformations from integrals of motion. Thus, suppose $y_i \to y_i + \epsilon g_i$, $t \to t + \epsilon g_0$ is a symmetry transformation of the equations of motion $\Gamma_{ij}\dot{y}_j = \partial U_0/\partial y_i = U_{0i}$. Then writing

$$\Gamma_{ij}(y_l + \epsilon g_l)\frac{d(y_j + \epsilon g_j)}{d(t + \epsilon g_0)} = U_{0i}(y_l + \epsilon g_l)$$

to terms of first order in ϵ gives the so-called variational equations[14]

$$\dot{g}_j\Gamma_{ij} + g_l\frac{\partial\Gamma_{ij}}{\partial y_l}\dot{y}_j = g_l\frac{\partial^2 U_0}{\partial y_i\partial y_l} + \dot{g}_0\Gamma_{ij}\dot{y}_j \quad (27)$$

for g_α. Now if $\varphi(y_i, t)$ is any integral of motion, a solution to the variational equations is

$$\Gamma_{ij}g_j = \frac{\partial\varphi}{\partial y_i}, \quad g_0 = 0. \quad (28)$$

For, using Eq. (28) and the equations of motion, the first term in Eq. (27) may be written as

$$D\frac{\partial\varphi}{\partial y_i} - \dot{\Gamma}_{ij}y_j \quad \left(D = \frac{\partial}{\partial t} + \dot{y}_i\frac{\partial}{\partial y_i}\right),$$

the second as [using the identity $(\partial\Gamma_{ij}/\partial y_l) + (\partial\Gamma_{li}/\partial y_j) + (\partial\Gamma_{jl}/\partial y_i) = 0$]

$$g_l\dot{\Gamma}_{il} + g_l\frac{\partial\Gamma_{lj}}{\partial y_i}\dot{y}_j,$$

and the third as

$$\frac{\partial}{\partial y_i}\cdot\frac{\partial\varphi}{\partial t} - \frac{\partial g_l}{\partial y_i}\cdot\frac{\partial U_0}{\partial y_l},$$

giving for Eq. (27) upon rearrangements and cancellations,

$$\frac{\partial\dot{y}_m}{\partial y_i}\left[\frac{\partial\varphi}{\partial y_m} - \Gamma_{ml}g_l\right] = 0,$$

so that Eq. (27) is indeed satisfied by Eq. (28).

In view of Eq. (25), we can say that for every integral of motion there is a canonically represented symmetry transformation described by Eq. (28).

ACKNOWLEDGMENTS

It is a pleasure to thank Dr. Robert N. Hill for stimulating discussions throughout this work and the National Science Foundation for its partial support of the program of study around it.

APPENDIX: ELEMENTARY PROOF OF ZERO-INTERACTION THEOREM

The theorem states that only free-particle motions are possible for a system of relativistic particles described in a Hamiltonian dynamics giving invariant world lines, when: (a) Lorentz transformations are canonical transformations, and (b) when physical particle positions are taken to be canonical coordinates. The manner in which (b) particularly forces zero interaction has been brought out especially cogently by Hill.[10] It is in fact just by giving up (b) that one can open up the development of a canonical scheme for relativistic interacting-particle dynamics on an instantaneous action-at-a-distance basis.[4,8]

In the present proof, let us begin by supposing an ordinary Hamilton's principle for the motion of two particles in one dimension,

$$\delta\int L(\chi_1, \chi_2, \dot{\chi}_1, \dot{\chi}_2, t)\,dt = 0$$

which is just as good as starting with Hamiltonian equations of motion, and of course incorporates the canonicity of particle coordinates χ_1, χ_2. If an infinitesimal time translation $t \to t + \epsilon$ is to be canonical, we must have $L(t + \epsilon) - L(t)$ be, up to first order in ϵ, an exact derivative,

$$\frac{\partial L}{\partial t} = DE,$$

where E can depend at most on χ_1, χ_2, t. Placing $E = \partial E_1/\partial t$, this integrates to

$$L = DE_1 + E_2,$$

where E_α can depend on $\chi_1, \chi_2, \dot{\chi}_1, \dot{\chi}_2$ but not on t. As L is always indeterminate up to an added exact derivative, it will be no loss to assume L is at the outset $L(\chi_1, \chi_2, \dot{\chi}_1, \dot{\chi}_2)$, i.e., that time-translation invariance is manifest.

In similar fashion, for an infinitesimal space translation $\chi_1 \to \chi_1 + \epsilon$, $\chi_2 \to \chi_2 + \epsilon$ to be canonical, it must be demanded that

$$\frac{\partial L}{\partial \chi_1} + \frac{\partial L}{\partial \chi_2} = DF(\chi_1, \chi_2),$$

[16] E. L. Hill, Rev. Mod. Phys. 23, 253 (1951).

or with

$$F = [(\partial/\partial\chi_1) + (\partial/\partial\chi_2)]F_1(\chi_1, \chi_2)$$

that

$$L = DF_1 + F_2(\chi_1 - \chi_2, \dot{\chi}_1, \dot{\chi}_2),$$

and therefore, with impunity, L can be taken to be dependent only on the relative coordinate $\chi \equiv \chi_1 - \chi_2$ and on $\dot{\chi}_1$ and $\dot{\chi}_2$.

For an infinitesimal pure Lorentz transformation in which, say, a point P on orbit 1 (to the right of orbit 2) in one frame is simultaneous with Q on orbit 2, but in another frame moving at infinitesimal speed v is simultaneous with point Q' on orbit 2, slightly earlier than Q, we have, up to first order in v [as in the discussion preceding Eq. (5)],

$$\begin{aligned} t &\to t - v\chi_1, \\ \chi &\to \chi + \chi\dot{\chi}_2 v, \\ \dot{\chi}_1 &\to \dot{\chi}_1 - (1 - \dot{\chi}_1^2)v, \\ \dot{\chi}_2 &\to \dot{\chi}_2 - (1 - \dot{\chi}_2^2)v - \chi\ddot{\chi}_2 v. \end{aligned}$$

The last term in the last line is needed just because of the shift in world point $Q \to Q'$ on orbit 2 that speaks to the noncovariance of simultaneity. Under this transformation, the action integrand $L(\chi, \dot{\chi}_1, \dot{\chi}_2)\, dt$ goes into

$$L(\chi, \dot{\chi}_1, \dot{\chi}_2)\, dt + v\left\{\chi\dot{\chi}_2 \frac{\partial L}{\partial \chi_1} - (1 - \dot{\chi}_1^2)\frac{\partial L}{\partial \dot{\chi}_1} - (1 - \dot{\chi}_2^2)\frac{\partial L}{\partial \dot{\chi}_2} - \chi\ddot{\chi}_2 \frac{\partial L}{\partial \dot{\chi}_2} - \dot{\chi}_1 L\right\},$$

and in order that the infinitesimal Lorentz transformation be canonical, the { } must be an exact derivative,

$$\chi v_2 \frac{\partial L}{\partial \chi} - (1 - v_1^2)\frac{\partial L}{\partial v_1} - (1 - v_2^2)\frac{\partial L}{\partial v_2} - \chi\dot{v}_2 \frac{\partial L}{\partial v_2} - v_1 L = DG. \quad \text{(A1)}$$

It should be noted that the occurrence of the higher derivative $\dot{v}_2$ in the { } is signifying that under infinitesimal Lorentz transformation the action integrand is rather drastically altered: The variation of the transformed action gives Euler–Lagrange equations of motion of *higher order* than the original action, unless the transformation is indeed canonical. Under finite Lorentz transformation, the transformed action would lead to *infinite order* equations of motion if the transformation was not canonical. The infinitesimal increment v\{ \} is not different, as a calculation shows, from what is obtained at the corresponding stage in the proofs[7,10] that proceed directly in a canonical framework.

As the left side of Eq. (A1) does not involve $\dot{v}_1$, the function G can at most depend on χ, v_2. Writing $G = -\chi G_2(\chi, v_2)$ gives in Eq. (A1) the terms

$$\chi\dot{v}_2 \frac{\partial}{\partial v_2}(L - G_2)$$

that alone involve $\dot{v}_2$. Hence $L - G_2$ must be a function solely of χ, v_1, i.e., $L = G_1(\chi, v_1) + G_2(\chi, v_2)$. After some cancellations, what remains of Eq. (A1) is then

$$\chi v_2 \frac{\partial G_1}{\partial \chi} + \chi v_1 \frac{\partial G_2}{\partial \chi} - v_1 G_1 - v_2 G_2 - (1 - v_1^2)\frac{\partial G_1}{\partial v_1} - (1 - v_2^2)\frac{\partial G_2}{\partial v_2} = 0. \quad \text{(A2)}$$

Now dissect this equation with the operator $\partial^2/\partial v_1 \partial v_2$ to give

$$\frac{\partial^2 G_1}{\partial v_1 \partial \chi} + \frac{\partial^2 G_2}{\partial v_2 \partial \chi} = 0,$$

whereupon, since G_1 involves v_1 but not v_2, and G_2 conversely,

$$\begin{aligned} \frac{\partial^2 G_1}{\partial v_1 \partial \chi} &= \frac{\partial A(\chi)}{\partial \chi}, \\ G_1 &= v_1 A(\chi) + B_1(\chi) + \theta_1(v_1), \\ \frac{\partial^2 G_2}{\partial v_2 \partial \chi} &= -\frac{\partial A(\chi)}{\partial \chi}, \\ G_2 &= -v_2 A(\chi) + B_2(\chi) + \theta_2(v_2), \end{aligned}$$

(A, B, θ_i arbitrary functions) so that L is now of the form

$$L = (v_1 - v_2)A(\chi) + B_1 + B_2 + \theta_1 + \theta_2.$$

Here the first term may be discarded as it is only trivially adding to L, the exact derivative $D \int A(\chi)\, d\chi$.

This brings Eq. (A2) to the form (primes meaning derivatives)

$$\chi v_2 B_1' + \chi v_1 B_2' - (1 - v_1^2)\theta_1' - (1 - v_2^2)\theta_2' - v_1(B_1 + \theta_1) - v_2(B_2 + \theta_2) = 0.$$

Now dissecting with $\partial/\partial\chi$, yields

$$v_1[(\chi B_2')' - B_1'] + v_2[(\chi B_1')' - B_2'] = 0$$

from which, each [ ] having to vanish separately, B_1 and B_2 work out simply as

$$\begin{aligned} B_1 &= \alpha_1 + \tfrac{1}{2}\beta\chi + \gamma/\chi, \\ B_2 &= \alpha_2 + \tfrac{1}{2}\beta\chi - \gamma/\chi \end{aligned}$$

with α, β, γ arbitrary constants, of which α_1, α_2, and γ can plainly be ignored. Thence the remainder

$$(1 - v_1^2)\theta_1' + v_1\theta_1 + (1 - v_2^2)\theta_2' + v_2\theta_2 = 0$$

splits into (λ being a separation constant)

$$\begin{aligned} (1 - v_1^2)\theta_1' + v_1\theta_1 &= \lambda, \\ (1 - v_2^2)\theta_2' + v_2\theta_2 &= -\lambda, \end{aligned}$$

or

$$\theta_1 = \lambda v_1 + \sigma_1(1 - v_1^2)^{\frac{1}{2}},$$
$$\theta_2 = -\lambda v_2 + \sigma_2(1 - v_2^2)^{\frac{1}{2}}$$

with arbitrary constants σ. Ignoring the trivial λ terms, the Lagrangian is boiled down to

$$L = \sigma_1(1 - v_1^2)^{\frac{1}{2}} + \sigma_2(1 - v_2^2)^{\frac{1}{2}} + \beta\chi.$$

None other is allowed that lets Lorentz transformations be canonical.

One small fish slips through the net. The interaction term[17] $\beta\chi$ gives equal and opposite constant forces acting upon the particles, placing them in hyperbolic motion. This corresponds to the physically realizable case of uniformly charged parallel planes in motion along the direction perpendicular to them. With this one exception, we have shown that only a sum of free-particle terms can make up the Lagrangian under the hypotheses (a) and (b). The proof for two particles in three dimensions is readily constructed along completely parallel lines, and need not be set out in detail, starting from a Lagrangian that is manifestly time- and space-translation invariant and also rotationally invariant through dependence on only the three-scalars $\boldsymbol{\chi}^2$, $\mathbf{v}_1^2$, $\mathbf{v}_2^2$, $\mathbf{v}_1 \cdot \boldsymbol{\chi}$, $\mathbf{v}_2 \cdot \boldsymbol{\chi}$, $\mathbf{v}_1 \cdot \mathbf{v}_2$ (as well as $\boldsymbol{\chi} \cdot \mathbf{v}_1 \times \mathbf{v}_2$ if desired). A sequence of straightforward dissections in the same vein as above leads to the strict zero-interaction result without any exceptions.

The extension not only to many particles but to Lagrangians that at the beginning contain higher derivatives than just velocities also proceeds in the same fashion. In the latter case, curved world lines are readily possible, but they do not come about from *coupled* equations of motion for the particles, but only from separated equations of motion, one for each particle, consequent from uncoupled single-particle terms in the Lagrangian.

[17] A slightly more general analysis permits an arbitrary linear combination of x_1, x_2.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Particle Motion and Interaction in Nonlinear Field Theories

G. H. DERRICK AND WAN KAY-KONG
Department of Theoretical Physics, School of Physical Sciences, University of St. Andrews, St. Andrews, Fife, Scotland

(Received 16 June 1967)

A variational method is given for determining the motion and interaction of particles associated with fields governed by nonlinear differential equations. For field equations derived from Lagrangian densities of the type $g^{\iota\kappa}(\partial\theta/\partial x^\iota)(\partial\theta/\partial x^\kappa) + f(\theta)$, one obtains an attractive inverse-square law of force between like particles, provided $f(\theta)$ vanishes more rapidly than (constant) θ^4 for $\theta \to 0$.

1. INTRODUCTION

Since the early years of this century there have been numerous discussions of fields which satisfy nonlinear differential equations. These include Einstein's General Theory of Relativity,[1] modifications of Maxwell's equations,[2–8] and models of elementary particles.[9–19]

One motive behind many of these investigations has been the desire to overcome the divergences which plague conventional field theory, both classical and quantized. Another has been dissatisfaction with the arbitrary nature of present-day quantum field theory, where new field variables have to be added in

[1] A. Einstein, *The Meaning of Relativity* (Methuen and Company, Ltd., London, 1956), 6th ed.
[2] G. Mie, Ann. Physik **37**, 511 (1912); **39**, 1 (1912); **40**, 1 (1913).
[3] M. Born, Nature **132**, 282 (1933); M. Born and L. Infeld, Nature **132**, 1004 L (1933); Proc. Roy. Soc. (London) **A143**, 410 (1934); **A144**, 425 (1934); **147**, 522 (1934); **150**, 141 (1935).
[4] J. Frenkel, Proc. Roy. Soc. (London) **A146**, 930 (1934); E. Feenberg, Phys. Rev. **47**, 148 (1935); M. Born, Proc. Indian Acad. Sci. 3, 8, 85 (1936); M. H. L. Pryce, Proc. Roy. Soc. (London) **A155**, 597 (1936).
[5] B. Hoffman and L. Infeld, Phys. Rev. **51**, 765 (1937).
[6] N. Rosen, Phys. Rev. **55**, 94 (1939); A .C. Menius and N. Rosen, Phys. Rev. **62**, 436 (1942).
[7] H. Schiff, Proc. Roy. Soc. (London) **A269**, 277 (1962).
[8] P. A. M. Dirac, Proc. Roy. Soc. (London) **A209**, 291 (1951); **A212**, 330 (1952); **A223**, 438 (1954); E. Schrödinger, Nature **169**, 538 (1952).
[9] N. Rosen and H. Rosenstock, Phys. Rev. **85**, 257 (1952); H. Rosenstock, Phys. Rev. **93**, 331 (1954).
[10] R. Finkelstein, Phys. Rev. **75**, 1079 (1949); R. Finkelstein, R. Le Levier, and M. Ruderman, Phys. Rev. **83**, 326 (1951); **81**, 655 (A) (1951); N. V. Mitskevich, Zh. Eksp. Teor. Fiz. **29**, 354 (1955) [English transl.: Soviet Phys.—JETP **2**, 197 (1956)]; D. Ivanenko, Nuovo Cimento Suppl. **6**, 349 (1957); D. F. Kurdgelaidze, Zh. Eksp. Teor. Fiz. **32**, 1156 (1957); **34**, 1587 (1958); **35**, 1572 (1958); **36**, 842 (1959); **38**, 462 (1960) [English transl.: Soviet Phys.—JETP **5**, 941 (1957); **7**, 1093 (1958); **8**, 1099 L (1959); **9**, 594 (1959); **11**, 339 (1960)]; V. B. Glasko, F. Leruist, and Ia P. Terletskii, Zh. Eksp. Teor. Fiz. **35**, 452 (1958) [Eng. transl.: Soviet Phys.— JETP **8**, 312 (1959)]; D. D. Betts, H. Schiff, and W. B. Strickfaden, J. Math. Phys. **4**, 334 (1963).
[11] L. I. Schiff, Phys. Rev. **80**, 137A (1950); **83**, 239A (1951); **84**, 1, 10 (1951); **92**, 766 (1953); P. T. Matthews, Phil. Mag. **42**, 221 (1951); A. Salam, Phys. Rev. **82**, 217 (1951); B. J. Malenka, Phys. Rev. **85**, 686 (L) (1952); **86**, 68 (1952); R. O. Fornaguera, Nuovo Cimento **1**, 132 (1955); W. E. Thirring, Z. Naturforsch. **7a**, 63 (1952).

or

$$\theta_1 = \lambda v_1 + \sigma_1(1 - v_1^2)^{\frac{1}{2}},$$
$$\theta_2 = -\lambda v_2 + \sigma_2(1 - v_2^2)^{\frac{1}{2}}$$

with arbitrary constants σ. Ignoring the trivial λ terms, the Lagrangian is boiled down to

$$L = \sigma_1(1 - v_1^2)^{\frac{1}{2}} + \sigma_2(1 - v_2^2)^{\frac{1}{2}} + \beta\chi.$$

None other is allowed that lets Lorentz transformations be canonical.

One small fish slips through the net. The interaction term[17] $\beta\chi$ gives equal and opposite constant forces acting upon the particles, placing them in hyperbolic motion. This corresponds to the physically realizable case of uniformly charged parallel planes in motion along the direction perpendicular to them. With this one exception, we have shown that only a sum of free-particle terms can make up the Lagrangian under the hypotheses (a) and (b). The proof for two particles in three dimensions is readily constructed along completely parallel lines, and need not be set out in detail, starting from a Lagrangian that is manifestly time- and space-translation invariant and also rotationally invariant through dependence on only the three-scalars $\boldsymbol{\chi}^2$, $\mathbf{v}_1^2$, $\mathbf{v}_2^2$, $\mathbf{v}_1 \cdot \boldsymbol{\chi}$, $\mathbf{v}_2 \cdot \boldsymbol{\chi}$, $\mathbf{v}_1 \cdot \mathbf{v}_2$ (as well as $\boldsymbol{\chi} \cdot \mathbf{v}_1 \times \mathbf{v}_2$ if desired). A sequence of straightforward dissections in the same vein as above leads to the strict zero-interaction result without any exceptions.

The extension not only to many particles but to Lagrangians that at the beginning contain higher derivatives than just velocities also proceeds in the same fashion. In the latter case, curved world lines are readily possible, but they do not come about from *coupled* equations of motion for the particles, but only from separated equations of motion, one for each particle, consequent from uncoupled single-particle terms in the Lagrangian.

[17] A slightly more general analysis permits an arbitrary linear combination of x_1, x_2.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Particle Motion and Interaction in Nonlinear Field Theories

G. H. DERRICK AND WAN KAY-KONG
Department of Theoretical Physics, School of Physical Sciences, University of St. Andrews, St. Andrews, Fife, Scotland

(Received 16 June 1967)

A variational method is given for determining the motion and interaction of particles associated with fields governed by nonlinear differential equations. For field equations derived from Lagrangian densities of the type $g^{\iota\kappa}(\partial\theta/\partial x^\iota)(\partial\theta/\partial x^\kappa) + f(\theta)$, one obtains an attractive inverse-square law of force between like particles, provided $f(\theta)$ vanishes more rapidly than (constant) θ^4 for $\theta \to 0$.

1. INTRODUCTION

Since the early years of this century there have been numerous discussions of fields which satisfy nonlinear differential equations. These include Einstein's General Theory of Relativity,[1] modifications of Maxwell's equations,[2–8] and models of elementary particles.[9–19]

One motive behind many of these investigations has been the desire to overcome the divergences which plague conventional field theory, both classical and quantized. Another has been dissatisfaction with the arbitrary nature of present-day quantum field theory, where new field variables have to be added in

[1] A. Einstein, *The Meaning of Relativity* (Methuen and Company, Ltd., London, 1956), 6th ed.

[2] G. Mie, Ann. Physik **37**, 511 (1912); **39**, 1 (1912); **40**, 1 (1913).

[3] M. Born, Nature **132**, 282 (1933); M. Born and L. Infeld, Nature **132**, 1004 L (1933); Proc. Roy. Soc. (London) **A143**, 410 (1934); **A144**, 425 (1934); **147**, 522 (1934); **150**, 141 (1935).

[4] J. Frenkel, Proc. Roy. Soc. (London) **A146**, 930 (1934); E. Feenberg, Phys. Rev. **47**, 148 (1935); M. Born, Proc. Indian Acad. Sci. 3, 8, 85 (1936); M. H. L. Pryce, Proc. Roy. Soc. (London) **A155**, 597 (1936).

[5] B. Hoffman and L. Infeld, Phys. Rev. **51**, 765 (1937).

[6] N. Rosen, Phys. Rev. **55**, 94 (1939); A.C. Menius and N. Rosen, Phys. Rev. **62**, 436 (1942).

[7] H. Schiff, Proc. Roy. Soc. (London) **A269**, 277 (1962).

[8] P. A. M. Dirac, Proc. Roy. Soc. (London) **A209**, 291 (1951); **A212**, 330 (1952); **A223**, 438 (1954); E. Schrödinger, Nature **169**, 538 (1952).

[9] N. Rosen and H. Rosenstock, Phys. Rev. **85**, 257 (1952); H. Rosenstock, Phys. Rev. **93**, 331 (1954).

[10] R. Finkelstein, Phys. Rev. **75**, 1079 (1949); R. Finkelstein, R. Le Levier, and M. Ruderman, Phys. Rev. **83**, 326 (1951); **81**, 655 (A) (1951); N. V. Mitskevich, Zh. Eksp. Teor. Fiz. **29**, 354 (1955) [English transl.: Soviet Phys.—JETP **2**, 197 (1956)]; D. Ivanenko, Nuovo Cimento Suppl. **6**, 349 (1957); D. F. Kurdgelaidze, Zh. Eksp. Teor. Fiz. **32**, 1156 (1957); **34**, 1587 (1958); **35**, 1572 (1958); **36**, 842 (1959); **38**, 462 (1960) [English transl.: Soviet Phys.—JETP **5**, 941 (1957); **7**, 1093 (1958); **8**, 1099 L (1959); **9**, 594 (1959); **11**, 339 (1960)]; V. B. Glasko, F. Leruist, and Ia P. Terletskii, Zh. Eksp. Teor. Fiz. **35**, 452 (1958) [Eng. transl.: Soviet Phys.— JETP **8**, 312 (1959)]; D. D. Betts, H. Schiff, and W. B. Strickfaden, J. Math. Phys. **4**, 334 (1963).

[11] L. I. Schiff, Phys. Rev. **80**, 137A (1950); **83**, 239A (1951); **84**, 1, 10 (1951); **92**, 766 (1953); P. T. Matthews, Phil. Mag. **42**, 221 (1951); A. Salam, Phys. Rev. **82**, 217 (1951); B. J. Malenka, Phys. Rev. **85**, 686 (L) (1952); **86**, 68 (1952); R. O. Fornaguera, Nuovo Cimento **1**, 132 (1955); W. E. Thirring, Z. Naturforsch. **7a**, 63 (1952).

an *ad hoc* manner from time to time to allow for the ever increasing number of "elementary" particles. This has led several authors[12] to search for a single universal field, which, it is hoped, would contain all the diverse particles and fields observed in nature as special cases. The field equations in such a theory must necessarily be nonlinear, for a linear theory would allow superposition of different solutions and thereby exclude interaction between particles. Indeed, in nonlinear theories the force law between particles and their equations of motion in general follow as a consequence of the field equations themselves and do not need to be introduced as separate postulates. One thinks of the remarkable achievement of Einstein, Infeld, and Hoffmann[20] in deriving the equations of motion of gravitating bodies from Einstein's field equations.

Thus, from the point of view of particle interaction, nonlinear field theory is logically simpler than a linear one in that we need fewer postulates. However, if we accept a nonlinear system of field equations, we are faced with the problem of how to represent a "particle." Since we would like to insist that a satisfactory field variable should satisfy the field equations everywhere, we cannot regard particles as point singularities as in linear theory. Instead we might picture the representation of an N-particle system in an ideal classical theory somewhat as follows: there would exist classes of regular solutions of the field equations with N distinct local maxima of energy density (or some such quantity) at any given time. As time develops, these maxima would trace out timelike world lines, about which the energy remains localized. If the particles were unstable, these "tubes" of energy would be expected to dissipate in a time of the order of the half-life. To be satisfactory such a theory would have to answer the following questions:

(i) How do these local maxima move, i.e., what are the equations of motion giving the time development of the positions of the maxima? Solution of this problem would automatically give the interaction between particles.

(ii) Under what conditions do these N maxima remain distinct and conserved in number? One could imagine two maxima merging to form a single particle, or new maxima being formed, or the energy being dissipated by radiation.

In the next section a review is given of results obtained to date on the first problem above, that of particle interaction and equations of motion. The second problem has been approached from a topological standpoint by Finkelstein, Misner, and Rubinstein,[21] who derive certain homotopic conservations laws. For the simplest case, a single particle, some partial results have been obtained on the stability of the solution. Thus Hobart and one of the present authors (G. H. D.)[14] demonstrated that the scalar field equation

$$(1/c^2)(\partial^2\theta/\partial t^2) - \nabla^2\theta = F(\theta, \theta^*) \qquad (1.1)$$

has no stable, time-independent, localized solutions of finite energy for any $F(\theta, \theta^*)$. Rosen[15] has shown, however, that such an equation can have a metastable solution, where the energy dissipates itself with a decay rate small compared with the characteristic de Broglie frequency (energy)/h. This result would seem to correspond to the actual situation in nature, where the observed particles of zero spin, the π and K mesons, are all metastable.

Rosen discussed the equation

$$(1/c^2)(\partial^2\theta/\partial t^2) - \nabla^2\theta = 3g\theta^5 \qquad (1.2)$$

for a real field θ and showed that it possesses the metastable solution

$$\theta = Z(Z^4g + r^2)^{-\frac{1}{2}} \qquad (1.3)$$

with mass

$$\frac{1}{8\pi c^2}\int[(\nabla\theta)^2 - g\theta^6]\,d^3\mathbf{r} = \pi/(16g^{\frac{1}{2}}c^2)$$

and half-life of the order $Z^2g^{\frac{1}{2}}/c$. Here g and Z are constants, the latter being an arbitrary constant of integration.

The object of this present paper is to give new methods for finding the equations of motion of

[12] A. Einstein, Ref. 1, Appendix II; W. Heisenberg, Rev. Mod. Phys. **29**, 269 (1957), and papers reviewed there; H. P. Dürr, W. Heisenberg, H. Mitter, S. Schlieder, and K. Yamazaki, Z. Naturforsch. **14a**, 441 (1959); W. Heisenberg, in *Proceedings of the 1960 Annual International Conference on High-Energy Physics at Rochester* (Interscience Publishers, Inc., New York, 1960), p. 851; F. Gürsey, Nuovo Cimento **3**, 988 (1956); G. A. Sokolik, Dokl. Akad. Nauk SSSR **106**, 429 (1956) [English transl.: Soviet Phys.—Doklady **1**, 57 (1956)]; R. J. Finkelstein, C. Fronsdal, and P. Kaus, Phys. Rev. **103**, 1571 (1956); T. H. R. Skyrme, Proc. Roy. Soc. (London) **A247**, 260 (1958); **A252**, 236 (1959); **A260**, 127 (1961); **A262**, 237 (1961).

[13] S. D. Drell, Phys. Rev. **79**, 220 (A) (1950); S. P. Lloyd, Phys. Rev. **77**, 757 (A) (1950); F. A. Kaempffer, Phys. Rev. **99**, 1614 (L) (1955); I. Bialynicki-Birula, Phys. Rev. **130**, 465 (1963).

[14] R. H. Hobart, Proc. Phys. Soc. (London) **82**, 201 (1963); **85**, 610 (1965); G. H. Derrick, J. Math. Phys. **5**, 1252 (1964).

[15] G. Rosen, J. Math. Phys. **6**, 1269 (1965).

[16] G. Rosen, J. Math. Phys. **7**, 2066, 2071 (1966).

[17] J. K. Perring and T. H. R. Skyrme, Nucl. Phys. **31**, 550 (1962).

[18] A. Seeger and A. Kochendörfer, Z. Physik **127**, 533 (1950); **130**, 321 (1951); U. Enz, Phys. Rev. **131**, 1392 (1963).

[19] R. H. Hobart (private communication).

[20] A. Einstein, L. Infeld, and B. Hoffman, Ann. Math. **39**, 65 (1938).

[21] D. Finkelstein and C. Misner, Ann. Phys. (N.Y.) **6**, 230 (1959); D. Finkelstein, J. Math. Phys. **7**, 1218 (1966); J. Rubinstein, Ph.D. thesis, Yeshiva University (1967).

particles in nonlinear field theories. In particular, we investigate how Rosen particles of type (1.3) move in the presence of one another and hence determine the interaction between such particles. First a nonrelativistic variation method will be given in Sec. 3 and then the relativistic motion discussed in Sec. 4. Before embarking on this program, we first review the work already done on the equations of motion of particles in nonlinear field theories.

2. REVIEW OF EARLIER WORK ON THE MOTION AND INTERACTION OF PARTICLES IN NONLINEAR FIELD THEORIES

The most successful derivation of the equations of motion of particles in a nonlinear field theory is without doubt that of Einstein, Infeld, and Hoffman[20] for the gravitational case. However, since the gravitational field equations are not really self-contained but need an external source term $T^{\iota\kappa}$, this work lies outside the scope of the present paper and is not treated further. We confine ourselves here to a discussion of self-contained field equations whose solutions are required to be free of singularities for all space and time.

The first attempts of Born and Infeld[3] to derive the Lorentz equations for the motion of charged particles in their nonlinear electromagnetic theory proved to be unsatisfactory, since singularities were still present and additional postulates were found to be necessary[4]—the field equations by themselves were insufficient to determine the motion of particles. Modifications to the original Born–Infeld theory were introduced by Hoffmann and Infeld,[5] Rosen,[6] and Schiff[7] in order to remove the remaining singularities from the theory, and these authors were able to show that the field equations were then sufficient to imply the static Coulomb-force law between charged particles. The technique adopted was essentially to integrate the normal component of the stress tensor over a surface enclosing the particle, which then yields the force on that particle. To illustrate the method, consider a scalar equation of type (1.1) which possesses a static, particlelike solution $\theta_0(\mathbf{r})$ localized about the point $\mathbf{r}=0$. Since Eq. (1.1) is of second order in the time, we completely determine the field $\theta(\mathbf{r}, t)$ if we specify the values of θ and $\partial\theta/\partial t$ at $t=0$. Suppose we have the initial conditions

$$\theta(\mathbf{r}, 0) = \theta_0(\mathbf{r}-\mathbf{r}_1) + \theta_0(\mathbf{r}-\mathbf{r}_2),$$
$$(\partial\theta/\partial t)_{t=0} = 0, \tag{2.1}$$

i.e., we hold two particles at positions $\mathbf{r}_1$ and $\mathbf{r}_2$ and "let them go" at time $t=0$. If $T^{\iota\kappa}$ is the stress–energy tensor, the ith component ($i=1,2,3$) of force on a volume V enclosed by a surface S is

$$\frac{1}{c}\frac{d}{dt}\int_V T^{i0}\, d^3\mathbf{r} = -\int_S T^{ik} n^k\, dS, \tag{2.2}$$

where n^k is the normal to the surface. Let S be a surface surrounding the point $\mathbf{r}_1$ but excluding $\mathbf{r}_2$. If $|\mathbf{r}_1-\mathbf{r}_2|$ is much greater than the size of the particles, it will be possible to choose S so that it contains almost all the energy due to the first term of (2.1) and excludes most of the energy from the second. Under these conditions (2.2) will give the force on particle 1 due to particle 2 when the two particles are instantaneously at rest at a separation $|\mathbf{r}_1-\mathbf{r}_2|$.

In a later paper[9] Rosen and Rosenstock applied the same method to establish the form of the static interaction in a nonlinear modification of the Klein–Gordon equation. They showed that an equation of type (1.1) with

$$F(\theta, \theta^*) = -\sigma^2\theta + g(\theta^*\theta)^n\theta$$

in general implies an attractive Yukawa interaction between like particles, provided σ^2, g, and n are all positive.

The above procedure gives only the static interaction, while the initial conditions (2.1) are somewhat contrived. We have merely superimposed two one-particle solutions, whereas one would expect in the actual two-particle solution to find a change in shape of the particles from the free form $\theta_0(\mathbf{r})$. To write the initial conditions simply as a superposition of two free particles may impose strains on the system which might not be present in the correct two-particle solution. Moreover, we have no guarantee that there will not be a pulse of radiation emitted as a consequence of taking the initial conditions (2.1). If this were to be the case, then some of the force given by (2.2) might more properly be interpreted as that of the radiation acting on the particle, and not as a direct action at a distance by the second particle.

Seeger and Kochendörfer,[18] and Perring and Skyrme[17] have found exact two-particle solutions of a particular field equation with only one space dimension x:

$$\partial^2\theta/\partial x^2 - (1/c^2)(\partial^2\theta/\partial t^2) = \sin\theta. \tag{2.3}$$

Their analytical solutions (in closed form!) describe the collision of two unlike or of two like particles, and also the bound state of two unlike particles. For the scattering solutions the energy density has two local maxima, which initially travel towards each other, then collide and partially merge, and finally move apart. When the separation of the maxima is large, they interact with an exponential force law, repulsive for like particles and attractive for unlike.

Equation (2.3) has also been considered by Hobart[19] using a piecewise solution method which takes account of the change of shape of accelerated particles. Unfortunately the techniques of these authors for the simple one-dimensional equation (2.3) are difficult to apply in the three-dimensional case.

3. NONRELATIVISTIC MOTION OF ROSEN PARTICLES

A. Motion of Two Particles

In this section we seek solutions of Eq. (1.2) which describe the motion of two particles whose distance apart is much greater than their sizes and whose velocities are much less than the velocity of light. When the particle separation approaches infinity, the field near each particle is required to reduce to the form of Eq. (1.3) in the rest frame of that particle.

The approximation method adopted is based on the variation principle

$$\delta\int_{t_1}^{t_2} L\,dt = 0, \tag{3.1}$$

with

$$L = \frac{1}{8\pi}\int\left(\frac{\dot\theta^2}{c^2} - (\nabla\theta)^2 + g\theta^6\right)d^3\mathbf{r}$$

for variations $\delta\theta$ which vanish at $t = t_1$, t_2, and at spatial infinity.[22] Let us substitute in Eq. (3.1) the trial wavefunction

$$\theta = \theta_1 + \theta_2, \tag{3.2}$$

where

$$\theta_A = Z_A\{Z_A^4 g + [\mathbf{r} - \mathbf{r}_A(t)]^2\}^{-\frac{1}{4}}, \quad A = 1, 2. \tag{3.3}$$

The function (3.2) simulates the motion of two particles of type (1.3) along the paths $\mathbf{r} = \mathbf{r}_1(t)$ and $\mathbf{r} = \mathbf{r}_2(t)$. The variation "parameters" $\mathbf{r}_1(t)$ and $\mathbf{r}_2(t)$ are to be determined by the variation principle (3.1), with the variations $\delta\mathbf{r}_A(t)$ being required to vanish at $t = t_1$, t_2. One would expect (3.2) to be a good trial function for slowly moving particles whose overlap is small, i.e., $|\dot{\mathbf{r}}_A(t)| \ll c$, $|\mathbf{r}_1(t) - \mathbf{r}_2(t)| \gg Z_1^2 g^{\frac{1}{4}}$ and $Z_2^2 g^{\frac{1}{4}}$. We assume the validity of these inequalities in the evaluation of the Lagrangian L in Eq. (3.1). On inserting Eq. (3.2) in (3.1), we obtain

$$L = L_1 + L_2 - V,$$

where

$$\begin{aligned}
L_A &= \frac{1}{8\pi}\int\left(\frac{\dot\theta_A^2}{c^2} - (\nabla\theta_A)^2 + g\theta_A^6\right)d^3\mathbf{r} \\
&= -mc^2 + \tfrac{1}{2}m\dot{\mathbf{r}}_A^2(t), \\
m &= \tfrac{1}{16}\pi g^{-\frac{1}{4}}c^{-2}, \qquad\qquad (3.4)\\
V &= \int[-2\dot\theta_1\dot\theta_2/c^2 + 2\nabla\theta_1\cdot\nabla\theta_2 - g(6\theta_1^5\theta_2 + 15\theta_1^4\theta_2^2 \\
&\quad + 20\theta_1^3\theta_2^3 + 15\theta_1^2\theta_2^4 + 6\theta_1\theta_2^5)]\,d^3\mathbf{r}.
\end{aligned}$$

[22] The factor $1/8\pi$ has been inserted in Eq. (3.1) in order to exploit the analogy with electromagnetic theory.

The details of the calculation of the effective potential energy V are given in Appendix A, where the following result is obtained:

$$\begin{aligned}V = &-Z_1Z_2/r_{12} \\ &+ \text{order}\,[1/r_{12}^2, (\log r_{12})/r_{12}^3, \dot{\mathbf{r}}_1\cdot\dot{\mathbf{r}}_2/c^2r_{12}],\end{aligned} \tag{3.5}$$

where $r_{12} = |\mathbf{r}_1(t) - \mathbf{r}_2(t)|$. Thus we have to determine $\mathbf{r}_1(t)$ and $\mathbf{r}_2(t)$ from the variation principle

$$\delta\int_{t_1}^{t_2}[-2mc^2 + \tfrac{1}{2}m\dot{\mathbf{r}}_1^2(t) + \tfrac{1}{2}m\dot{\mathbf{r}}_2^2(t) + Z_1Z_2/r_{12}]\,dt = 0. \tag{3.6}$$

This is simply the Lagrangian principle specifying the motion of two particles of mass m which have a mutual potential energy $-Z_1Z_2/r_{12}$. Hence we conclude that the force between the two particles is given by an inverse-square law, and is attractive if Z_1 and Z_2 have the same sign. This is in contrast to the repulsive electrostatic interaction between like charged particles.

B. Motion of n Particles

The extension of the above formalism to a system of n particles is trivial. Instead of (3.2) we take

$$\theta = \sum_{A=1}^{n}\theta_A$$

with each θ_A of form (3.3). On calculating L we find, in addition to the one and two-body terms obtained in the previous section, that there occur 3, 4, $\cdots$ body interactions arising from the integrals

$$\int[\theta_1^4\theta_2\theta_3, \theta_1^3\theta_2^2\theta_3, \theta_1^2\theta_2^2\theta_3^2, \theta_1^2\theta_2^2\theta_3\theta_4, \cdots]\,d^3\mathbf{r}.$$

However, if all interparticle distances are greater than or of the order of some large distance D which is much greater than all particle sizes, then all the many-body interactions are readily seen to be less than or of the order of $1/D^2$ or $(\log D)/D^3$. Hence these terms are negligible compared with the two-body interactions of order $1/D$. Retaining only the leading terms for large particle separations thus yields the Lagrangian

$$L = \sum_A[-mc^2 + \tfrac{1}{2}m\dot{\mathbf{r}}_A^2(t)] + \sum_{A<B}Z_AZ_B/|\mathbf{r}_A(t) - \mathbf{r}_B(t)|.$$

Hence, to the accuracy of our variation approximation, each pair of particles A, B interacts via a mutual potential energy $-Z_AZ_B/|\mathbf{r}_A(t) - \mathbf{r}_B(t)|$.

4. RELATIVISTIC MOTION

Since action-at-a-distance is not a tenable concept in relativistic theory, we first seek the equations of motion of a single particle in a weak external field. The n-particle problem is then treated by evaluating the external field at each particle due to all the other particles. As in electromagnetic theory we encounter retardation effects arising from the finite velocity of light c. Despite this, it turns out to be possible to find an action-at-a-distance Lagrangian correct to order $1/c^2$ just as with charged particles. We again use a variational technique, but we have to modify both the principle (3.1) and the trial wavefunction in order to obtain manifestly covariant equations.

A. Single Particle in a Weak External Field

We want to find a solution of Eq. (1.2) which describes a particle moving along a timelike path $x^\lambda = z^\lambda(p)$ in an external field θ_e.[23] For the moment we leave the choice of the parameter p open, but shall later set it equal to the proper time along the path. But how is one to divide the total field θ into an "external" part θ_e and a "particle" part $\theta - \theta_e$? One possible separation, which is Lorentz invariant, is achieved as follows. Given a solution θ of Eq. (1.2), define the "retarded part" θ_r by

$$\theta_r = \frac{3g}{4\pi}\int d^3\mathbf{r}'\,\frac{\theta^5(\mathbf{r}', t - |\mathbf{r} - \mathbf{r}'|/c)}{|\mathbf{r} - \mathbf{r}'|} \tag{4.1}$$

(assuming θ is localized enough for the integral to converge). Then $g^{\iota\kappa}\theta_{r,\iota\kappa} = 3g\theta^5$, and $\theta_e = \theta - \theta_r$ is a solution of the homogeneous equation $g^{\iota\kappa}\theta_{e,\iota\kappa} = 0$. Conversely, suppose we are given an arbitrary solution θ_e of $g^{\iota\kappa}\theta_{e,\iota\kappa} = 0$. Then any solution θ of the integral equation

$$\theta = \theta_e + \frac{3g}{4\pi}\int d^3\mathbf{r}'\,\frac{\theta^5(\mathbf{r}', t - |\mathbf{r} - \mathbf{r}'|/c)}{|\mathbf{r} - \mathbf{r}'|} \tag{4.2}$$

will also be a solution of Eq. (1.2). Hence, if θ is highly localized about a timelike path $x^\lambda = z^\lambda(p)$, it would seem reasonable to interpret θ_r as the field of the particle and θ_e as the external field. For this reason we adopt trial wavefunctions of the form $\theta = \theta_e + \theta_0$, where θ_e is an arbitrary, given solution of $g^{\iota\kappa}\theta_{e,\iota\kappa} = 0$ and θ_0 is an approximation to the correct solution θ_r, to be optimized variationally. In what follows we assume θ_e is weak enough for its quadratic and higher powers to be discarded.

The trial wavefunction proposed is

$$\theta = \theta_e + \theta_0,$$

where

$$\theta_0 = Z(Z^4 g + R^2)^{-\frac{1}{4}} \tag{4.3}$$

[23] Notation: $x^0 = ct$, $(x^1, x^2, x^3) = \mathbf{r}$ or $\mathbf{x}$. Greek indices take the values 0, 1, 2, 3, Latin indices the values 1, 2, 3. $g_{\iota\kappa}$ is the metric tensor with nonzero components $g_{00} = -g_{11} = -g_{22} = -g_{33} = 1$. A comma denotes differentiation, thus $\theta_{,\kappa} = \partial\theta/\partial x^\kappa$, and a dot over a symbol signifies its time derivative. $\epsilon_{\iota\kappa\lambda\mu}$ is the Levi-Civita permutation symbol, equal to ± 1 accordingly as $\iota\kappa\lambda\mu$ is an even or odd permutation of 0 1 2 3, and zero otherwise. The summation convention for repeated indices is adopted.

and R is a retarded distance defined below by analogy with a similar quantity which arises in the theory of Lienard–Weichert potentials in electromagnetic theory.[24] We recall that if a point charge e moves along the timelike path $x^\lambda = z^\lambda(p)$, then the electromagnetic potential is $A^\lambda(x) = eu^\lambda(p_r)/R$ with retarded quantities p_r, R, $u^\lambda(p_r)$ (functions of x^λ) defined by

$$\begin{aligned} x^0 - z^0(p_r) &= |\mathbf{x} - \mathbf{z}(p_r)|, \\ u^\lambda(p) &= dz^\lambda/(g_{\iota\kappa}\,dz^\iota\,dz^\kappa)^{\frac{1}{2}}, \\ R &= [x^\lambda - z^\lambda(p_r)]u_\lambda(p_r). \end{aligned} \tag{4.4}$$

The physical significance of these equations is that a light signal emitted by the particle at the parameter value $p = p_r$ will reach the point $\mathbf{x}$ at time x^0/c. If we choose an inertial frame such that the particle is instantaneously at rest at $p = p_r$, then R is the instantaneous (three-dimensional) distance from the particle to the point $\mathbf{x}$, measured in this frame. In this instantaneous frame our trial function (4.3), with R given by (4.4), is of the free-particle form (1.3), immersed in the given external field θ_e. The path $z^\lambda(p)$ is to be varied to optimize some variation principle.

What principle should we adopt in place of (3.1)? The general form of variation principle which leads to Eq. (1.2) is

$$\delta\left\{\frac{1}{8\pi c}\int_\Sigma [g^{\iota\kappa}\theta_{,\iota}\theta_{,\kappa} + g\theta^6]\,d^4x\right\} = \frac{1}{4\pi c}\int_S \delta\theta\theta_{,\sigma}\,dS^\sigma. \tag{4.5}$$

Here Σ is an arbitrary 4-volume bounded by the closed three-dimensional surface S, and the surface element dS^σ is defined by

$$dS^\sigma = (1/3!)g^{\sigma\iota}\epsilon_{\iota\kappa\lambda\mu}[\partial(x^\kappa, x^\lambda, x^\mu)/\partial(\alpha, \beta, \gamma)]\,d\alpha\,d\beta\,d\gamma,$$

where α, β, γ is any set of three coordinates which parameterizes the surface. The obvious choice for Σ is the region $p_1 < p_r < p_2$ or, equivalently,

$$z^0(p_1) + |\mathbf{x} - \mathbf{z}(p_1)| < x^0 < z^0(p_2) + |\mathbf{x} - \mathbf{z}(p_2)| \tag{4.6}$$

in place of the domain $t_1 < t < t_2$ of Eq. (3.1), i.e., the region between the forward light cones at two arbitrary points p_1 and p_2 lying on the path $x^\lambda = z^\lambda(p)$.

To summarize our variational procedure: We take the trial function (4.3) with R given by (4.4) and θ_e an arbitrary, given solution of $g^{\iota\kappa}\theta_{e,\iota\kappa} = 0$. This is then substituted in the principle (4.5), taking Σ as the region (4.6) and the path $x^\lambda = z^\lambda(p)$ optimized variationally keeping the endpoints $z^\lambda(p_1)$ and $z^\lambda(p_2)$ fixed. A set of differential equations for $z^\lambda(p)$ results,

[24] L. Landau and E. Lifshitz, *The Classical Theory of Fields* (Addison-Wesley Publishing Company, Inc., Reading, Mass., 1951), Chap. 8.

which should describe the motion of a particle in an external field θ_e, provided θ_e is weak enough for (4.3) to be a good approximation.

On substituting (4.3) into (4.5) and retaining only first-order terms in θ_e, one obtains

$$\delta\left\{\frac{1}{8\pi c}\int[g^{\iota\kappa}\theta_{0,\iota}\theta_{0,\kappa}+g\theta_0^6+6g\theta_0^5\theta_e+2g^{\iota\kappa}\theta_{0,\iota}\theta_{e,\kappa}]\,d^4x\right\}$$
$$=\frac{1}{4\pi c}\left[\int\delta\theta_0\theta_{0,\sigma}\,dS^\sigma+\delta\int\theta_0\theta_{e,\sigma}\,dS^\sigma\right].\quad(4.7)$$

In Appendix B $\int\delta\theta_0\theta_{0,\sigma}\,dS^\sigma$ is shown to vanish for variations $\delta z^\lambda(p)$ which leave the endpoints $z^\lambda(p_1)$ and $z^\lambda(p_2)$ fixed, even though $\delta\theta_0$ itself does not vanish on the surface S. The second term on the right-hand side of Eq. (4.7) cancels the last term on the left-hand side:

$$\int\theta_0\theta_{e,\sigma}\,dS^\sigma=\int g^{\iota\kappa}(\theta_{0,\iota}\theta_{e,\kappa}+\theta_0\theta_{e,\iota\kappa})\,d^4x$$
$$=\int g^{\iota\kappa}\theta_{0,\iota}\theta_{e,\kappa}\,d^4x.$$

Hence the variation principle reduces to the form

$$\delta\left[\frac{1}{8\pi c}\int(g^{\iota\kappa}\theta_{0,\iota}\theta_{0,\kappa}+g\theta_0^6+6g\theta_0^5\theta_e)\,d^4x\right]=0.\quad(4.8)$$

The integrations involved in Eq. (4.8) are performed in Appendix C. The result is

$$\delta\int_1^2[-mc^2+Z\theta_e(z^\lambda)]\,ds=0\quad(4.9)$$

with

$$ds=(g_{\iota\kappa}\,dz^\iota\,dz^\kappa)^{\frac{1}{2}}=[g_{\iota\kappa}(dz^\iota/dp)(dz^\kappa/dp)]^{\frac{1}{2}}\,dp.$$

Carrying out the variation in the standard way then yields the equations of motion (taking as parameter p the proper time s)[25]:

$$m\,d^2z_\lambda(s)/ds^2=(Z/c^2)(u_\lambda\theta_{e,\kappa}-u_\kappa\theta_{e,\lambda})u^\kappa,\quad(4.10)$$

where $u^\lambda=dz^\lambda(s)/ds$, and it is understood that θ_e and its derivatives are to be evaluated at the point $x^\lambda=z^\lambda(s)$.

B. Motion of n Particles

It is reassuring to observe that, for $R^2\gg Z^4g$, Eq. (4.3) reduces to $\theta=\theta_e+Z/R$, which is an *exact* solution of $g^{\iota\kappa}\theta_{,\iota\kappa}=0$ for quite arbitrary timelike paths $x^\lambda=z^\lambda(p)$. Since by assumption θ_e is weak enough for its quadratic and higher powers to be neglected, (4.3) is asymptotically a solution of Eq. (1.2) for large R, i.e., at large distances from the particle. This means that a second particle located at position $\mathbf{x}$ at time x^0/c will experience an external field Z/R due to the particle moving along $x^\lambda=z^\lambda(p)$, provided $R^2\gg Z^4g$. Of course, at closer distances there may be multipole terms of higher order in $1/R$, but we are interested only in the leading term at large particle separations.

[25] For consistency we omit the second-order term $\theta_e\,d^2z_\lambda(s)\,ds^2$ from Eq. (4.10).

Suppose now we have n particles, moving along paths $x^\lambda=z_A^\lambda(p_A)$, $A=1,2,\cdots,n$. If the concept of an n-particle solution is to have any validity at all, the energy density must be large near these paths and small elsewhere. Under these conditions we can write the source term in Eq. (4.1) as $(3g/4\pi)\theta^5=\sum_A\sigma_A$, where σ_A is large in the vicinity of path number A but small everywhere else. Then Eq. (4.2) becomes $\theta=\theta_e+\sum_A\theta_A$ with

$$\theta_A=\int d^3\mathbf{r}'\,\frac{\sigma_A(\mathbf{r}',t-|\mathbf{r}-\mathbf{r}'|/c)}{|\mathbf{r}-\mathbf{r}'|}.$$

If θ_A is interpreted as the field due to the Ath particle, then

$$\theta_{eA}=\theta-\theta_A=\theta_e+\sum_{B\neq A}\theta_B$$

must be regarded as the external field of this particle. θ_e represents incoming radiation from infinity, and would be omitted if the only fields present are those originating on particles. Let us now go to the limit of very highly concentrated particles separated by distances much greater than their sizes. Then, neglecting higher multipole terms, we can replace θ_B by its asymptotic form $Z_B/R_B(x^\lambda)$, where for each space–time point x^λ we define the retarded distance $R_B(x^\lambda)$ from particle B by equations analogous to (4.4). The external field at particle A is then

$$\theta_{eA}(z_A^\lambda)=\theta_e(z_A^\lambda)+\sum_{B\neq A}Z_B/R_B(z_A^\lambda).\quad(4.11)$$

Hence (4.11) is the appropriate expression to substitute for θ_e in (4.10) in order to determine the path of particle A. We thus obtain a set of n coupled equations in which the forces on each particle depend on the positions and velocities of the other particles at appropriate retarded times. This situation is analogous to that pertaining with a set of interacting charged particles,[24] except that the present approximation fails to yield any self-force terms.

C. Lagrangian to Order $1/c^2$

As in the electromagnetic case there exists a Lagrangian which reproduces the equations of motion for n particles correct to order $1/c^2$. Let us take the time t as parameter instead of the proper time s, and denote the path of particle number A by $\mathbf{r}=\mathbf{r}_A(t)$, so that the correspondence with the notation of the previous section is $z_A^\lambda(p_A)=[ct,\mathbf{r}_A(t)]$. According to Eq. (4.9), the Lagrangian for the motion of particle A in the external field θ_{eA}, Eq. (4.11), is

$$L_A = (-mc^2 + Z_A\theta_{eA})(1 - \mathbf{v}_A^2/c^2)^{\frac{1}{2}}, \quad (4.12)$$

writing $\mathbf{v}_A = d\mathbf{r}_A(t)/dt$. Let us assume that there is no incoming radiation from infinity, so that the first term on the right of Eq. (4.11) is absent. To proceed further we must expand θ_{eA} in powers of $1/c$. The expansion for the retarded distance $R_B(x^\lambda)$ is

$$R_B(x^\lambda) = \rho_B + [(\boldsymbol{\rho}_B \cdot \mathbf{v}_B)^2/\rho_B + \rho_B(\boldsymbol{\rho}_B \cdot \dot{\mathbf{v}}_B)]/2c^2 + \text{order}\,(1/c^3), \quad (4.13)$$

where

$$\boldsymbol{\rho}_B = \mathbf{r} - \mathbf{r}_B(t), \quad \rho_B = |\boldsymbol{\rho}_B|.$$

Inserting this expression into Eqs. (4.11), (4.12) yields the following expansion for L_A:

$$L_A = T_A - \sum_{B \neq A} [V_{AB} + (d/dt)(Z_A Z_B \mathbf{n}_{AB} \cdot \mathbf{v}_B)], \quad (4.14)$$

where

$$\begin{aligned} T_A &= -mc^2(1 - \mathbf{v}_A^2/c^2)^{\frac{1}{2}} \\ &= -mc^2 + \tfrac{1}{2}m\mathbf{v}_A^2 + m(\mathbf{v}_A^2)^2/8c^2 + \text{order}\,(1/c^4), \\ m &= \pi/(16g^{\frac{1}{2}}c^2), \\ \mathbf{n}_{AB} &= \text{unit vector along } \mathbf{r}_A(t) - \mathbf{r}_B(t), \\ V_{AB} &= -[Z_A Z_B/|\mathbf{r}_A(t) - \mathbf{r}_B(t)|][1 - \{\mathbf{v}_A^2 + \mathbf{v}_B^2 - \mathbf{v}_A \cdot \mathbf{v}_B \\ &\qquad + (\mathbf{n}_{AB} \cdot \mathbf{v}_A)(\mathbf{n}_{AB} \cdot \mathbf{v}_B)\}/2c^2]. \end{aligned} \quad (4.15)$$

On discarding the irrelevant time derivative from Eq. (4.14), it is evident that the equations of motion to order $1/c^2$ may be derived from the total Lagrangian

$$L = \sum_A T_A - \sum_{A<B} V_{AB}, \quad (4.16)$$

with the effective pair interaction V_{AB} given by Eq. (4.15). It is interesting to note that the same Lagrangian (4.16) can be obtained directly from the variation principle (3.1) by taking the trial wavefunction $\theta = \sum_A Z_A [Z_A^4 g + R_A^2(x^\lambda)]^{-\frac{1}{2}}$ and retaining only terms up to order $1/c^2$.

5. DISCUSSION

As a check on our result for the interaction let us find the static force between two particles by the method of Rosen described in Sec. 2. Consider two particles held at points $\mathbf{r}_1 = (0, 0, -\frac{1}{2}r_{12})$ and $\mathbf{r}_2 = (0, 0, \frac{1}{2}r_{12})$ on the Z axis at times $t \leq 0$; the separation r_{12} is assumed much larger than the particle sizes. The wavefunction and its derivative at $t = 0$ are then of the form given in Eq. (2.1):

$$\begin{aligned} \theta(\mathbf{r}, 0) &= Z_1\{Z_1^4 g + (\mathbf{r} - \mathbf{r}_1)^2\}^{-\frac{1}{2}} \\ &\quad + Z_2\{Z_2^4 g + (\mathbf{r} - \mathbf{r}_2)^2\}^{-\frac{1}{2}}, \\ [\partial\theta/\partial t]_{t=0} &= 0. \end{aligned}$$

To find the force on particle 2 we have to compute the surface integral, Eq. (2.2), of the stress tensor over a surface S enclosing this particle, but excluding particle 1. Following Rosen, we take for S the surface enclosing the infinite hemisphere $z > 0$, $x^2 + y^2 + z^2 < \infty$. By symmetry the force on particle 2 lies along the Z axis, so that the only nonvanishing component is

$$F^3 = \int_{z=0} T^{33}\, dx\, dy,$$

where

$$T^{33} = \frac{1}{8\pi}\left[\frac{\dot{\theta}^2}{c^2} - \left(\frac{\partial\theta}{\partial x}\right)^2 - \left(\frac{\partial\theta}{\partial y}\right)^2 + \left(\frac{\partial\theta}{\partial z}\right)^2 + g\theta^6\right].$$

This integral is trivial to compute in the limit of large r_{12}, when it yields the expected result $F^3 = -Z_1 Z_2/r_{12}^2$ in agreement with the potential energy $-Z_1 Z_2/r_{12}$ found in Sec. 3.

The variational method applied in this paper to the motion of the Rosen particle is of quite general application. For example, consider a real field θ obeying the field equation

$$\delta\left(\frac{1}{8\pi c}\right)\int [g^{\iota\kappa}\theta_{,\iota}\theta_{,\kappa} + f(\theta)]\, d^4x = 0,$$

and suppose $f(\theta)$ vanishes rapidly enough for $\theta \to 0$ to allow a time-independent regular solution $\theta_0(\mathbf{r})$ with the asymptotic behavior Z/r for large r.[26] On carrying through the analysis of Sec. IV with $f(\theta)$ replacing $g\theta^6$, one again obtains the Lagrangian principle, Eq. (4.9), with

$$m = \frac{1}{8\pi c^2}\int [(\nabla\theta_0)^2 - f(\theta_0)]\, d^3\mathbf{r},$$

$$Z = \frac{1}{8\pi}\int f'(\theta_0)\, d^3\mathbf{r}.$$

The field equation $\nabla^2\theta_0 = -\frac{1}{2}f'(\theta_0)$ enables the integral for Z to be transformed into the surface integral

$$-\frac{1}{4\pi}\int \nabla\theta_0 \cdot d\mathbf{S}$$

on an infinitely distant sphere, so that Z is indeed the coefficient of $1/r$ in the asymptotic form of $\theta_0(\mathbf{r})$ for large r. Thus we again obtain the equations of motion (4.10).

Note added in proof: Many of the results of Sec. 3 have been obtained in a paper by G. Rosen [J. Math. Phys. **8**, 573 (1967)], which appeared after the present paper went to press.

[26] A sufficient condition is that $f(\theta)$ behaves like (constant) θ^n as $\theta \to 0$ with $n > 4$.

ACKNOWLEDGMENT

One of us (Wan K.-K.) would like to thank the Carnegie Trust for the Universities of Scotland for a research scholarship.

APPENDIX A: PROOF OF EQ. (3.5)

In evaluating V, Eq. (3.5), one encounters four distinct types of integral:

$$I_1 = \int \nabla\theta_1 \cdot \nabla\theta_2 \, d^3\mathbf{r} = \int 3g\theta_1^5\theta_2 \, d^3\mathbf{r} = \int 3g\theta_1\theta_2^5 \, d^3\mathbf{r},$$

$$I_2 = \int \dot\theta_1\dot\theta_2 \, d^3\mathbf{r},$$

$$I_3 = \int \theta_1^4\theta_2^2 \, d^3\mathbf{r},$$

$$I_4 = \int \theta_1^3\theta_2^3 \, d^3\mathbf{r}.$$

Let us consider these integrals in turn.

$$I_1 = Z_1Z_2 \int \mathbf{R}_1 \cdot \mathbf{R}_2 (R_1^2 + a_1^2)^{-\frac{3}{2}} (R_2^2 + a_2^2)^{-\frac{3}{2}} \, d^3\mathbf{r}$$
$$= \frac{Z_1Z_2}{r_{12}} \int \boldsymbol{\rho} \cdot (\boldsymbol{\rho} + \mathbf{n}) \left(\rho^2 + \frac{a_1^2}{r_{12}^2}\right)^{-\frac{3}{2}} \times ([\boldsymbol{\rho} + \mathbf{n}]^2 + a_2^2/r_{12}^2)^{-\frac{3}{2}} \, d^3\boldsymbol{\rho},$$

where

$$\mathbf{R}_1 = \mathbf{r} - \mathbf{r}_1, \qquad \mathbf{R}_2 = \mathbf{r} - \mathbf{r}_2,$$
$$a_1 = Z_1^2 g^{\frac{1}{2}}, \qquad a_2 = Z_2^2 g^{\frac{1}{2}},$$
$$\mathbf{n} = (\mathbf{r}_1 - \mathbf{r}_2)/r_{12}, \quad \boldsymbol{\rho} = \mathbf{R}_1/r_{12}.$$

Hence, for large r_{12},

$$I_1 \to (Z_1Z_2/r_{12}) \int \boldsymbol{\rho} \cdot (\boldsymbol{\rho} + \mathbf{n}) \rho^{-3} |\boldsymbol{\rho} + \mathbf{n}|^{-3} \, d^3\boldsymbol{\rho}$$
$$= 4\pi Z_1Z_2/r_{12}.$$

A similar argument gives

$$I_2 \to (2\pi Z_1Z_2/r_{12})[\dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_2 - (\dot{\mathbf{r}}_1 \cdot \mathbf{n})\dot{\mathbf{r}}_2 \cdot \mathbf{n})].$$

However, for I_3 and I_4, the above technique fails as it yields divergent integrals. Instead we may proceed as follows:

$$I_3 = \frac{Z_1^2Z_2^2}{g^{\frac{1}{2}} r_{12}^2} \int \frac{a_1}{r_{12}} \left(\rho^2 + \frac{a_1^2}{r_{12}^2}\right)^{-2} \left([\boldsymbol{\rho} + \mathbf{n}]^2 + \frac{a_2^2}{r_{12}^2}\right)^{-1} d^3\boldsymbol{\rho}.$$

Now

$$\lim_{r_{12}\to\infty} (a_1/r_{12})(\rho^2 + a_1^2/r_{12}^2)^{-2} = 0, \quad \boldsymbol{\rho} \neq 0,$$
$$= \infty, \quad \boldsymbol{\rho} = 0,$$

and

$$\int \left(\frac{a_1}{r_{12}}\right) \left(\rho^2 + \frac{a_1^2}{r_{12}^2}\right)^{-2} d^3\boldsymbol{\rho} = \pi^2.$$

Hence

$$\lim_{r_{12}\to\infty} (a_1/r_{12})(\rho^2 + a_1^2/r_{12}^2)^{-2} = \pi^2\delta(\boldsymbol{\rho}),$$
$$I_3 \to \pi^2 Z_1^2Z_2^2/g^{\frac{1}{2}} r_{12}^2 \quad \text{for large} \quad r_{12}.$$

To evaluate I_4 we write it as the sum of three integrals:

$$I_4 = I_4' + I_4'' + I_4''',$$

where

$$I_4' = \frac{Z_1^3Z_2^3}{r_{12}^3} \int_{\rho<\frac{1}{2}} \left(\rho^2 + \frac{a_1^2}{r_{12}^2}\right)^{-\frac{3}{2}} \left([\boldsymbol{\rho} + \mathbf{n}]^2 + \frac{a_2^2}{r_{12}^2}\right)^{-\frac{3}{2}} d^3\boldsymbol{\rho}$$

and I_4'' and I_4''' have the same integrand as I_4', but are taken over the domains $|\boldsymbol{\rho} + \mathbf{n}| < \frac{1}{2}$ and $\rho > \frac{1}{2}$, $|\boldsymbol{\rho} + \mathbf{n}| > \frac{1}{2}$, respectively.

I_4' may be written as the sum of the two integrals:

$$\frac{Z_1^3Z_2^3}{r_{12}^3} \int_{\rho<\frac{1}{2}} \left(\rho^2 + \frac{a_1^2}{r_{12}^2}\right)^{-\frac{3}{2}} d^3\boldsymbol{\rho} = (4\pi Z_1^3Z_2^3/r_{12}^3)$$
$$\times [\sinh^{-1}(r_{12}/2a_1) - (1 + 4a_1^2/r_{12}^2)^{-\frac{1}{2}}]$$
$$\to (4\pi Z_1^3Z_2^3/r_{12}^3) \log r_{12} + \text{order}\,(1/r_{12}^3),$$
$$\text{for large} \quad r_{12},$$

and

$$\frac{Z_1^3Z_2^3}{r_{12}^3} \int_{\rho<\frac{1}{2}} \left(\rho^2 + \frac{a_1^2}{r_{12}^2}\right)^{-\frac{3}{2}} \left[\left([\boldsymbol{\rho} + \mathbf{n}]^2 + \frac{a_2^2}{r_{12}^2}\right)^{-\frac{3}{2}} - 1\right] d^3\boldsymbol{\rho}$$
$$\to \frac{Z_1^3Z_2^3}{r_{12}^3} \int_{\rho<\frac{1}{2}} \rho^{-3}(|\boldsymbol{\rho} + \mathbf{n}|^{-3} - 1) \, d^3\boldsymbol{\rho}$$
$$= (2\pi Z_1^3Z_2^3/r_{12}^3) \log \tfrac{4}{3}.$$

Whence

$$I_4' \to (4\pi Z_1^3Z_2^3/r_{12}^3) \log r_{12} + \text{order}\,(1/r_{12}^3)$$
$$\text{for large} \quad r_{12}.$$

A similar argument yields the same limiting form for I_4'', while I_4''' is of order $1/r_{12}^3$. Thus, for large r_{12}, $I_4 \to (8\pi Z_1^3Z_2^3/r_{12}^3) \log r_{12}$. Collecting together all the various integrals then yields Eq. (3.5).

APPENDIX B: PROOF THAT $\int \delta\theta_0\theta_{0,\sigma} \, dS^\sigma$ VANISHES

We wish to show that $\int \delta\theta_0\theta_{0,\sigma} \, dS^\sigma$, Eq. (4.7), vanishes for variations $\delta z^\lambda(p)$ which leave the endpoints $z^\lambda(p_1)$ and $z^\lambda(p_2)$ fixed. To facilitate the integrations involved in this and the next section let us replace the four coordinates x^λ by a new set p_r, R, Θ, ϕ, defined below. First we introduce a null vector e^λ with components $(1, \sin\Theta\cos\phi, \sin\Theta\sin\phi, \cos\Theta)$ and define the retarded functions p_r, R, $u^\lambda(p_r)$ as in Eq. (4.4). Then the equations

$$y^\lambda = x^\lambda - z^\lambda(p_r) = Re^\lambda/e^\mu u_\mu(p_r)$$

define x^λ as functions of the four variables p_r, R, Θ, ϕ. In the new coordinates the volume integral becomes

$$\int_\Sigma d^4x = \int dp_r \, dR \, d\Theta \, d\phi \, \frac{\partial(x^0, x^1, x^2, x^3)}{\partial(p_r, R, \Theta, \phi)}$$
$$= \int_{p_1}^{p_2} s_r' \, dp_r \int_0^\infty R^2 \, dR \int d\Omega, \tag{B1}$$

where

$$s_r' = [g_{\iota\kappa}(dz^\iota(p_r)/dp_r)(dz^\kappa(p_r)/dp_r)]^{\frac{1}{2}}$$

and the Lorentz-invariant angular integration over a unit sphere is

$$\int d\Omega = \int d\Theta\, d\phi\, \frac{\sin\Theta}{[e^\mu u_\mu(p_r)]^2}.$$

In what follows we need the two results

$$\int 1 \cdot d\Omega = 4\pi,$$

$$\int [\gamma^\lambda - Ru^\lambda(p_r)]\, d\Omega = 0, \tag{B2}$$

which are readily proved by choosing a special inertial system in which $u^\lambda(p_r) = (1, 0, 0, 0)$. The surface element dS^λ on $p_r = \text{const}$ is

$$\begin{aligned} dS^\sigma &= (1/3!)g^{\sigma\iota}\epsilon_{\iota\kappa\lambda\mu}\partial(x^\kappa, x^\lambda, x^\mu)/\partial(R, \Theta, \phi)\, dR\, d\Theta\, d\phi \\ &= g^{\sigma\iota}(\partial p_r/\partial x^\iota) \\ &\quad \times \partial(x^0, x^1, x^2, x^3)/\partial(p_r, R, \Theta, \phi)\, dR\, d\Theta\, d\phi \\ &= R^2\, dR(\gamma^\sigma/R)\, d\Omega. \end{aligned}$$

Let us now consider a variation of the path $z^\lambda(p) \to z^\lambda(p) + \delta z^\lambda(p)$, keeping the end points $z^\lambda(p_1)$ and $z^\lambda(p_2)$ fixed. On the surfaces $p_r = p_1$ and p_2 we have

$$\begin{aligned} \delta z^\lambda(p_r) &= 0, \\ \delta R &= [\gamma_\lambda - Ru_\lambda(p_r)][d\delta z^\lambda(p_r)/dp_r]/s_r', \\ \delta\theta_0\theta_{0,\sigma}\, dS^\sigma &= (d\theta_0/dR)^2R^2\, dR(\gamma^\sigma/R)(\partial R/\partial x^\sigma)\delta R\, d\Omega. \end{aligned}$$

From the definition of R, Eq. (4.4), we have

$$\partial R/\partial x^\sigma = u_\sigma(p_r) + (\xi - 1)\gamma_\sigma/R,$$

$$(\gamma^\sigma/R)\partial R/\partial x^\sigma = 1, \tag{B3}$$

where

$$\xi = [\gamma_\lambda - Ru_\lambda(p_r)][du^\lambda(p_r)/dp_r](s_r')^{-1},$$

whence

$$\int \delta\theta_0\theta_{0,\sigma}\, dS^\sigma = \int R^2\left(\frac{d\theta_0}{dR}\right)^2 \int \delta R\, d\Omega,$$

which vanishes on account of Eq. (B2).

APPENDIX C: PROOF OF EQ. (4.9)

The integrations of Eq. (4.8) are most readily effected in the variables p_r, R, Θ, ϕ introduced in Appendix B above. We have

$$\begin{aligned} \frac{1}{8\pi}\int R^2\, dR \int d\Omega[g^{\iota\kappa}\theta_{0,\iota}\theta_{0,\kappa} + g\theta_0^6] &= \frac{1}{8\pi}\int R^2\, dR\int d\Omega\left[(2\xi - 1)\left(\frac{d\theta_0}{dR}\right)^2 + g\theta_0^6\right] \\ &= \frac{1}{2}\int R^2\, dR\left[\left(\frac{d\theta_0}{dR}\right)^2 + g\theta_0^6\right] \\ &= \pi/16g^{\frac{1}{2}}. \end{aligned}$$

[See Eqs. (B2) and Eq. (B3).]

To evaluate the integral of the interaction term $6g\theta_0^5\theta_e$, let us assume that the external field θ_e does not vary appreciably over distances of the order of the particle size $Z^2g^{\frac{1}{2}}$. Then

$$\begin{aligned} \frac{1}{8\pi}\int R^2\, dR\int d\Omega 6g\theta_0^5\theta_e &\approx \theta_e(z^\lambda(p_r))\int R^2\, dR\int d\Omega 6g\theta_0^5 \\ &= Z\theta_e(z^\lambda(p_r)). \end{aligned}$$

On carrying out the final integration

$$\int_{p_1}^{p_2} s_r'\, dp_r$$

[see Eq. (B1)], we obtain Eq. (4.9).

Infinite-Spin Ising Model in One Dimension

COLIN J. THOMPSON*
Applied Mathematics Department, The University of New South Wales, Kensington, N.S.W., Australia.

(Received 12 June 1967)

The partition function z, the pair correlation function ρ, and the zero-field susceptibility χ for the one-dimensional Ising model with infinite spin, are expressed in terms of the eigenvalues and eigenfunctions of an integral equation. The eigenfunctions of the integral equation are shown to be the oblate spheroidal wavefunctions, and, from known asymptotic expansions, high- and low-temperature expansions are given for z, ρ, and χ. It is shown that the low-temperature behavior of z, ρ, and χ differs qualitatively from the corresponding behavior for all finite spin.

1. INTRODUCTION

The one-dimensional Ising model with spin one-half was solved by Ising in his dissertation of 1925,[1] but it was not until recently that the Ising model with general spin S was considered. For arbitrary S the problem can be set up as a matrix problem in the same way as Kramers and Wannier[2] set up the $S = \frac{1}{2}$ problem. Suzuki *et al.*[3] have done this for the one-dimensional model and have obtained solutions (numerical for $S > 1$) for $S = 1, \frac{3}{2}, \cdots, \frac{7}{2}$. They also obtained high-temperature expansions for the partition function z and zero-field susceptibility χ for arbitrary spin, S, and low-temperature expansions for z and χ for $S = 1$ and $\frac{3}{2}$. Their high-temperature expansions for arbitrary S agree with those obtained previously by Domb and Sykes and others,[4] who have obtained expansions in two and three dimensions as well as in one dimension.

The aim of obtaining such expansions is to attempt, by extrapolation techniques, to determine the dependence, if any, of critical behavior on spin. Domb and Sykes[4] suggest that the high-temperature critical behavior of the Ising and Heisenberg models is independent of spin, although for the Heisenberg model this has recently been questioned.[5]

Low-temperature expansions are more difficult to obtain because these are determined by the excitation spectrum which is nontrivial for general spin, and, as far as the author knows, no low-temperature expansion for general S, apart from the one-dimensional $S = 1$ and $\frac{3}{2}$ expansions of Suzuki *et al.*, has been developed.

To leading order it is clear, however, that the low-temperature behavior of the partition function is independent of S (apart from trivial normalization factors) for all finite S. Thus in one dimension, for example, if one defines the energy in a given configuration (assuming cyclic boundary conditions) by

$$E = -\frac{J}{S^2}\sum_{i=1}^{N} S_i^z S_{i+1}^z, \quad S_i^z = S, S-1, \cdots, -S, \tag{1}$$

so that the ground state energy is $E_0 = -JN$ for all S, it follows immediately that, to leading order,

$$z = \lim_{N\to\infty} (Z_N)^{1/N} \sim e^K \quad \text{as} \quad T \to 0^+ \quad \text{for all } \textit{finite } S, \tag{2}$$

where Z_N is the partition function and $K = J/kT$. In particular, for $S = \frac{1}{2}$, $z = e^K[1 + e^{-2K}]$, and for $S = 1$ and $\frac{3}{2}$, respectively,[3] $z \sim e^K[1 + 3e^{-2K} + \cdots]$ and $z \sim e^K[1 + e^{-4K/3} + 3e^{-2K} + \cdots]$. Equation (2) is true in any dimension and, in general, the multiplicative correction to (2) is $1 + O[\exp(-qK/S)]$, with q a positive constant.

Low-temperature expansions for infinite spin could be obtained from the finite-spin expansions, but it must be remembered that the multiplicative correction to (2), which is a sum for finite S, becomes an integral in the limit $S \to \infty$ (or, in other words, the excitation spectrum becomes continuous in the limit $S \to \infty$), so that the leading term could conceivably, and almost surely will, have a different temperature dependence. For any finite S, of course, there will be a range of temperature ($K \leqslant S$) for which the finite-spin partition function will closely approximate the infinite-spin partition function, but at sufficiently low temperatures [$K \geqslant S$, where the sum comprising the multiplicative correction to (2) can no longer be approximated by an integral], one can expect the infinite-spin model to show qualitatively different temperature behavior than the finite-spin model.

---

* Address from September, 1968: Applied Mathematics Department, Massachusetts Institute of Technology, Cambridge, Mass.
[1] E. Ising, Z. Physik **31**, 253 (1925).
[2] H. A. Kramers and G. Wannier, Phys. Rev. **60**, 252 (1941).
[3] M. Suzuki, B. Tsujiyama, and S. Katsura, J. Math. Phys. **8**, 124 (1967).
[4] C. Domb and M. F. Sykes, Phys. Rev. **128**, 168 (1962), and references quoted therein.
[5] H. E. Stanley and T. A. Kaplan, J. Appl. Phys. **38**, 977 (1967).

We consider here the infinite spin limit of (1) where the matrix problem reduces to an integral equation whose eigenfunctions and eigenvalues can be obtained exactly. We find that at low temperatures

$$z \sim (e^K/2K)\left[1 + \frac{1}{4K} + \frac{3}{16K^2} + \frac{7}{32K^3} + O(K^{-4})\right] \quad \text{as} \quad T \to 0^+, \quad (3)$$

which indeed shows qualitatively different temperature behavior than the corresponding finite spin expressions above.

High-temperature expansions ($K \leqslant 1$) for z of the one-dimensional infinite spin model are given in the following section, and in Sec. 3 high- and low-temperature expansions are given for the pair correlation function ρ and the zero field susceptibility χ. The low-temperature expansions for ρ and χ again show qualitatively different temperature behavior than the corresponding finite spin expansions.

In conclusion, it is perhaps interesting to note that Eq. (3) is equivalent asymptotically (as $T \to 0^+$) to the corresponding expression for the one-dimensional isotropic Heisenberg model with infinite spin,[6] for which

$$(Z_N)^{1/N} = K^{-1} \sinh K. \quad (4)$$

2. PARTITION FUNCTION

For infinite spin the partition function of the one-dimensional chain of N spins with interaction energy (1) is given by

$$Z_N = \int_{-1}^{+1} \cdots \int_{-1}^{+1} dx_1 \cdots dx_N \exp\left(K \sum_{i=1}^{N} x_i x_{i+1}\right) \quad (x_{N+1} = x_1). \quad (5)$$

In terms of the Nth iterate $K^{(N)}(x, y)$ of the integral operator $K(x, y) = \exp(Kxy)$,

$$Z_N = \int_{-1}^{+1} dx K^{(N)}(x, x) = \sum_{n=0}^{\infty} \lambda_n^N \sim \lambda_0^N \quad \text{as} \quad N \to \infty, \quad (6)$$

where $\lambda_0 > \lambda_1 > \lambda_2 > \cdots$ are the eigenvalues of the integral equation

$$\int_{-1}^{+1} e^{Kxy} \phi(y)\, dy = \lambda \phi(x). \quad (7)$$

From (6) the free energy per spin ψ is given by

$$-\psi/kT = \lim_{N\to\infty} \frac{1}{N} \log Z_N = \log \lambda_0. \quad (8)$$

The integral equation (7) is familiar in the theory of spheroidal wavefunctions (Ref. 7, Ch. 4). It has also occurred in the problem of determining the distribution function of level spacing for a random matrix.[8] Thus it is easily verified[8] that if $\phi(x)$ is a solution of

$$L\phi = \mu\phi, \quad (9)$$

where the differential operator L is defined by

$$L = (x^2 - 1)(d^2/dx^2) + 2x(d/dx) - (Kx)^2, \quad (10)$$

then

$$\chi(x) = \int_{-1}^{+1} e^{Kxy} \phi(y)\, dy \quad (11)$$

is also a solution of (9), provided that

$$(1 - x^2)\phi(x) = (1 - x^2)\phi'(x) = 0 \quad \text{for} \quad |x| = 1. \quad (12)$$

The regularity conditions (12) determine the eigenvalues of L and, since $\chi(x)$ is necessarily regular at $x = 1$, Eq. (11) implies that χ is proportional to ϕ; it follows that L and the integral operator $\exp(Kxy)$ commute.

The eigenfunctions of the differential equation (9) [and, consequently, of the integral equation (7)] are the oblate spheroidal wave functions $\phi_q(x)$ [in the notation of Ref. 7 $S_{0q}(-iK, x)$] where $\phi_q(x)$ is an even (or odd) function of x for even (or odd) integral q. The corresponding eigenvalues λ_q of (7) are found, for example, by substituting $\phi_q(x)$ for $\phi(x)$ in (7), putting $x = 0$ for q even, and differentiating and putting $x = 0$ for q odd, i.e.,

$$\lambda_{2q} = \int_{-1}^{+1} \phi_{2q}(x)\, dx / \phi_{2q}(0),$$

$$\lambda_{2q+1} = K \int_{-1}^{+1} x \phi_{2q+1}(x)\, dx / \phi'_{2q+1}(0), \quad q = 0, 1, 2, \cdots. \quad (13)$$

Equations (13) are convenient for numerical computation and also for determining expansions for small K (i.e., high temperatures). To obtain low-temperature expansions it is more convenient, as we will see in a moment, to substitute $x = 1$ in (7), i.e.,

$$\lambda_q = \int_{-1}^{+1} e^{Ky} \phi_q(y)\, dy / \phi_q(1), \quad q = 0, 1, 2, \cdots. \quad (14)$$

High-temperature expansions for λ_q are best obtained using the representation [Ref. 7, Eq. (3.1.36)]

$$\phi_q(x) = \sum_{r=0,1}^{\infty}{}' d_r^q P_r(x), \quad (15)$$

where the primed sum is over even (or odd) values for r when q is even (or odd). $P_r(x)$ is the Legendre

[6] M. E. Fisher, Am. J. Phys. 32, 343 (1964).

[7] C. Flammer, *Spheroidal Wave Functions* (Stanford University Press, Stanford, California, 1957).

[8] M. Gaudin, Nucl. Phys. 25, 447 (1961). I am grateful to Professor F. J. Dyson for drawing my attention to this fact.

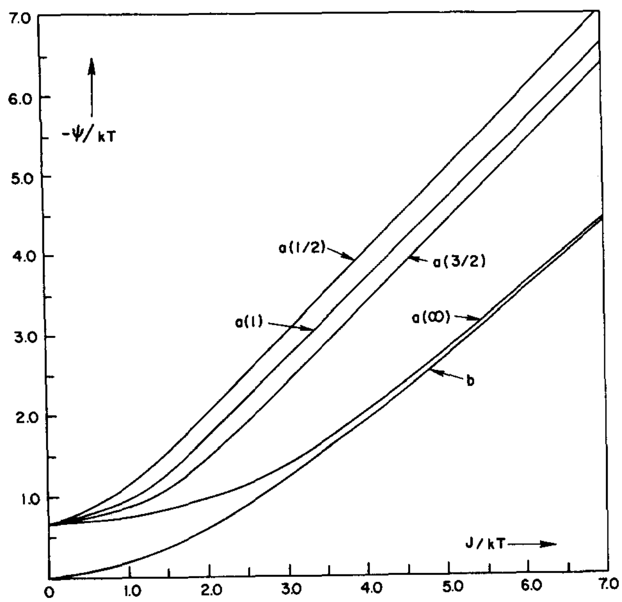


FIG. 1. Free energy $-\psi/kT$ versus J/kT for $a(s)$, the Ising chain, with spin s (normalized to log 2 at $J/kT = 0$), and (b) the infinite-spin Heisenberg chain.

polynomial and d_r^q are tabulated coefficients (for K in the range 0 to 5 in Ref. 7 and 0 to 8 in Ref. 9). Substituting (15) into (13), we find that

$$\lambda_{2q} = 2d_0^{2q}\left[\sum_{p=0}^{\infty}(-1)^p c_p d_{2p}^{2q}\right]^{-1},$$
$$\lambda_{2q-1} = Kd_1^{2q-1}\left[3\sum_{p=1}^{\infty}(-1)^{p-1}pc_p d_{2p-1}^{2q-1}\right]^{-1}, \tag{16}$$

where

$$c_p = (-1)^p P_{2p}(0) = (2p)!/2^{2p}(p!)^2. \tag{17}$$

Series expansions for d_p^q are given in Ref. 7. Tables 3 and 4, for example, state that

$$d_2^0/d_0^0 = K^2/9 + 2K^4/567 + \cdots,$$
$$d_4^0/d_0^0 = K^4/525 + \cdots,$$
$$d_3^1/d_1^1 = K^2/25 + \cdots,$$

from which it follows that

$$z = \lim_{N\to\infty}(Z_N)^{1/N} = \lambda_0$$
$$= 2 + K^2/9 + 67K^4/8100 + \cdots,$$

and

$$\lambda_1 = 2K/3 + K^3/25 + \cdots. \tag{18}$$

We remark that the expansion for z agrees with the infinite spin limit of the expression for z of Suzuki *et al.* Numerical values for $\log \lambda_0 = -\psi/kT$ [obtained from (16) and tabulated values for d_p^q] are shown in Fig. 1, where, for comparison, we have also shown $-\psi/kT$ for the isotropic infinite-spin Heisenberg chain [log (sinh K/K)] and for the spin $\frac{1}{2}$, 1, and $\frac{3}{2}$ Ising chain (normalized to log 2 at $K = 0$).

For large K the appropriate expansion for $\phi_q(x)$ is [Ref. 7, Eq. (8.2.9)]

$$\phi_q(x) = \sum_{s=-\nu}^{\infty} A_s^q\{e^{-K(1-x)}L_{s+\nu}[2K(1-x)]$$
$$+ (-1)^q e^{-K(1+x)}L_{s+\nu}[2K(1+x)]\}, \tag{19}$$

where

$$\nu = \begin{cases} q/2 & q \text{ even}, \\ (q-1)/2 & q \text{ odd}, \end{cases}$$

and $L_n(z)$ are Laguerre polynomials. Expansions for A_s^q in inverse powers of K are given in Ref. 7 [Eq. (8.2.15)]; for example,

$$A_1^0/A_0^0 = -\tfrac{1}{4}K^{-1} - \tfrac{1}{4}K^{-2} - \tfrac{23}{64}K^{-3} - \cdots,$$
$$A_2^0/A_0^0 = \tfrac{1}{8}K^{-2} + \tfrac{5}{16}K^{-3} + \cdots,$$
$$A_3^0/A_0^0 = -\tfrac{3}{32}K^{-3} - \cdots,$$

and

$$A_q^1/A_0^1 = A_q^0/A_0^0.$$

In general, $A_{\pm s}^q/A_0^q \sim K^{-|s|}$ and hence to leading order $\sum_{s=0}^{\infty} A_s^0 L_s(2K)$, which, from (19), is proportional to $\phi_0(0)$ and is a constant which cannot be simply calculated. It follows then that Eq. (13) is unsuited for calculating λ_0. Equation (19), on the other hand, is ideally suited to calculation of $\phi_0(1)$, and, by substituting (19) into (14), it is straightforward to show that

$$\lambda_0 = (e^K/2K)$$
$$\times [1 + \tfrac{1}{4}K^{-1} + \tfrac{3}{16}K^{-2} + \tfrac{7}{32}K^{-3} + O(K^{-4})] \tag{20}$$

and

$$\lambda_1/\lambda_0 = 1 - \alpha K e^{-2K}, \quad \alpha = O(1). \tag{21}$$

The constant α is difficult to calculate directly from (19), but if one uses the numerical values of λ_0 and λ_1 [computed from (13) and tables in Refs. 7 and 8 for K in the range 5 to 8], one finds that $\alpha \approx 25$.

High- and low-temperature expansions for the free energy can be obtained straightforwardly from (8), (18), and (20), and corresponding expansions for the internal energy and specific heat can then be obtained by differentiating.

3. CORRELATION FUNCTIONS AND SUSCEPTIBILITY

The pair correlation function can be written as

$$\rho_N(r) = \langle x_j x_{j+r}\rangle$$
$$= Z_N^{-1}\int_{-1}^{+1}\cdots\int_{-1}^{+1} dx_1 \cdots dx_N x_j x_{j+r}$$
$$\times \exp\left(K\sum_{j=1}^{N} x_j x_{j+1}\right)$$
$$= Z_N^{-1}\int_{-1}^{+1}\int_{-1}^{+1}\int_{-1}^{+1} dx\, dy\, dz K^{(j-1)}(x, y) y K^{(r)}(y, z)$$
$$\times z K^{(N-j-r+1)}(z, x), \tag{22}$$

[9] J. A. Stratton, P. M. Morse, L. J. Chu, J. D. C. Little, and F. J. Corbato, *Spheroidal Wave Functions* (Technology Press, Cambridge, Mass., 1956).

where, as before, $K^{(s)}(x,y)$ is the sth iterate of the integral operator exp (Kxy). Using the representation

$$K^{(s)}(x,y)=\sum_{j=0}^{\infty}\lambda_j^s\phi_j(x)\phi_j(y), \quad (23)$$

in terms of the eigenvalues λ_j and corresponding *normalized* eigenfunctions $\phi_j(x)$ of $K(x,y)$, we then see that

$$\rho_N(r)=Z_N^{-1}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\lambda_l^{N-r}\lambda_m^r\left[\int_{-1}^{+1}dx\phi_l(x)x\phi_m(x)\right]^2. \quad (24)$$

Using Eq. (6), we have finally (keeping r fixed) that

$$\rho(r)=\lim_{N\to\infty}\rho_N(r)=\sum_{m=0}^{\infty}\left(\frac{\lambda_m}{\lambda_0}\right)^r\left(\int_{-1}^{+1}dx\phi_0(x)x\phi_m(x)\right)^2. \quad (25)$$

Notice that, since $\phi_m(x)$ is even when m is even, the integral in (25) vanishes for even m. Hence only odd values of m contribute to the sum in Eq. (25).

The susceptibility in zero field (assuming each spin has unit magnetic moment) from the fluctuation relation and Eq. (24) is found to be

$$\begin{aligned}\chi&=\lim_{N\to\infty}(NkT)^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\langle x_ix_j\rangle\\&=(kT)^{-1}\sum_{m=0}^{\infty}[(\lambda_0+\lambda_{2m+1})(\lambda_0-\lambda_{2m+1})^{-1}]\\&\qquad\times\left[\int_{-1}^{+1}dx\phi_0(x)x\phi_{2m+1}(x)\right]^2. \quad (26)\end{aligned}$$

High- and low-temperature expansions for $\rho(r)$ and χ can be obtained using the results of the previous section. Thus, at high temperatures, using (15), one obtains

$$\int_{-1}^{+1}dxx^2\phi_0^2(x)=\tfrac{1}{3}[1+4K^2/45+O(K^4)] \quad (27)$$

and

$$\left[\int_{-1}^{+1}dx\phi_0(x)x\phi_1(x)\right]^2=\tfrac{1}{3}[1+4K^2/45+O(K^4)]. \quad (28)$$

If we use Parseval's theorem, then it follows that

$$\begin{aligned}\sum_{m=1}^{\infty}&\left(\int_{-1}^{+1}dx\phi_0(x)x\phi_{2m+1}(x)\right)^2\\&=\int_{-1}^{+1}dxx^2\phi_0^2(x)-\left[\int_{-1}^{+1}dx\phi_0(x)x\phi_1(x)\right]^2\\&=O(K^4), \quad (29)\end{aligned}$$

and hence, from (25), that

$$\rho(r)=\left(\frac{\lambda_1}{\lambda_0}\right)^r\left\{\left[\int_{-1}^{+1}dx\phi_0(x)x\phi_1(x)\right]^2+R\right\}, \quad (30)$$

where, from (25) and (29)

$$\begin{aligned}R&=\sum_{m=1}^{\infty}\left(\frac{\lambda_{2m+1}}{\lambda_1}\right)^r\left[\int_{-1}^{+1}dx\phi_0(x)x\phi_{2m+1}(x)\right]^2\\&\le\sum_{m=1}^{\infty}\left[\int_{-1}^{+1}dx\phi_0(x)x\phi_{2m+1}(x)\right]^2=O(K^4). \quad (31)\end{aligned}$$

Substituting (28) and the expansions (18) for λ_0 and λ_1 into (30) then gives

$$\rho(r)=\left\{\frac{K}{3}\left[1+\frac{K^2}{225}+O(K^4)\right]\right\}^r\times\left\{\frac{1}{3}\left[1+\frac{4K^2}{45}+O(K^4)\right]\right\}. \quad (32)$$

Similarly, from (26), one sees that

$$\chi=(K/3J)[1+2K/3+14K^2/45+92K^3/675+O(K^4)]. \quad (33)$$

Note that $\rho(1)$ obtained from (32) is proportional to the internal energy (as it should be) computed from Eqs. (8) and (18), and that Eq. (33) agrees to leading order with the infinite-spin limit of the formula given by Suzuki *et al.* We remark also that, by subtracting off any finite number of terms in the sums (25) and (26) and using Parseval's theorem as above to estimate the remainder, expansions at both high and low temperature for $\rho(r)$ and χ can be obtained to any order.

At low temperatures one proceeds exactly as above. Thus, from (19), one finds that

$$\int_{-1}^{+1}dxx^2\phi_0^2(x)=1-K^{-1}+O(K^{-2}) \quad (34)$$

and

$$\left(\int_{-1}^{+1}dx\phi_0(x)x\phi_1(x)\right)^2=1-K^{-1}+O(K^{-2}). \quad (35)$$

Using (25), (26), and Parseval's theorem (as above) to estimate the remainder, one then obtains

$$\rho(r)=(1-\alpha Ke^{-2K})^r\{1-K^{-1}+O(K^{-2})\} \quad (36)$$

and

$$\chi=\frac{2e^{2K}}{\alpha J}[1-K^{-1}+O(K^{-2})]. \quad (37)$$

The corresponding formulas for the spin one-half Ising model (the asymptotic formulas are the same apart from trivial normalization factors for all finite S^5) are, respectively,

$$\rho(r)=(\tanh K)^r\sim(1-2e^{-2K})^r \quad \text{as} \quad T\to O^+, \quad (38)$$

and

$$\chi = (kT)^{-1}[(1 + \tanh K)/(1 - \tanh K)] \sim \frac{K}{J} e^{2K} \quad \text{as} \quad T \to O^+.$$

For the infinite-spin isotropic Heisenberg chain[6] we have

$$\rho(r) = \tfrac{1}{3}[(\coth K - K^{-1})]^r \sim \tfrac{1}{3}(1 - K^{-1})^r \quad \text{as} \quad T \to O^+ \tag{39}$$

and

$$\chi = (kT)^{-1} \times [(1 + \coth K - K^{-1})/(1 - \coth K + K^{-1})] \sim \frac{2}{J}(1 - K^{-1}) \quad \text{as} \quad T \to O^+.$$

4. CONCLUSION

It has been shown that the partition function z, the pair correlation function ρ, and the zero-field susceptibility χ can be expressed in terms of the eigenvalues and eigenfunctions of the integral equation (7). High-temperature expansions for z, ρ, and χ are given in Eqs. (18), (32), and (33), respectively; it is noted that these expansions can be obtained straightforwardly from the corresponding finite-spin expansions by letting $S \to \infty$. Low-temperature expansions for z, ρ, and χ are given in Eqs. (20), (36), and (37), respectively, which to leading order behave differently as functions of temperature than the corresponding finite-spin expansions, Eq. (2) and (38), respectively.

In conclusion we remark that, although the partition functions for the infinite-spin Ising and Heisenberg models have the same asymptotic behavior in the limit $T \to O^+$, the pair correlation functions and zero-field susceptibilities have a completely different temperature dependence in this limit [cf. Eqs. (37) and (39)].

Note added in proof: G. S. Joyce[10] has independently discovered most of the results presented here and has generalized them to include the classical anisotropic Heisenberg chain, which reduces to the present model in the limit of extreme anisotropy. Joyce's formula (9) for $E/2J_{11}$, which was obtained from a formula of Sips,[11] Eq. (19.1) on p. 363 of Ref. 11

$$\lambda_0 = \frac{e^K}{2K}\left[1 + \frac{1}{4K} + \frac{5}{48K^2} + O(K^{-3})\right],$$

disagrees in the K^{-2} term with our formula (20), which was obtained from formulas given in Ref. 7. Although we have been unable to muster sufficient energy to check the basic asymptotic formulas, comparison with numerical values (computed from tables in Ref. 9) suggests that (20) is correct and that Sips's formula is incorrect.

ACKNOWLEDGMENTS

The author is grateful to Professor J. M. Blatt and N. E. Frankel for valuable discussions.

The support of the Commonwealth Government through the award of a Queen Elizabeth II Post Doctoral Fellowship is gratefully acknowledged.

[10] G. S. Joyce, Phys. Rev. Letters **19**, 581 (1967).
[11] R. Sips, Trans. Am. Math. Soc. **90**, 340 (1959).

2V Sector of the Lee Model

LEONARD M. SCARFONE
Department of Physics, University of Vermont, Burlington, Vermont

(Received 20 March 1967)

The Lehmann–Symanzik–Zimmermann (LSZ) formalism is further used to analyze the $2V$ sector of the Lee model with boson sources at zero separation. Unlike previous LSZ investigations of the model, it is found that the solutions to two singular integral equations solve the entire sector. Two scattering amplitudes, a production amplitude, and an equation for the determination of the $2V$ potential energy are obtained. Similarities and differences between this sector and the $V + \theta$ subspace are pointed out.

I. INTRODUCTION

In a previous paper,[1] the Lehmann–Symanzik–Zimmermann approach to the Lee model[2] was used as a calculational technique to obtain an equation for the determination of the $V + N$ potential energy. This equation was established by setting the denominator of the $V + N$ propagator equal to zero at the $V + N$ bound-state pole, and the corresponding residue led to the normalization constant for this state. To simplify the procedure in that problem, we conveniently set the separation parameter equal to zero and imposed commutation relations on all the field operators. These conditions are also required in the present paper which deals with the much more complicated $2V$ sector of the Lee model. The $V + N$ propagator was found by solving one simple algebraic equation; the $V + N \rightleftarrows 2N + \theta$ vertex function and the $2N + \theta$ scattering amplitude followed straightaway. These characteristics of the first nontrivial two-heavy-particle sector are completely analogous to the first nontrivial one-heavy-particle sector.[2] The vanishing of the denominator of the V-particle propagator at the V pole yields the mass renormalization constant δm_V, and the residue leads to the normalization constant of the physical V-particle state. In this case the propagator is also obtained from one simple algebraic equation. The $V \rightleftarrows N + \theta$ vertex function and the $N + \theta$ scattering amplitude are then readily found.

The basic elements of calculation in the LSZ formalism are the τ functions which are vacuum expectation values of time-ordered products of Heisenberg operators. In the V and $V + N$ sectors there are four such quantities; the symmetry properties of those representing the Green's functions for scattering are revealed at once by the Matthews–Salam equations.[3] It is well known that the V sector has been analyzed with the conventional eigenvalue-equations technique[4] and with the methods of dispersion theory.[5] It has also been shown[6] that these methods of solution go through with equal ease for the $V + N$ case. In these approaches one must solve a separable integral equation or a Low or Omnes type of singular integral equation. Furthermore, these simple sectors are promptly solved with the S-matrix diagrammatic techniques. Early efforts at solving the more complicated $V + \theta$ sector in a standard way were thwarted by the appearance of a nonclassical type of singular integral equation, the so-called Källén–Pauli equation.[4] It turned out, therefore, that new results were first obtained with the methods of dispersion theory,[7] which led to the exact amplitude for $V + \theta$ elastic scattering and the production process $V + \theta \rightarrow N + 2\theta$. Additional considerations[8] showed that these methods could be extended to yield a particular solution of the Källén–Pauli equation. The same solution was also obtained directly through a renewed effort with the eigenvalue equations formulation,[9] which further required the solution of another integral equation for the $N + 2\theta$ elastic scattering amplitude. This amplitude also forms a distinct problem in the dispersion relations approach. A modified dispersion method,[10] a diagrammatic technique,[11] and the N-quantum approximation[12] have also been used to solve the $V + \theta$ sector. In addition, the literature[13] now contains various other methods for solving the Källén–Pauli equation or generalizations thereof.

[1] L. M. Scarfone, J. Math. Phys. 7, 159 (1966).
[2] M. S. Maxon and R. B. Curtis, Phys. Rev. 137, B996 (1965).
[3] P. T. Matthews and A. Salam, Proc. Roy. Soc. (London) A221, 128 (1953).
[4] T. D. Lee, Phys. Rev. 95, 1329 (1954); G. Källén and W. Pauli, Kgl. Danske Videnskab. Selskab, Mat.-Fys. Medd. 30, No. 7 (1955).
[5] M. L. Goldberger and S. B. Treiman, Phys. Rev. 113, 1163 (1959); P. DeCelles and G. Feldman, Nucl. Phys. 14, 517 (1959).
[6] S. Weinberg, Phys. Rev. 102, 285 (1955); L. M. Scarfone, Nucl. Phys. 39, 658 (1962).
[7] R. D. Amado, Phys. Rev. 122, 696 (1961).
[8] R. P. Kenschaft and R. D. Amado, J. Math. Phys. 5, 1340 (1964).
[9] A. Pagnamenta, J. Math. Phys. 6, 955 (1965).
[10] M. T. Vaughn, Nuovo Cimento 40, 803 (1965).
[11] J. C. Howard, Ann. Inst. Henri Poincaré 2A, 105 (1965).
[12] A. Pagnamenta, Ann. Phys. (New York) 39, 453 (1966).
[13] E. Kazes, J. Math. Phys. 6, 1772 (1965); C. M. Sommerfield J. Math. Phys. 6, 1170 (1965); J. B. Bronzan, Phys. Rev. 139, B751 (1965).

The LSZ treatment[14] of this sector has appeared in a recent publication, and the solution to one singular integral equation leads to the two elastic-scattering amplitudes and the production amplitude. The $V + \theta$ discrete state problem has also been studied from several points of view.[14–16] Besides being an off-the-energy shell method of calculation, the LSZ formalism in the Lee model has an impressive mathematical compactness that holds even in the $2V$ sector. The authors[2] of this approach to the model have stated that "it elucidates the basic structure of the Lee model in the most natural way." Another author[17] gives a functional formulation of the model, but he feels that this approach "provides a more general way."

In this paper we investigate the LSZ method of solution in the $2V$ sector with the heavy particles treated as bosons with zero separation. Unlike the previous LSZ problems, we must here solve two singular integral equations in order to solve the entire sector. The first of these is an equation for the Fourier transform of the τ function representing the $2V \rightleftarrows V + N + \theta$ vertex. Its solution also provides the $2V \rightleftarrows 2N + 2\theta$ vertex function and the $2V$ propagator. An equation for the determination of the $2V$ potential energy is then developed by studying the analytic properties of the propagator. The second integral equation is an equation for the Fourier transform of the τ function corresponding to the elastic scattering of a θ particle by the $V + N$ bound system. This function will also make available the functions that describe $2N + 2\theta$ elastic scattering and the production process $V + N + \theta \rightarrow 2N + 2\theta$. It is interesting to note that the amplitudes for these transitions split into two terms, one of which resembles an amplitude in the $V + \theta$ case, the other is proportional to the $2V$ propagator. We do not leave these amplitudes in this form, although they do first come that way.

II. $2V$ PROPAGATOR AND VERTEX FUNCTIONS

In this section we consider the τ functions associated with the propagation of two V particles. These functions form two equivalent sets of Matthews–Salam equations which yield a singular integral equation for the Fourier transform of the τ function representing the $2V \rightleftarrows V + N + \theta$ vertex function. The solution of this equation then leads to the Fourier transform of the $2V \rightleftarrows 2N + 2\theta$ τ function and the $2V$ propagator. Before carrying out this development, it is convenient to review some results found in the $V + N$ sector.

The Fourier transform of the $V + N$ propagator, evaluated at the energy $W + 2m$, is given by the inverse of the function

$$G^+(W) \equiv G(W + i\epsilon) = W + \frac{W^2}{2\pi}\int_\mu^\infty \frac{\operatorname{Im} G^+(\omega)\, d\omega}{\omega^2(\omega - W - i\epsilon)} + \frac{1}{2\pi}\int_\mu^\infty \frac{\operatorname{Im} G^+(\omega)\, d\omega}{\omega - W - i\epsilon}, \quad (1)$$

in which

$$\operatorname{Im} G^+(\omega) = 2\pi\gamma f^2(\omega)(\omega^2 - \mu^2)^{\frac{1}{2}}, \quad (2)$$

and where ϵ is a positive number to be treated as infinitesimally small. The cutoff factor $f(\omega)$ secures the convergence of all integrals and is only a function of the relativistic θ-particle energy $\omega = (k^2 + \mu^2)^{\frac{1}{2}}$. The symbol γ is an abbreviation for $(g/2\pi)^2$, where g is the renormalized coupling constant and the heavy particles V and N are taken with equal mass m. When the energy parameter W is set equal to the $V + N$ potential energy ω_0 where $\omega_0 < \mu$, we are led to the condition

$$G(\omega_0) = 0. \quad (3)$$

This equation determines ω_0 (at zero separation) as a function of the renormalized coupling constant. When this constant is less than its critical value (no ghosts), ω_0 is real (negative) and single-valued.[18] By subtracting $G(\omega_0)$ from (1), we can express $G^+(W)$ in the useful form

$$G^+(W) = (W - \omega_0)\alpha^+(W), \quad (4)$$

where

$$\alpha^+(W) = G'(\omega_0) + \frac{(W - \omega_0)}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G^+(\omega)\, d\omega}{(\omega - \omega_0)^2(\omega - W - i\epsilon)} \quad (5)$$

and $G'(\omega_0)$ is the derivative of $G^+(W)$ evaluated at ω_0. In this way we have extracted the zero in $G^+(W)$ and obtained the function $\alpha^+(W)$ which has a cut for $\mu < W < \infty$ and no zero or poles in the cut plane. The Fourier transform of the single V propagator, evaluated at the energy $W + m$, is also given by the inverse of a function that has a zero and the same cut. Because of this the $V + N$ propagator in the $2V$ sector plays a role very similar to that of the V-particle propagator in the $V + \theta$ sector. We can show by contour integrations that the functions (4) and

[14] M. S. Maxon, Phys. Rev. **149**, 1273 (1966).
[15] A. Pagnamenta, J. Math. Phys. 7, 356 (1966).
[16] T. Muta, Progr. Theoret. Phys. (Kyoto) 33, 666 (1965).
[17] H. M. Fried, J. Math. Phys. 7, 583 (1966).
[18] Reference 6.

(5) have the following integral representations:

$$\frac{1}{G^+(W)} = \frac{1}{(W-\omega_0)\alpha(\omega_0)} + \frac{1}{\pi}\int_\mu^\infty \mathrm{Im}\left[\frac{1}{G^+(\omega)}\right]\frac{d\omega}{\omega - W - i\epsilon}, \tag{6}$$

$$\alpha^+(W) = Z + \frac{1}{\pi}\int_\mu^\infty \frac{\mathrm{Im}\,\alpha^+(\omega)\,d\omega}{\omega - W - i\epsilon}, \tag{7}$$

where Z, the V particle field operator renormalization constant, is chosen to lie between zero and unity. We also note that

$$\delta m_V = (2\pi Z)^{-1}\int_\mu^\infty \frac{\mathrm{Im}\,G^+(\omega)\,d\omega}{\omega}, \tag{8}$$

$$Z = 1 - \frac{1}{2\pi}\int_\mu^\infty \frac{\mathrm{Im}\,G^+(\omega)\,d\omega}{\omega^2}. \tag{9}$$

If Z_0 is called the normalization constant of the $V+N$ bound-state vector, then $Z_0^{-2} = \alpha(\omega_0)$, and from (5) or (7) we have

$$\alpha(Z') \underset{|Z'|\to\infty}{\longrightarrow} Z\ . \tag{10}$$

Let us now introduce the τ functions representing the $2V$ propagator and the two vertices. These are given by

$$\tau_1(s) = \langle 0|\, T[\psi_V(s)\psi_V(s)\psi_V^+\psi_V^+]\,|0\rangle, \tag{11a}$$

$$\tau_2(s;\omega) = X^{-1}(\omega) \times \langle 0|\, T[\psi_V(s)\psi_N(s)a_k(s)\psi_V^+\psi_V^+]\,|0\rangle, \tag{11b}$$

$$\tau_3(s;\omega,\omega') = X^{-1}(\omega)X^{-1}(\omega') \times \langle 0|\, T[\psi_N(s)\psi_N(s)a_k(s)a_{k'}(s)\psi_V^+\psi_V^+]\,|0\rangle, \tag{11c}$$

$$\tau_4(s;\omega) = X^{-1}(\omega) \times \langle 0|\, T[\psi_V(s)\psi_V(s)\psi_V^+\psi_N^+a_k^+]\,|0\rangle, \tag{11d}$$

$$\tau_5(s;\omega,\omega') = X^{-1}(\omega)X^{-1}(\omega') \times \langle 0|\, T[\psi_V(s)\psi_V(s)\psi_N^+\psi_N^+a_k^+a_{k'}^+]\,|0\rangle, \tag{11e}$$

where $X(\omega)$ is an abbreviation for $f(\omega)/(2\omega)^{\frac{1}{2}}$ and we are quantizing in a box of unit volume. Using the equal time commutation relations and the field equations [as given in Ref. 1 by Eqs. (3) and (18)], we are led to the Matthews–Salam equations:

$$\left(i\frac{d}{ds} - 2m_0\right)\tau_1(s) = \frac{2i}{Z^2}\delta(s) + \frac{2g}{Z}\sum_k X^2(\omega)\begin{Bmatrix}\tau_2(s;\omega)\\ \tau_4(s;\omega)\end{Bmatrix}, \quad \begin{Bmatrix}12a\\13a\end{Bmatrix}$$

$$\left(i\frac{d}{ds} - m_0 - m - \omega\right)\begin{Bmatrix}\tau_2(s;\omega)\\ \tau_4(s;\omega)\end{Bmatrix} = g\tau_1(s) + \frac{g}{Z}\sum_{k'} X^2(\omega')\begin{Bmatrix}\tau_3(s;\omega,\omega')\\ \tau_5(s;\omega,\omega')\end{Bmatrix}, \quad \begin{Bmatrix}12b\\13b\end{Bmatrix}$$

$$\left(i\frac{d}{ds} - 2m - \omega - \omega'\right)\begin{Bmatrix}\tau_3(s;\omega,\omega')\\ \tau_5(s;\omega,\omega')\end{Bmatrix} = 2g\begin{Bmatrix}\tau_2(s;\omega)+\tau_2(s;\omega')\\ \tau_4(s;\omega)+\tau_4(s;\omega')\end{Bmatrix}, \quad \begin{Bmatrix}12c\\13c\end{Bmatrix}$$

where $m_0 = m + \delta m_V$. We see that (12a), (12b), and (12c) form one set of coupled equations for the first three τ functions, and that (13a), (13b), and (13c) form an identical set for the first, fourth, and fifth τ functions. At this stage in the $V+N$ sector only two simple Matthews–Salam equations are needed to obtain the $V+N$ propagator from an algebraic relation. Here we are forced to solve a singular integral equation for the Fourier transform of $\tau_2(s;\omega)$ or, equivalently, $\tau_4(s;\omega)$ prior to finding the Fourier transform of the $2V$ propagator from an algebraic equation. The solution of this singular integral equation will not, however, solve the entire sector, since we have not yet presented all the appropriate τ functions in the $2V$ sector.

Introducing the Fourier representations

$$\tau_\beta(s;\omega,\omega',\cdots) = \frac{i}{2\pi}\int_{-\infty}^{\infty} dW e^{-iWs}\hat{\tau}_\beta(W;\omega,\omega',\cdots), \tag{14a}$$

$$\hat{\tau}_\beta(W;\omega,\omega',\cdots) = \frac{1}{i}\int_{-\infty}^{\infty} ds e^{iWs}\tau_\beta(s;\omega,\omega',\cdots), \tag{14b}$$

and transforming to continuous space, we obtain

$$(W-2m_0)\hat{\tau}_1(W) = 2Z^{-2} + (\pi g Z)^{-1} \times \int_\mu^\infty \mathrm{Im}\,[G^+(\omega)]\hat{\tau}_2(W;\omega)\,d\omega, \tag{15a}$$

$$(W-m_0-m-\omega)\hat{\tau}_2(W;\omega) = g\hat{\tau}_1(W) + (2\pi g Z)^{-1} \times \int_\mu^\infty \mathrm{Im}\,[G^+(\omega')]\hat{\tau}_3(W;\omega,\omega')\,d\omega', \tag{15b}$$

$$(W-2m-\omega-\omega')\hat{\tau}_3(W;\omega,\omega') = 2g[\hat{\tau}_2(W;\omega)+\hat{\tau}_2(W;\omega')], \tag{15c}$$

and an identical set of equations in which $\hat{\tau}_4(W;\omega)$ and $\hat{\tau}_5(W;\omega,\omega')$ replace $\hat{\tau}_2(W;\omega)$ and $\hat{\tau}_3(W;\omega,\omega')$, respectively. From (15c)

$$\hat{\tau}_3(W;\omega,\omega') = \frac{2g[\hat{\tau}_2(W;\omega)+\hat{\tau}_2(W;\omega')]}{W-2m-\omega-\omega'+i\epsilon}. \tag{16}$$

Substituting this expression into (15b), we arrive at the following integral equation:

$$G^+(W-2m-\omega)\hat{\tau}(W;\omega) = gZ\hat{\tau}_1(W) + \frac{1}{\pi}\int_\mu^\infty \frac{\mathrm{Im}\,[G^+(\omega')]\hat{\tau}(W;\omega')\,d\omega'}{W-2m-\omega-\omega'+i\epsilon}, \tag{17}$$

where $\hat{\tau}(W;\omega)$ denotes either $\hat{\tau}_2(W;\omega)$ or $\hat{\tau}_4(W;\omega)$. The inhomogeneous term in (17) is, of course, at this point unknown, but we consider this integral equation as an equation in the variable ω for fixed W, so that $\hat{\tau}_1(W)$ appears as an unknown constant. In carrying out the solution of (17), we obtain the same result for $\hat{\tau}_2(W;\omega)$ and $\hat{\tau}_4(W;\omega)$. This result involves the propagator as an over-all factor and, upon substituting into (15a), we can subsequently remove $\hat{\tau}_1(W)$ from the integral and determine it by solving the resulting algebraic relation. The equality of $\hat{\tau}_2(W;\omega)$ and $\hat{\tau}_4(W;\omega)$ leads to the equality of $\tau_2(s;\omega)$ and $\tau_4(s;\omega)$, which are known from their definitions to coincide at $s=0$. It then follows from (12c) and (13c) that $\tau_3(s;\omega,\omega')=\tau_5(s;\omega,\omega')$, since $\tau_3(0;\omega,\omega')=0=\tau_5(0;\omega,\omega')$.

Letting $W\to W+2m$, introducing the definition

$$J^-(W;\omega)\equiv J(W;\omega-i\epsilon)=\frac{G^+(W-\omega)\hat{\tau}(W+2m;\omega)}{gZ\hat{\tau}_1(W+2m)}, \quad (18)$$

and extending ω into the complex Z' plane, we see that Eq. (17) becomes

$$J(W;Z')=1-\frac{1}{\pi}\int_\mu^\infty\frac{\operatorname{Im}[G^+(\omega')]J^-(W;\omega')\,d\omega'}{(\omega'-W+Z')G^+(W-\omega')}, \quad (19)$$

where

$$J^-(W;\omega)\equiv\lim_{Z'\to\omega-i\epsilon}J(W;Z'). \quad (20)$$

The singular integral equation that arises in the eigenvalue approach to the $2V$ bound-state problem can be deduced from Eq. (19) by letting W become the $2V$ potential energy. We shall consider this formulation of the $2V$ sector in a future communication. Equations of the type (19) with a perfectly general inhomogeneous term have been studied by Maxon.[14] In our case, however, we are involved with the $V+N$ function G rather than the V particle function (also called G in Ref. 14). Applying his method of solution, we find

$$J(W;Z')=\frac{G(W-Z')}{Z[1+\alpha(\omega_0)G^+(W-\omega_0)I_W^+(W-\omega_0)]}\times\left\{\frac{1}{W-\omega_0-Z'}+\frac{\alpha(\omega_0)G^+(W-\omega_0)}{Z'-\omega_0}\times[I_W(W-Z')-I_W^+(W-\omega_0)]\right\}, \quad (21)$$

where

$$I_W(Z')=\frac{1}{\pi}\int_\mu^\infty\operatorname{Im}\left[\frac{1}{G^+(\omega)}\right]\frac{d\omega}{(\omega-Z')\alpha^+(W-\omega)}. \quad (22)$$

An integral corresponding to $I_W(Z')$ was first introduced in the $V+\theta$ sector,[9] and, in analogy with that case, we find the identity

$$(\omega_0-Z')I_W(Z')+(\omega_0-W+Z')I_W(W-Z')=\frac{W-2\omega_0}{\alpha(\omega_0)G^+(W-\omega_0)}+(2\omega_0-W)I_W^+(W-\omega_0)-\frac{(Z'-\omega_0)(W-Z'-\omega_0)}{G(Z')G(W-Z')}. \quad (23)$$

This relation is useful in verifying that Eq. (21) does indeed satisfy Eq. (19) and in simplifying the matrix elements that arise in Sec. III. From Eqs. (18), (20), and (21) we are led to the solution of Eq. (17):

$$\hat{\tau}(W+2m;\omega)=\hat{\tau}_2(W+2m;\omega)=\hat{\tau}_4(W+2m;\omega)=\frac{g\hat{\tau}_1(W+2m)}{[1+\alpha(\omega_0)G^+(W-\omega_0)I_W^+(W-\omega_0)]}\times\left\{\frac{1}{W-\omega-\omega_0+i\epsilon}+\frac{\alpha(\omega_0)G^+(W-\omega_0)}{\omega-\omega_0}\times[I_W^+(W-\omega)-I_W^+(W-\omega_0)]\right\}. \quad (24)$$

When this result is combined with Eq. (15a), we solve for the $2V$ propagator and express it in the form (see Appendix)

$$\tau_1(W+2m)=\frac{4[1+\alpha(\omega_0)G^+(W-\omega_0)I_W^+(W-\omega_0)]}{D(W)}, \quad (25)$$

where

$$D(W)\equiv Z^2(W-2\delta m_V)\times[1+\alpha(\omega_0)G^+(W-\omega_0)I_W^+(W-\omega_0)]+\alpha(\omega_0)G^+(W-\omega_0). \quad (26)$$

The calculation of the $\hat{\tau}$ functions in this section is now complete. We have resolved the $2V$ propagator in terms of (a) quantities and functions that arise in the V and $V+N$ sectors and (b) the integral I_W. This result finishes off the other τ functions, as shown by Eqs. (24) and (16)

In order to obtain an eigenvalue equation for the determination of the $2V$ potential energy, it is necessary to consider the analytic properties of the $2V$ propagator. For this purpose we insert Eq. (11a) and a complete set of intermediate states into Eq. (14b) to get

$$\hat{\tau}_1(W)=\frac{2Z_B^2}{W-E_B+i\epsilon}+\sum_n\frac{|\langle 0|\,\psi_V\psi_V\,|n\rangle|^2}{W-E_n+i\epsilon}, \quad (27)$$

where Z_B is the normalization constant of the $2V$ bound state (assuming only one such state) and $|n\rangle$

refers to all other states having the same quantum numbers as two V particles. We see that the propagator has a simple pole at the total bound-state energy $W = E_B \equiv 2m + \omega_B$. Thus it follows from Eq. (26) that

$$D(\omega_B) = Z^2(\omega_B - 2\delta m_V)[1 - \alpha(\omega_0)G(\omega_B - \omega_0) \times A(\omega_B - \omega_0)] + \alpha(\omega_0)G(\omega_B - \omega_0) = 0, \quad (28)$$

where, more generally,

$$A(\omega) \equiv -I^{+}_{\omega+\omega_0}(\omega) = \frac{1}{\pi}\int_{\mu}^{\infty} \mathrm{Im}\left[\frac{1}{G^{+}(\omega')}\right] \frac{d\omega'}{G^{+}(\omega + \omega_0 - \omega')}. \quad (29)$$

Equation (28) determines the $2V$ potential energy ω_B (recall the lack of recoil in the model) as a function of the renormalized coupling constant, and, of course, the $V + N$ potential energy ω_0 is obtained from (3). In Eq. (28) $G(\omega_B - \omega_0)$ and $A(\omega_B - \omega_0)$ are real integrals, since the stability of the $2V$ system requires that $\omega_B < \omega_0 + \mu < 2\mu$. We see from Eq. (29) that $A(\omega)$ has the branch cuts $\mu \leq \omega \leq \infty$ and $2\mu - \omega_0 \leq \omega \leq \infty$. These cuts originate from the $V + N + \theta$ and $2N + 2\theta$ intermediate states, respectively. At the pole $W = E_B$, $\hat{\tau}_1(W)$ has the residue $2Z_B^2$, and consequently we can use

$$2Z_B^2 = \lim_{W \to \omega_B} (W - \omega_B)\hat{\tau}_1(W + 2m) \quad (30)$$

to find Z_B. The $\hat{\tau}$ functions that have been derived in this section can also be used to obtain the other expansion coefficients in the $2V$ bound-state vector; but since the chief purpose of this paper is to pursue the formal LSZ solution of the $2V$ sector, we do not calculate these coefficients nor do we go into a detailed study of Eq. (28). These problems will also be present in the eigenvalue treatment of this sector, and we reserve them for that investigation. To simplify the actual analysis of Eq. (28) one could set the cutoff function equal to unity and use nonrelativistic θ particles. It should also be possible to analyze the $2V$ bound-state spectrum when the probability Z is less than zero.

III. SCATTERING AND PRODUCTION PROCESSES

So far we have seen that the $2V$ propagator and two vertex functions form a distinct problem in the LSZ approach to the $2V$ sector. In this section we consider another set of τ functions that require a separate procedure. These functions give rise to a new system of coupled Matthews–Salam equations, and the solution to one singular integral equation yields the $\hat{\tau}$ functions corresponding to $V + N + \theta$ and

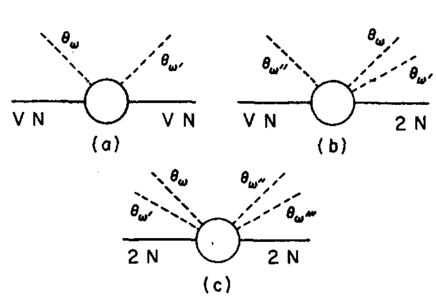


FIG. 1. Diagrams corresponding to (a) $V + N + \theta$ elastic scattering, (b) Production process $V + N + \theta \to 2N + 2\theta$, and (c) $2N + 2\theta$ elastic scattering.

$2N + 2\theta$ elastic scattering and the production process $V + N + \theta \to 2N + 2\theta$. The S-matrix elements for these transitions are given by

$$S^{V+N+\theta}_{k'k} = \delta_{kk'} - \alpha(\omega_0)X(\omega)X(\omega')\int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} dt \times e^{i(2m+\omega_0+\omega)t'}\vec{\mathfrak{D}}(t';\omega)\tau_6(t'-t;\omega,\omega') \times \overleftarrow{\mathfrak{D}}^*(t;\omega')e^{-i(2m+\omega_0+\omega')t}, \quad (31)$$

$$P_{k'';kk'} = -\tfrac{1}{2}\alpha^{\frac{1}{2}}(\omega_0)X(\omega)X(\omega')X(\omega'')\int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} dt \times e^{i(2m+\omega+\omega')t'}\vec{\mathfrak{D}}(t';\omega-\omega_0+\omega')\tau_8(t'-t;\omega,\omega',\omega'') \times \overleftarrow{\mathfrak{D}}^*(t;\omega'')e^{-i(2m+\omega_0+\omega'')t}, \quad (32)$$

$$S^{2N+2\theta}_{k''k''';k'k} = \frac{\delta_{kk''}\delta_{k'k'''} + \delta_{kk'''}\delta_{k'k''}}{2} - \frac{X(\omega)X(\omega')X(\omega'')X(\omega''')}{4}\int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} dt \times e^{i(2m+\omega+\omega')t'}\vec{\mathfrak{D}}(t';\omega-\omega_0+\omega') \times \tau_9(t'-t;\omega,\omega',\omega'',\omega''') \times \overleftarrow{\mathfrak{D}}^*(t;\omega''-\omega_0+\omega''')e^{-i(2m+\omega''+\omega''')t}, \quad (33)$$

where

$$\mathfrak{D}(t;\omega) = [i(d/dt) - 2m - \omega_0 - \omega].$$

The factors $\alpha(\omega_0)$ and $\alpha^{\frac{1}{2}}(\omega_0)$ originate in the bound-state contractions. These processes are depicted in Fig. 1 where the solid lines in the diagrams are used to symbolize two heavy particles. Actually, there are four appropriate τ functions involved here and they are given as follows:

$$\tau_6(s;\omega,\omega') = X^{-1}(\omega)X^{-1}(\omega') \times \langle 0|\, T[\psi_V(s)\psi_N(s)a_k(s)\psi_V^{+}\psi_N^{+}a_{k'}^{+}]\, |0\rangle, \quad (34a)$$

$$\tau_7(s;\omega,\omega',\omega'') = X^{-1}(\omega)X^{-1}(\omega')X^{-1}(\omega'') \times \langle 0|\, T[\psi_V(s)\psi_N(s)a_k(s)\psi_N^{+}\psi_N^{+}a_{k'}^{+}a_{k''}^{+}]\, |0\rangle, \quad (34b)$$

$$\tau_8(s;\omega,\omega',\omega'') = X^{-1}(\omega)X^{-1}(\omega')X^{-1}(\omega'') \times \langle 0|\, T[\psi_N(s)\psi_N(s)a_k(s)a_{k'}(s)\psi_V^{+}\psi_N^{+}a_{k''}^{+}]\, |0\rangle, \quad (34c)$$

$$\tau_9(s;\omega,\omega',\omega'',\omega''') = X^{-1}(\omega)X^{-1}(\omega')X^{-1}(\omega'')X^{-1}(\omega''') \times \langle 0|\, T[\psi_N(s)\psi_N(s)a_k(s)a_{k'}(s)\psi_N^{+}\psi_N^{+}a_{k''}^{+}a_{k'''}^{+}]\, |0\rangle. \quad (34d)$$

As before we use the equal time commutation relations and the field equations to establish the corresponding Matthews–Salam equations:

$$\left(i\frac{d}{ds} - m_0 - m - \omega\right)\tau_6(s;\omega,\omega') = \frac{i}{Z}\,\delta(s)\,\delta_{kk'}X^{-2}(\omega) + g\tau_4(s;\omega') + \frac{g}{Z}\sum_{k''} X^2(\omega'')\tau_8(s;\omega'',\omega,\omega'), \quad (35a)$$

$$\left(i\frac{d}{ds} - 2m - \omega' - \omega''\right)\tau_7(s;\omega,\omega',\omega'') = 2g[\tau_6(s;\omega,\omega') + \tau_6(s;\omega,\omega'')], \quad (35b)$$

$$\left(i\frac{d}{ds} - 2m - \omega - \omega'\right)\tau_8(s;\omega,\omega',\omega'') = 2g[\tau_6(s;\omega,\omega'') + \tau_6(s;\omega',\omega'')], \quad (35c)$$

$$\left(i\frac{d}{ds} - 2m - \omega - \omega'\right)\tau_9(s;\omega,\omega',\omega'',\omega''') = 2i\delta(s)X^{-2}(\omega)X^{-2}(\omega')[\delta_{kk''}\delta_{k'k'''} + \delta_{k'k''}\delta_{kk'''}] + 2g[\tau_7(s;\omega,\omega'',\omega''') + \tau_7(s;\omega',\omega'',\omega''')]. \quad (35d)$$

In the LSZ solution of the V and $V + N$ sectors, it was found from the Matthews–Salam equations that the τ functions for $N + \theta$ and $2N + \theta$ scattering were symmetric under interchange of the initial and final ω variables. In both of these cases, this basically follows from the very close relationship between the scattering amplitude, the propagator and the vertex function—a relationship that is characteristic of the sectors $nN + \theta$ ($n = 1, 2, 3, \cdots$) when the separation parameters are all taken equal to zero.[19] Since the interaction effects are more complicated in the $2V$ sector, we do not expect equally simple connections between the $V + N + \theta$ scattering amplitude, the $2V$ propagator, and the vertex functions. The expression for $\hat{\tau}_6$ [see (43)] shows that it is not symmetric in ω and ω'. It is because of this that the only obvious symmetry properties of τ_7, τ_8, and τ_9 are those due to the interchange of ω variables in the initial or final states. A similar situation exists in the $V + \theta$ case which has a system of Matthews–Salam equations analogous to those given above except for the τ_4 term in Eq. (35a). This term embodies the $2V$ interaction and its effects will appear in all three transition amplitudes. Carrying out the Fourier transformation as before, we have

$$(W - m_0 - m - \omega)\hat{\tau}_6(W;\omega,\omega') = \frac{\delta_{kk'}}{Z}X^{-2}(\omega) + g\hat{\tau}_4(W;\omega') + \frac{g}{Z}\sum_{k''} X^2(\omega'')\hat{\tau}_8(W;\omega'',\omega,\omega'), \quad (36a)$$

$$(W - 2m - \omega - \omega')\hat{\tau}_8(W;\omega,\omega',\omega'') = 2g[\hat{\tau}_6(W;\omega,\omega'') + \hat{\tau}_6(W;\omega',\omega'')], \quad (36b)$$

$$(W - 2m - \omega' - \omega'')\hat{\tau}_7(W;\omega,\omega',\omega'') = 2g[\hat{\tau}_6(W;\omega,\omega') + \hat{\tau}_6(W;\omega,\omega'')], \quad (36c)$$

$$(W - 2m - \omega - \omega')\hat{\tau}_9(W;\omega,\omega',\omega'',\omega''') = 2X^{-2}(\omega)X^{-2}(\omega')[\delta_{kk''}\delta_{k'k'''} + \delta_{k'k''}\delta_{kk'''}] + 2g[\hat{\tau}_7(W;\omega,\omega'',\omega''') + \hat{\tau}_7(W;\omega',\omega'',\omega''')]. \quad (36d)$$

From Eq. (36b) we get the relation

$$\hat{\tau}_8(W;\omega'',\omega,\omega') = \frac{2g[\hat{\tau}_6(W;\omega'',\omega') + \hat{\tau}_6(W;\omega,\omega')]}{W - 2m - \omega - \omega'' + i\epsilon}. \quad (37)$$

Substituting Eq. (37) into Eq. (36a) and letting $W \to W + 2m$, we are led to the following singular integral equation for $\hat{\tau}_6$:

$$G^+(W-\omega)\hat{\tau}_6(W+2m;\omega,\omega') = \delta_{kk'}X^{-2}(\omega) + gZ\hat{\tau}_4(W+2m;\omega') + 2g^2\sum_{k''}\frac{X^2(\omega'')\hat{\tau}_6(W+2m;\omega'',\omega')}{W - \omega'' - \omega - i\epsilon}. \quad (38)$$

Next, we introduce the definitions

$$L^-(W;\omega) \equiv L(W;\omega - i\epsilon) \equiv (Z/2g)G^+(W-\omega)\hat{\tau}_4(W+2m;\omega), \quad (39)$$

$$T^-(W;\omega,\omega') \equiv T(W;\omega - i\epsilon,\omega') \equiv (1/2g^2)G^+(W-\omega') \times [G^+(W-\omega)\hat{\tau}_6(W+2m;\omega,\omega') - \delta_{kk'}X^{-2}(\omega')], \quad (40)$$

and transform to continuous space to obtain (38) in the form

$$T(W;Z',\omega') = L^-(W;\omega') + \frac{1}{W - Z' - \omega'} - \frac{1}{\pi}\int_u^\infty \frac{\operatorname{Im}[G^+(\omega'')]T^-(W;\omega'',\omega')\,d\omega''}{(\omega'' + Z' - W)G^+(W-\omega'')}, \quad (41)$$

where

$$T^-(W;\omega,\omega') = \lim_{Z'\to\omega-i\epsilon} T(W;Z',\omega'). \quad (42)$$

With W and ω' fixed, we consider this equation as an equation in Z'. Thus the first term on the right-hand side of Eq. (41) is treated as a known constant. The singular integral equation that originates in the eigenvalue approach to $V + N + \theta$ scattering is readily deduced from (41) by letting $W \to \omega' + \omega_0$ and $Z' \to \omega - i\epsilon$. Under these conditions there is a direct relationship between $T(W;Z',\omega')$ and the expansion coefficients of the bare $V + N + \theta$ states.

[19] L. M. Scarfone, Nuovo Cimento **48**, 84 (1967).

$L^-(W;\omega')$ essentially becomes the coefficient of the $2V$ bare state that occurs in the physical $V+N+\theta$ scattering states.

Equation (41) is solved with the same method used for (19) and then from Eqs. (39), (40), and (42) we obtain the result

$$\begin{aligned}
\hat{\tau}_6(W+2m;\omega,\omega') &= \frac{\delta_{kk'}X^{-2}(\omega)}{G^+(W-\omega')} + \frac{2g^2}{(W-\omega-\omega_0)(W-\omega'-\omega_0)} \\
&\times \Bigg(\frac{1}{W-\omega-\omega'+i\epsilon}\left[(\omega_0-\omega')I_W^+(\omega') + \frac{1}{\alpha^+(\omega')\alpha^+(W-\omega')}\right] \\
&+ \frac{(W-\omega'-\omega_0)^2 I_W^+(W-\omega')}{(\omega_0-\omega')(\omega-\omega'-i\epsilon)} - \frac{(W-\omega-\omega_0)^2(W-2\omega')I_W^+(W-\omega)}{(\omega_0-\omega')(\omega-\omega'-i\epsilon)(W-\omega-\omega'+i\epsilon)} \\
&\times \frac{-\alpha(\omega_0)G^+(W-\omega_0)}{(\omega_0-\omega)(\omega_0-\omega')}\left[\frac{(W-\omega-\omega_0)I_W^+(W-\omega)+(2\omega_0-W)I_W^+(W-\omega_0)}{1+\alpha(\omega_0)G^+(W-\omega_0)I_W^+(W-\omega_0)}\right] \\
&\times \left[(\omega_0-\omega')I_W^+(\omega') + (W-\omega'-\omega_0)I_W^+(W-\omega') + \frac{1}{\alpha^+(\omega')\alpha^+(W-\omega')}\right] \\
&+ \frac{(W-\omega'-\omega_0)\hat{\tau}_4(W+2m;\omega')}{2g} \\
&\times \left\{1 - \frac{\alpha(\omega_0)G^+(W-\omega_0)}{\omega_0-\omega}\left[\frac{(W-\omega-\omega_0)I_W^+(W-\omega)+(2\omega_0-W)I_W^+(W-\omega_0)}{1+\alpha(\omega_0)G^+(W-\omega_0)I_W^+(W-\omega_0)}\right]\right\}\Bigg).
\end{aligned} \tag{43}$$

It is evident, as stated previously, that $\hat{\tau}_6$ is not symmetric in ω and ω'. A comparison between our solution (43) and the one given in Ref. 14 for $\hat{\tau}^5$ shows an obvious correspondence of terms, except, of course, for those involving $\hat{\tau}_4$ which do not occur in the latter case. The entire expression for $\hat{\tau}_6$ is formulated in terms of simple energy factors, the integral I_W, and functions and quantities that arise in the V and $V+N$ sectors. From Eqs. (36c) and (36d) we see that

$$\hat{\tau}_7(W+2m;\omega,\omega',\omega'') = 2g\,\frac{[\hat{\tau}_6(W+2m;\omega,\omega')+\hat{\tau}_6(W+2m;\omega,\omega'')]}{W-\omega'-\omega''+i\epsilon}, \tag{44}$$

$$\begin{aligned}
\hat{\tau}_9(W+2m;\omega,\omega',\omega'',\omega''') &= \frac{2X^{-2}(\omega)X^{-2}(\omega')[\delta_{kk''}\delta_{k'k'''}+\delta_{k'k''}\delta_{kk'''}]}{W-\omega-\omega'+i\epsilon} \\
&+ 4g^2\frac{[\hat{\tau}_6(W+2m;\omega,\omega'')+\hat{\tau}_6(W+2m;\omega,\omega''')+\hat{\tau}_6(W+2m;\omega',\omega'')+\hat{\tau}_6(W+2m;\omega',\omega'')]}{(W-\omega-\omega'+i\epsilon)(W-\omega''-\omega'''+i\epsilon)}.
\end{aligned} \tag{45}$$

This completes the determination of the $\hat{\tau}$ functions, and now we proceed to calculate the S-matrix elements. For $V+N+\theta$ elastic scattering we use Eqs. (14a), (31), and (43) to get

$$\begin{aligned}
S_{k'k}^{V+N+\theta} &= \delta_{kk'} + 4\pi i g^2\delta(\omega-\omega')X^2(\omega)\alpha(\omega_0) \\
&\quad\left\{\frac{1}{\alpha(\omega_0)G^+(\omega)}\left[\frac{1+\alpha(\omega_0)G^+(\omega)A(\omega)}{1-\alpha(\omega_0)G^+(\omega)A(\omega)}\right] - \frac{2}{[1-\alpha(\omega_0)G^+(\omega)A(\omega)]D(\omega+\omega_0)}\right\} \\
&= \delta_{kk'} + 4\pi i g^2\delta(\omega-\omega')X^2(\omega)\left[\frac{Z^2(\omega+\omega_0-2\,\delta m_V)[1+\alpha(\omega_0)G^+(\omega)A(\omega)]-\alpha(\omega_0)G^+(\omega)}{G^+(\omega)D(\omega+\omega_0)}\right].
\end{aligned} \tag{46}$$

In carrying out the integrations in Eq. (31), we find that $W\to\omega+\omega_0=\omega'+\omega_0$, which means that terms in $\hat{\tau}_6$ containing the factors $(W-\omega-\omega_0)$ and $(W-\omega'-\omega_0)$ give no contribution. Note, however, by Eq. (24), that $(W-\omega'-\omega_0)\hat{\tau}_4(W+2m;\omega')$ is nonvanishing. The first term in the curly brackets of Eq. (46) resembles the result found for $V+\theta$ elastic scattering and the second term is proportional to the $2V$ propagator as shown by Eq. (25). In the $V+\theta$ case, a denominator similar to $1-\alpha G^+A$ leads to a condition for the $V+\theta$ bound state. Here, however, this denominator is eliminated when Eq. (26) is used to obtain the second equality in Eq. (46).

Turning to the production amplitude (32), we employ Eqs. (14a), (37), and (43) to derive

$$P_{k'';kk'} = 8\pi i g^3 X(\omega)X(\omega')X(\omega'')\delta(\omega_0 + \omega'' - \omega - \omega')\alpha^{\frac{1}{2}}(\omega_0)$$
$$\times \left\{\frac{1}{G^+(\omega)G^+(\omega')[1 - \alpha(\omega_0)G^+(\omega'')A(\omega'')]} - \frac{\alpha(\omega_0)G^+(\omega'')}{G^+(\omega)G^+(\omega')D(\omega'' + \omega_0)[1 - \alpha(\omega_0)G^+(\omega'')A(\omega'')]}\right\}$$
$$= \frac{8Z^2\pi i\alpha^{\frac{1}{2}}(\omega_0)g^3X(\omega)X(\omega')X(\omega'')\delta(\omega_0 + \omega'' - \omega - \omega')(\omega'' + \omega_0 - 2\delta m_V)}{G^+(\omega)G^+(\omega')D(\omega'' + \omega_0)}. \quad (47)$$

In arriving at this result, we have used $W = \omega'' + \omega_0 = \omega + \omega'$ which gives less simplification in $\hat{\tau}_6$ than the W of the previous case; thus more algebra is required. The first term in the curly brackets is similar to the production amplitude in the $V + \theta$ sector. The final form of Eq. (47) comes from using Eq. (26), and it contains a factor for the $2V$ interaction effects and a factor $[G^+(\omega)G^+(\omega')]^{-1}$ describing the final state scattering of the θ particles by the $2N$ sources.

Finally, we consider the scattering process (33) which, with the help of Eqs. (14a), (43), and (45), becomes

$$S^{2N+2\theta}_{k'''k'';k'k} = \frac{\delta_{kk''}\delta_{k'k'''} + \delta_{kk'''}\delta_{k'k''}}{2} - 2\pi i g^2\delta(\omega + \omega' - \omega'' - \omega''')X(\omega)X(\omega')X(\omega'')X(\omega''')$$
$$\times \left[\frac{X^{-2}(\omega)}{G^+(\omega')}(\delta_{kk''} + \delta_{kk'''}) + \frac{X^{-2}(\omega')}{G^+(\omega)}(\delta_{k'k''} + \delta_{k'k'''})\right]$$
$$+ 8\pi i g^4\delta(\omega + \omega' - \omega'' - \omega''')X(\omega)X(\omega')X(\omega'')X(\omega''')$$
$$\times \left\{\frac{\alpha(\omega_0)G^+(\omega + \omega' - \omega_0)}{G^+(\omega)G^+(\omega')G^+(\omega'')G^+(\omega''')[1 - \alpha(\omega_0)G^+(\omega + \omega' - \omega_0)A(\omega + \omega' - \omega_0)]}\right.$$
$$\left. - \frac{[\alpha(\omega_0)G^+(\omega + \omega' - \omega_0)]^2}{G^+(\omega)G^+(\omega')G^+(\omega'')G(\omega''')D(\omega + \omega')[1 - \alpha(\omega_0)G^+(\omega + \omega' - \omega_0)A(\omega + \omega' - \omega_0)]}\right\}. \quad (48)$$

The algebraic manipulations involved in reaching this result are more complicated than in the other two cases because of the lack of any initial simplification in (43) for $W = \omega + \omega' = \omega'' + \omega'''$. The Kronecker deltas in the first part of Eq. (48) denote the occurrence of no scattering at all and the next part (in square brackets) refers to the elastic scattering of one θ particle by the $2N$ sources while the other θ is unscattered. The first term in the curly brackets has the same form as the $N + 2\theta$ scattering amplitude, and when Eq. (26) is used the entire curly-bracketed expression becomes

$$\frac{Z^2(\omega + \omega' - 2\delta m_V)\alpha(\omega_0)G^+(\omega + \omega' - \omega_0)}{G^+(\omega)G^+(\omega')G^+(\omega'')G^+(\omega''')D(\omega + \omega')}. \quad (49)$$

Once again, the denominator $1 - \alpha G^+A$ is eliminated. In Eq. (49) there are factors describing the initial and final scattering of the θ particles by $2N$ particles and a factor for the intermediate $2V$ effects.

IV. CONCLUDING REMARKS

We have found that two singular integral equations arise in the LSZ approach to the $2V$ sector of the Lee model with boson sources and that their solutions determine all the appropriate $\hat{\tau}$ functions. These functions separate into two distinct, yet coupled, sets, one of which contains the $2V$ propagator and related vertex functions, the other embodies the $\hat{\tau}$ functions corresponding to scattering and production processes. Each function in the latter set gets coupled to the $2V$ interaction through the vertex $2V \rightleftarrows V + N + \theta$. Accordingly, all three amplitudes (46), (47), and (48) received the factor D^{-1} where D is the denominator of the $2V$ propagator. There are poles in these amplitudes attributed to the vanishing of D and G in the denominators and branch cuts due to G and A. The interpretation of the amplitudes parallels that of the $V + \theta$ subspace. The appearance of D^{-1} in the final form of each amplitude in place of a denominator like $1 - \alpha G^+A$ is due to the formation of an intermediate $2V$ state rather than a $V + N + \theta$ discrete state.

It should be reasonably straightforward to extend our considerations to the higher sectors $2V + nN$ ($n = 1, 2, 3, \cdots,$) since these also contain only two θ particles. Now that we have found the solution to the scattering processes $V + N + \theta$ and $2N + 2\theta$, we can consider the next most complicated two source case, namely $2V + \theta$ scattering and other related processes. We can also think of using heavy-particle sources with nonzero separation parameters. In this regard we have already carried out the exactly soluble and instructive problem of the scattering of a single θ by two noncoincident N particles. Finally, we propose to investigate a dispersion relations calculation of the $2V$ bound-state energy condition (25) starting with the vertex function $\langle V|f_V|B\rangle$ where f_V is the V particle current operator, $|V\rangle$ the physical one particle V state, and $|B\rangle$ the $2V$ bound state. Blankenbecker and Cook[20] have pointed out that many of the properties of the bound state can be studied in this way, and our example may be the first instance of an exact dispersion-relations calculation involving the exchange of two mesons.

APPENDIX

To derive (25), we substitute (24) into (15a) and obtain

$$\begin{aligned}\hat{\tau}_1(W+2m) &= (2/Z)[1+\alpha(\omega_0)G^+(W-\omega_0)\\ &\times I_W^+(W-\omega_0)]\Big\{Z(W-2\delta m_V)[1+\alpha(\omega_0)\\ &\times G^+(W-\omega_0)I_W^+(W-\omega_0)]\\ &+\frac{1}{\pi}\int_\mu^\infty\frac{\operatorname{Im} G^+(\omega)\,d\omega}{\omega-(W-\omega_0)-i\epsilon}\\ &+\alpha(\omega_0)G^+(W-\omega_0)I_W^+(W-\omega_0)\\ &\times\frac{1}{\pi}\int_\mu^\infty\frac{\operatorname{Im} G^+(\omega)\,d\omega}{\omega-\omega_0}-\alpha(\omega_0)\omega^+(W-\omega_0)\\ &\times\frac{1}{\pi}\int_\mu^\infty\frac{\operatorname{Im}[G^+(\omega)]I_W^+(W-\omega)\,d\omega}{\omega-\omega_0}\Big\}^{-1}. \qquad \text{(A1)}\end{aligned}$$

The first two integrals in the curly brackets are evaluated with the relation

$$\frac{1}{\pi}\int_\mu^\infty\frac{\operatorname{Im} G^+(\omega)\,d\omega}{\omega-W-i\epsilon} = G^+(W)+Z\delta m_V - ZW, \qquad \text{(A2)}$$

which follows from Eqs. (1), (8), and (9). Thus, with the help of Eq. (3), the denominator in (A1) becomes

$$\begin{aligned}&Z(W-\omega_0-\delta m_V)\alpha(\omega_0)G^+(W-\omega_0)I_W^+(W-\omega_0)\\ &+Z(\omega_0-\delta m_V)+G^+(W-\omega_0)\\ &-\alpha(\omega_0)G^+(W-\omega_0)\\ &\times\frac{1}{\pi}\int_\mu^\infty\frac{\operatorname{Im}[G^+(\omega)]I_W^+(W-\omega)\,d\omega}{\omega-\omega_0}. \qquad \text{(A3)}\end{aligned}$$

If we let $C(W)$ be an abbreviation for the integral in Eq. (A3), including the factor $(1/\pi)$, then from Eqs. (22) and (7), and one elementary contour integration we obtain

$$C(W) = \frac{Z-\alpha(\omega_0)}{Z\alpha(\omega_0)} - \frac{Z}{\pi}\int_\mu^\infty \operatorname{Im}\left[\frac{1}{G^+(\omega)}\right]\frac{d\omega}{\alpha^+(W-\omega)}. \qquad \text{(A4)}$$

Now by doing a partial fraction decomposition in $I_W^+(W-\omega_0)$ we learn that

$$\begin{aligned}-\frac{1}{\pi}\int_\mu^\infty \operatorname{Im}\left[\frac{1}{G^+(\omega)}\right]\frac{d\omega}{\alpha^+(W-\omega)} &= (W-2\omega_0)I_W^+(W-\omega_0)\\ &+\frac{1}{\pi}\int_\mu^\infty \operatorname{Im}\left[\frac{1}{\alpha(\omega)}\right]\frac{d\omega}{G^+(W-\omega)}. \qquad \text{(A5)}\end{aligned}$$

The integral on the left-hand side of Eq. (A5) is the integral that appears in Eq. (A4). At this point, we make use of Eq. (6) in the integral on the right-hand side of Eq. (A5), and after some straightforward contour integrations, Eq. (A5) becomes

$$\begin{aligned}-\frac{2}{\pi}\int_\mu^\infty \operatorname{Im}\left[\frac{1}{G^+(\omega)}\right]\frac{d\omega}{\alpha^+(W-\omega)} &= (W-2\omega_0)I_W^+(W-\omega_0)\\ &+\frac{\alpha(\omega_0)\alpha^+(W-\omega_0)-Z^2}{Z^2\alpha(\omega_0)\alpha^+(W-\omega_0)}. \qquad \text{(A6)}\end{aligned}$$

Returning to Eq. (A4) with this result, and then back to Eqs. (A3) and (A1), we get Eq. (25).

[20] R. Blankenbecker and L. F. Cook, Phys. Rev. 119, 1745 (1960).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 JANUARY 1968

High-Energy Scattering for Yukawa Potentials

H. J. W. MÜLLER* AND K. SCHILCHER†
Department of Physics, American University of Beirut, Beirut, Lebanon

(Received 22 May 1967)

A reexamination is made of nonrelativistic scattering by a superposition of Yukawa potentials in the frame of a high-energy perturbation method developed previously by one of the authors. It is found that for Yukawa potentials expandable in ascending powers of r, the S matrix may be written in a remarkably simple form. A high-energy asymptotic expansion is obtained for the phase shift. The residue function at the Regge poles is found to have the same general form as for a Coulomb potential and is formally independent of any high-energy approximation.

1. INTRODUCTION

In recent years some high-energy experiments have motivated an intense study of the high-energy behavior of scattering amplitudes. Particular attention has always been paid to potential theory because there the scattering amplitude can be calculated more easily than in field theory or S-matrix theory. Moreover, potential theory arises frequently as a simplified version of a more ambitious theory and often yields results which, even in relativistic cases, are as good as any other approximations. Also, considerable insight into the analytic nature of scattering amplitudes has been gained from a thorough study of potential theory.

The Yukawa potential is of particular interest because, of all well-known potentials, it is the only one which corresponds directly to the exchange of an elementary particle. It is also the only potential for which (so far) a Mandelstam representation of the scattering amplitude has been shown to exist.

In previous papers[1,2] one of us used a powerful perturbation procedure for an examination of the scattering by a superposition of Yukawa potentials. Following Mandelstam[3] and Lovelace and Masson,[4] the over-all potential was assumed to be expandable in a series of ascending powers of r. However, some of the results obtained—as, for instance, the expression for the nonrelativistic S matrix—look clumsy and complicated. Here we show by a reexamination that these complications are only apparent. Expressed more precisely, we show by an investigation of the first seven terms of the high-energy expansions of the solutions and eigenvalues of the Schrödinger equation, that these complications do in fact cancel out. Moreover, it becomes obvious that this cancellation must occur to any order of approximation.

* Present address: Sektion Physik, Universität München, Lehrstuhl Prof. Bopp, München.

† Partially supported by Research Corporation grant.

[1] H. J. W. Müller, Ann. Physik **15**, 395 (1965).

[2] H. J. W. Müller, Ann. Physik **16**, 255 (1965); Physica **31**, 688 (1965); Z. Physik **183**, 402 (1965); **186**, 79 (1965); **198**, 59 (1967).

[3] S. Mandelstam, Ann. Phys. (N.Y.) **19**, 254 (1962).

[4] C. Lovelace and D. Masson, Nuovo Cimento **26**, 472 (1962).

In Sec. 2 we introduce our notation and basic definitions. In Sec. 3 we recalculate the fundamental Jost solution from which the Jost function, and hence the S matrix, may be derived. We present these results and the expansion for the eigenvalues in sufficient detail to facilitate programming for computers. In Sec. 4 we derive the S matrix and calculate phase shifts and residues. These results are found to be remarkably simple and completely analogous to corresponding formulas for the Coulomb potential.[5] In particular, the expression for the residue is found to be independent of any high-energy approximation.

2. SOLUTIONS OF THE RADIAL SCHRÖDINGER EQUATION

The radial Schrödinger equation for the partial wave $\psi_l(k, r)$ corresponding to a potential $V(r)$ is

$$\left[d^2/dr^2 + k^2 - \frac{l(l+1)}{r^2} - V(r)\right]\psi_l(k, r) = 0, \quad (2.1)$$

where $E = k^2$ and $\hbar = c = 1 = 2m$, m being the reduced mass of the system. Previously[1] we derived four high-energy asymptotic solutions of this equation subject to the boundary conditions

$$\left.\begin{aligned}\psi_1^N(l, k; r) &\sim (kr)^{l+1}\\ \psi_2^N(l, k; r) &\sim (kr)^{-l}\end{aligned}\right\} \quad \text{for} \quad |r| \to \infty \quad (2.2a)$$

and

$$\left.\begin{aligned}\psi_3^N(l, k; r) &\sim e^{ikr}\\ \psi_4^N(l, k; r) &\sim e^{-ikr}\end{aligned}\right\} \quad \text{for} \quad |r| \to \infty. \quad (2.2b)$$

As is well known, these solutions may be conveniently expressed in terms of each other by the relation

$$\psi_1^N(l, k; r) = \left(\frac{2l+1}{2ik}\right) f_l^{-1}(-k) \times [f(l, k)\psi_3^N(l, k; r) - (-)^l f(l, -k)\psi_4^N(l, k; r)], \quad (2.3)$$

where the Jost function $f(l, -k)$ is given by[6]

$$f(l, -k) \equiv \lim_{r\to 0} \frac{\Gamma(l+1)}{\Gamma(2l+1)} (-2ikr)^l \psi_3^N(l, k, r). \quad (2.4)$$

[5] V. Singh, Phys. Rev. **127**, 632 (1962).

[6] D. I. Fivel and A. Klein, J. Math. Phys. **4**, 274 (1960).

The asymptotic form of the regular solution ψ_1^N is then found to be

$$\psi_1^N(l, k; r) \xrightarrow{r\to\infty} \frac{i^l}{k}(2l+1)\exp[i\delta(l,k)] \times \sin[kr - \tfrac{1}{2}l\pi + \delta(l,k)], \quad (2.5)$$

with

$$S(l,k) = \exp[2i\delta(l,k)] = \frac{f(l,k)}{f(l,-k)}. \quad (2.6)$$

Thus, to determine the scattering matrix S or the phase shift δ, it is first necessary to find the Jost function f or, equivalently, the Jost solution ψ_3^N. Previously[1] we derived high-energy asymptotic expansions for all solutions (2.2). However, for a better understanding of the subsequent section, we rederive the solution ψ_3^N.

Throughout we consider a generalized Yukawa potential $V(r)$ which can be expanded as a power series in r:

$$V(r) = \sum_{i=-1}^{\infty} M_{i+1}(-r)^i, \quad (2.7)$$

where, for the real potentials we consider here, all M_i are real and independent of k. Then, as pointed out before[1] and discussed in detail by Bethe and Kinoshita,[7] a countably infinite number of Regge poles may be defined in the region of large negative energies, by requiring $l + n + 1$, for $n = 0, 1, 2, \cdots$, to be of order $1/k$, i.e.,

$$l + n + 1 = -\frac{\Delta(K)}{2K}, \quad |K| \to \infty, \quad (2.8)$$

where $K = ik$ and Δ is an expansion in descending powers of K. The function Δ is obtained from the secular equation of the eigenvalue problem, and it is known[1,2] up to (and including) the term in $1/K^7$. It may be written down in the two forms:

$$\begin{aligned}\Delta_n(K) = M_0 &- \frac{1}{2K^2}[n(n+1)M_2 + M_1M_0] \\ &- \frac{(2n+1)M_0M_2}{4K^3} + \frac{1}{8K^4} \\ &\times [3M_4(n-1)n(n+1)(n+2) \\ &+ 2M_3M_0(3n^2+3n-1) + 6M_2M_1n(n+1) \\ &+ 2M_2M_0^2 + 3M_1^2M_0] + \frac{(2n+1)}{8K^5} \\ &\times [3M_4M_0(n^2+n-1) + 3M_3M_0^2 \\ &+ M_2^2n(n+1) + 4M_2M_1M_0] + O(1/K^7) \quad (2.9)\end{aligned}$$

and

$$\begin{aligned}\Delta_l(K) = M_0 &- \frac{1}{2K^2}[l(l+1)M_2 + M_0M_1] + \frac{1}{8K^4} \\ &\times [3(l-1)l(l+1)(l+2)M_4 + 2M_3M_0(3l^2+3l-1) \\ &+ 6l(l+1)M_2M_1 + 3M_2M_0^2 + 3M_1^2M_0] + O(1/K^6), \quad (2.10)\end{aligned}$$

Eq. (2.10) being obtained by a trivial inversion of (2.8). In particular we observe that $\Delta_l(K)$ has the property

$$\Delta_l(K) = \Delta_l(-K). \quad (2.11)$$

3. JOST SOLUTIONS

To obtain the Jost solution ψ_3^N, we set $z = -2Kr$ and

$$\psi_3^N(l, k; z) = c \cdot e^{-z/2}z^{l+1}\chi_3(l, k; z), \quad (3.1)$$

c being a constant independent of r. Then χ_3 is a solution of the equation

$$D_n\chi = \frac{1}{2K}(M_0 - \Delta(K))\chi + \frac{1}{2K}\sum_{i=1}^{\infty}\left(\frac{z}{2K}\right)^i M_i\chi, \quad (3.2)$$

where

$$D_n = zd^2/dz^2 + (b-z)\,d/dz - a \quad (3.3)$$

and [cf. (2.8)]

$$a = l + 1 + \frac{\Delta(K)}{2K} = -n, \quad (3.4)$$

$$b = 2l + 2 = -2n - \frac{\Delta(K)}{K}. \quad (3.5)$$

Since, however, the right-hand side of (3.2) is of order $1/K$, we have to a first approximation

$$D_n\chi^{(1)} = 0. \quad (3.6)$$

A glance at the differential operator (3.3) shows that (3.6) is a confluent hypergeometric equation. A particular solution of this equation is the function[8]

$$\Psi(a, b; z) = \frac{\Gamma(1-b)}{\Gamma(a-b+1)}\Phi(a, b; z) + \frac{\Gamma(b-1)}{\Gamma(a)}z^{1-b}\Phi(a-b+1, 2-b; z), \quad (3.7)$$

where Φ is the well-known hypergeometric function defined by the Kummer series. It is now convenient to set the constant c in (3.1) equal to $1/\Gamma(a)$ and to define the first approximation of χ_3 as

$$\chi_3^{(1)} = \Gamma(a)\Psi(a, b; z) \equiv \tilde{\Psi}(a, b; z). \quad (3.8)$$

The solution $\tilde{\Psi}$ so defined then satisfies the recurrence relation

$$z\tilde{\Psi}(a) = (a, a+1)\tilde{\Psi}(a+1) + (a, a)\tilde{\Psi}(a) + (a, a-1)\tilde{\Psi}(a-1), \quad (3.9)$$

where

$$\tilde{\Psi}(a+j) \equiv \tilde{\Psi}(a+j, b; z)$$

[7] H. A. Bethe and T. Kinoshita, Phys. Rev. 128, 1418 (1962).

[8] *Higher Transcendental Functions*, A. Erdléyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, Eds. (McGraw-Hill Book Company, Inc., New York, 1955), Vol. III.

and

$$(a, a+1) = a - b + 1 \equiv \alpha,$$
$$(a, a) = b - 2a = \beta - 2\alpha, \quad \beta \equiv 2 - b, \quad (3.10)$$
$$(a, a-1) = a - 1 = \alpha - \beta.$$

Note: These coefficients correspond to the coefficients $(a, a+j)^*$ defined in our earlier paper.[1]

By a repeated application of the recurrence relation (3.9) we obtain

$$z^m \tilde{\Psi}(a) = \sum_{j=-m}^{m} S_m(a, j) \tilde{\Psi}(a+j). \quad (3.11)$$

The coefficients $S_m(a, j)$ may then be computed from a recurrence relation which follows from the coefficients (3.10):

$$\begin{aligned} S_m(a, j) = {} & (a+j-1, a+j) S_{m-1}(a, j-1) \\ & + (a+j, a+j) S_{m-1}(a, j) \\ & + (a+j+1, a+j) S_{m-1}(a, j+1). \quad (3.12) \end{aligned}$$

The corresponding boundary conditions are:

A. $S_0(a, 0) = 1$, all other $S_0(a, i \neq 0) = O$.
B. All $S_m(a, j)$ for $|j| > m$ are zero.

If we now substitute the first approximation (3.8) into the right-hand side of (3.2), the latter may be written

$$\frac{1}{2K} [a, a]_1 \tilde{\Psi}(a) + \sum_{i=1}^{\infty} \frac{1}{(2K)^{i+1}} \sum_{j=-i}^{i} [a, a+j]_{i+1} \tilde{\Psi}(a+j), \quad (3.13)$$

where

$$[a, a]_1 = M_0 - \Delta(K),$$
$$[a, a+j]_{i+1} = M_i S_i(a, j), \quad 0 \leq |j| \leq \text{i}. \quad (3.14)$$

The usefulness of this notation is now seen in the ease with which it permits the calculation of any number of higher-order perturbation terms. For, any term $\mu \tilde{\Psi}(a+n)$ on the right-hand side of (3.2) [e.g., for the second approximation in (3.13)] may be cancelled out by adding a contribution $\mu \tilde{\Psi}(a+n)/n$ to the previous approximation, except, of course, when $n = 0$. This follows simply from the fact that

$$D_n \left[\mu \frac{\tilde{\Psi}(a+n)}{n} \right] = \mu \tilde{\Psi}(a+n). \quad (3.15)$$

The coefficient of the sum of all the remaining terms in $\tilde{\Psi}(a)$ is then set equal to zero and yields the secular equation from which $\Delta(K)$ is determined. Reordering successive approximations in powers of $1/K$, one finally obtains the expansion

$$\chi_3(a, b; z) = \tilde{\Psi}(a, b; z) + \sum_{i=2}^{\infty} \frac{1}{(2K)^i} \sum_{j=-(i-1)}^{i-1}{}' P_i(a, j) \tilde{\Psi}(a+j, b; z), \quad (3.16)$$

where the prime on the second sum implies $j \neq 0$. The coefficients $P_i(a, j)$ are easily found, e.g.,

$$P_2(a, 1) = \frac{[a, a+1]_2}{1}, \quad P_2(a, -1) = \frac{[a, a-1]_2}{-1},$$
$$P_3(a, 2) = \frac{[a, a+2]_3}{2},$$
$$P_3(a, 1) = \frac{[a, a+1]_3}{1} + \frac{[a, a+1]_2}{1} \times \frac{[a+1, a+1]_1}{1}. \quad (3.17)$$

Again these coefficients may be obtained from a recurrence relation which follows from (3.14):

$$tP_r(a, t) = \sum_{i=1}^{r} \sum_{j=-(i-1)}^{i-1} [a+t-j, a+t]_i \times P_{r-1}(a, t-j). \quad (3.18)$$

The corresponding boundary conditions are

A. $P_0(a, 0) = 1, \quad P_0(a, t \neq 0) = 0$.
B. $P_r(a, 0) = 0 \quad$ for $\quad r > 1$.
C. $P_1(a, t) = 0, \quad P_r(a, t) = 0 \quad$ for $\quad |t| \geq r$.

The expansion for the eigenvalues was previously[1] found to be

$$\begin{aligned} 0 = {} & \frac{1}{2K} [a, a]_1 + \frac{1}{(2K)^2} [a, a]_2 + \frac{1}{(2K)^3} [a, a]_3 \\ & + \frac{1}{(2K)^4} \left\{ [a, a]_4 + \frac{[a, a+1]_2}{1} [a+1, a]_2 \right. \\ & \left. + \frac{[a, a-1]_2}{-1} [a-1, a]_2 \right\} + \cdots. \quad (3.19) \end{aligned}$$

The general term of this expansion may be derived as described in a mathematical paper of one of us.[9] We first define coefficients $Q_s(a, t)$ as, for instance,

$$Q_2(a, 1) = \frac{[a+1, a]_2}{1},$$
$$Q_2(a, -1) = \frac{[a-1, a]_2}{-1}, \quad (3.20)$$
$$Q_3(a, 2) = \frac{[a+2, a]_3}{2},$$

which satisfy a recurrence relation similar to (3.18):

$$tQ_s(a, t) = \sum_{i=1}^{s} \sum_{j=-(i-1)}^{i-1} [a+t, a+t-j]_i \times Q_{s-i}(a, t-j). \quad (3.21)$$

[9] H. J. W. Müller, J. Reine Angew. Math. **211**, 179 (1962).

The corresponding boundary conditions are:

A. $Q_0(a, 0) = 1, \quad Q_0(a, t \neq 0) = 0.$
B. $Q_s(a, 0) = 0 \quad$ for $\quad s > 1.$
C. $Q_1(a, t) = 0, \quad Q_s(a, t) = 0 \quad$ for $\quad |t| \geq s.$

Each term in the expansion (3.19) may now conveniently be considered as the product of a contribution from a sequence of r coefficients (starting from a and ending at $a + t$) and a contribution from a sequence of s coefficients (starting at $a + t$ and ending at a), i.e., as a product $tP_r(t)Q_s(t)$. Then, if we rewrite (3.19) as

$$\Delta(K) = M_0 + \sum_{i=1}^{\infty} \frac{1}{(2K)^i} M^{(i)}, \qquad (3.22)$$

we have (for $r, s \geq 2$)

$$\begin{aligned} M^{(r+s)} &= \sum_{t=1}^{r-1} tP_r(t)Q_s(t) + \sum_{t=-1}^{-s+1} tP_r(t)Q_s(t) + [a, a]_{r+s} \\ &= \sum_{t=1}^{r-1} tP_r(t)Q_s(t) - \sum_{t=1}^{s-1} tP_r(-t)Q_s(-t) \\ &\qquad + [a, a]_{r+s}. \qquad (3.23) \end{aligned}$$

Hence

$$M^{(2r)} = \sum_{t=1}^{r-1} t[P_r(t)Q_r(t) - P_r(-t)Q_r(-t)] + [a, a]_{2r} \qquad (3.24)$$

and

$$\begin{aligned} M^{(2r+1)} &= \sum_{t=1}^{r-1} t[P_r(t)Q_{r+1}(t) - P_r(-t)Q_{r+1}(-t)] \\ &\qquad - rP_r(-r)Q_{r+1}(-r) + [a, a]_{2r+1}. \qquad (3.25) \end{aligned}$$

These coefficients together with the recurrence relations (3.18) and (3.21) permit a joint computer calculation of eigenvalues and eigenfunctions.

Returning now to the solution (3.16), we first observe that (3.1) has the asymptotic behavior as claimed by (2.2b). For, setting $V = 0$, we obtain

$$\psi_3^N(l, k; r) = e^{i(\pi/2)(l+1)} \left(\frac{\pi kr}{2}\right)^{\frac{1}{2}} H^{(1)}_{l+\frac{1}{2}}(kr), \qquad (3.26)$$

where

$$\left(\frac{\pi kr}{2}\right)^{\frac{1}{2}} H^{(1)}_n(kr) \sim e^{i(kr - \frac{1}{2}n\pi - \frac{1}{4}\pi)}$$

for $r \to \infty$. The Jost function $f(l, -k)$ therefore follows from (2.4) and is readily found to be

$$f(l, -k) = \frac{\Gamma(l+1)}{\Gamma\{l + 1 + [\Delta(K)/2K]\}} \times \left[1 + \sum_{i=2}^{\infty} \frac{1}{(2K)^i} \sum_{j=-i+1}^{i-1}{}' P_i(a, j)\right] \qquad (3.27)$$

for Re $l > -\frac{1}{2}$ (and for Re $l < \frac{1}{2}$ by analytic continuation). Since the coefficients $P_i(a,j)$ can easily be computed, as explained above, the Jost function (3.27) can also be calculated to any required number of terms. We find

$$\begin{aligned} f(l, -k) &= \frac{\Gamma(l+1)}{\Gamma[l + 1 + (\Delta(K)/2K)]} \\ &\times \Big\{1 + \frac{(n+1)M_1}{2K^2} + \frac{M_0M_3}{4K^3} - \frac{1}{3(2K)^4} \\ &\times [9(2n+3)M_0M_2 - 3(n-2)(n+1)M_1^2 \\ &+ 2(n+1)(n+2)(8n-3)M_3] \\ &+ \frac{1}{3(2K)^5}[3(2n-5)M_0M_1^2 - 18M_0^2M_2 \\ &-2(24n^2 + 42n + 7)M_0M_3 - 12n(n+1)M_1M_2] \\ &+ \frac{1}{60(2K)^6}[120M_0^2M_1^2 + 120(2n+3)M_0^2M_3 \\ &- 120(3n^2 - 7n - 22)M_0M_1M_2 \\ &+ 100(22n^3 + 69n^2 + 17n - 45)M_0M_4 \\ &- 40(n-1)(n+1)(n+6)M_1^3 \\ &- 40(n+1)(n+2)(7n^2 - 45n + 18)M_1M_3 \\ &+ 5(n+1)(3n^3 + 235n^2 + 422n - 120)M_2^2 \\ &+ 8(n+1)(n+2)(n+3)(128n^2 - 223n + 60)M_5] \\ &+ O\left(\frac{1}{K^7}\right)\Big\}. \qquad (3.28) \end{aligned}$$

The Regge poles were defined originally in the region of negative energies. In the region of positive energies they are defined by analytic continuation. The corresponding scattering solutions are obtained by replacing n by $-l - 1 - [\Delta_l(K)/2K]$, where $\Delta_l(K)$ is given by (2.10). We then obtain

$$\begin{aligned} f(l, -k) &= \frac{\Gamma(l+1)}{\Gamma\{l + 1 + [\Delta_l(K)/2K]\}} \\ &\times \Big\{1 - \frac{M_1 l}{2K^2} + \frac{1}{3(2K)^4} \\ &\times [9(2l-1)M_0M_2 + 3l(l+3)M_1^2 \\ &+ 2l(l-1)(8l+11)M_3] + \frac{1}{60(2K)^6} \\ &\times [60M_0^2M_1^2 - 1200(l-1)M_0^2M_3 \\ &- 120(3l^2 + 17l - 10)M_0M_1M_2 \\ &- 100(22l^3 - 3l^2 - 55l + 15)M_0M_4 \\ &+ 40l(l+2)(l-5)M_1^3 \\ &- 40l(l-1)(7l^2 + 59l + 70)M_1M_3 \\ &+ 5l(3l^3 - 226l^2 - 39l + 310)M_2^2 \\ &- 8l(l-1)(l-2)(128l^2 + 479l + 411)M_5] \\ &+ O\left(\frac{1}{K^8}\right)\Big\}. \qquad (3.29) \end{aligned}$$

The asymptotic behavior of $f(l, -k)$ is now easily found to be

$$f(l, -k) \sim 1 - \frac{M_0}{2K}\,\psi(l+1), \tag{3.30}$$

in agreement with the well-known limit

$$\lim_{|k|\to\infty} f(l, -k) = 1. \tag{3.31}$$

The function ψ in (3.30) is defined by

$$\psi(z) = \frac{d \ln \Gamma(z)}{dz}. \tag{3.32}$$

4. S MATRIX, PHASE SHIFTS, AND RESIDUES

Having calculated the Jost function $f(l, -k)$, the S matrix now follows from (2.6). We find

$$S(l, K) = \frac{\Gamma\{l + 1 + [\Delta_l(K)/2K]\}}{\Gamma\{l + 1 - [\Delta_l(K)/2K]\}}. \tag{4.1}$$

This remarkably simple expression for S is due to the fact that the additional factors in the Jost functions are either independent of K or else contain only powers of K^2 and hence drop out in the ratio $f(l, K)/f(l, -K)$. Of course $\Delta_l(K)$ is also only an even function of K by (2.11).

The formal expression of the S matrix (4.2) was derived earlier by one of us.[1] However, it still contained these additional factors which cannot easily be seen to cancel out without explicit calculation.

We observe that the S matrix is now completely determined by the eigenvalues or Regge poles (2.8).

The phase shift $\delta(l, K)$ as defined by (2.6) now becomes

$$\delta(l, K) = \frac{1}{2i} \times \left[\ln \Gamma\left(l + 1 + \frac{\Delta(K)}{2K}\right) - \ln \Gamma\left(l + 1 - \frac{\Delta(K)}{2K}\right)\right]. \tag{4.2}$$

Taylor expansions of the Γ functions around $l + 1$ yield

$$\delta(l, K) = \frac{1}{2i}\sum_{n=1}^{\infty} \frac{1}{n!}\,\psi^{(n-1)}(l+1)\left(\frac{\Delta_l(K)}{2K}\right)^n [1 - (-1)^n], \tag{4.3}$$

where $\psi^{(i)}$ is the ith derivative of the Gaussian function defined by

$$\psi(x) = \frac{d}{dx}[\ln \Gamma(x)] = C - \sum_{n=0}^{\infty}\left(\frac{1}{x+n} - \frac{1}{n+1}\right), \tag{4.4}$$

where C is Euler's constant. The $(i - 1)$th derivative may be reexpressed in terms of the Riemann zeta function:

$$\psi^{(i-1)}(\chi) = (-1)^i i!\,\zeta(i, x) = (-1)^i i! \sum_{n=0}^{\infty} \frac{1}{(x+n)^i}. \tag{4.5}$$

We now insert in (4.3) the expansion (2.10) for $\Delta(k)$ and return to the variable k defined earlier by $K = ik$. Retaining only terms of order up to and including $1/k^6$, we obtain

$$\begin{aligned}\delta(l, k) = &-\frac{1}{2k} M_0 \psi(l+1) + \frac{1}{48k^3}\{M_0^3 \psi^{(2)}(l+1) \\ &- 12[l(l+1)M_2 + M_0 M_1]\psi(l+1)\} - \frac{1}{3840k^5} \\ &\times \{M_0^5 \psi^{(4)}(l+1) - 120 M_0^2[l(l+1)M_2 + M_0 M_1] \\ &\times \psi^{(2)}(l+1) + 240\{3(l-1)l(l+1)(l+2)M_4 \\ &+ 3M_0 M_1^2 + 3M_0^2 M_2 + 2(3l^2 + 3l - 1)M_0 M_3 \\ &+ 6l(l+1)M_1 M_2]\psi(l+1)\} + O\left(\frac{1}{k^7}\right).\end{aligned} \tag{4.6}$$

The residues of the S matrix at the poles, i.e., where the argument of the Γ function in the numerator is a negative integer, may also be determined from (4.1). One readily finds

$$\beta_n(K) = \frac{(-1)^n}{n!\,\Gamma(-n + 2i \operatorname{Im} \alpha_n(K))}. \tag{4.7}$$

where $\alpha_n(k) = l$ is given by (2.8). We observe, however, that the form (4.7) is independent of any high-energy approximation and thus is valid also for other approximations of $\alpha_n(k)$. In particular, if $\operatorname{Im} \alpha_n(k)$ is small, as for resonances which occur just above threshold, we may expand (4.7) and obtain

$$\beta_n(K) \simeq -\frac{2i}{(n!)^2} \operatorname{Im} \alpha_n(K). \tag{4.8}$$

This approximation is also valid for $|k| \to \infty$. In the high-energy region we may substitute for $\alpha_n(k)$ from (2.8) and (2.9). Then

$$\beta_n(k) \simeq -\frac{i}{(n!)^2}\cdot\frac{M_0}{k}, \quad k^2 \to \infty. \tag{4.9}$$

Thus the residue function is zero for negative energies, where $\alpha_n(k)$ is real (this is in fact the bound-state condition equivalent to the vanishing of the wave-function renormalization constant in field theory). Above the threshold it increases from zero as in (4.8). With increasing energy it decreases and becomes zero again for $k^2 \to \infty$.

Regge trajectories have been discussed earlier by one of us.[1,2]

Asymptotic Theory of Cerenkov Radiation in Inhomogeneous Media*

R. A. HANDELSMAN† AND R. M. LEWIS
Courant Institute of Mathematical Sciences, New York University, New York, N.Y.

(Received 11 May 1967)

The problem of Cerenkov radiation in infinite inhomogeneous media is considered. The mathematical description of this phenomenon is given by the integro-differential system of equations for the electromagnetic field in a dispersive medium. The leading term of the asymptotic expansion of the electromagnetic field is obtained by applying an expansion procedure called the "ray method." In this method all the functions that appear in the expansion satisfy ordinary differential equations along certain space-time curves called rays. The source which gives rise to the radiation is taken to be quite general. In fact, it is shown that any multipole moving along an arbitrary trajectory is a special case of the general source considered. From the expansion of the fields an expression for the total energy of the radiation is determined. Then, as an example, the case of plane-stratified media is treated in detail.

1. INTRODUCTION

Since 1940 a great deal of theoretical research has been carried out in the field of Cerenkov radiation. The problems considered in recent years are far more complex than the original problem treated by Frank and Tamm.[1] One area, however, which has not received much attention—undoubtedly due to its complexity—is the problem of Cerenkov radiation in inhomogeneous media. Ter-Mikaelyan,[2] Feinberg and Khizhnyak,[3] and others have investigated this subject under very restrictive conditions. In this paper we develop an asymptotic theory of Cerenkov radiation in isotropic inhomogeneous media. No restrictions, however, are placed on the nature of the inhomogeneity.

The mathematical description of Cerenkov radiation is given by the time-dependent form of Maxwell's equations for dispersive media. For such media, the constituitive equation takes the form of a convolution integral. Therefore we are led to seek the asymptotic solution of an integro-differential system of equations. That the medium is inhomogeneous is reflected in the fact that the index of refraction of the medium is a function of spatial position. This, as we shall see, significantly complicates the asymptotic analysis.

As is well known, Cerenkov radiation can occur only from sources moving with great speed. In Sec. 2 it is shown that the type of source to be considered in this paper is quite general. In fact, we find that any multipole moving along an arbitrary trajectory is a special case of the general source that is treated. Furthermore, the source is allowed to have an "oscillatory factor" so that the Cerenkov-Doppler effect can be considered.

In Sec. 2 the expansion parameter λ is introduced as a characteristic frequency of the medium. However, to understand better the meaning of our asymptotic expansion, an equivalent dimensionless parameter λ_0 must be found. If dimensionless variables are introduced throughout the problem, we find that $\lambda_0 = a\lambda/c$ where a is a characteristic dimension of the problem and c is the speed of light in a vacuum. The correct interpretation of our expansion is that it is valid for $1 \ll \lambda_0$. The length a can be thought of as a distance over which the properties of the medium are significantly altered. Alternatively, a can be taken as the distance from the source trajectory to the point in space at which the solution is obtained. This latter interpretation can be used to obtain a far-field expansion of the solution.

In order to obtain the asymptotic expansion, a procedure called the "ray method" is applied. This method was introduced by Keller[4] in connection with the study of certain diffraction problems for the reduced wave equation and has been extended by Lewis[5] to include integro-differential equations of the type to be considered in this paper. In applying this method we assume that the solution is given by an asymptotic power series in λ^{-1} which involves a "phase function" and an infinite sequence of amplitude functions. The phase function and the zeroth-order amplitude function are determined by inserting this series into the integro-differential system of equations. It is found that certain space-time curves called "rays" play a central role. They satisfy a first-order system of ordinary differential equations. In turn, the phase function and the amplitude function

* This research was supported by the Office of Naval Research under Contract No. NONR 285(48).

† Present address: Division of Applied Mathematics, Brown University, Providence, R.I.

[1] I. M. Frank and Ig. Tamm, C. R. Acad. Sci. (Doklady) 14, 109 (1937).

[2] M. L. Ter-Mikaelyan, Dokl. Akad. Nauk SSSR, 134, 318 (1960) [Sov. Physics—Dokl. 5, 1015 (1961)].

[3] Ya. B. Fainberg and N. A. Khizhnyak, Zh. Eksperim. i Teor. Fiz. 32, 833 (1957) [Sov. Phys.—JETP 32, 726 (1957)].

[4] J. B. Keller, J. Opt. Soc. Am. 52, 116 (1962).

[5] R. M. Lewis, Arch. Rational Mech. Analysis 20, (3), 191 (1965).

are found to satisfy ordinary differential equations along the rays. In order to obtain the asymptotic solution, initial conditions for these equations are required. Some of these conditions can be determined from the source data. However, other conditions must be obtained by an "indirect method."

The indirect method requires the asymptotic solution of a "canonical problem," which is an appropriately constructed problem for homogeneous media. The canonical problem is, of course, much simpler to solve than the original problem. The required initial conditions for the original problem are then determined from the solution of the canonical problem. Previously[6] the authors have determined the asymptotic expansion for the case of homogeneous media. From the results obtained there, the solution of the canonical problem is found.

In Sec. 3 an asymptotic expression for the total energy of the radiation is determined. As an example, in Sec. 4, results obtained in the previous sections are applied to the special case of stratified media, i.e., to media whose properties vary in only one space direction. This example is selected because the asymptotic expansion of the solution can also be obtained by an exact method. This exact method, however, is much more difficult to apply and for more general problems fails entirely. Upon comparing the results of the two procedures, it is found that they are in perfect agreement. This agreement serves as a partial justification for the validity of the application of the ray method.

2. ELECTROMAGNETIC EQUATIONS FOR DISPERSIVE MEDIA

In Gaussian units Maxwell's equations are given by

$$\mathbf{D}_t - c\nabla \times \mathbf{H} = -4\pi\mathbf{J}, \quad \mathbf{B}_t + c\nabla \times \mathbf{E} = 0, \tag{2.1}$$

$$\nabla \cdot \mathbf{D} = 4\pi\rho, \quad \nabla \cdot \mathbf{B} = 0. \tag{2.2}$$

Here $\mathbf{D}$, $\mathbf{B}$, $\mathbf{E}$, $\mathbf{H}$, and $\mathbf{J}$ are 3-vectors which depend on t and $\mathbf{X} = (x_1, x_2, x_3)$. The source functions ρ and $\mathbf{J}$ satisfy the continuity equation

$$\rho_t + \nabla \cdot \mathbf{J} = 0. \tag{2.3}$$

It then follows from (2.1) and (2.3) that

$$\frac{\partial}{\partial t}(\nabla \cdot \mathbf{D} - 4\pi\rho) = 0 \quad \text{and} \quad \frac{\partial}{\partial t}(\nabla \cdot \mathbf{B}) = 0. \tag{2.4}$$

Thus, if Eqs. (2.2) are satisfied at any time t, they are satisfied for all time. We shall assume that the source and fields are identically zero for $t < 0$ and that the source is "switched on" at $t = 0$. We then seek the fields for $t > 0$.

[6] R. A. Handelsman and R. M. Lewis, J. Math. Phys. 7, 1982 (1966).

Dispersive media are characterized by the fact that the constitutive equation is given by a convolution integral of the form

$$\mathbf{v}(t, \mathbf{X}) = \int_0^\infty \mathcal{F}(\tau, \mathbf{X})\mathbf{u}(t - \tau, \mathbf{X})\, d\tau. \tag{2.5}$$

In (2.5) we have introduced the column vectors $\mathbf{u}$ and $\mathbf{v}$ defined by

$$\mathbf{u} = [\mathbf{E}, \mathbf{H}] = (E_1, E_2, E_3, H_1, H_2, H_3),$$
$$\mathbf{v} = [\mathbf{D}, \mathbf{B}] = (D_1, D_2, D_3, B_1, B_2, B_3). \tag{2.6}$$

(We shall often represent column vectors with 6 components by an ordered pair of two 3-vectors.) $\mathcal{F}(\tau, \mathbf{X})$ is a 6×6 matrix which is a real function of time and space. Furthermore, we assume that the causality condition, $\mathcal{F}(t, \mathbf{X}) \equiv 0$ for $t < 0$, is satisfied.

We now define $\hat{\mathcal{E}}(\hat{\omega}, \mathbf{X})$ as the Fourier transform of $\mathcal{F}(t, \mathbf{X})$ with respect to time. That is, we set

$$\hat{\mathcal{E}}(\hat{\omega}, \mathbf{X}) = \int_{-\infty}^\infty e^{i\hat{\omega}t}\mathcal{F}(t, \mathbf{X})\, dt. \tag{2.7}$$

Since in Gaussian units the elements of $\hat{\mathcal{E}}$ are dimensionless and $\hat{\omega}$ has dimensions of frequency, $\hat{\mathcal{E}}$ must be a function of the dimensionless variable $\omega = \hat{\omega}/\lambda$. Here λ is a characteristic frequency of the medium. Therefore we may write

$$\hat{\mathcal{E}}(\hat{\omega}, \mathbf{X}) = \tilde{\mathcal{E}}(\omega, \mathbf{X}; \lambda). \tag{2.8}$$

It then follows from (2.7) that

$$\mathcal{F}(t, \mathbf{X}; \lambda) = \frac{\lambda}{2\pi}\int_{-\infty}^\infty e^{-i\lambda\omega t}\tilde{\mathcal{E}}(\omega, \mathbf{X}; \lambda)\, d\omega. \tag{2.9}$$

We shall seek the asymptotic expansion of the solution of the integro-differential system of Eqs. (2.1) and (2.5) for $\lambda \to \infty$.

We now assume that the matrix $\tilde{\mathcal{E}}(\omega, \mathbf{X}; \lambda)$ is of the form

$$\tilde{\mathcal{E}}(\omega, \mathbf{X}; \lambda) = \begin{bmatrix} \tilde{\epsilon}(\omega, \mathbf{X}; \lambda)I & 0 \\ 0 & \tilde{\mu}(\omega, \mathbf{X}; \lambda)I \end{bmatrix}$$
$$= \mathcal{E}(\omega, \mathbf{X}) + \frac{i}{\lambda}\mathcal{D}(\omega, \mathbf{X}) + O\left(\frac{1}{\lambda^2}\right). \tag{2.10}$$

Here,

$$\mathcal{E}(\omega, \mathbf{X}) = \begin{bmatrix} \epsilon(\omega, \mathbf{X})I & 0 \\ 0 & \mu(\omega, \mathbf{X})I \end{bmatrix},$$
$$\mathcal{D}(\omega, \mathbf{X}) = \begin{bmatrix} d_1(\omega, \mathbf{X})I & 0 \\ 0 & d_2(\omega, \mathbf{X})I \end{bmatrix}, \tag{2.11}$$

where ϵ, μ, d_1, and d_2 are real and I is the 3×3 unit matrix. A matrix of this form represents what we call an "inhomogeneous weakly dissipative isotropic medium" (weak dissipation is the simplest type of dissipation which can be treated by our asymptotic methods).

We now define the function

$$n(\omega, \mathbf{X}) = [\epsilon(\omega, \mathbf{X})\mu(\omega, \mathbf{X})]^{\frac{1}{2}} \qquad (2.12)$$

for all $\mathbf{X}$ and real ω such that $\epsilon(\omega, \mathbf{X})\mu(\omega, \mathbf{X}) \geq 0$. It is clear from Eqs. (2.10)–(2.12) that $n(\omega, \mathbf{X}) = \lim_{\lambda\to\infty} (\tilde{\epsilon}\tilde{\mu})^{\frac{1}{2}}$. Thus, we shall call n the *asymptotic index of refraction* of the medium. From Eq. (2.7) and the fact that $\mathcal{F}(t, \mathbf{X})$ is real, it follows that

$$\tilde{\mathcal{E}}(-\omega, \mathbf{X}; \lambda) = \tilde{\mathcal{E}}(\omega, \mathbf{X}; \lambda).$$

This then implies that $n(-\omega, \mathbf{X}) = n(\omega, \mathbf{X})$.

It is convenient to write (2.1) in matrix form. To do this we introduce the antisymmetric matrix $(\mathbf{L})$, corresponding to any 3-vector $\mathbf{L}$, given by

$$(\mathbf{L}) = \begin{pmatrix} 0 & -L_1 & L_2 \\ L_3 & 0 & -L_1 \\ -L_2 & L_1 & 0 \end{pmatrix}. \qquad (2.13)$$

Then if $\mathbf{W}$ is an arbitrary 3-vector, $(\mathbf{L})\mathbf{W} = \mathbf{L} \times \mathbf{W}$. We also define the three 6×6 matrices A^1, A^2, and A^3 by

$$k_\nu A^\nu = \begin{bmatrix} 0 & -c(\mathbf{K}) \\ c(\mathbf{K}) & 0 \end{bmatrix}, \quad \mathbf{K} = (k_1, k_2, k_3). \quad (2.14)$$

(Here, and in what follows, the summation convention with respect to repeated indices from 1 to 3 is used.) The matrix A^1 is obtained by setting $k_1 = 1$, $k_2 = k_3 = 0$ in (2.14). A^2 and A^3 are determined in a similar manner. Using (2.14), (2.1) can be written in the compact form

$$\frac{\partial \mathbf{v}}{\partial t} + A^\nu \frac{\partial \mathbf{u}}{\partial x_\nu} = f. \qquad (2.15)$$

In (2.15) $\mathbf{f}$ is the column vector with components $-4\pi(J_1, J_2, J_3, 0, 0, 0)$.

To conclude this section we shall describe in detail the nature of the source function $\mathbf{f}$ that is considered in this paper. Our concern here will be with moving sources. The current $\mathbf{J}$ corresponding to a charged particle moving along the trajectory $\mathbf{X} = \mathbf{Y}(t)$ is given by

$$\begin{aligned}\mathbf{J}(t, \mathbf{X}) &= e\dot{\mathbf{Y}}(t)\delta[\mathbf{X} - \mathbf{Y}(t)] \\ &= e\,\dot{\mathbf{Y}}(t)\delta[x_1 - y_1(t)]\delta[x_2 - y_2(t)]\delta[x_3 - y_3(t)].\end{aligned} \qquad (2.16)$$

For greater generality we consider source functions of the form

$$\mathbf{f}(t, \mathbf{X}; \lambda) = \lambda^d \mathbf{g}\{t, \lambda[\mathbf{X} - \mathbf{Y}(t)]\} \cos[\lambda q(t)]. \quad (2.17)$$

Here d is a real number. The vector $\mathbf{g}(t, \mathbf{X})$ is taken to be real and to have for each value of t, compact support in $\mathbf{X}$. The term $\cos[\lambda q(t)]$ is called the *oscillatory factor*. We observe that, as $\lambda \to \infty$, the support of f shrinks to the moving point $\mathbf{X} = \mathbf{Y}(t)$. Therefore, (2.17) can be used to describe an oscillatory current which is nonzero only in a small neighborhood of the moving point $\mathbf{X} = \mathbf{Y}(t)$.

By using the well-known relation $\delta(x) = \lambda\delta(\lambda x)$ and remembering that $\mathbf{J}$ represents the first three components of f, it is easy to see that (2.16) is a special case of (2.17) with $q(t) \equiv 0$. Moreover it can be seen that (2.17) may be used to describe the current associated with any moving multipole source, i.e., any source whose current is given by a linear combination of partial derivatives of the three-dimensional δ function.

It follows from the fact that $\mathbf{f}$ is real, the solution $\mathbf{u}$ itself is real. We may then set

$$\mathbf{f}(t, \mathbf{X}; \lambda) = \lambda^d \mathbf{g}\{t, \lambda[\mathbf{X} - \mathbf{Y}(t)]\} e^{i\lambda q(t)} \qquad (2.18)$$

if $\mathbf{u}$ is taken to be the real part of the resulting solution.

3. ASYMPTOTIC SOLUTION OF THE INTEGRO-DIFFERENTIAL SYSTEM OF EQUATIONS

A. Asymptotic Expansion

We assume that away from the source trajectory $\mathbf{X} = \mathbf{Y}(t)$, Eqs. (2.5) and (2.15) have a solution given by an asymptotic power series of the form

$$\mathbf{u}(t, \mathbf{X}) \sim \exp\{i\lambda s(t, \mathbf{X})\} \sum_{m=0}^{\infty} (i\lambda)^{-m} \mathbf{z}_m(t, \mathbf{X}). \quad (3.1)$$

This assumption is motivated by the form of the asymptotic expansion of $\mathbf{u}$ for the case of homogeneous media obtained in Ref. 6. The *phase function* $s(t, \mathbf{X})$ and the *amplitude function* $\mathbf{z}(t, \mathbf{X}) = \mathbf{z}_0(t, \mathbf{X})$ are determined by inserting (3.1) into (2.5) and (2.15). (In principle, the lower-order terms $\mathbf{z}_1, \mathbf{z}_2, \cdots$ may also be obtained. However, we will limit our considerations here, to the determination of $\mathbf{z}_0$.) In order to describe the functions s and $\mathbf{z}$, we introduce the quantities

$$k_\nu = \frac{\partial s}{\partial x_\nu}, \quad \omega = -\frac{\partial s}{\partial t}, \quad \mathbf{K} = (k_1, k_2, k_3),$$
$$k^2 = k_\nu k_\nu. \quad (3.2)$$

Our object is to insert (3.1) into (2.5) and (2.15) and to equate to zero in the result, the coefficients of like powers of λ. However, before this procedure can be carried out, certain computations must be performed. These computations are somewhat involved and will not be described here. The reader is referred to Ref. 5 for a detailed description of this analysis. In Ref. 5 it is found that Eqs. (2.5) and (2.15) are satisfied up to $O(\lambda^{-1})$ if

$$G\mathbf{z} = 0 \qquad (3.3)$$

and

$$G\mathbf{z}_1 + A^0\mathbf{z}_t + A^\nu \mathbf{z}_{x_\nu} + \tfrac{1}{2}A^0_t\mathbf{z} + \omega\mathfrak{D}\mathbf{z} = 0. \quad (3.4)$$

Here the matrices G and A^0 are defined by

$$G = k_\nu A^\nu - \omega \mathcal{E}(\omega, \mathbf{X}), \tag{3.5}$$

$$A^0 = (\omega \mathcal{E})_\omega . \tag{3.6}$$

A^0 is called the energy matrix and is assumed to be positive definite for ω real. It follows from (3.3) that $\mathbf{z}$ is nontrivial only if

$$\det G = 0. \tag{3.7}$$

For real values of ω, $\mathbf{X}$, and $\mathbf{K}$, Eq. (7) defines a functional relation between these quantities which we call the *dispersion relation*. It can be shown using (2.11) and (2.14) that the dispersion relation is given by

$$k = m(\omega, \mathbf{X}) = \frac{|\omega|}{c}\, n(\omega, X). \tag{3.8}$$

B. Dispersion Equation and the Ray Equations

From the definitions of ω and k given above, we see that (3.8) is also a first-order partial equation and is called the *dispersion equation* for $s(t, \mathbf{X})$. This equation may be solved by the "method of characteristics," described by Courant and Hilbert.[7] Therefore we introduce the "characteristic equations" which, as can be shown, take the form

$$\frac{dx_\nu}{dt} = g_\nu = \frac{k_\nu}{mm_\omega}, \quad \frac{dk_\nu}{dt} = \frac{m_{x_\nu}}{m_\omega}; \quad \nu = 1, 2, 3. \tag{3.9}$$

Each solution of the system (3.9) defines a curve $[t, \mathbf{X}(t)]$ in space–time which we shall call a *ray*.

We shall require that all rays emanate from the source trajectory with increasing t. Let $t = \tau$ denote the time of emission of a ray. It is clear that $\mathbf{X}_0$, the initial value of $\mathbf{X}$ along a ray, is given by

$$\mathbf{X}_0 = \mathbf{Y}(\tau). \tag{3.10}$$

(We shall often refer to the initial values of various quantities along a ray which are obtained by evaluating these quantities at time $t = \tau$.) Then once $\mathbf{K}_0$, the initial value of $\mathbf{K}$, is known, Eqs. (3.9) may be solved to determine a ray. By differentiating (3.8) with respect to t and making use of Eqs. (3.9), we find that along this ray ω is constant. Moreover, from Eqs. (3.2) and (3.9) we obtain

$$\frac{ds}{dt} = \frac{\partial s}{\partial t} + \frac{\partial s}{\partial x_\nu}\frac{dx_\nu}{dt} = l, \tag{3.11}$$

where

$$l = \left[\frac{m(\omega, \mathbf{X})}{m_\omega(\omega, \mathbf{X})} - \omega\right]. \tag{3.12}$$

[7] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1962), Vol. II.

Therefore s may be determined from (3.11) by integration, once its initial value s_0 is known.

It is reasonable to assume that the phase at the source trajectory is equal to $q(\tau)$, the oscillation frequency of the source itself. This assumption is further motivated by the fact that $s_0 = q(\tau)$ for the case of homogeneous media, as is shown in Ref. 6. That is,

$$s_0 = s[\tau, \mathbf{Y}(\tau)] = q(\tau). \tag{3.13}$$

Differentiation of (3.13) with respect to τ yields

$$k_{\nu 0}\dot{y}_\nu(\tau) = \omega + \dot{q}(\tau); \quad \mathbf{K}_0 = (k_{10}, k_{20}, k_{30}). \tag{3.14}$$

We now let $\mathbf{T}$ represent the unit vector in the direction of the group velocity vector $\mathbf{G} = (g_1, g_2, g_3)$, and $\mathbf{A} = (\alpha_1, \alpha_2, \alpha_3)$ represents the unit vector in the direction of $\mathbf{K}$. We see from (3.9) that $\mathbf{T}$ is tangent to the space projection of the ray and that

$$\mathbf{T} = \operatorname{sgn}\,[m_\omega(\omega, \mathbf{X})]\mathbf{A}. \tag{3.15}$$

We now define θ to be the angle which the vector $\mathbf{G}$ makes with the vector $\dot{\mathbf{Y}}(\tau)$. If by θ_0 we denote the initial value of θ, Eqs. (3.8), (3.14), and (3.15) yield

$$\cos\theta_0 = \operatorname{sgn}\left[\frac{\partial m_0(\omega, \tau)}{\partial \omega}\right]\left[\frac{\omega + \dot{q}(\tau)}{v(\tau)m_0(\omega, \tau)}\right]. \tag{3.16}$$

Here $v(\tau) = |\dot{\mathbf{Y}}(\tau)|$ and $m_0(\omega, \tau) = |\mathbf{K}_0| = m[\omega, \mathbf{Y}(\tau)]$. In the special case of homogeneous media, the function m is independent of $\mathbf{X}$. We then see from (3.9) that in this case $\mathbf{K} \equiv \mathbf{K}_0$ and θ is constant along a ray. Furthermore, when $\dot{q}(\tau) \equiv 0$ and the medium is homogeneous, (3.16) is the well-known "Cerenkov condition" and $\theta = \theta_0$ is the Cerenkov angle. When $\dot{q}(\tau) \not\equiv 0$, (3.16) is usually referred to as the "Cerenkov–Doppler condition." Returning to the general case, we shall call the angle θ defined above the Cerenkov–Doppler angle for inhomogeneous media.

We now introduce along the source trajectory $\mathbf{X} = \mathbf{Y}(\tau)$, the orthonormal set $\mathbf{T}^*$, $\mathbf{N}^*$, and $\mathbf{B}^*$ consisting of the unit tangent, principal normal, and binormal vectors, respectively. If the trajectory is a straight line, $\mathbf{N}^*$ and $\mathbf{B}^*$ can be any two unit vectors such that $\mathbf{T}^*$, $\mathbf{N}^*$, and $\mathbf{B}^*$ form a right-handed orthonormal set. We also define γ to be the angle which the projection of $\mathbf{T}$ into the $\mathbf{N}^*$, $\mathbf{B}^*$ plane makes with $\mathbf{N}^*$ as measured in a counterclockwise direction from $\mathbf{N}^*$. We then obtain

$$\mathbf{K} = m(\omega, \mathbf{X})\mathbf{A} = m \operatorname{sgn}\,[m_\omega]\{\cos\theta \mathbf{T}^* + \sin\theta\cos\gamma \mathbf{N}^* + \sin\theta\sin\gamma \mathbf{B}^*\}. \tag{3.17}$$

In (3.17) θ is restricted to lie between 0 and π, and γ is allowed to vary between 0 and 2π. The angles θ

and γ will, in general, vary along a ray. To obtain expressions for them it is, of course, necessary to solve the ray equation (3.9). In general the ray equations cannot be solved in closed form. In Sec. 5, however, we shall treat a case for which such a solution can be obtained.

If we denote the initial value of γ by γ_0 we have

$$\mathbf{K}_0 = m_0\mathbf{A}_0 = m_0 \operatorname{sgn}[(m_0)_\omega]\{\cos\theta_0\mathbf{T}^* + \sin\theta_0\cos\gamma_0\mathbf{N}^* + \sin\theta_0\sin\gamma_0\mathbf{B}^*\}. \quad (3.18)$$

We see from (3.10) and (3.18) that the initial values $\mathbf{X}_0$ and $\mathbf{K}_0$ have been expressed as functions of the parameters $\mathbf{P} = (\tau, \omega, \gamma_0)$. By using these initial values we may, in principle, solve the ray equations to obtain the solution

$$\mathbf{X} = \mathbf{X}(t; \mathbf{P}), \quad \mathbf{K} = \mathbf{K}(t; \mathbf{P}). \quad (3.19)$$

The values of $\mathbf{P}$ lie in a *parameter space* $\mathcal{T}$ which is defined by the inequalities

$$0 \le \tau, 0 \le \gamma_0 < 2\pi, 0 \le \left[\frac{\omega + \dot{q}(\tau)}{v(\tau)m_0(\omega,\tau)}\right]^2 \le 1. \quad (3.20)$$

The last of conditions (3.20) follows from the fact that for real $\mathbf{K}_0$ the angle θ_0 must be real. When $\dot{q}(\tau) \equiv 0$, this condition states that $v(\tau) \ge |\omega|/m_0 = \{c/n[\omega, Y(\tau)]\}$. Thus, for Cerenkov radiation to occur at a given point along the source trajectory and at a given frequency ω, the source speed at that point must be greater than the phase speed corresponding to that value of ω.

For $\mathbf{P}$ in $\mathcal{T}$ and $t \ge \tau$, the equation $(t, \mathbf{X}) = [t, \mathbf{X}(t; \mathbf{P})]$ defines a 3-parameter family of rays. We see from (3.12) that along each ray,

$$l(t; \mathbf{P}) = l[t, \mathbf{X}(t; \mathbf{P})] = \left[\frac{m[\omega, \mathbf{X}(t; \mathbf{P})]}{m_\omega(\omega, \mathbf{X}(t; \mathbf{P}))} - \omega\right]. \quad (3.21)$$

Then Eq. (3.11) yields by integration

$$s(t; \mathbf{P}) = s[t, \mathbf{X}(t; \mathbf{P})] = q(\tau) + \int_\tau^t l(t'; \mathbf{P})\, dt'. \quad (3.22)$$

C. Ray Transformation

For each fixed value of t, $\mathbf{X} = \mathbf{X}(t; \mathbf{P})$ defines a transformation from $\mathcal{T}$ to $\mathbf{X}$ space. We denote the Jacobian of this *ray transformation* by $j(t; \mathbf{P})$. It is defined by

$$j(t; \mathbf{P}) = \frac{\partial(x_1, x_2, x_3)}{\partial(\tau, \omega, \gamma_0)} = \frac{\partial \mathbf{X}}{\partial \tau} \cdot \frac{\partial \mathbf{X}}{\partial \omega} \times \frac{\partial \mathbf{X}}{\partial \gamma_0}. \quad (3.23)$$

To compute this Jacobian the solution of the ray equations is required. We can, however, obtain information about the behavior of j near the source trajectory directly from the ray equations without solving them. This information is needed below when an expression for $\mathbf{z}$ is determined.

We expand $\mathbf{X}(t; \mathbf{P})$ for small values of $(t - \tau)$. From Eqs. (3.9) and (3.10) we obtain

$$\mathbf{X}(t; \mathbf{P}) = \mathbf{Y}(\tau) + \frac{(t-\tau)}{(m_0)_\omega}\mathbf{A}_0(\tau, \omega, \gamma_0) + O[(t-\tau)^2]. \quad (3.24)$$

Here $\mathbf{A}_0$ is expressed in terms of $\mathbf{P}$ through equations (3.16) and (3.18). By differentiating (3.24), we find that

$$\frac{\partial \mathbf{X}}{\partial \tau} = \dot{\mathbf{Y}}(\tau) - \frac{\mathbf{A}_0}{(m_0)_\omega} + O[(t-\tau)], \quad (3.25)$$

$$\frac{\partial \mathbf{X}}{\partial \omega} = \frac{(t-\tau)}{(m_0)_\omega}\left[\frac{\partial \mathbf{A}_0}{\partial \omega} - \frac{(m_0)_{\omega\omega}}{(m_0)_\omega}\mathbf{A}_0\right] + O[(t-\tau)^2], \quad (3.26)$$

$$\frac{\partial \mathbf{X}}{\partial \gamma_0} = \frac{(t-\tau)}{(m_0)_\omega}\frac{\partial \mathbf{A}_0}{\partial \gamma_0} + O[(t-\tau)^2]. \quad (3.27)$$

The expressions for $\partial\mathbf{A}_0/\partial\gamma_0$ and $\partial\mathbf{A}_0/\partial\omega$ may be obtained from Eqs. (3.16) and (3.18). The Jacobian is then determined by inserting Eqs. (3.25)–(3.27) into (3.23). The result is

$$j(t; \mathbf{P}) = (t-\tau)^2\tilde{j}(\mathbf{P}) + O[(t-\tau)^3], \quad (3.28)$$

where

$$\tilde{j}(\mathbf{P}) = \frac{1}{v}\left\{\frac{(m_0)_{\omega\omega}}{(m_0)_\omega}\left[v^2 - \left(\frac{\omega + \dot{q}}{m_0}\right)^2\right] - \frac{1}{m_0(m_0)_\omega}\left[\frac{1}{(m_0)_\omega} - \frac{(\omega + \dot{q})}{m_0}\right]^2\right\}. \quad (3.29)$$

Equation (3.29) yields the expected result that the source trajectory is a caustic of the ray family, i.e., $j(t; \mathbf{P})$ vanishes at $t = \tau$.

D. Determination of the Amplitude Function z

In order to describe $\mathbf{z}$ we introduce $\mathbf{r}^1(t)$ and $\mathbf{r}^2(t)$, the two linearly independent null eigenvectors of the singular matrix $G = m(\omega, \mathbf{X})\alpha_\nu A^\nu - \omega\mathcal{E}(\omega, \mathbf{X})$. Here, $\mathbf{r}^j(t) = \mathbf{r}^j(t; \mathbf{P})$. (In what follows we do not explicitly exhibit the dependence of various functions on $\mathbf{P}$ when this dependence is obvious.) These vectors are orthonormalized by the condition

$$(\mathbf{r}^i, A^0\mathbf{r}^j) = \delta_{ij}. \quad (3.30)$$

The inner product of any two column vectors $\mathbf{a}$ and $\mathbf{b}$ having 6 components is defined by

$$(\mathbf{a}, \mathbf{b}) = \sum_{j=1}^{6} \bar{a}_j b_j.$$

It can be shown from (2.11), (2.14), and (3.5) that $\mathbf{r}^1$ and $\mathbf{r}^2$ are given by

$$\mathbf{r}^1 = (\zeta)^{\frac{1}{2}}\left[\frac{\hat{\mathbf{B}}}{(\epsilon)^{\frac{1}{2}}}, -\frac{\hat{\mathbf{N}}}{(\mu)^{\frac{1}{2}}}\right], \quad \mathbf{r}^2 = (\zeta)^{\frac{1}{2}}\left[\frac{\hat{\mathbf{N}}}{(\epsilon)^{\frac{1}{2}}}, \frac{\hat{\mathbf{B}}}{(\mu)^{\frac{1}{2}}}\right]. \tag{3.31}$$

Here

$$\zeta = \frac{\epsilon(\omega, \mathbf{X})\mu(\omega, \mathbf{X})}{\epsilon(\omega\mu)_\omega + \mu(\omega\epsilon)_\omega} \tag{3.32}$$

and $\hat{\mathbf{N}}$ and $\hat{\mathbf{B}}$ are any real unit vectors such that $\mathbf{T}$, $\hat{\mathbf{N}}$, and $\hat{\mathbf{B}}$ form a right-handed orthonormal set.

From (3.3) we see that the vector $\mathbf{z}$ lies in the null space of the matrix G and hence must be a linear combination of the vectors $\mathbf{r}^1$ and $\mathbf{r}^2$. Therefore we may write

$$\mathbf{z} = \sigma_1(t, \mathbf{X})\mathbf{r}^1 + \sigma_2(t, \mathbf{X})\mathbf{r}^2. \tag{3.33}$$

It is shown by Lewis in Ref. 5 that, by taking the inner product of (3.4) with $\mathbf{r}^1$ and $\mathbf{r}^2$ successively, a system of two first-order ordinary differential equations for the coefficients σ_1 and σ_2 can be derived. He furthermore obtains an explicit solution for this system. In order to describe his results we introduce the quantities $\eta(t)$ and $\delta(t)$ defined by

$$(\mathbf{r}^i, \omega\mathfrak{D}\mathbf{r}^j) = \eta\delta_{ij}, \quad i, j = 1, 2, \tag{3.34}$$

$$\delta(t) = -\int_\tau^t \hat{\mathbf{N}}(t')\frac{d\hat{B}(t')}{dt'}\,dt'. \tag{3.35}$$

By inserting (2.11) and (3.31) into (3.34) we have

$$\eta = \omega\zeta\left(\frac{d_1}{\epsilon} + \frac{d_2}{\mu}\right). \tag{3.36}$$

Using the results obtained in Ref. 5, we find that σ_1 and σ_2 are given parametrically by

$$\sigma_{1,2}(t; \mathbf{P}) = \left|\frac{\tilde{j}(\mathbf{P})}{j(t, \mathbf{P})}\right|^{\frac{1}{2}} \exp\left[-\int_\tau^t \eta(t')\,dt'\right]\tilde{\beta}_j(t); \quad j = 1, 2. \tag{3.37}$$

In (3.37)

$$\tilde{\beta}_1(t) = \beta_1(\tau)\cos\delta(t) - \beta_2(\tau)\sin\delta(t), \tag{3.38}$$

$$\tilde{\beta}_2(t) = \beta_1(\tau)\sin\delta(t) + \beta_2(\tau)\cos\delta(t), \tag{3.39}$$

where

$$\beta_j(t) = (\mathbf{r}^j[\tau], A^0[\tau]\tilde{\mathbf{z}}[\tau]) \tag{3.40}$$

and

$$\tilde{\mathbf{z}} = \lim_{(t-\tau)\to 0_+} (t-\tau)\mathbf{z}(t; P). \tag{3.41}$$

Equation (3.37) is valid only in the interval $\tau < t < \tau'$ where τ' is the location of the next caustic point on the ray $[t, \mathbf{x}(t; \mathbf{P})]$. More precisely, $t = \tau'$ is the smallest value of $t > 0$ for which the Jacobian j vanishes. For $t > \tau'$, (3.37) holds only when an appropriate phase-shift rule has been applied (see Ref. 5, Appendix F).

To complete the asymptotic solution $\tilde{\mathbf{z}}$ must be determined. Because we are dealing with linear equations it is reasonable to expect that

$$\tilde{\mathbf{z}} = \mathcal{C}\mathbf{g}, \tag{3.42}$$

where $\mathcal{C}$ is a linear operator and $\mathbf{g}$ is the vector portion of the source function $\mathbf{f}$. The operator $\mathcal{C}$ plays a role in this theory analogous to the role played by the "diffraction coefficient" introduced by Keller in his "geometrical theory of diffraction" (see Ref. 4). Therefore we shall call $\mathcal{C}$ the *Cerenkov radiation coefficient.*

The value of $\tilde{\mathbf{z}}$ may be determined by the following *indirect method.* We assume that $\tilde{\mathbf{z}}$ is determined only by the local properties of the medium at the source. Therefore we specialize our problem to one for homogeneous media where the constant value chosen for $\mathcal{E}$ is equal to $\mathcal{E}(\omega, \mathbf{Y}[\tau])$. It is clear that $\tilde{\mathbf{z}}$ is indedependent of η. Therefore in our specialized problem we set $\mathfrak{D}$, and hence η, equal to zero. In this manner we obtain at each point along the source trajectory a corresponding problem for homogeneous media.

The problem of nondissipative homogeneous isotropic media has been previously treated by the authors in Ref. 6. There an exact expression for the leading term of this *canonical problem* has been obtained. (The term "canonical problem" and the idea of the "indirect method," were introduced by Keller in Ref. 4.) It can be shown that this expression is identical to the result obtained above (and specialized to nondissipative homogeneous media), except that $\tilde{\mathbf{z}}$ is given explicitly. The value of $\tilde{\mathbf{z}}$ determined from the canonical problem may be inserted into (3.40) to complete the solution of our original problem.

From the results obtained in Ref. 6, Sec. 2, we find that

$$\tilde{\mathbf{z}} = \frac{\lambda^{d-2}}{2\pi}\left\{\frac{m_0\,|(m_0)_\omega|}{v\,|j|}\right\}^{\frac{1}{2}} \exp\left\{\frac{\pi i}{4}[\operatorname{sgn} j - \operatorname{sgn}(m_0)_\omega]\right\} \times \sum_{j=1}^{2}(\hat{\mathbf{g}}, \mathbf{r}^j[\tau])\mathbf{r}^j[\tau]. \tag{3.43}$$

In (3.43) $\tilde{j}$ is given by (3.29) and $\hat{\mathbf{g}}(\mathbf{P})$ is defined by

$$\hat{\mathbf{g}}(\mathbf{P}) = \int_{-\infty}^{\infty} \exp\{-i\mathbf{K}_0\cdot\mathbf{Q}\}\mathbf{g}(\tau, \mathbf{Q})\,d\mathbf{Q}; \quad d\mathbf{Q} = dq_1\,dq_2\,dq_3. \tag{3.44}$$

Thus, we see that $\hat{\mathbf{z}}$ is indeed of the form (3.42). In fact, we find that the Cerenkov radiation coefficient $\mathcal{C}$ is given by

$$\mathcal{C} = \psi DF. \tag{3.45}$$

Here

$$\psi = \frac{\lambda^{d-2}}{2\pi}\left\{\frac{m_0\,|(m_0)_\omega|}{v\,|\tilde{j}|}\right\}^{\frac{1}{2}} \exp\left\{\frac{\pi i}{4}\,[\operatorname{sgn}\tilde{j} - \operatorname{sgn}(m_0)_\omega]\right\}, \tag{3.46}$$

and D and F are linear operators defined by

$$F\mathbf{g} = \hat{\mathbf{g}}, \tag{3.47}$$

$$D\hat{\mathbf{g}} = \sum_{j=1}^{2}(\hat{\mathbf{g}}, \mathbf{r}^j[\tau])\mathbf{r}^j[\tau]. \tag{3.48}$$

We see from (3.44) that F is a Fourier transform and, from (3.48), that D is a dyadic. By inserting the expression for $\tilde{\mathbf{z}}$ just obtained into (3.40), we have

$$\tilde{\beta}_j(\tau) = \psi(\hat{\mathbf{g}}, \mathbf{r}^j[\tau]);\quad j = 1, 2. \tag{3.49}$$

Equations (3.38) and (3.39) then yield the values of $\tilde{\beta}_j(t)$ and, finally, (3.33) and (3.39) determine $\tilde{\mathbf{z}}(t; \mathbf{P})$.

The leading term of the asymptotic expansion of $\mathbf{u}$ is obtained by inserting the relations for $s(t; \mathbf{P})$ and $\mathbf{z}(t; \mathbf{P})$ derived above into (3.1) and taking the real part of the result. Thus

$$\mathbf{u} \sim \operatorname{Re}\,[\exp\{i\lambda s(t; \mathbf{P})\}\mathbf{z}(t; \mathbf{P})]. \tag{3.50}$$

More precisely, (3.50) and the rays

$$\mathbf{X} = \mathbf{X}(t; \mathbf{P}) \tag{3.51}$$

yield a parametric representation of the asymptotic expansion. To obtain $\mathbf{u}$ at a given space–time point $(t, \mathbf{X})$, (3.50) is to be summed over all values of $\mathbf{P}$ which lie in the domain $\mathcal{T}$ and satisfy (3.51). That is, we sum over all rays which pass through $(t, \mathbf{X})$.

4. ENERGY OF CERENKOV RADIATION FOR INHOMOGENEOUS ISOTROPIC MEDIA

In this section we shall obtain an expression for $W(t, \tau_1)$, the total energy, measured at time $t > \tau_1$, radiated from the source as it traverses the portion of trajectory defined by $0 \le \tau \le \tau_1$. With this purpose in mind we introduce the "average asymptotic energy density" $w(t, \mathbf{X})$ defined by

$$w(t, \mathbf{X}) = \frac{1}{16\pi}(\mathbf{z}[t; \mathbf{P}], A^0\mathbf{z}[t; \mathbf{P}]). \tag{4.1}$$

It is understood that the right-hand side of (4.1) is summed over all rays which pass through the point $(t, \mathbf{X})$; i.e., over all values of $\mathbf{P} = (\tau, \omega, \gamma_0)$ which, for fixed $(t, \mathbf{X})$, satisfy (3.19). It can be shown that

$$w_t + \nabla\cdot\langle\mathbf{S}\rangle = -2\omega\eta, \tag{4.2}$$

where $\langle\mathbf{S}\rangle$ is the average over a small time interval of the Poynting vector $\mathbf{S} = (c/4\pi)\mathbf{E}\times\mathbf{H}$. If our system is conservative, $\eta = 0$ and (4.2) is the well-known equation of energy conservation. This justifies our designation of w as an energy density.

The total energy $W(t, \tau_1)$ defined above is then given by

$$W(t, \tau_1) = \int w(t, \mathbf{X})\,d\mathbf{X} = \frac{1}{16\pi}\int(\mathbf{z}, A^0\mathbf{z})\,d\mathbf{x}. \tag{4.3}$$

The integrand in the last term of (4.3) is summed over all values of $\mathbf{P}$ such that the corresponding ray passes through $(t, \mathbf{X})$ and $0 \le \tau \le \tau_1$.

The ray transformation (3.19) maps the parameter space $\mathcal{T}$ in a one-to-one manner on a multiple $\mathbf{X}$ space which consists of κ replicas of physical $\mathbf{X}$ space. Here κ is the maximum number of rays which pass through any point $\mathbf{X}$ at time t. The change of variables from $\mathbf{X}$ to $\mathbf{P}$ yields the simple result

$$W(t, \tau_1) = \frac{1}{16\pi}\int_{\mathcal{T}_1}|j(t; \mathbf{P})|\,(\mathbf{z}[t; \mathbf{P}], A^0\mathbf{z}[t; \mathbf{P}])\,d\mathbf{P}. \tag{4.4}$$

Here $\mathcal{T}_1$ is the domain defined by (3.20) with the added restriction $\tau \le \tau_1$. From Eqs. (3.33), (3.37), (3.40), (3.43), and the orthonormality condition (3.30), we obtain

$$(\mathbf{z}[t; \mathbf{P}], A^0\mathbf{z}[t; \mathbf{P}]) = \frac{\lambda^{2(d-2)}}{4\pi^2}\frac{m_0\,|(m_0)_\omega|}{v\,|j(t; \mathbf{P})|} \times \exp\left[-2\int_t^t \eta(t')\,dt'\right]\sum_{j=1}^{2}|(\hat{\mathbf{g}}, \mathbf{r}^j)|^2. \tag{4.5}$$

Then, by inserting (4.5) into (4.4) and noting the definition of $\mathcal{T}_1$, we have

$$W(t, \tau_1) = \frac{1}{64\pi^3}\int_0^{2\pi} d\gamma_0\int_0^{\tau} d\tau\int d\omega\,\frac{m_0\,|(m_0)_\omega|}{v(\tau)} \times \exp\left\{-2\int_\tau^t \eta(t')\,dt'\right\}\sum_{j=1}^{2}|(\hat{\mathbf{g}}, \mathbf{r}^j[\tau])|^2.$$

$$0 \le \left\{\frac{\omega + \dot{q}(\tau)}{v(\tau)m_0(\omega, \tau)}\right\}^2 \le 1. \tag{4.6}$$

We observe that the total energy W is not conserved due to the presence in (4.6) of the dissipative factor $\exp\{-2\int_\tau^t \eta(t')\,dt'\}$. Moreover, we note that, in order to evaluate this exponential, it is necessary to first solve the ray equations (3.9). For nondissipative media, however, $\eta \equiv 0$; as a result W may be obtained without solving the ray equations. This is an important fact because, as we have pointed out in Sec. 3B, the rays in general cannot be determined. We can conclude, therefore, that for conservative inhomogeneous media, the total energy can always be obtained, whereas in many cases the fields cannot be completely determined.

Let us suppose that the medium is nondissipative. By

$$W^*(\tau_1) = \frac{1}{v(\tau_1)}\frac{dW(\tau_1)}{d\tau_1},$$

we denote the energy radiated per unit path length of the source trajectory. After setting $\eta = 0$ in (4.6) we obtain

$$W^*(\tau_1) = \frac{1}{v^2(\tau_1)64\pi^3}\int_0^{2\pi} d\gamma_0 \int d\omega m_0(\omega, \tau_1)$$
$$\times \frac{\partial}{\partial\omega}[m_0(\omega, \tau_1)]\sum_{j=1}^{2}|(\hat{\mathbf{g}}, \mathbf{r}^j[\tau_1])|^2,$$
$$0 \le \left\{\frac{\omega + \dot{q}(\tau_1)}{v(\tau_1)m_0(\omega, \tau_1)}\right\}^2 \le 1. \quad (4.7)$$

Equation (4.1) holds for arbitrary, inhomogeneous, nondissipative media. If we specialize to homogeneous media by choosing the constant value of $\mathcal{E}$ equal to $\mathcal{E}(\omega, Y[\tau_1])$, the value of $W^*(\tau_1)$ is unaltered. This is another expression of the fact that for conservative media the total energy depends on the behavior of $\mathcal{E}$ only at the source trajectory, i.e., it is independent of the rays.

5. PLANE-STRATIFIED MEDIA

As was pointed out in Sec. 3B the ray equations (3.9) cannot, in general, be solved in closed form. In this section we shall consider the simplest inhomogeneous media for which such a solution can be obtained. In particular, we shall assume that the matrix $\tilde{\mathcal{E}}$, defined by Eqs. (2.10) and (2.11), is a function only of ω, x_1, and λ. A matrix $\tilde{\mathcal{E}}$ of this form represents what we call a *plane-stratified*, weakly dissipative, isotropic medium. (Here, we have chosen, with no loss of generality, the x_1 direction as the direction of stratification.)

The dispersion relation (3.8) now takes the form

$$k = m(\omega, x_1) = \frac{|\omega|}{c}n(\omega, x_1)$$
$$= \frac{|\omega|}{c}[\epsilon(\omega, x_1)\mu(\omega, x_1)]^{\frac{1}{2}}. \quad (5.1)$$

The analysis of the rays is greatly simplified if we avoid the occurrence of what are called "turned rays." (In this case the turning point of a ray occurs when $dx_1/dt = 0$.) With this purpose in mind we impose the following conditions. We assume that for fixed x_1, n_ω is positive for $\omega > 0$ (i.e., the dispersion is normal), and that for fixed ω, $n(\omega, x_1)$ is a monotonically increasing function of x_1. We must also place conditions on the source function $\mathbf{f}$. We assume that $q(t) \equiv 0$ and that the source trajectory in is the direction of the positive x_1 axis. That is,

$$\mathbf{Y}(t) = y(t)\hat{\mathbf{X}}_1; \quad \dot{y}(t) = v(t) > 0. \quad (5.2)$$

We see below that, as a result of these conditions, dx_1/dt is always positive and hence there will be no turned rays. For a treatment of the case of turned rays see Bleistein.[8]

We now define $\mathbf{u}_\pm$ to be the contribution to the asymptotic expansion of $\mathbf{u}$ corresponding to positive (negative) values of ω. Thus

$$\mathbf{u} \sim \mathbf{u}_+ + \mathbf{u}_-. \quad (5.3)$$

These quantities may be determined separately. We first consider $\mathbf{u}_+$; therefore in what follows we restrict ω to be positive. The assumption $n_\omega > 0$ for $\omega > 0$ implies that $m_\omega > 0$ for $\omega > 0$. We see from (5.2) that $\mathbf{T}^* = \hat{\mathbf{X}}_1$, and therefore we may select $\mathbf{N}^* = \hat{\mathbf{X}}_2$ and $\mathbf{B}^* = \hat{\mathbf{X}}_3$. Equation (3.17) then yields

$$\mathbf{K} = m(\omega, x_1)[\cos\theta, \sin\theta\cos\gamma, \sin\theta\sin\gamma]. \quad (5.4)$$

The initial value of $\cos\theta$ along a ray is determined by setting $q(\tau)$ equal to zero in Eq. (3.16). This yields

$$\cos\theta_0 = \frac{\omega}{v(\tau)m_0(\omega, \tau)}. \quad (5.5)$$

It follows from (3.9) that because m is independent of x_2 and x_3, k_2 and k_3 are constant along a ray. This, in turn, implies that the angle γ is constant along a ray, an expression of the fact that the space projections of the rays are plane curves. Equations (5.4) and (5.5) yield

$$k_{2,3} = k_{20,30} = \left\{m_0^2 - \frac{\omega^2}{v^2}\right\}^{\frac{1}{2}}(\cos\gamma, \sin\gamma) \quad (5.6)$$

and

$$\sin\theta = \frac{\{m_0^2 - (\omega^2/v^2)\}^{\frac{1}{2}}}{m(\omega, x_1)}. \quad (5.7)$$

Then, from Eqs. (5.1) and (5.6), we obtain

$$k_1 = \left\{m^2(\omega, x_1) - m_0^2(\omega, \tau) + \frac{\omega^2}{v^2(\tau)}\right\}^{\frac{1}{2}} = \Phi(x_1; \tau, \omega). \quad (5.8)$$

The positive square root is chosen because, as seen from (5.4) and (5.5), $k_{10} = \omega/v(\tau)$.

We now complete the determination of the rays. By inserting (5.8) into the first of Eqs. (3.9) we have

$$\frac{dx_1}{dt} = \frac{\Phi(x_1; \tau, \omega)}{m(\omega, x_1)m_\omega(\omega, x_1)}. \quad (5.9)$$

[8] N. Bleistein, Ph.D. thesis, New York University (1965).

It is easily seen from (5.8) and (5.9) that the assumptions made above concerning the behavior of $n(\omega, x_1)$ imply that dx_1/dt is always positive. Equation (5.7) then shows that $\sin\theta$ decreases monotonically along a ray. Therefore the space projections of the rays continuously bend toward the positive x_1 axis with increasing t.

Equation (5.9) may be integrated to obtain

$$(t-\tau) = \int_{y(\tau)}^{x_1} [\Phi(\xi; \tau, \omega)]^{-1} m(\omega, \xi) m_\omega(\omega, \xi)\, d\xi; \qquad t \geq \tau. \quad (5.10)$$

This defines x_1 implicitly as a function of t, τ, and ω. x_2 and x_3 may also be expressed in terms of integrals. In fact, Eqs. (3.9) and (5.6) yield

$$x_{2,3} = (\cos\gamma, \sin\gamma)\left[m_0^2(\omega, \tau) - \frac{\omega^2}{v^2(\tau)}\right]^{\frac{1}{2}} \times \int_{y(\tau)}^{x_1} [\Phi(\xi; \tau, \omega)]^{-1}\, d\xi. \quad (5.11)$$

Equation (5.11) shows that, once x_1 is obtained from (5.10), x_2 and x_3 are expressed in terms of t and the parameters $\mathbf{P} = (\tau, \omega, \gamma)$. Thus the rays $(t, \mathbf{X}) = [t, \mathbf{X}(t; \mathbf{P})]$ are completely determined by Eqs. (5.10) and (5.11). The values of $\mathbf{P}$ lie in the domain $\mathcal{T}_+$ defined by

$$0 \leq \tau,\ 0 < \omega,\ 0 \leq \gamma < 2\pi,\ 0 \leq \frac{\omega^2}{v^2(\tau) m_0^2(\omega, \tau)} \leq 1. \quad (5.12)$$

The phase, $s(t; \mathbf{P})$, is given along a ray by Eqs. (3.21) and (3.22). However, by making use of (5.9), we arrive at the more convenient expression

$$s(t; \mathbf{P}) = \int_{y(\tau)}^{x_1(t;\mathbf{P})} m^2(\omega, \xi)[\Phi(\xi; \tau, \omega)]^{-1}\, d\xi - \omega(t-\tau). \quad (5.13)$$

The Jacobian $j(t; \mathbf{P})$ may be obtained from Eqs. (5.10) and (5.11). The result is

$$j(t; \mathbf{P}) = \frac{\Phi I_3}{m m_\omega}\left\{I_3 l_1\left(\frac{v^2}{\omega} l_1 - I_2 l_2\right) + l_2 I_3 I_2 + \left(m_0^2 - \frac{\omega^2}{v^2}\right)\left[l_2 I_2 I_4 - \frac{v^2 I_2}{\omega} + (l_1 I_4 - I_1)\left(\frac{v^2}{\omega} l_1 - l_2 I_1\right)\right]\right\}. \quad (5.14)$$

Here

$$l_1 = \left[m_0(m_0)_\omega - \frac{\omega}{v^2}\right], \quad l_2 = \left[m_0(m_0)_\tau + \frac{\omega^2 \dot{v}}{v^3}\right] \quad (5.15)$$

and

$$I_1(t; \mathbf{P}) = \int_{y(\tau)}^{x_1(t;\mathbf{P})} m(\omega, \xi) m_\omega(\omega, \xi)[\Phi(\xi; \tau, \omega)]^{-3}\, d\xi, \quad (5.16)$$

$$I_2(t; \mathbf{P}) = \int_{y(\tau)}^{x_1(t;\mathbf{P})} \{m(\omega, \xi) m_\omega(\omega, \xi)[\Phi(\xi; \tau, \omega)]^{-1}\}_\omega\, d\xi, \quad (5.17)$$

$$I_3(t; \mathbf{P}) = \int_{y(\tau)}^{x_1(t;\mathbf{P})} [\Phi(\xi, \tau, \omega)]^{-1}\, d\xi, \quad (5.18)$$

$$I_4(t; \mathbf{P}) = \int_{y(\tau)}^{x_1(t;\mathbf{P})} [\Phi(\xi; \tau, \omega)]^{-3}\, d\xi. \quad (5.19)$$

We now determine the amplitude function $\mathbf{z}$. It follows from (5.4) that we may select

$$\hat{\mathbf{N}} = [\sin\theta, -\cos\theta\cos\gamma, -\cos\theta\sin\gamma] \quad \text{and } \hat{\mathbf{B}} = [0, \sin\gamma, -\cos\gamma]. \quad (5.20)$$

By inserting (5.20) into (3.31), expressions for $\mathbf{r}^1(t)$ and $\mathbf{r}^2(t)$ are obtained. We have seen above that γ is constant along a ray. Differentiation of the second of Eqs. (5.20) with respect to t then yields $d\hat{\mathbf{B}}/dt \equiv 0$. We see from (3.35) that this in turn implies $\delta(t) \equiv 0$. It then follows from Eqs. (3.38) and (3.39) that $\beta_j(t) = \beta_j(\tau)$, $j = 1, 2$. By inserting the quantities $\beta_j(\tau)$ as given by (3.46) and (3.49) into (3.37), we obtain

$$\mathbf{z}(t; \mathbf{P}) = \frac{\lambda^{d-2}}{2\pi}\exp\left\{-\int_\tau^t \eta(t')\, dt' + \frac{\pi i}{4}(\operatorname{sgn}\tilde{j} - 1)\right\} \times \left\{\frac{m_0(m_0)_\omega}{v|j|}\right\}^{\frac{1}{2}} \sum_{j=1}^{2} (\hat{\mathbf{g}}, \mathbf{r}^j[\tau])\mathbf{r}^j[t]. \quad (5.21)$$

Here $\eta(t')$, $\tilde{j}$, and j are given by (3.36), (3.29), and (5.14), respectively.

The parametric representation of $\mathbf{u}_+$ is determined by inserting Eqs. (5.13) and (5.21) into

$$\mathbf{u}_+(t, \mathbf{X}) = \operatorname{Re}\,[\exp\{i\lambda s(t; \mathbf{P})\}\mathbf{z}(t; \mathbf{P})], \quad (5.22)$$

where, for fixed $(t, \mathbf{X})$, we sum the right-hand side of (5.22) over all values of $\mathbf{P}$ which lie in $\mathcal{T}_+$ and satisfy (5.10) and (5.11). To complete the determination of the asymptotic expansion of $\mathbf{u}$, we must obtain $\mathbf{u}_-$. By repeating the above analysis with $\omega < 0$, it can be shown that the parametric representation of $\mathbf{u}_-$ is identical to the parametric representation of $\mathbf{u}_+$. Therefore, it follows from (5.3) that

$$\mathbf{u} \sim 2\mathbf{u}_+. \quad (5.23)$$

This completes the asymptotic solution of the problem we have selected.

One reason we have selected this particular problem is that our results can be checked by another method. The alternative method is much more difficult than the one we have presented above, and for more general inhomogeneous media it fails entirely. Rather than give the analysis, which is long and computationally awkward, we shall briefly outline the procedure and quote the results.

We introduce (in the standard way) the potential functions $\mathbf{A} = (a_1, a_2, a_3)$ and ϕ. We then take Fourier transforms of all relevant quantities with respect to t, x_2, and x_3, where the transformation variables are ω, k_2 and k_3. We denote the transform of $\mathbf{A}$ by $\hat{\mathbf{A}} = (\hat{a}_1, \hat{a}_2, \hat{a}_3)$. It follows from the symmetry of the problem that $\hat{a}_1$ is the only nonzero component of $\hat{\mathbf{A}}$. If we set $\Psi = \hat{a}_1/m(\omega, x_1)$, it can be shown that Ψ must satisfy an ordinary differential equation of the form

$$(d^2\Psi/dx_1^2) + \lambda^2[m^2(\omega, x_1) - k_2^2 - k_3^2]\Psi + b(x_1)\Psi = r(x_1; \lambda), \quad (5.24)$$

and obey certain radiation conditions at $x_1 = \pm\infty$.

The asymptotic expansion of Ψ can be obtained using the WKB method. It is then a simple matter to obtain integral expressions of the Fourier type for the asymptotic expansions of the electromagnetic fields. These integrals, in turn, may be evaluated asymptotically by the method of stationary phase to obtain the leading term of the expansion of $\mathbf{u}$. The results of this analysis are in perfect agreement with those obtained in this section by the ray method. Moreover, comparison of the two procedures shows that the ray method is significantly simpler in its application.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Einstein Tensor and Spherical Symmetry

JERZY PLEBANSKI*
Centro de Investigacion y de Estudios, Avanzados del I.P.N., Mexico
AND
JOHN STACHEL
Boston University, Boston, Massachusetts

(Received 3 April 1967)

The classification of symmetric second-rank tensors in Minkowski space and its application to the Einstein tensor is reviewed. It is shown that, for spherically symmetric metrics, the Einstein tensor always has a spacelike double eigenvector; and the possible types of Einstein tensor that this degeneracy allows are discussed. A complete classification of all spherically symmetric metrics with two double eigenvalues is given. A study of the timelike eigencongruence, in the case when one timelike and two spacelike eigenvectors exist, is carried out. Canonical forms for the metric, the Einstein tensor, and the Weyl tensor (which is always of type D) are given for each of the various possible types.

1. INTRODUCTION

As is well known, the invariant characterization of the gravitational field in general relativity is best given in terms of the Riemann tensor. The Riemann tensor at a point itself is decomposable into three irreducible objects under Lorentz transformations in the tangent space; the conformal curvature tensor or Weyl tensor, the traceless Ricci tensor, and the curvature scalar R. Petrov and others[1] have shown how an algebraic classification of the Weyl tensor may be carried out, and its usefulness in the physical interpretation of Einstein's theory has been amply demonstrated.[2] In empty space, the other two objects vanish, of course, and no further classification is possible or necessary. However, when the Einstein tensor does not vanish, its algebraic structure can be classified and, as we hope to show by example in this paper, may prove helpful in the physical interpretation of the corresponding metric. Such classifications were given by Churchill,[3] and independently and in more

* On leave of absence from the University of Warsaw.

[1] A. Z. Petrov, Kazan. Gos. Univ. Uc. Zap. **114**, Series 8, 55 (1954); J. Geheniau, Compt. Rend. **244**, 723 (1957); R. Penrose, Ann. Phys. **10**, 179 (1960); R. K. Sachs, Proc. Roy. Soc. (London) **A264**, 309 (1961).

[2] F. Pirani, Phys. Rev. **105**, 1089 (1957); R. K. Sachs, Proc. Roy. Soc. (London) **A270**, 103 (1962). A useful summary of the use of the Petrov classification in gravitational radiation theory is given in the series of lectures on "Gravitational Radiation Theory" by F. Pirani in *Brandeis Summer Institute in Theoretical Physics 1964, Volume One: Lectures on General Relativity* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1965).

[3] R. V. Churchill, Trans. Am. Math. Soc. **34**, 784 (1932).

is that our results can be checked by another method. The alternative method is much more difficult than the one we have presented above, and for more general inhomogeneous media it fails entirely. Rather than give the analysis, which is long and computationally awkward, we shall briefly outline the procedure and quote the results.

We introduce (in the standard way) the potential functions $\mathbf{A} = (a_1, a_2, a_3)$ and ϕ. We then take Fourier transforms of all relevant quantities with respect to t, x_2, and x_3, where the transformation variables are ω, k_2 and k_3. We denote the transform of $\mathbf{A}$ by $\hat{\mathbf{A}} = (\hat{a}_1, \hat{a}_2, \hat{a}_3)$. It follows from the symmetry of the problem that $\hat{a}_1$ is the only nonzero component of $\mathbf{A}$. If we set $\Psi = \hat{a}_1/m(\omega, x_1)$, it can be shown that Ψ must satisfy an ordinary differential equation of the form

$$(d^2\Psi/dx_1^2) + \lambda^2[m^2(\omega, x_1) - k_2^2 - k_3^2]\Psi + b(x_1)\Psi = r(x_1; \lambda), \quad (5.24)$$

and obey certain radiation conditions at $x_1 = \pm\infty$.

The asymptotic expansion of Ψ can be obtained using the WKB method. It is then a simple matter to obtain integral expressions of the Fourier type for the asymptotic expansions of the electromagnetic fields. These integrals, in turn, may be evaluated asymptotically by the method of stationary phase to obtain the leading term of the expansion of $\mathbf{u}$. The results of this analysis are in perfect agreement with those obtained in this section by the ray method. Moreover, comparison of the two procedures shows that the ray method is significantly simpler in its application.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Einstein Tensor and Spherical Symmetry

JERZY PLEBANSKI*
Centro de Investigacion y de Estudios, Avanzados del I.P.N., Mexico
AND
JOHN STACHEL
Boston University, Boston, Massachusetts

(Received 3 April 1967)

The classification of symmetric second-rank tensors in Minkowski space and its application to the Einstein tensor is reviewed. It is shown that, for spherically symmetric metrics, the Einstein tensor always has a spacelike double eigenvector; and the possible types of Einstein tensor that this degeneracy allows are discussed. A complete classification of all spherically symmetric metrics with two double eigenvalues is given. A study of the timelike eigencongruence, in the case when one timelike and two spacelike eigenvectors exist, is carried out. Canonical forms for the metric, the Einstein tensor, and the Weyl tensor (which is always of type D) are given for each of the various possible types.

1. INTRODUCTION

As is well known, the invariant characterization of the gravitational field in general relativity is best given in terms of the Riemann tensor. The Riemann tensor at a point itself is decomposable into three irreducible objects under Lorentz transformations in the tangent space; the conformal curvature tensor or Weyl tensor, the traceless Ricci tensor, and the curvature scalar R. Petrov and others[1] have shown how an algebraic classification of the Weyl tensor may be carried out, and its usefulness in the physical interpretation of Einstein's theory has been amply demonstrated.[2] In empty space, the other two objects vanish, of course, and no further classification is possible or necessary. However, when the Einstein tensor does not vanish, its algebraic structure can be classified and, as we hope to show by example in this paper, may prove helpful in the physical interpretation of the corresponding metric. Such classifications were given by Churchill,[3] and independently and in more

* On leave of absence from the University of Warsaw.

[1] A. Z. Petrov, Kazan. Gos. Univ. Uc. Zap. **114**, Series 8, 55 (1954); J. Geheniau, Compt. Rend. **244**, 723 (1957); R. Penrose, Ann. Phys. **10**, 179 (1960); R. K. Sachs, Proc. Roy. Soc. (London) **A264**, 309 (1961).

[2] F. Pirani, Phys. Rev. **105**, 1089 (1957); R. K. Sachs, Proc. Roy. Soc. (London) **A270**, 103 (1962). A useful summary of the use of the Petrov classification in gravitational radiation theory is given in the series of lectures on "Gravitational Radiation Theory" by F. Pirani in *Brandeis Summer Institute in Theoretical Physics 1964, Volume One: Lectures on General Relativity* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1965).

[3] R. V. Churchill, Trans. Am. Math. Soc. **34**, 784 (1932).

detail by one of the authors[4] from the purely algebraic point of view. Just as the algebraic classification of the Weyl tensor at a point together with certain properties of metric in the large has proved helpful in the study of empty-space solutions of the Einstein theory; so it is hoped that additional study of the algebraic structure of the Ricci tensor (or the equivalent Einstein tensor), when combined with global properties of the metric, may be useful in the study of metrics representing nonempty spaces. In this paper, we take this approach in what is perhaps the simplest possible case. We apply the classification scheme to the class of spherically symmetric metrics. We distinguish the various possible subclasses and investigate them in more or less detail, singling out the cases where a physical interpretation of metrics of the subclass has been found possible. A number of well-known metrics are seen to emerge naturally from the classification scheme, independently of direct application of any field equations, and canonical forms of the metric are given for a number of cases where solutions of a given physical type may be expected to occur. Thus, we hope to demonstrate the usefulness of the classification of the Einstein tensor (in conjunction with that of the Weyl tensor) as a tool in the search for interesting solutions to the nonempty-space field equations and in their physical interpretation.

In Sec. 2 of the paper, we briefly review the general classification scheme for the Einstein tensor. Section 3 applies this scheme to the case of spherical symmetry. We then discuss the various subclasses in more detail, giving useful canonical forms of the metric wherever possible, eigenvalues of the Einstein tensor, the invariant of the Weyl tensor, etc., as well as particular metrics of physical interest. A summary follows with some indications of remaining problems and other possible applications of the methods discussed here. Finally, we include two Appendices, giving the Einstein tensor, Weyl tensor, and other useful results for a number of canonical forms of spherically symmetric metrics.

2. CLASSIFICATION OF THE EINSTEIN TENSOR

The Einstein tensor is, of course, a second-rank symmetric tensor and the basic problem is therefore the classification of a second-rank symmetric tensor in a Riemann space. The Churchill classification depends on the study of the invariant planes associated with the tensor and is more geometrical. The method of Ref. 4 is based on the study of the eigenvectors of the matrix associated with the tensor and is more algebraic. In this section we outline both basic classifications and show the relationship between the two. No proofs are given; they may be found in Refs. 3 and 4.

If we are given a second-rank symmetric tensor $A^{\mu\nu}$ and a metric tensor at a point, we may use the metric tensor to put the symmetric tensor into the mixed form $A_\mu{}^\nu$. In this form it may be looked upon as an operator which operates on a contravariant vector to produce another one:

$$w^\nu = A_\mu{}^\nu v^\mu. \tag{2.1}$$

If we think of it as a matrix, we may study the eigenvalues and eigenvectors of this matrix independent of their reality. These eigenvectors, if they are real, define invariant directions left unchanged by the operator. We may also look for invariant planes (i.e., planes, or 2-flats, as Synge[5] calls them, in the tangent space) such that the operator always takes a vector in the invariant plane into another vector in this plane. Every mixed second-rank tensor, symmetric or not, has at least one such invariant plane.[6]

If the metric tensor is positive (or negative) definite in character, a real orthogonal transformation always exists which diagonalizes the matrix, so that four real eigenvectors always exist, defining the Ricci principal directions.[7] The planes which contain any pair of these eigenvectors are then invariant planes; and, apart from degeneracy of some of the eigenvalues which can make any vector in an invariant plane an eigenvector, no further classification with respect to eigenvectors or invariant planes is possible.

However, with an indefinite metric, and in particular with the Minkowski metric, the tensor may be situated in various nonequivalent ways with respect to the light cone of null directions; and, the various possibilities that arise here give rise to distinct types even before the degeneracy of eigenvalues is taken into account. Churchill showed there are four such distinct types with respect to the properties of the invariant planes of the tensor. In Ref. 4, the same four basic types were obtained through a study of the

[4] J. Plebanski, Acta Phys. Polon. **26**, 963 (1964). An algebraic classification of second-rank symmetric tensors in Minkowski space has also been given by A. Z. Petrov; details of this classification together with references to earlier work on this subject may be found in A. Z. Petrov, *New Methods in the General Theory of Relativity* (in Russian) (Science Publishers, Moscow, 1966); English transl. of the first ed. of this work: A. Z. Petrov, *Einstein Spaces* (Pergamon Press, Inc., London, 1964). Roger Penrose has carried out a classification of traceless symmetric second-rank tensors as a special case of a general classification scheme for arbitrary spinors (private communication).

[5] J. L. Synge, *Relativity: The Special Theory* (North-Holland Publishing Company, Amsterdam, 1964).

[6] G. Y. Rainich, Trans. Am. Math. Soc. **27**, 106 (1925).

[7] G. Ricci, Atti R. Istituto Veneto (VIII) **6**, 1233 (1904). See, for example, J. A. Schouten, *Ricci-Calculus* (Springer-Verlag, Berlin, 1954).

eigenvectors; and by study of the possible degeneracies within each type, a much more detailed classification of symmetric second-rank tensors in Minkowski space followed.

We discuss the Churchill approach[3] first. Since every second-rank tensor has at least one invariant plane, the first question is how many distinct types of planes can occur in Minkowski space. There are three kinds: timelike planes, which cut the null cone in two null directions; spacelike planes which do not cut the null cone at all, and therefore contain no null directions; and null planes, which are tangent to the null cone and therefore contain one null direction. Suppose the tensor has a spacelike invariant plane; using the fact that it is a symmetric tensor it can be shown that the orthogonal plane, which is timelike, is also invariant. The converse also holds: If the invariant plane is timelike, the orthogonal spacelike plane is also invariant. Since the spacelike plane has a positive definite metric, there always exist two real spacelike eigenvectors in this case. The timelike plane, having an indefinite metric, is more complicated and three possibilities can occur: two real orthogonal eigenvectors exist, in which case one is spacelike and the other timelike; no real eigenvectors exist; or, one double null eigenvector exists. Churchill calls these Case Ia, Ib and Ic, and Id, respectively—b and c being merely designations for alternate canonical forms for the same case. If the invariant plane is null, it is unique, and Churchill calls this Case II. He shows that, by appropriate choice of the orthonormal tetrad, the physical components of $A_{\hat{\mu}}{}^{\nu}$ in each of these cases may be brought to the corresponding canonical form shown in Table I.

TABLE I. Classification of Einstein tensor by invariant planes (Churchill). Subclassifications may be made on the basis of coincidence of various eigenvalues, see Table III.

Type	Canonical Form	Geometric Characterization
I		Two orthogonal invariant planes, one spacelike, one timelike; two eigenvectors in the spacelike plane.
I(a)	$\begin{pmatrix} \omega_0 & & & \\ & \omega_1 & & \\ & & \omega_2 & \\ & & & \omega_3 \end{pmatrix}$	(a): One timelike, one spacelike eigenvector in the timelike plane.
I(b)	$\begin{pmatrix} 0 & b & & \\ -b & a & & \\ & & \omega_2 & 0 \\ & & 0 & \omega_3 \end{pmatrix}$	(b) [or (c)]: No (real) eigenvector in the timelike plane.
I(d)	$\begin{pmatrix} 1 & 1 & & \\ -1 & -1 & & \\ & & \omega_2 & 0 \\ & & 0 & \omega_3 \end{pmatrix}$	(d): One double null eigenvector in the timelike plane; null eigenvalue = 0.
II	$\begin{pmatrix} \omega_0 & 0 & -1 & 0 \\ 0 & \omega_0 & 1 & 0 \\ 1 & 1 & \omega_0 & 0 \\ 0 & 0 & 0 & \omega_3 \end{pmatrix}$	An invariant null plane; one triple null vector, one spacelike eigenvector.

Reference 4 considers the matrix corresponding to $A_{\hat{\mu}}{}^{\nu}$. which we shall symbolize by $\mathbf{A}$ when we are considering it as a matrix. Let its *distinct* eigenvalues be symbolized by $A_1', A_2', \cdots, A_k'$ with respective multiplicities $n_1, n_2, \cdots, n_k$ (of course $n_1 + n_2 + \cdots + n_k = 4$). $\mathbf{A}$ will always obey a minimal equation of the form

$$(\mathbf{A} - \mathbf{I}A_1')^{m_1}(\mathbf{A} - \mathbf{I}A_2')^{m_2} \cdots (\mathbf{A} - \mathbf{I}A_k')^{m_k} = 0, \quad (2.2)$$

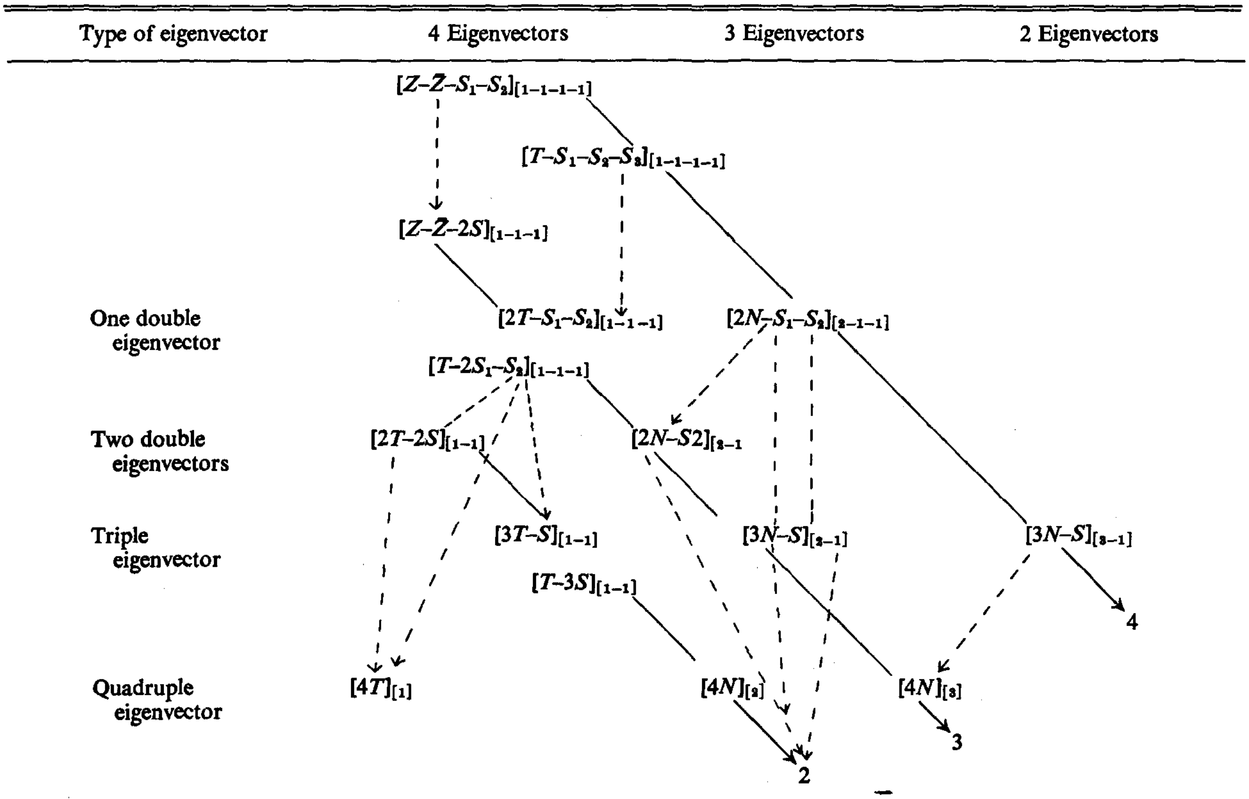
with $m_i \leq n_i$, of course. Each eigenvalue naturally has a corresponding eigenspace, spanned by the eigenvectors, whose dimensionality depends on the multiplicity of the eigenvalue. If the eigenvalue is real, the eigenspace is real as well. If this eigenspace contains a (real) timelike vector, we symbolize the eigenvalue by T. If its eigenspace does not contain a timelike vector but does contain a (real) null vector, we symbolize the eigenvalue by N. If its eigenspace contains only (real) spacelike eigenvectors, we symbolize the eigenvalue by S. Finally, if the eigenvalue is complex, we symbolize it by Z. The following symbol serves to sum up this information:

$$[n_1A_1' - n_2A_2' - \cdots - n_kA_k']_{[m_1-m_2-\cdots-m_k]}, \quad (2.3)$$

where A_i' is to be replaced by T, N, S, or Z as appropriate for that eigenvalue. The various possible symbols then correspond to the various possible types of symmetric second-rank tensors. Table II is a chart showing all the various possible algebraic types in terms of these symbols, together with the various possible degenerations between those of higher- and lower-order minimal equations. Particular examples of many of the types are given in later sections for the case of spherically symmetric metrics.

The relationship between the two classifications is summarized in Table III. Churchill's types correspond to the four nondegenerate types with minimal equation of order four. By making the indicated equalities of eigenvalues in Churchill's canonical forms, the degenerate types of minimal order three or lower in the classification of Ref. 4 arise. We shall use the

TABLE II. Classification of the Einstein tensor by eigenvectors and character of eigenspaces (Ref. 4). The symbol T, N, or S is used if the eigenspace of an eigenvalue contains a timelike vector, no timelike but a null vector, or only spacelike vectors. Dotted lines indicate degeneracies of higher forms. The heavy lines connect types which have minimal equations of the same order.



symbols (2.3) to describe the various types in later sections.

This classification scheme may naturally be applied to the Ricci tensor, and was so applied in Ref. 4. We apply it to the Einstein tensor $G_\mu{}^\nu$; since this is the tensor which occurs in the field equations, its eigenvalues G_i' might therefore be expected to have more immediate physical significance. Of course, since the Ricci and Einstein tensors only differ by a scalar multiple of the metric tensor, the use of one or the other leads to exactly the same type for a given metric, and merely shifts all eigenvalues by an equal amount.

3. SPHERICALLY SYMMETRIC METRICS

The condition of spherical symmetry imposes a number of restrictions on the algebraic structure of the Riemann tensor at a point. In the first place, as is well known, the Weyl tensor must be of Petrov type ID, or [2–2] in the Penrose notation[1]; which means that there are two doubly-degenerate Debever vectors. Of course, the Weyl tensor may degenerate to the extent that it vanishes altogether; this case of conformally flat spherically symmetric metrics was

TABLE III. Comparison of the Churchill classification and the classification of Ref. 4. (i, j, k is any permutation of 1, 2, 3.)

	Churchill	Ref. 4
I(a)	$\omega_0 \neq \omega_1 \neq \omega_2 \neq \omega_3$	$[T\text{-}S_1\text{-}S_2\text{-}S_3]_{[1-1-1-1]}$
	$\omega_i = \omega_0$ $\omega_j \neq \omega_k \neq \omega_i$	$[2T\text{-}S_1\text{-}S_2]_{[1-1-1]}$
	$\omega_i = \omega_j \neq \omega_0 \neq \omega_k$	$[T\text{-}2S_1\text{-}S_2]_{[1-1-1]}$
	$\omega_i = \omega_0$ $\omega_j = \omega_k \neq \omega_i$	$[2T\text{-}2S]_{[1-1]}$
	$\omega_i = \omega_j = \omega_k \neq \omega_0$	$[T\text{-}3S]_{[1-1]}$
	$\omega_i = \omega_j = \omega_0 \neq \omega_k$	$[3T\text{-}S]_{[1-1]}$
	$\omega_i = \omega_j = \omega_k = \omega_0$	$[4T]_{[1]}$
I(b) or (c)	$\omega_2 \neq \omega_3$	$[Z\text{-}\bar{Z}\text{-}S_1\text{-}S_2]_{[1-1-1-1]}$
	$\omega_2 = \omega_3$	$[Z\text{-}\bar{Z}\text{-}2S]_{[1-1-1]}$
I(d)	$\omega_2 \neq \omega_3$	$[2N\text{-}S_1\text{-}S_2]_{[2-1-1]}$
	$\omega_2 = \omega_3 \neq 0$	$[2N\text{-}2S]_{[2-1]}$
	$\omega_2 = 0 \quad \omega_3 \neq 0$ or $\omega_3 = 0 \quad \omega_2 \neq 0$	$[3N\text{-}S]_{[2-1]}$
	$\omega_2 = \omega_3 = 0$	$[4N]_{[2]}$
II	$\omega_3 \neq \omega_0$	$[3N\text{-}S]_{[3-1]}$
	$\omega_3 = \omega_0$	$[4N]_{[3]}$

recently examined by one of the authors with another collaborator.[8] At any rate, the Weyl tensor is characterized in the spherically symmetric case by just one invariant, the vanishing of which signifies that the space is conformally flat.

The possible types of the Einstein tensor are also limited by the spherical symmetry of the metric, as we shall now show. In this section we give a geometrical discussion based on the symmetry properties themselves; an analytic discussion based on an explicit form for the metric is given in Appendix A.

We shall define a spherically symmetric metric as one which has a group of symmetries (isometries) which has the special orthogonal group $SO(3)$ as a subgroup (not necessarily proper), such that the orbit (or minimal invariant variety) of this group at each point shall be a two-dimensional manifold with spacelike tangent plane (the inner geometry of these surfaces is that of the two-sphere). Now we prove that the tangent plane must be an invariant plane of the Einstein tensor at each point.

The isotropy group at any point (i.e., the subgroup of the rotation group which leaves the point fixed) consists of the rotations in the tangent plane. Pick any two perpendicular directions in the spacelike tangent plane and consider unit vectors in these directions, ${}_2e^\mu$ and ${}_3e^\mu$. Then there exists a rotation $\mathcal{R}$ in the isotropy group which takes ${}_2e^\mu$ into ${}_3e^\mu$ and ${}_3e^\mu$ into $-{}_2e^\mu$ (i.e., a rotation through 90°):

$$\mathcal{R}\,{}_2e^\mu = {}_3e^\mu, \quad \mathcal{R}\,{}_3e^\mu = -{}_2e^\mu. \tag{3.1}$$

Now consider the effect of $G_\mu{}^\nu$ on ${}_2e^\mu$ and ${}_3e^\mu$:

$$G_\mu{}^\nu{}_2e^\mu = u^\nu, \quad G_\mu{}^\nu{}_3e^\mu = v^\nu; \tag{3.2}$$

if we can show that u^μ and v^ν lie in the tangent plane, then any vector in the tangent plane (which can always be expressed as a linear combination of ${}_2e^\mu$ and ${}_3e^\mu$) is also taken into a vector of the tangent plane by $G_\mu{}^\nu$; and the tangent plane is invariant. Suppose u^μ and v^μ have projections $u^\mu_\perp$ and $v^\mu_\perp$ in the plane perpendicular to the tangent plane (which is timelike, since the tangent plane is spacelike). Then we can decompose u^μ v^μ into their components with respect to these two planes:

$$u^\mu = u^\mu_\top + u^\mu_\perp, \quad v^\mu = v^\mu_\top + v^\mu_\perp. \tag{3.3}$$

Now perform the rotation $\mathcal{R}$. Since $\mathcal{R}$ is an isometry, $G_\mu{}^\nu$ is unaffected by the rotation and Eq. (3.2) is transformed into

$$G_\mu{}^\nu{}_3e^\mu = \mathcal{R}u^\nu, \quad G_\mu{}^\nu{}_2e^\mu = -\mathcal{R}v^\nu. \tag{3.4}$$

[8] R. Boyer and J. Plebanski, "Conformal Curvature and Spherical Symmetry" (report of work prior to publication).

Comparing (3.4) and (3.2), we see that

$$\mathcal{R}u^\mu = v^\mu, \quad \mathcal{R}v^\mu = -u^\mu. \tag{3.5}$$

But $u^\mu_\perp$ and $v^\mu_\perp$ are unaffected by a rotation in the tangent plane, so that (3.3) becomes

$$\mathcal{R}u^\mu = \mathcal{R}u^\mu_\top + u^\mu_\top, \quad \mathcal{R}v^\mu = \mathcal{R}v^\mu_\top + v^\mu_\perp; \tag{3.6}$$

and since $\mathcal{R}u^\mu_\top$ and $\mathcal{R}v^\mu_\top$ must still lie in the tangent plane, comparison of (3.5) and (3.6) shows that $u^\mu_\perp$ and $v^\mu_\perp$ must vanish. Thus, u^μ and v^μ do lie in the tangent plane, which must be invariant. But the existence of a spacelike invariant plane means that there must be two spacelike eigenvectors in the plane, and the equivalence of all directions in the plane under the isotropy group means all directions must be invariant. Thus, there must exist a double eigenvalue corresponding to this eigenspace. So only those types of the Einstein tensor which have a $2S$ in them, or are obtainable from these by further degeneration, can be connected with a spherically symmetric metric. Explicitly, this means the following "parent types" may be expected to occur with minimum degeneracy (i.e., minimal equation of third order):

$$[Z\text{-}\bar{Z}\text{-}2S]_{[1-1-1]},$$
$$[T\text{-}2S_1\text{-}S_2]_{[1-1-1]},$$
$$[2N\text{-}2S]_{[2-1]}.$$

The further degeneracies that may occur are shown in Table V, together with the criteria for their occurrence in terms of the explicit form of the metric used in Appendix A.

4. TWO DOUBLE EIGENVALUES

The preceding section and Appendix A show that the spherically symmetric Einstein tensor always has one double eigenvalue which we call G_1'. A look at Table V shows that there are four subclasses in which there is an additional double eigenvalue. In the $[2T\text{-}2S]_{[1-1]}$ and $[2N\text{-}2S]_{[1-1]}$ cases, these are distinct, while in the $[4T]_{[1]}$ and $[4N]_{[2]}$ cases, a further coincidence of the pairs of eigenvalues takes place. We shall investigate this class of metrics in some detail, showing that they contain a number of known models of spherically symmetric spaces of physical interest, as well as generalizations of some of these models.

First we develop a canonical form for the line element of this class of metrics, using a null coordinate related to one of the Debever vectors of the Weyl tensor (if the metric is not conformally flat), in terms of which the Einstein tensor becomes exceedingly simple. Consider the spherically symmetric line element in the standard spherical coordinates (called

TABLE IV. Canonical forms for the metric when the timelike-eigenvector congruence of the Einstein tensor has indicated properties. The Einstein tensor for each form is given in Appendix B. In addition, condition (5.2) is assumed to hold where not automatically satisfied by conditions in column two.

Properties of timelike congruence	Conditions	Canonical form of metric
Expansion free ($\theta = 0$)	$\beta + \frac{2R_0}{R} = 0, \quad R_0 \neq 0$	$(R_0R^2\,dx^0)^2 - \frac{1}{R^4}\,dr^2 - R^2\,d\omega^2$ $R = R(x^0, r)$
Rigid in Born sense	$\beta_0 = 0, \quad R_0 = 0$	$e^{2\alpha}(dx^0)^2 - e^{2\beta}\,dr^2 - r^2\,d\omega^2$ $\alpha = \alpha(x^0, r), \quad \beta = \beta(r)$
Killing vector (static metric)	$\beta_0 = 0, \quad R_0 = 0,$ $\alpha = f(x^0) - \lambda(r)$	$e^{2\lambda}(dx^0)^2 - e^{2\beta}\,dr^2 - r^2\,d\omega^2$ $\lambda = \lambda(r), \quad \beta = \beta(r)$
Geodesic	$\alpha_1 = 0$	$(dx^0)^2 - R_1^2\,dr^2 - R^2\,d\omega^2$ $R = R(x^0, r)$
Shear-free, nongeodesic	$B_0 = \frac{R_0}{R}, \quad \alpha_1 \neq 0$	$\left(\frac{R_0}{R}\right)^2(dx^0)^2 - R^2(dr^2 + d\omega^2)$ $R = R(x^0, r), \quad R_0 \neq 0$
Shear-free, geodesic	$\beta_0 = \frac{R_0}{R}, \quad \alpha_1 = 0$	$(dx^0)^2 - R^2(dr^2 + d\omega^2)$ $R = f(x^0)g(r)$

curvature coordinates by Synge[9]):

$$ds^2 = e^{\nu}(dx^0)^2 - e^{\lambda}\,dr^2 - r^2\,d\omega^2. \tag{4.1}$$

The nonvanishing components of the Einstein tensor are[9] (with x^0, x^1, x^2, x^3 equalling x^0, r, θ, φ, respectively)

$$\begin{aligned}
-G_0^0 &= e^{-\lambda}\left(\frac{-\lambda'}{r} + \frac{1}{r^2}\right) - \frac{1}{r^2},\\
-G_1^1 &= e^{-\lambda}\left(\frac{\nu'}{r} + \frac{1}{r^2}\right) - \frac{1}{r^2},\\
-G_2^2 &= -G_3^3 = \tfrac{1}{2}e^{-\lambda}\left(\nu'' + \tfrac{1}{2}\nu'^2 + \frac{\nu' - \lambda'}{r} - \tfrac{1}{2}\nu'\lambda'\right),\\
-G_0^1 &= e^{-\lambda}\dot{\lambda}/r, \quad -G_1^0 = e^{-\nu}\dot{\lambda}/r,
\end{aligned} \tag{4.2}$$

and

$$-\tfrac{1}{2}e^{-\nu}(\ddot{\lambda} + \tfrac{1}{2}\dot{\lambda}^2 - \tfrac{1}{2}\dot{\lambda}\dot{\nu}),$$

The eigenvalues of the Einstein tensor are

$$\begin{aligned}
G_1' &= G_2^2\,(\text{DOUBLE}),\\
G_2' &= \tfrac{1}{2}(G_0^0 + G_1^1) + [\tfrac{1}{4}(G_0^0 - G_1^1)^2 + G_1^0G_0^1]^{\frac{1}{2}},\\
G_3' &= \tfrac{1}{2}(G_0^0 + G_1^1) - [\tfrac{1}{4}(G_0^0 - G_1^1)^2 + G_1^0G_0^1]^{\frac{1}{2}}.
\end{aligned} \tag{4.3}$$

The condition for an additional double root is

$$\tfrac{1}{4}(G_0^0 - G_1^1)^2 + G_1^0G_0^1 = \frac{1}{r^2}\,e^{-3\lambda-\nu}\left[\left(\frac{\partial}{\partial r}\,e^{\frac{1}{2}(\lambda+\nu)}\right)^2 - \left(\frac{\partial}{\partial x^0}\,e^{\lambda}\right)^2\right] = 0. \tag{4.4}$$

Thus, a double root exists if and only if

$$\frac{\partial}{\partial r}\,e^{\frac{1}{2}(\lambda+\nu)} = \frac{\partial}{\partial x^0}\,e^{\lambda}. \tag{4.5}$$

[9] J. L. Synge, *Relativity: The General Theory* (North-Holland Publishing Company, Amsterdam, 1960). Metrics mentioned in this paper for which no special reference is given may be found in this book.

[In principle, a plus or a minus sign could occur on the right-hand side of (4.5). But, in the case of a minus sign, the coordinate transformation $x^0 \to -x^0$ changes the sign to the positive one.] Now (4.5) is the integrability condition for the existence of a function $u(x^0, r)$ such that

$$e^{\frac{1}{2}(\lambda+\nu)} = -\dot{u}, \quad e^{\lambda} = -u'. \tag{4.6}$$

(Of course, both $\dot{u}$ and u' must be nonvanishing.) Solving for e^{ν} and e^{λ} and inserting these expressions into (4.1), we find

$$ds^2 = (-1/u')\,du^2 + 2\,du\,dr - r^2\,d\omega^2. \tag{4.7}$$

But, since $\dot{u} \neq 0$, we can invert $u(x^0, r)$ and find $x^0 = x^0(u, r)$. Consequently, u can be used as a new coordinate, easily seen to be null, and the line element takes the form

$$ds^2 = F(u, r)\,du^2 + 2\,du\,dr - r^2\,d\omega^2, \tag{4.8}$$

which is thus a canonical form for a spherically symmetric metric with two double eigenvalues. It proves more convenient to represent this function F in the form $F = 1 - f(u, r)/r$ so that the canonical form is finally given by[10]

$$ds^2 = [1 - f(u, r)/r]\,du^2 + 2\,du\,dr - r^2\,d\omega^2. \tag{4.9}$$

[10] H. Bondi, Proc. Roy. Soc. (London) **A281**, 39 (1964) has considered a general form for the spherically symmetric metric based upon the use of one null or "radiation" coordinate and involving two arbitrary functions. When one of these functions (β) is set equal to zero, the metric of Eq. (4.9) results. A. Papapetrou, Ref. 11, has considered this metric along with a number of special cases and shows that it may describe the field of a radiating star in certain cases.

TABLE V. Algebraic types of Einstein tensor in spherically symmetric case. The heavy lines connect types which have minimal equations of the same order. The parameters characteristic of each type are invariantly defined, since $\frac{1}{4}(a + c)^2 - b^2 = \frac{1}{4}(G_2' - G_3')^2$, $\Delta = (G_2' - G_1')(G_3' - G_1')$, and $a - c = (G_2 - G_1') + (G_3' - G_1')$. Although b is not an invariant as such; its vanishing or nonvanishing provides a convenient criterion for distinguishing between types with the other criteria the same, but a different order of minimal equation.

Criteria			
$\frac{1}{4}(a + c)^2 - b^2 < 0$, $\Delta > 0$	$[Z\text{-}\bar{Z}\text{-}2S]_{[1-1-1]}$		
$\frac{1}{4}(a + c)^2 - b^2 > 0$, $\Delta \neq 0$		$[T\text{-}2S_1\text{-}S_2]_{[1-1-1]}$	
$\frac{1}{4}(a + c)^2 - b^2 = 0$, $\Delta \neq 0$	$[2T\text{-}2S]_{[1-1]}$ $(b = 0)$		$[2N\text{-}2S]_{[2-1]}$ $(b \neq 0)$
$a - c \neq 0$, $\Delta = 0$		$[3T\text{-}S]_{[1-1]}$ $[T\text{-}3S]_{[1-1]}$	
$a - c = 0$, $\Delta = 0$	$[4T]_{[1]}$ $(b = 0)$	$[4N]_{[2]}$ $(b \neq 0)$	
	1	2	3

The Weyl and Einstein tensors for this metric are easily computed. The null vector formed from the gradient of the null coordinate u is one of the double Debever (or principal null) vectors of the conformal curvature tensor, if the latter does not vanish. Thus, an additional element of degeneracy is present in these subcases; the null vector is also an eigenvector of the Einstein tensor. Thus, the invariant angle between an eigenvector of the Einstein tensor and an eigenvector of the conformal tensor vanishes. The conformal invariant of this class of metric is given by

$$C = \pm\tfrac{1}{8}(3C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta})^{\frac{1}{2}} = \frac{r}{8}(r^{-2}f)_{,rr}. \quad (4.10)$$

The Einstein tensor has this remarkably simple structure:

$$G^\alpha_\beta = \begin{pmatrix} \frac{1}{r^2}f_{,r} & 0 & 0 & 0 \\ \frac{1}{r^2}f_{,u} & \frac{1}{r^2}f_{,r} & 0 & 0 \\ 0 & 0 & \frac{1}{2r}f_{,rr} & 0 \\ 0 & 0 & 0 & \frac{1}{2r}f_{,rr} \end{pmatrix}. \quad (4.11)$$

Thus, the curvature scalar is given by

$$R = -G^\alpha_\alpha = -r^{-2}(rf)_{,rr}. \quad (4.12)$$

Note the simplicity of these results: all curvature invariants are linear in the single structural function $f(u, r)$; the only quantity which depends on a u derivative is G^u_r; no second derivatives of u enter the curvature. As a consequence, at any given retarded time u (assuming we choose u as a retarded null coordinate), the conformal tensor has just the same structure as it would have if $f(u, r)$ were to keep its instantaneous value for all time (i.e., if the metric were "time-independent"); the effect of time variation on the curvature tensor is an "induction" type of effect (of course, a true radiation type of effect in the Weyl tensor, which represents the propagated part of the gravitational field, is impossible in any spherically symmetric case).

Matrix (4.11) is already in a form which makes the identification of our four subcases simple. The various cases are distinguished by whether or not $f_{,u}$ vanishes, which determines whether the matrix is diagonalizable, and whether or not $(r^{-2}f_{,r})_{,r}$ vanishes, which determines whether the two double eigenvalues are equal or not.

Also, we may classify the metrics according to whether or not the conformal invariant and the Ricci scalar vanish, leading to a total of sixteen subcases of spherically symmetric spaces with two pairs of double eigenvalues. We symbolize these cases by adjoining to the previous symbol, the parenthesis (R, C) which indicates whether or not the Ricci scalar and conformal invariant vanish. Examining all sixteen possibilities, we get the following results (first come the most degenerate types):

$[4T]_{[1]}(0, 0)$ $ds^2 = du^2 + 2\,du\,dr - r^2\,d\omega^2$, Flat space;

$[4T]_{[1]}(0, C)$ $ds^2 = (1 - (2m/r))\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $m = \text{const} \neq 0$, external Schwarzchild, $C = \frac{3}{2}mr^{-3}$;

$[4T]_{[1]}(R, 0)$ $ds^2 = (1 - (\frac{1}{3}\Lambda r^2))\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $\Lambda = \text{const} \neq 0$, $R = -4\Lambda$, de Sitter space;

$[4T]_{[1]}(R, C)$ $ds^2 = (1 - (2m/r) - (\frac{1}{3}\Lambda r^2))\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $m \neq 0$, $\Lambda \neq 0$, Schwarzchild solution with cosmological const, $R = -4\Lambda$, $C = \frac{3}{2}mr^{-3}$.

The cases of type $4N$ are more interesting. $[4N]_{[2]}(0, 0)$ is empty. The remaining possible types are

$[4N]_{[2]}(0, C)$ $ds^2 = [1 - (2m(u)/r)]\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $m(u)_{,u} \neq 0$, Vadya's radiating metric,[11] $C = \frac{3}{2}m(u)r^{-3}$;

$[4N]_{[2]}(R, 0)$ $ds^2 = [1 - (\frac{1}{3}\Lambda(u)r^2)]\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $\Lambda(u)_{,u} \neq 0$," radiating" de Sitter Space $R = -4\Lambda(u)$;

$[4N]_{[2]}(R, C)$ $ds^2 = [1 - (2m(u)/r) - (\frac{1}{3}\Lambda(u)r^2)]\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $m_{,u} \neq 0$, $\Lambda_{,u} \neq 0$, generalized Vadya metric in "radiating" De sitter space, $C = \frac{3}{2}m(u)r^{-3}$, $R = -4\Lambda(u)$.

The family of type $[2T\text{–}2S]_{[1-1]}$ represents a family of static space–times. The case $[2T\text{–}2S]_{[1-1]}(0, 0)$ is empty. The remaining possible types are:

$[2T\text{–}2S]_{[1-1]}(0, C)$ $ds^2 = (1 - (2m/r) + (e^2/r^2))\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $e^2 \neq 0$, Reissner–Nordstrom solution[12] for a point charge in Maxwell electrodynamics plus a point mass, $C = \frac{3}{2}(m/r^3 - e^2/r^4)$. We make the integration const e^2 positive for the physical interpretation.

$[2T\text{–}2S]_{[1-1]}(R, 0)$ $ds^2 = (1 + 2ar + br^2)\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $a \neq 0$, $R = 12(a/r + b)$. A physical interpretation is unknown. The conformal factor for this metric is given in Appendix B.

$[2T\text{–}2S]_{[1-1]}(R, C)$ $ds^2 = (1 - (f(r)/r))\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $R = -r^{-2}(rf)_{,rr} \neq 0$, $C = (r/8)(r^{-2}f)_{,rr} \neq 0$, static point charges in nonlinear Born–Infeld type electrodynamics.[13]

Finally, the most general type $[2N\text{–}2S]_{[2-1]}$ represents a family of nonstatic space–times. The case $[2N\text{–}2S]_{[2-1]}(0, 0)$ is empty. The remaining possible types are

$[2N\text{–}2S]_{[2-1]}(0, C)$ $ds^2 = [(1 - 2m(u)/r) + (e^2(u)/r^2)]\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $e^2 \neq 0$, $m^2_{,u} + e^2_{,u} \neq 0$, the "radiating" Reissner–Nordstrom solution, $C = \frac{3}{2}(m(u)/r^3 - e^2(u)/r^4)$.

$[2N\text{–}2S]_{[2-1]}(R, 0)$ $ds^2 = (1 + 2a(u)r) + (b(u)r^2)\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $a^2_{,u} + b^2_{,u} \neq 0$, $R = 12(a(u)/r + b(u))$. A physical interpretation is unknown. The conformally flat representation of this metric is given in Appendix B.

$[2N\text{–}2S]_{[2-1]}(R, C)$ $ds^2 = [1 - (f(u, r)/r)]\,du^2 + 2\,du\,dr - r^2\,d\omega^2$, $(rf)_{,rr} \neq 0$, $(r^{-2}f)_{,rr} \neq 0$, $(r^{-2}f_{,r})_{,r} \neq 0$, $f_{,u} \neq 0$. A physical interpretation is unknown (see Papapetrou, Ref. 11).

Among the 13 nonempty cases listed above, we have found a number of familiar metrics, as well as some unexplored ones.

Birkhoff's theorem for empty spaces, and a number of generalizations of it to nonempty cases may easily be obtained from this classification by noting that all metrics of types $4T$ and $4N$ are necessarily static.

5. ONE TIMELIKE AND THREE SPACELIKE EIGENVALUES

To treat the cases $[T\text{–}2S_1\text{–}S_2]_{[1-1-1]}$ (and its degenerate forms $[T\text{–}3S]_{[1-1]}$ and $[3T\text{–}S]_{[1-1]}$), we start from a very general form of the spherically symmetric metric, always locally valid[9]:

$$ds^2 = e^{2\alpha}(dx^0)^2 - e^{2\beta}\,dr^2 - R^2\,d\omega^2; \qquad (5.1)$$

α, β, and R are arbitrary functions of x^0 and $r = x^1(\theta = x^2, \varphi = x^3)$, and $d\omega^2 = d\theta^2 + \sin^2\theta\,d\varphi^2$. In general, it is possible to impose one additional condition on the three functions without restricting the type of the Einstein tensor, since only two arbitrary functions are needed to describe spherically symmetric space–times. Various standard coordinate conditions give rise to the familiar polar, Gaussian, curvature,

[11] P. C. Vadya, Nature **171**, 260 (1953); see also R. W. Linquist, R. A. Schwartz, and C. W. Misner, Phys. Rev. **137**, B1364 (1965); A. Papapetrou, Compt. Rend. **262A**, 162 (1966).

[12] H. Reissner, Ann. Physik **50**, 106 (1916); G. Nordstrom, Verhandel. Koninkl. Ned. Akad. Wetenschap. Afdel. Natuurk. Sect. II, **26**, 1201 (1918).

[13] R. Pellicer, Ph.D. thesis, Centro de Investigacion y de Estudios Avanzados del I.P.N., Mexico (1968).

isothermal, and isotropic coordinates, as discussed for example, by Synge. However, as we shall see in a moment, by example, the imposition of certain relations among the three functions may restrict the possible algebraic type of the corresponding Einstein tensor. Once we have restricted the algebraic type of the Einstein tensor, a certain condition may not further restrict the range of metrics, even though it is suited only to that particular type or types. This may serve as an example of the need for caution in setting conditions on the coordinates in general relativity. These conditions must be investigated to see whether or not they imply restrictions on the class of metrics that can satisfy them.

The Einstein and conformal curvature tensors of the above metric are given in Appendix B. We note that $e^{2\beta}G_0^1 = -e^{2\alpha}G_1^0$, so that if we impose the coordinate condition

$$R_{10} - R_1\beta_0 - R_0\alpha_1 = 0, \quad R \neq 0, \tag{5.2}$$

the Einstein tensor will already be diagonalized with real eigenvalues; and one timelike, and three spacelike eigenvectors along the coordinate directions. The eigenvalues are given by

$$\begin{aligned}
G_1' &= G_2^2 = G_3^3 \\
&= e^{-2\beta}\left[-\frac{R_{11}}{R} - \alpha_{,11} - \alpha_{,1}^2 - \frac{\alpha_1 R_1}{R} + \frac{\beta_1 R_1}{R} + \beta_1\alpha_1\right] \\
&\quad + e^{-2\alpha}\left[\frac{R_{00}}{R} + \beta_{00} + \beta_0^2 + \frac{\beta_0 R_0}{R} - \frac{R_0\alpha_0}{R} - \alpha_0\beta_0\right], \\
G_2' &= G_0^0 = e^{-2\beta}\left[-\frac{2R_{11}}{R} - \frac{R_1^2}{R^2} + \frac{2R_1\beta_1}{R}\right] \\
&\quad + \frac{1}{R^2} + e^{-2\alpha}\left[\frac{R_0^2}{R^2} + \frac{2R_0\beta_0}{R}\right], \\
G_3' &= G_1^1 = e^{-2\beta}\left[-\frac{R_1^2}{R^2} - \frac{2R_1\alpha_1}{R}\right] \\
&\quad + \frac{1}{R^2} + e^{-2\alpha}\left[\frac{2R_{00}}{R} + \frac{R_0^2}{R^2} - \frac{2R_0\alpha_0}{R}\right].
\end{aligned} \tag{5.3}$$

Condition (5.2) is understood to hold as well. Apart from possible degeneracies, discussed below, it will then always be of type $[T\text{-}2S_1\text{-}S_2]_{[1-1-1]}$. Thus, the metric (5.1) together with the coordinate condition (5.2) represents a canonical form for spherically symmetric Einstein tensors with one timelike and three spacelike eigenvalues—one double, of course. Since the vector ${}_0e^{\mu} = e^{-\alpha}\delta_0^{\mu}$ always represents the unit timelike eigenvector under these conditions, this system of coordinates represents a natural generalization of the idea of "comoving coordinates"[14] to the general case where no assumption is made about the nature of the sources of the gravitational field, except that the stress–energy tensor has one timelike and three spacelike eigenvectors.

The coordinate condition may be interpreted in several ways. If α and R are given, it is an equation determining β_0,

$$\beta_0 = \frac{R_{10} - R_0\alpha_1}{R_1}. \tag{5.4a}$$

We see this interpretation is always possible, unless $R_1 = 0$. If β and R are given, it is an equation determining α_1,

$$\alpha_1 = \frac{R_{10} - R_1\beta_0}{R_0}, \tag{5.4b}$$

and we see that this interpretation is always possible, unless $R_0 = 0$. Finally, if α and β are given, it can be regarded as an equation for R, which is seen to be of the standard form for a linear hyperbolic equation of second order in two variables. Thus, one can in general always find a unique solution locally if additionally one gives either the values of R and its normal derivative along any noncharacteristic curve in the (x^0, r) plane, or gives the value of R along two intersecting characteristic curves (i.e., along r = const and x^0 = const).[15] This formulation may prove particularly useful in the problem of matching two solutions along some boundary.

Now we enumerate the possible degeneracies. If $G_1' = G_3' \neq G_2'$, the Einstein tensor will be of type $[T\text{-}3S]_{[1-1]}$. If $G_2' = G_3' \neq G_1'$, the Einstein tensor will be of type $[3T\text{-}S]_{[1-1]}$. If $G_1' = G_2' \neq G_3'$, the Einstein tensor will be of type $[2T\text{-}2S]_{[1-1]}$, and if $G_1' = G_2' = G_3'$, the Einstein tensor will be of type $[4T]_{[1]}$. The latter two cases of pairs of double eigenvalues have been discussed in the last section, where a more useful canonical form was given.

The type $[T\text{-}3S]_{[1-1]}$ will include as a subclass (certainly proper) those metrics which admit of a standard macroscopic hydrodynamical interpretation. In addition to being of this algebraic type, these, in addition, must satisfy such conditions to assure that the rest-mass density be positive at every point in any local Lorentz frame and that the pressure be

[14] See, for example, R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Oxford at the Clarendon Press, Oxford, England, 1934); or L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Addison-Wesley Publishing Company, Inc., Reading, Mass., 1962), 2nd ed., for discussions of comoving coordinates.

[15] See, for example, R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1962), Vol. II, for a good discussion of hyperbolic partial differential equations in two independent variables.

nonnegative. These conditions could easily be formulated in our coordinate system, but as they would not be particularly illuminating (or easily satisfied) we omit them. The type $[T\text{-}2S_1\text{-}S_2]_{[1-1-1]}$ will include all other reasonable types of macroscopic, nonzero rest mass, matter with differential pressures (or compressive forces) in the radial direction and the two directions normal to the radial one—such as might arise from spherically symmetric shells of varying matter. Whenever such an interpretation is possible, the timelike eigenvector of the Einstein tensor may be looked upon as the generalized velocity field of the material sources of the Einstein tensor. The properties of the timelike congruence everywhere tangent to the timelike eigenvector are therefore of potential physical interest.

These properties are particularly simple when expressed in our canonical coordinate system. Since the congruence is spherically symmetric, it is always hypersurface orthogonal and therefore the rotation of the congruence vanishes.[16] Of the remaining six possible invariants of the congruence, three also vanish, so that only three remain: the expansion, one component of the acceleration along the radial eigenvector of the shear tensor, and one independent eigenvalue of the shear tensor. They are given by

$$\begin{aligned} \theta &= \text{expansion} = e^{-\alpha}\left(\beta_0 + \frac{2R_0}{R}\right), \\ a &= \text{acceleration} = -\alpha_1 e^{-\beta}, \\ \sigma &= \text{shear invariant} = e^{-\alpha}\left(\beta_0 - \frac{R_0}{R}\right). \end{aligned} \tag{5.5}$$

We now investigate various special cases when one or another of these invariants vanishes, and we find canonical forms for the metric in these cases.

First, suppose the congruence is expansion-free. Then we have $\beta_0 + (2R_0/R) = 0$, together with the coordinate condition (5.2). For the moment, we assume $R_0 \neq 0$. Then we can solve these equations to get

$$R^2 = f(r)e^{-\beta}, \quad R_0R^2 = g(x^0)e^{\alpha}, \tag{5.6}$$

but we still have the coordinate freedom remaining to let $r' = r'(r)$ and $x^{0\prime} = x^{0\prime}(x^0)$ without changing the form of our metric or violating our coordinate condition. We may also use this freedom to make $f(r)$ and $g(x^0)$ equal to one so that the canonical form of a spherically symmetric metric with an expansion-free timelike eigencongruence of the Einstein tensor is

$$ds^2 = (R_0R^2\,dx^0)^2 - 1/R^4\,dr^2 - R^2\,d\omega^2, \tag{5.7}$$

where we have dropped the primes on x^0 and r, and R is an arbitrary nonvanishing function of (x^0, r).

If we demand that the congruence be rigid in the sense of Born,[9] it is easily seen that this requires that $\beta_0 = 0$ and $R_0 = 0$ both hold, so that this is just the case we neglected above. We can then use the coordinate freedom of r mentioned above to set $R = r$, so that the canonical form of the metric in this case is

$$ds^2 = e^{2\alpha(x^0,r)}(dx^0)^2 - e^{2\beta(r)}\,dr^2 - r^2\,d\omega^2. \tag{5.8}$$

Note that the coordinate condition is automatically fulfilled.

If we additionally demand that the congruence be Killing, i.e., that the metric be static, we get the additional restriction that $\alpha = f(x^0) - \lambda(r)$, and using the coordinate freedom on x^0, we get the well-known result that a static spherically symmetric metric can be represented in the form,

$$ds^2 = e^{2\lambda(r)}(dx^0)^2 - e^{2\beta(r)}\,dr^2 - r^2\,d\omega^2 \tag{5.9}$$

with δ^μ_0 obviously the Killing vector.

If we demand that the timelike congruence be geodesic, this means that $\alpha_1 = 0$ in addition to the coordinate condition (5.2). In this case, the metric may, by similar methods, be reduced to any of several standard forms, for example, the geodesic normal form,

$$ds^2 = (dx^0)^2 - R_1^2\,dr^2 - R^2\,d\omega^2, \tag{5.10}$$

where R is any nonvanishing function of (x^0, r).

If we demand that the congruence be shear-free, but nongeodesic, then we have $\beta_0 = R_0/R$, and $\alpha_1 \neq 0$ in addition to (5.2). Assuming R_0 does not vanish (which would also make β_0 vanish and thus get us back to the rigid congruence case), we can then put the metric in the canonical form,

$$ds^2 = (R_0/R)^2(dx^0)^2 - R^2\,dr^2 - R^2\,d\omega^2, \tag{5.11}$$

where R is again any nonvanishing function of (x^0, r) with nonvanishing x^0 derivative.

If we demand that the congruence be shear-free and geodesic, the requirements are that $\beta_0 = R_0/R$ and $\alpha_1 = 0$ in addition to (5.2). The freedom of x^0 transformations can here be used to make the coefficient of $(dx^0)^2$ equal to unity, so that the metric takes the form,

$$ds^2 = (dx^0)^2 - R^2(dr^2 + d\omega^2), \tag{5.12}$$

where now $R = f(x^0)g(r)$ and all freedom of transformations of x^0 and r has been used.

If we additionally demand that $G_2' = G_3'$ so that we are in the type $[T\text{-}3S]_{[1-1]}$ where a hydrodynamical model is possible, the form of $g(r)$ is fully determined:

$$R = f(x^0)/(Ae^r + Be^{-r}), \tag{5.13}$$

[16] A discussion of first-order differential invariants of a timelike congruence may be found in the article by J. Ehlers and W. Kundt, in *Gravitation: An Introduction to Current Research*, L. Witten, Ed. (John Wiley & Sons, Inc., New York, 1962).

where $f(x^0)$ is an arbitrary function of x^0 and A and B are arbitrary constants. This class of metrics is seen, by direct substitution into the expression for the conformal invariant Eq. (B.11), to be conformally flat. Now suppose neither A nor B vanishes. Then, the substitution $\rho = 2e^r(|A/B|)^{\frac{1}{2}}$ transforms the metric into the form,

$$ds^2 = (dx^0)^2 - \frac{[f(x^0)/2(|AB|)^{\frac{1}{2}}]^2}{1 \pm (\rho^2/4)}(d\rho^2 + \rho^2\, d\omega^2), \tag{5.14}$$

one of the standard forms of the Robertson–Walker metric.[17] If either A or B vanishes (it makes no difference which we take as nonvanishing, since the transformation $r \to -r$ takes one case into the other), say $A = 0$, then the substitution $\rho = e^r$ takes the metric into the form,

$$ds^2 = (dx^0)^2 - [f(x^0)/B]^2(d\rho^2 + \rho^2\, d\omega^2), \tag{5.15}$$

which is the "flat" Robertson–Walker metric.

It is well known that the Friedmann metrics can be characterized as Robertson–Walker type solutions to the field equations for incoherent matter with a velocity vector which is shear free and geodesic. We see that the assumption about the nature of the matter tensor is irrelevant, as long as it has a shear-free and geodesic velocity vector field and a triple spacelike eigenvalue.

The results of the various assumptions about the nature of the timelike congruence of the Einstein tensor, derived above, are summarized in Table IV. The form of the Einstein tensor in each case is given in Appendix B.

We also note that the assumption that the expansion θ is negative at a point in space–time is easily seen to be equivalent to the statement that the volume of a shell of matter at that point is decreasing at that moment. Thus, the assumption that the expansion is everywhere negative can be used for a generalized definition of collapsing matter, and the resulting inequalities may enable a discussion of some aspects of spherical gravitational collapse to be carried out with very general assumptions about the nature of the matter tensor.

6. TWO COMPLEX AND TWO SPACELIKE EIGENVALUES

The only remaining case to be discussed is that of an Einstein tensor of the form $[Z\text{-}\bar{Z}\text{-}2S]_{[1-1-1]}$, i.e., where there do not exist any real eigenvectors in the timelike invariant plane of the Einstein tensor. This case may also be treated by the use of the form (5.1) for the metric of a spherically symmetric Riemann space. As noted in the last section, $e^{2\beta}G_0^1 = -e^{2\alpha}G_1^0$. By picking unit vectors ${}_0e^\mu$ and ${}_1e^\mu$ along the (orthogonal) coordinate directions t and r, respectively, it is easy to show that this implies that $G_{(0)}^{(1)} = -G_{(1)}^{(0)}$, where the parenthesis around an index means projection onto the tetrad vector of that index so that these are physical components of the Einstein tensor. By looking at the canonical form of the Einstein tensor in Case I(b) or (c) (Table 1), we see that if either $G_{(0)}^{(0)}$ or $G_{(1)}^{(1)}$ vanishes, the Einstein tensor will already be in canonical form with respect to a tetrad composed of unit vectors along the four coordinate axes. If we choose $R = R(r)$ only, then $G_{(1)}^{(1)}$ will vanish if

$$e^{2\beta} = R_1^2 + 2R_1R\alpha_1; \tag{6.1}$$

indeed, we could choose $R = r$ (standard Schwarzschild r, or curvature coordinates in Synge's terminology[9]), if it were not for the requirement that $e^{2\beta} > 0$, which might otherwise be violated for arbitrary α. However, this class of solutions in canonical form depends "essentially" on only one arbitrary function in the sense that α may be picked arbitrarily, and then any choice of $R(r)$ made to satisfy the inequality

$$R_1^2 + 2R_1R\alpha_1 > 0. \tag{6.2}$$

β is then fixed by Eq. (6.1). The physical components of the Einstein tensor will then be in canonical form I(c). No matter tensors corresponding to this case have been studied.

7. CONCLUSIONS

We have seen that in the spherically symmetric case, the subclassification of metrics by type of Einstein tensor leads to the possibility of canonical forms for the various types, which, particularly in the case of two doubly degenerate eigenvalues, are closely related to possible physical interpretations of the solutions. This suggests, more generally, that when looking for solutions of a certain physical type, the structure of the matter tensor may limit the type of the Einstein tensor in such a way that a more restrictive canonical form for the metric can be used than symmetry alone indicates without the loss of any solutions of the desired kind. Thus, the method may serve to simplify the search for exact solutions of certain types.

In addition, if an Einstein tensor is of a certain type, only certain kinds of matter tensors will in general be compatible with this type, so that the study of the type of the Einstein tensor can help in the physical interpretation of metrics. For example, in the case of an

[17] A discussion of the Robertson–Walker line element with references to the original papers and a detailed historical treatment is found in J. D. North, *The Measure of the Universe: A History of Modern Cosmology* (Clarendon Press, Oxford, England 1965).

electromagnetic field, it is clear from the structure of the stress–energy tensor $T_\mu{}^\nu = f_{\mu\kappa}f^{\kappa\nu} - \delta_\mu{}^\nu f_{\kappa\lambda}f^{\lambda\kappa}$, that eigenvectors and invariant planes of the field tensor $f^{\mu\nu}$ are also eigenvectors and invariant planes of $T_\mu{}^\nu$. The study of the eigenvectors and invariant planes of the Maxwell tensor (as found in Ref. 5, for example) then shows that the only possible types for electromagnetic stress–energy tensors are: $[4N]_{[2]}$ for the null Maxwell tensor, and $[2T\text{–}2S]_{[1-1]}$ for all others (and possible degeneracies of these, of course).

Certain canonical forms may prove useful in general investigations where assumptions are made about the matter tensor sufficient to limit the type of the Einstein tensor, but not detailed enough to fix it completely. For example, it was suggested in Sec. V that it might be possible to study certain problems of gravitational collapse with no more definite assumptions about the matter tensor than that it had a timelike eigenvector (velocity field) whose divergence is everywhere negative.

We have found a number of metrics with two doubly degenerate eigenvalues in Sec. IV which do not seem to have been previously investigated. In particular, the "radiating" de Sitter metric, the "radiating" generalized Vaidya solution, and the "radiating" Reissner–Nordstrom solutions would seem worthy of study for possible cosmological or other applications.

Finally, we add that the demonstrated usefulness for general relativity of the study of the classification of the Weyl tensor and of the traceless Ricci tensor suggests that the "mixed" parts of the Riemann tensor be studied and classified. Two traceless symmetric tensors can be formed from the Weyl tensor and the traceless Ricci tensor

$$D_{\alpha\beta} = C_{\alpha\mu\nu\beta}S^{\mu\nu} \quad\text{and}\quad D^*_{\alpha\beta} = C^*_{\alpha\mu\nu\beta}S^{\mu\nu}, \tag{7.1}$$

where $C_{\alpha\beta\mu\nu}$ is the Weyl tensor, $S_{\mu\nu}$ is the traceless Ricci tensor $R_{\mu\nu} - \frac{1}{4}g_{\mu\nu}R$, and $C^*_{\alpha\mu\nu\beta}$ is the dual of the Weyl tensor. Six of the fourteen second-order differential invariants of the Riemann tensor are associated with these two tensors, and may therefore be called "mixed" invariants of the Weyl and traceless Ricci tensors.[18] As far as we know, the study of these "mixed" invariants has not been carried out. The classification of $D_{\alpha\beta}$ and $D^*_{\alpha\beta}$ as traceless symmetric tensors is, of course, similar to the classification of the Ricci tensor. But the relation between its properties and those of the Weyl and Ricci tensors for various classes of metrics has not been investigated. In the spherically symmetric case, there is too much symmetry for anything new to arise, but in more general cases it should prove interesting to study this question. The Weyl tensor describes the gravitational field proper, algebraically independent of its sources; the Einstein tensor tells us about that part of the gravitational field that is tied directly to the sources of the field. It is natural to speculate that $D_{\alpha\beta}$ and $D^*_{\alpha\beta}$ should tell us something about the interaction of sources and gravitational field proper.

[18] See R. Debever, Cahiers Phys. **18**, 303 (1964) for an extensive review of the algebraic decomposition of the Riemann tensor and discussion of its invariants, including the definitions of $D_{\alpha\beta}$ and $D^*_{\alpha\beta}$.

ACKNOWLEDGMENTS

The authors would like to pay tribute to the late Dr. Richard Boyer, who participated in several helpful discussions on this work. One of us (J. S.) would like to thank Dr. Arturo Rosenblueth, Director of the Centro de Investigacion, for his hospitality during a visit last summer, when this work was started.

APPENDIX A

For the analytic study of the possible types of the Einstein tensor in the case of spherical symmetry, it proves useful to employ a conformal form of the metric, already discussed in Ref. 8. It was shown there that every space–time V_4 which is spherically symmetric can be described, locally at least, in isotropic coordinates by the following metric form with two structural functions:

$$ds^2 = \varphi^{-2}(x^0, w)[\psi^2(x^0, w)(dx^0)^2 - dx^a\,dx^a], \tag{A1}$$

where $w = x^a x^a = r^2$, with the three-dimensional summation convention. The conformal curvature invariant in these coordinates is given by

$$C^2 = \tfrac{3}{64}(C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta}) = (w\varphi^2\psi^{-1}\psi_{,ww})^2. \tag{A2}$$

Of course C^2 is an invariant and as mentioned above, it characterizes the Weyl tensor in this case. All further information about the curvature of the spherically symmetric spaces is contained in the Einstein tensor G^α_β. When computed in the above isotropic coordinates, this has the following form[8]:

$$G^\alpha_\beta = A^\alpha_\beta - \delta^\alpha_\beta\chi; \tag{A3}$$

A^α_β has the following structure:

$$\begin{aligned} &A^0_0 = a, \quad A^0_a = \psi^{-1}bx^a/(w)^{\frac{1}{2}}, \\ &A^a_0 = -\psi b x^a/(w)^{\frac{1}{2}}, \quad A^a_b = -c\,\frac{x^a x^b}{w}\,; \end{aligned} \tag{A4}$$

the coefficients χ, a, b, and c are given by

$$\begin{aligned} \chi &= 4w\varphi^2\psi^{-1}\psi_{,ww} + 4\varphi^2\psi^{-1}\psi_{,w} - 8w\varphi\varphi_{,ww} \\ &\quad - 8\varphi\varphi_{,w} + 12w\varphi^2_{,w} - 3(\psi^{-1}\varphi_{,0})^2 \\ &\qquad + 2\varphi\psi^{-1}[(\psi^{-1}\varphi_{,0})_{,0} - 4w\psi_{,w}\varphi_{,w}], \\ a &= 4\varphi^2 w\psi^{-1}\psi_{,ww} + 4\varphi^2\psi^{-1}\psi_{,w} + 4\varphi\varphi_{,w} + 2\varphi\psi^{-1} \\ &\qquad \times [(\psi^{-1}\varphi_0)_{,0} - 4w\psi_{,w}\varphi_{,w}], \\ b &= 4(w)^{\frac{1}{2}}\varphi(\psi^{-1}\varphi_{,0})_{,w}, \\ c &= -4w\varphi^2\psi^{-1}\psi_{,ww} + 8w\varphi\varphi_{,ww}. \end{aligned} \tag{A5}$$

As we shall see shortly, χ is a scalar; therefore, A^α_β is a tensor and the study of its algebraic structure is equivalent to the study of that of the Einstein tensor.

First we determine the minimal equation of A^α_β, considered as a matrix, symbolized by $\mathbf{A}$. If we define $\Delta = b^2 - ac$, it is easily proved that

$$\mathbf{A}^3 - (a-c)A^2 + \Delta\mathbf{A} = 0. \tag{A6}$$

Thus, the minimal equation of the matrix $\mathbf{A}$ is of order less than or equal to three, so that at least one double root must occur, in contrast to the general situation (see Table II) where the minimal equation may be of fourth order with no degeneracies. This is, of course, a result of the spherical symmetry.

Further degeneration of the minimal equation occurs in two cases: when $\Delta = 0$, then $\mathbf{A}$ obeys a minimal equation of second order,

$$\mathbf{A}^2 - (a-c)\mathbf{A} = 0; \tag{A7a}$$

when $b = 0$, $a = -c$, then $\mathbf{A}$ obeys

$$\mathbf{A}^2 - a\mathbf{A} = 0, \tag{A7b}$$

while $\Delta = a^2$ need not vanish.

A little further algebra yields the eigenvalues of $\mathbf{A}$, and consequently of $\mathbf{G}$. These are

$$\begin{aligned} G_1' &= -\chi \text{ (the double root)} \\ G_2' &= -\chi + \tfrac{1}{2}(a-c) + [\tfrac{1}{4}(a-c)^2 - \Delta]^{\frac{1}{2}}, \\ G_3' &= -\chi + \tfrac{1}{2}(a-c) - [\tfrac{1}{4}(a-c)^2 - \Delta]^{\frac{1}{2}}. \end{aligned} \tag{A8}$$

These facts enable us to give the algebraic classification of the spherically symmetric Einstein tensors, which can have quite diversified structures, in contrast to the conformal curvature which has only two types, as we discussed earlier.

When

$$-\Delta + \tfrac{1}{4}(a-c)^2 = \tfrac{1}{4}(a+c)^2 - b^2 < 0, \tag{A9}$$

the eigenvalues G_2' and G_3' are complex conjugates. The eigenspace belonging to the double eigenvalue $G_1' = -\chi$ must then be spanned on two spacelike eigenvectors. Hence, when (A9) holds, $\mathbf{G}$ is of the type $[Z\text{-}\bar{Z}\text{-}2S_{[1-1-1]}]$. [Note that (A9) necessitates that $\Delta > 0$, $b \neq 0$.]

If

$$-\Delta + \tfrac{1}{4}(a-c)^2 = \tfrac{1}{4}(a+c)^2 - b^2 > 0, \tag{A10}$$

then the two eigenvalues G_2' and G_3' are real and distinct. If $\Delta \neq 0$, neither can coincide with G_1'. From this we infer that if (A10) holds and $\Delta \neq 0$, then $\mathbf{G}$ is of type $[T\text{-}2S_1\text{-}S_2]_{[1-1-1]}$. If

$$-\Delta + \tfrac{1}{4}(a-c)^2 = 0, \quad \text{and} \quad \Delta \neq 0, \tag{A11}$$

then $G_2' = G_3' \neq G_1'$. Two subcases are possible here. Suppose first that $b \neq 0$. Together with $\Delta \neq 0$, this rules out the possibility of a minimal equation of second order, (A7a). From this, one can infer that when (A.11) holds and $b \neq 0$, then $\mathbf{G}$ is of the type $[2N\text{-}2S]_{[2-1]}$, with $N = -\chi + \tfrac{1}{2}(a-c)$ and $S = -\chi$, of course.

The second possibility arises when $b = 0$. Together with (A11), this implies that $a = -c$, and minimal Eq. (A7b) holds, and $\mathbf{G}$ is diagonalizable. Thus, if (3.11) holds and $b = 0$ (or equivalently $a = -c$), then $\mathbf{G}$ is of type $[2T\text{-}2S]_{[1-1]}$, with $T = -\chi + \tfrac{1}{2}(a-c)$ and $S = -\chi$.

The next possibility of degeneration of the eigenvalues is the case when G_1' coincides with either of G_2' or G_3', which remain distinct. Since it is easily seen that $\Delta = (G_2' - G_1')(G_3' - G_1')$, this can only happen if $\Delta = 0$. In that case, minimal Eq. (A.7a) applies; the two different eigenvalues are $-\chi$ and $-\chi + (a-c)$. From this we infer that when $\Delta = 0$ and $(a-c) \neq 0$, then $\mathbf{G}$ is either of type $[3T\text{-}S]_{[1-1]}$, or type $[T\text{-}3S]_{[1-1]}$.

The final possible degeneracy arises if $\Delta = 0$, and $a = c$. Then $G_1' = -\chi$ becomes a quadruple eigenvalue of $\mathbf{G}$. The minimal Eq. (A7a) then becomes $(\mathbf{G} + \chi\mathbf{I})^2 = 0$. Two subcases again arise. If $b \neq 0$, the matrix $\mathbf{G} + \chi\mathbf{I}$ is nilpotent of order two. It thus follows that if $\Delta = 0$, $a = c$, $b \neq 0$, then $\mathbf{G}$ is of the type $[4N]_{[2]}$. If $b = 0$, $\mathbf{G} + \chi\mathbf{I}$ is nilpotent of order one, and consequently vanishes, so that the minimal equation is of first order. In this case, where $\Delta = 0$, $a = c$, $b = 0$, $\mathbf{G}$ is of the type $[4T]_{[1]}$.

The results of the above discussion are summarized in Table V, which gives a table of all the possible algebraic types of a spherically symmetric Einstein tensor.

APPENDIX B

In this Appendix we shall give some useful formulas for various canonical forms of spherically symmetric metrics mentioned in the text.

The Einstein tensor of metric (4.9) with two double eigenvalues has already been given in Eq. (4.11). If we introduce the following null tetrad:

$$l^\mu = \delta^\mu_r, \quad n^\mu = -\frac{1}{2}\left(1 - \frac{f}{r}\right)\delta^\mu_r + \delta^\mu_u,$$
$$m^\mu = \frac{1}{\sqrt{\frac{1}{2}}}\frac{1}{r}\left(\delta^\mu_\theta + \frac{1}{\sin\theta}\delta^\mu_\varphi\right), \tag{B1}$$

where l^μ and n^μ are two real null vectors, and m^μ is a complex null vector, then l^μ and n^μ are the two double principal null vectors of the metric, and the only surviving physical component of the Weyl tensor with respect to this tetrad is Ψ_2 in the Newman–Penrose notation[19]:

$$\Psi_2 = (-f/2r^3) + (f_r/3r^2) - (f_{rr}/12r). \quad \text{(B2)}$$

The metric is then conformally flat if (B2) vanishes, which has as its most general solution

$$1 - (f/r) = 1 + 2a(u)r + b(u)r^2. \quad \text{(B3)}$$

The conformal factor for this class of metrics is easily found. Let

$$r = \frac{r'}{A(u) + B(u)r'}, \quad \text{(B4)}$$

where r' is a new variable, and $A(u)$ and $B(u)$ remain to be chosen [we assume only that $A(u) \neq 0$]. We now require that $A(u)$ and $B(u)$ be solutions to the equations

$$A_u = A(B + a) \quad \text{and} \quad 2B_u = B^2 + 2aB + b, \quad \text{(B5)}$$

and then introduce a new variable $u' = \int A(u)\, du$. If we invert this relationship to get $u = u(u')$, we can then rewrite $fA(u)$ and $B(u)$ as functions of u':

$$\mathcal{A}(u') = A[u(u')], \quad \mathcal{B}(u') = B[u(u')]. \quad \text{(B6)}$$

Substitution of these new variables u' and r' into metric (3.9) with f given by (B3), then shows that it takes the form,

$$ds^2 = [\mathcal{A}(u') + \mathcal{B}(u')r']^{-2} \times [(du')^2 + 2\, du'\, dr' - (r')^2\, d\omega^2], \quad \text{(B7)}$$

which is already the conformally flat representation of the metric in null coordinates. The arbitrariness in choice of constants of integration in (B5) and (B6) corresponds to the freedom of choice in the conformal factor due to the possible transformations within the conformal group.[8]

If we make the Ansatz $A = A_0 w^{-2}$, where A_0 is any nonzero constant, this implies that $B = (-2w_u/w) - a$, and leads to the linear differential equation for w,

$$w_{uu} + \tfrac{1}{4}(2a_u - a^2 + b)w = 0. \quad \text{(B8)}$$

If a and b are constants, the integration is easily carried out, and the results may be expressed as follows. If we write $a = -e/l$, $b = (\epsilon/\Lambda^2) + (e/l^2)$ (where e, ϵ can take on the values 1, 0, or -1), which can always be done, then the conformal factor becomes

$$\left[1 + \frac{\epsilon}{(2\Lambda)^2} u'(u' + 2r') + \frac{e}{l} r'\right]^{-2}. \quad \text{(B9)}$$

For $e = 0$, this gives the representation of spaces of constant curvature (de Sitter spaces) in stereographic coordinates.

In the case of one timelike and three spacelike eigenvectors, we have used the metric (5.1); the Einstein and Riemann tensors for this metric may be found in Ref. 9. It is convenient to introduce the orthonormal tetrad adapted to the coordinates in this metric:

$$\begin{aligned} {}_0e^\mu &= e^{-\alpha}\delta^\mu_0, \quad {}_1e^\mu = e^{-\beta}\delta^\mu_1, \\ {}_2e^\mu &= (1/R)\delta^\mu_2, \quad {}_3e^\mu = \frac{1}{R\sin\theta}\,\delta^\mu_3. \end{aligned} \quad \text{(B10)}$$

If we then define a null tetrad by $l^\mu = {}_0e^\mu + {}_1e^\mu$, $n^\mu = \tfrac{1}{2}({}_0e^\mu - {}_1e^\mu)$, $m^\mu = 1/\sqrt{2}\,({}_2e^\mu + i\,{}_3e^\mu)$, then again only Ψ_2 will be nonvanishing. Thus l^μ and n^μ are the double principal vectors of the Weyl tensor. Ψ_2 is given by

$$\begin{aligned}\Psi_2 = \frac{1}{6}\Big\{ & e^{-2\alpha}\Big(-\frac{R_{00}}{R} + \frac{R_0\alpha_0}{R} + \beta_{00} + \beta_0^2 - \frac{\beta_0 R_0}{R} - \alpha_0\beta_0 + \frac{R_0^2}{R^2}\Big) \\ & + e^{-2\beta}\Big(\frac{R_{11}}{R} + \frac{R_1\alpha_1}{R} - \alpha_{11} - \alpha_1^2 - \frac{\beta_1 R_1}{R} + \alpha_1\beta_1 - \frac{R_1^2}{R^2}\Big) + \frac{1}{R^2}\Big\}. \end{aligned} \quad \text{(B11)}$$

The tetrad components of the Einstein tensor with respect to the orthonormal tetrad (B10) are given by

$$\begin{aligned} G_{(00)} &= e^{-2\beta}\Big(-\frac{2R_{11}}{R} - \frac{R_1^2}{R^2} + \frac{2R_1\beta_1}{R}\Big) + e^{-2\alpha}\Big(\frac{R_0^2}{R^2} + \frac{2R_0\beta_0}{R}\Big) + \frac{1}{R^2}, \\ G_{(01)} &= -e^{-(\beta+\alpha)}\Big(\frac{2R_{10}}{R} - \frac{2R_1\beta_0}{R} - \frac{2R_0\alpha_1}{R}\Big), \\ G_{(11)} &= -e^{-2\beta}\Big(-\frac{R_1^2}{R^2} - \frac{2R_1\alpha_1}{R}\Big) - e^{-2\alpha}\Big(\frac{2R_{00}}{R} + \frac{R_0^2}{R^2} - \frac{2R_0\alpha_0}{R}\Big) - \frac{1}{R^2}, \\ G_{(22)} &= G_{(33)} = -e^{-2\beta}\Big(-\frac{R_{11}}{R} - \alpha_{11} - \alpha_1^2 - \frac{\alpha_1 R_1}{R} + \frac{\beta_1 R_1}{R} + \beta_1\alpha_1\Big) \\ &\quad - e^{-2\alpha}\Big(\frac{R_{00}}{R} + \beta_{00} + \beta_0^2 + \frac{\beta_0 R_0}{R} - \frac{R_0\alpha_0}{R} - \alpha_0\beta_0\Big). \end{aligned} \quad \text{(B12)}$$

[19] See E. T. Newman and R. Penrose, J. Math. Phys. 3, 566 (1962); and A. Janis and E. T. Newman, J. Math. Phys. 6, 902 (1965), for a discussion of the projection of the Riemann tensor onto a null tetrad, the definition of the five complex Ψ's that result, and the relation between the Ψ's and the Petrov type. When the metric is type D, the use of the double principal null vectors always eliminates all components except $\Psi_2 = C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho m^\sigma$.

The Einstein tensor for metric (5.7) with an expansion-free timelike eigencongruence is given by

$$
\begin{aligned}
G_1^1 &= -5R_1^2R^2 - \frac{2R_1R_{01}}{R_0}R^3 + \frac{1}{R^2} - \frac{3}{R^6},\\
G_0^0 &= -5R_1^2R^2 - 2R_{11}R^3 + \frac{1}{R^2} - \frac{3}{R^6},\\
G_2^2 = G_3^3 &= -4R_1^2R^2 - \left(3R_{11} + \frac{3R_{01}R_1}{R_0}\right)R^3\\
&\quad + \left(\frac{R_{01}^2}{R_0^2} - \frac{R_{011}}{R_0}\right)R^4 + \frac{4}{R^6}.
\end{aligned}
\tag{B13}
$$

The Einstein tensor for metric (5.8) with a rigid timelike eigencongruence is given by

$$
\begin{aligned}
G_1^1 &= e^{-2\beta}\left(\frac{-1}{r^2} - \frac{2\alpha_1}{r}\right) + \frac{1}{r^2},\\
G_0^0 &= e^{-2\beta}\left(\frac{-1}{r^2} + \frac{2\beta_1}{r}\right) + \frac{1}{r^2},\\
G_2^2 = G_3^3 &= e^{-2\beta}\left(-\alpha_{11} - \alpha_1^2 - \frac{\alpha_1}{r} + \frac{\beta_1}{r} + \beta_1\alpha_1\right).
\end{aligned}
\tag{B14}
$$

We omit the Einstein tensor for the static metric (5.9), which may be found in Ref. 9, and any number of other texts.

The Einstein tensor for metric (5.10) with a geodesic timelike eigencongruence is given by

$$
\begin{aligned}
G_1^1 &= \frac{2R_{00}}{R} + \frac{3R_0^2}{R^2},\\
G_0^0 &= \frac{R_0^2}{R^2} + \frac{2R_{10}R_0}{RR_1},\\
G_2^2 = G_3^3 &= \frac{R_{00}}{R} + \frac{R_{100}}{R_1} + \frac{R_0R_{10}}{R_1R}.
\end{aligned}
\tag{B15}
$$

The Einstein tensor for metric (5.11) with a shear-free nongeodesic eigencongruence is given by

$$
\begin{aligned}
G_1^1 &= \frac{1}{R^2}\left(-\frac{2R_1R_{01}}{RR_0} + \frac{R_1^2}{R^2} + 1\right) + 3,\\
G_0^0 &= \frac{1}{R^2}\left(-\frac{2R_{11}}{R} + \frac{R_1^2}{R^2} + 1\right) + 3,\\
G_2^2 = G_3^3 &= \frac{1}{R^2}\left(-\frac{R_{011}}{R_0} + \frac{2R_{01}R_1}{R_0R} - \frac{R_1^2}{R^2}\right) + 3.
\end{aligned}
\tag{B16}
$$

The Einstein tensor for metric (5.12) with a shear-free geodesic timelike eigencongruence is given by

$$
\begin{aligned}
G_1^1 &= \frac{1}{R^2}\left(1 - \frac{R_1^2}{R^2} + R_0^2 + 2R_{00}R\right),\\
G_0^0 &= \frac{1}{R^2}\left(1 - \frac{2R_{11}}{R} + 3R_0^2 + R_1^2\right),\\
G_2^2 = G_3^3 &= \frac{1}{R^2}\left(-\frac{R_{11}}{R} + \frac{R_1^2}{R^2} + R_0^2 + 2R_{00}R\right).
\end{aligned}
\tag{B17}
$$

We omit the Einstein tensor for the Robertson–Walker metrics, which may also be found in innumerable textbooks.

Intrinsic Vector and Tensor Techniques in Minkowski Space with Applications to Special Relativity

C. P. LUEHR AND M. ROSENBAUM
TEMPO, *Center for Advanced Studies, General Electric Company, Santa Barbara, Calif.*

(Received 27 January 1967)

This paper describes an abstract formalism for tensor analysis in Minkowski space which entails considerable notational simplicity and calculational advantages, as evidenced when compared with the usual component techniques. The need to express tensor equations in component form is eliminated, and manipulations become formally the same as those in Euclidean space. The method is based on an extension to Minkowski space of the intrinsic concepts of vectors, differential operators, and polyadics in three-dimensional Euclidean space. Several examples from special relativity have been selected to illustrate the advantages of the formalism. In the first example, expressions in dyadic form for the Euler–Lagrange equations and canonical energy–momentum tensor are obtained and specialized to the electromagnetic field. From the electromagnetic-field dyadic, invariants and other useful relations are derived easily and economically. The dyadic form of the field equations is also shown to be particularly amenable for a derivation of the Dirac-like form of Maxwell's equations with the base elements of the Pauli algebra emerging in a most natural way. Further illustration of the practical utility of the method is given by considering several properties of the restricted homogeneous Lorentz transformations. Various dyadic expressions for these are obtained, and a detailed derivation of their eigenvalues and eigenvectors is given. By combining some of the results from the discussions of Lorentz transformations and the Dirac-like form of Maxwell's equations, it is shown how an isomorphism between the three-dimensional complex orthogonal group and the Lorentz group can be established in a simpler and different manner from other approaches appearing in the literature.

1. INTRODUCTION

The use of intrinsic vector algebra (in which vectors and their operations are defined axiomatically) in Euclidean spaces is recognized in modern mathematical physics to result in considerable notational simplification and calculational advantages, in comparison to the old-fashioned component techniques. Likewise, the intrinsic representation of tensors by means of polyadics has similar advantages over the representation in terms of components. Many authors[1] discuss dyadics in three-dimensional Euclidean space, and some[2] have applied them to various physical problems. On the other hand, in Minkowski space the component representation of tensors is used almost exclusively.[3] Within the framework of the component method there are, however, some useful simplifications which eliminate the need to distinguish between covariant and contravariant components. One of these is the introduction of the imaginary fourth component. Several objections to this approach are listed by Foldy,[4] who suggests a system of notation based on a modification of the Einstein summation convention combined with a modified differentiation with respect to tensor components.

The purpose of this paper is twofold: (i) to present a more abstract formalism based on an extension of the intrinsic concepts of vectors and tensors to Minkowski space, whereby manipulations become formally the same as those in Euclidean space, and the need to express tensor equations in the cumbersome component form is eliminated; and (ii) to illustrate the advantages implicit in these techniques by applying them to some specific problems in special relativity.

The discussion is divided into two main parts. The first part deals primarily with vector and dyadic algebra in Minkowski space. In addition a brief introduction to polyadics is given as well as some useful rules of differentiation.

The second part is devoted to the various applications. First, the Euler–Lagrange field equations are derived in dyadic form from a variational principle and then applied to obtain the canonical energy–momentum dyadic. These results are then specialized to the electromagnetic field leading to the dyadic form of the Maxwell field equations and the symmetric energy–momentum dyadic. By an alternate derivation

[1] See, e.g., J. W. Gibbs and E. B. Wilson, *Vector Analysis* (Dover Publications, Inc., New York, 1960); A. P. Wills, *Vector Analysis with an Introduction to Tensor Analysis* (Dover Publications, Inc., New York, 1958); C. E. Weatherburn, *Advanced Vector Analysis with Applications to Mathematical Physics* (G. Bell and Sons, London, 1944); T. B. Drew, *Handbook of Vector and Polyadic Analysis* (Reinhold Publishing Corporation, New York, 1961); L. Brand, *Vector and Tensor Analysis* (John Wiley & Sons, Inc., New York, 1947); P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953).

[2] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, Cambridge, 1960).

[3] However, vectors in flat Riemann spaces are sometimes represented intrinsically. See, e.g., A. Lichnerowicz, *Elements of Tensor Calculus* (Methuen and Company Ltd., London, 1962); D. E. Rutherford, *Vector Methods* (Oliver and Boyd, Edinburgh, 1954).

[4] L. L. Foldy, J. Math. Phys. **6**, 1871 (1965).

of these dyadic field equations, a form for the electromagnetic-field dyadic is obtained, which is then used to illustrate how easily the invariants of the field and other useful relations can be found. The dyadic form of the field equations is shown to be particularly amenable for a derivation of the Dirac-like form of Maxwell's equations[5,6] with the base elements of the Pauli algebra emerging in a most natural way. Next the techniques are applied to prove several properties of homogeneous restricted Lorentz transformations. Explicit dyadic forms for these are derived and some familiar results are presented in an interesting manner. In addition, a detailed derivation is given of their eigenvectors and eigenvalues, which are then used for expressing the transformations explicitly in terms of the eigenvectors. Finally, the results obtained in the discussion of the Dirac-like form of Maxwell's equations are used to show how the isomorphism between the Lorentz group and the three dimensional complex orthogonal group can be established by a simpler and different approach than others appearing in the literature.[7,8]

2. INTRINSIC TENSOR FORMALISM IN MINKOWSKI SPACE

A. Basic Definitions and Notation

Let $\underline{\mathbf{e}}_\mu$ denote[9] an orthonormal set of basis vectors of the Minkowski space $\mathcal{M}_4$ which satisfies

$$\underline{\mathbf{e}}_\mu \cdot \underline{\mathbf{e}}_\nu = g_{\mu\nu}, \tag{1}$$

where $g_{\mu\nu}$ is the metric

$$\begin{aligned} g_{\mu\nu} &= 0 \ (\mu \neq \nu), \\ g_{11} = g_{22} &= g_{33} = -g_{00} = 1. \end{aligned} \tag{2}$$

The reciprocal basis $\underline{\mathbf{e}}^\mu$, which satisfies

$$\underline{\mathbf{e}}^\mu \cdot \underline{\mathbf{e}}_\nu = \delta^\mu_\nu, \tag{3}$$

where δ^μ_ν is the Kronecker delta ($\delta^\mu_\nu = 1$ for $\mu = \nu$ and $\delta^\mu_\nu = 0$ for $\mu \neq \nu$), is given by

$$\begin{aligned} \underline{\mathbf{e}}^k &= \underline{\mathbf{e}}_k, \\ \underline{\mathbf{e}}^0 &= -\underline{\mathbf{e}}_0. \end{aligned} \tag{4}$$

In addition, by defining

$$g^{\mu\nu} = \underline{\mathbf{e}}^\mu \cdot \underline{\mathbf{e}}^\nu, \tag{5}$$

it follows that

$$g^{\mu\nu} = g_{\mu\nu}. \tag{6}$$

In terms of this basis any four-vector $\underline{\mathbf{a}}$ (i.e., a vector in $\mathcal{M}_4$) can be written as

$$\underline{\mathbf{a}} = a^\mu \underline{\mathbf{e}}_\mu = a_\mu \underline{\mathbf{e}}^\mu, \tag{7}$$

where a^μ and a_μ are the contravariant and covariant components of $\underline{\mathbf{a}}$, respectively. It is easily seen that

$$\begin{aligned} a^\mu &= \underline{\mathbf{e}}^\mu \cdot \underline{\mathbf{a}}, \\ a_\mu &= \underline{\mathbf{e}}_\mu \cdot \underline{\mathbf{a}}. \end{aligned} \tag{8}$$

The rules for raising and lowering indices are derived as follows:

$$\begin{aligned} a^\mu &= \underline{\mathbf{e}}^\mu \cdot \underline{\mathbf{a}} = \underline{\mathbf{e}}^\mu \cdot (a_\nu \underline{\mathbf{e}}^\nu) = \underline{\mathbf{e}}^\mu \cdot \underline{\mathbf{e}}^\nu a_\nu = g^{\mu\nu} a_\nu, \\ a_\mu &= \underline{\mathbf{e}}_\mu \cdot \underline{\mathbf{a}} = \underline{\mathbf{e}}_\mu \cdot (a^\nu \underline{\mathbf{e}}_\nu) = \underline{\mathbf{e}}_\mu \cdot \underline{\mathbf{e}}_\nu a^\nu = g_{\mu\nu} a^\nu. \end{aligned} \tag{9}$$

In view of the above results the product $\underline{\mathbf{a}} \cdot \underline{\mathbf{b}}$ can be written in terms of components in the alternate forms:

$$\begin{aligned} \underline{\mathbf{a}} \cdot \underline{\mathbf{b}} &= (a^\mu \underline{\mathbf{e}}_\mu) \cdot (b^\nu \underline{\mathbf{e}}_\nu) = \underline{\mathbf{e}}_\mu \cdot \underline{\mathbf{e}}_\nu a^\mu b^\nu = g_{\mu\nu} a^\mu b^\nu \\ &= a_\nu b^\nu = a^\mu b_\mu. \end{aligned} \tag{10}$$

Since the vectors $\underline{\mathbf{e}}_k$ span a three-dimensional Euclidean subspace $\mathcal{E}_3$ of $\mathcal{M}_4$, any four-vector can be expressed as the sum of a vector in $\mathcal{E}_3$ and a vector along $\underline{\mathbf{e}}_0$. In particular, the four-vector $\underline{\mathbf{x}}$, which locates an event with position $\mathbf{r} \in \mathcal{E}_3$ and time t in a given reference frame, can be written as

$$\underline{\mathbf{x}} = x^\mu \underline{\mathbf{e}}_\mu = \mathbf{r} + x^0 \underline{\mathbf{e}}_0 = \mathbf{r} + x_0 \underline{\mathbf{e}}^0, \tag{11}$$

where

$$x^0 = -x_0 = ct \tag{12}$$

and

$$\mathbf{r} = x^k \underline{\mathbf{e}}_k. \tag{13}$$

We now derive an expression for the four-gradient operator $\Box$ defined by the infinitesimal relation

$$df \equiv f(\underline{\mathbf{x}} + d\underline{\mathbf{x}}) - f(\underline{\mathbf{x}}) = d\underline{\mathbf{x}} \cdot \Box f, \tag{14}$$

where $f(\underline{\mathbf{x}})$ is an arbitrary function of the four-vector $\underline{\mathbf{x}}$ and $d\underline{\mathbf{x}}$ is an arbitrary infinitesimal change in $\underline{\mathbf{x}}$. First observe that f can be regarded as a function of the components x^μ and, therefore, it must also be true that

$$df = dx^\mu \partial_\mu f, \tag{15}$$

where

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu}.$$

Substitution of $dx^\mu = d\underline{\mathbf{x}} \cdot \underline{\mathbf{e}}^\mu$ into this equation gives

$$df = d\underline{\mathbf{x}} \cdot \underline{\mathbf{e}}^\mu \partial_\mu f. \tag{16}$$

Since $d\underline{\mathbf{x}}$ and f are arbitrary, comparison of the Eqs. (14) and (16) shows that

$$\Box = \underline{\mathbf{e}}^\mu \partial_\mu, \tag{17a}$$

---

[5] H. E. Moses, Nuovo Cimento Suppl. **7**, 1 (1958).
[6] J. S. Lomont, Phys. Rev. **111**, 1710 (1958).
[7] B. Kurşunoğlu, J. Math. Phys. **2**, 22 (1961).
[8] A. J. Macfarlane, J. Math. Phys. **3**, 1116 (1962).
[9] In the following discussion we use underlined boldface type to denote four-vectors in Minkowski space in order to distinguish them from vectors in Euclidean space (for which the conventional boldface roman type is used). Latin indices will take on the values 1, 2, 3, and Greek indices will take on the values 0, 1, 2, 3. Unless otherwise stated, the summation convention will apply to repeated indices in any expression.

which can also be written as

$$\square = \nabla + \underline{e}^0\partial_0 = \nabla - \underline{e}_0\partial_0, \tag{17b}$$

where ∇ is the usual gradient operator in three dimensions

$$\nabla = \underline{e}^k\partial_k = \underline{e}_k\partial_k. \tag{17c}$$

B. Summary of Dyadic Algebra

Here we review some of the basic definitions and operations of dyadic algebra[10] in the space $\mathcal{M}_4$. First recall that a dyad is a tensor product of two four-vectors $\underline{a}$ and $\underline{b}$, and is denoted by placing them in juxtaposition:

$$\underline{ab}.$$

A sum of dyads constitutes a dyadic:

$$\mathsf{A} = \sum_{l=1}^{l'} \underline{a}_l\underline{b}_l. \tag{18}$$

Two "dot" products of a dyadic A with a vector $\underline{x}$ can be formed according as $\underline{x}$ appears as the prefactor or the postfactor in the operation. Thus we have, respectively,

$$\underline{y} = \underline{x}\cdot\mathsf{A} = \underline{x}\cdot\sum_{l=1}^{l'}\underline{a}_l\underline{b}_l = \sum_{l=1}^{l'}(\underline{x}\cdot\underline{a}_l)\underline{b}_l, \tag{19a}$$

and

$$\underline{z} = \mathsf{A}\cdot\underline{x} = \sum_{l=1}^{l'}(\underline{a}_l\underline{b}_l)\cdot\underline{x} = \sum_{l=1}^{l'}\underline{a}_l(\underline{b}_l\cdot\underline{x}). \tag{19b}$$

It should be noted that the four-vectors $\underline{y}$ and $\underline{z}$ are, in general, different. Two dyadics A and B are regarded as equal if

$$\mathsf{A}\cdot\underline{x} = \mathsf{B}\cdot\underline{x}$$

for all four-vectors $\underline{x}$. From this definition of equality follows the uniqueness of the operations of addition of dyadics and of multiplication of dyadics by scalars.

Furthermore, the following distributive and associative laws hold for the above tensor and "dot" products:

(i) $\mathsf{A}\cdot(\underline{x}+\underline{y}) = \mathsf{A}\cdot\underline{x} + \mathsf{A}\cdot\underline{y}$,
$(\underline{x}+\underline{y})\cdot\mathsf{A} = \underline{x}\cdot\mathsf{A} + \underline{y}\cdot\mathsf{A}$;

(ii) $\underline{a}(\underline{b}+\underline{c}) = \underline{ab}+\underline{ac}$, $(\underline{b}+\underline{c})\underline{a} = \underline{ba}+\underline{ca}$; (20)

(iii) $\mathsf{A}\cdot(\lambda\underline{x}) = \lambda(\mathsf{A}\cdot\underline{x})$, $(\lambda\underline{x})\cdot\mathsf{A} = \lambda(\underline{x}\cdot\mathsf{A})$;

(iv) $(\lambda\underline{a})\underline{b} = \underline{a}(\lambda\underline{b}) = \lambda(\underline{ab})$.

Some additional operations are:

(1) The scalar of A, denoted by A_s, is defined as

$$\mathsf{A}_s = \left(\sum_l \underline{a}_l\underline{b}_l\right)_s = \sum_l \underline{a}_l\cdot\underline{b}_l. \tag{21}$$

(2) The vector of a dyadic A in $\mathcal{E}_3$ (i.e., $\mathbf{a}_l, \mathbf{b}_l \in \mathcal{E}_3$):

$$\mathsf{A}_\times = \left(\sum_l \mathbf{a}_l\mathbf{b}_l\right)_\times = \sum_l \mathbf{a}_l\times\mathbf{b}_l. \tag{22}$$

(3) The transpose of A (denoted by either $\tilde{\mathsf{A}}$ or A_T):

$$\tilde{\mathsf{A}} \equiv \mathsf{A}_T = \left(\sum_l \underline{a}_l\underline{b}_l\right)_T = \sum_l \underline{b}_l\underline{a}_l. \tag{23}$$

(4) The "dot" product of A with another dyadic $\mathsf{B} = \sum_k \underline{u}_k\underline{v}_k$, which yields a new dyadic:

$$\mathsf{C} = \mathsf{A}\cdot\mathsf{B} = \left(\sum_l \underline{a}_l\underline{b}_l\right)\cdot\left(\sum_k \underline{u}_k\underline{v}_k\right) = \sum_l\sum_k(\underline{b}_l\cdot\underline{u}_k)\underline{a}_l\underline{v}_k. \tag{24}$$

(5) "Cross" products of a vector and a dyadic both in $\mathcal{E}_3$ (for $\mathbf{a}_l, \mathbf{b}_l, \mathbf{v}\in\mathcal{E}_3$):

$$\mathsf{A}\times\mathbf{v} = \left(\sum_l \mathbf{a}_l\mathbf{b}_l\right)\times\mathbf{v} = \sum_l \mathbf{a}_l(\mathbf{b}_l\times\mathbf{v}), \tag{25a}$$

$$\mathbf{v}\times\mathsf{A} = \mathbf{v}\times\left(\sum_l \mathbf{a}_l\mathbf{b}_l\right) = \sum_l(\mathbf{v}\times\mathbf{a}_l)\mathbf{b}_l. \tag{25b}$$

(6) "Double dot" product of two dyadics:

$$\mathsf{A}:\mathsf{B} = \left(\sum_l \underline{a}_l\underline{b}_l\right):\left(\sum_k \underline{u}_k\underline{v}_k\right) = \sum_l\sum_k(\underline{a}_l\cdot\underline{u}_k)(\underline{b}_l\cdot\underline{v}_k). \tag{26}$$

(7) Exterior product[11] of two four-vectors $\underline{a}$ and $\underline{b}$:

$$\underline{a}\wedge\underline{b} = \underline{ab} - \underline{ba}. \tag{27}$$

A dyadic A is said to be symmetric if

$$\tilde{\mathsf{A}} = \mathsf{A}, \tag{28a}$$

and to be antisymmetric if

$$\tilde{\mathsf{A}} = -\mathsf{A}. \tag{28b}$$

Making use of the previous definitions we obtain the useful relations

$$\mathsf{A}:\mathsf{B} = (\tilde{\mathsf{A}}\cdot\mathsf{B})_s = (\mathsf{A}\cdot\tilde{\mathsf{B}})_s, \tag{29}$$

$$\underline{a}\cdot\tilde{\mathsf{A}} = \mathsf{A}\cdot\underline{a}, \quad \tilde{\mathsf{A}}\cdot\underline{a} = \underline{a}\cdot\mathsf{A}, \tag{30}$$

$$\underline{b}\wedge\underline{a} = -\underline{a}\wedge\underline{b}. \tag{31}$$

Antisymmetry of $\underline{a}\wedge\underline{b}$:

$$(\underline{a}\wedge\underline{b})_T = \underline{b}\wedge\underline{a} = -\underline{a}\wedge\underline{b}. \tag{32}$$

The covariant and contravariant components of a dyadic A with respect to the basis $\underline{e}_\mu$ are obtained as follows:

$$\begin{aligned} A_{\mu\nu} &= (\underline{e}_\mu\underline{e}_\nu):\mathsf{A} = \underline{e}_\mu\cdot\mathsf{A}\cdot\underline{e}_\nu, \\ A^{\mu\nu} &= (\underline{e}^\mu\underline{e}^\nu):\mathsf{A} = \underline{e}^\mu\cdot\mathsf{A}\cdot\underline{e}^\nu. \end{aligned} \tag{33}$$

---

[10] Dyadics will be denoted by sans serif type, while boldface Greek capitals are used for the few higher-order polyadics which appear in the text.

[11] For example, see Lichnerowicz, Ref. 3; also H. Flanders, *Differential Forms* (Academic Press Inc., New York, 1963).

Likewise the mixed components are given by

$$A^{\mu}{}_{\nu} = (\underline{e}^{\mu}\underline{e}_{\nu}):\mathsf{A} = \underline{e}^{\mu}\cdot\mathsf{A}\cdot\underline{e}_{\nu},$$
$$A_{\mu}{}^{\nu} = (\underline{e}_{\mu}\underline{e}^{\nu}):\mathsf{A} = \underline{e}_{\mu}\cdot\mathsf{A}\cdot\underline{e}^{\nu}. \tag{34}$$

In terms of these components the dyadic A can be expressed in the alternate ways:

$$\mathsf{A} = A^{\mu\nu}\underline{e}_{\mu}\underline{e}_{\nu} = A_{\mu\nu}\underline{e}^{\mu}\underline{e}^{\nu} = A^{\mu}{}_{\nu}\underline{e}_{\mu}\underline{e}^{\nu} = A_{\mu}{}^{\nu}\underline{e}^{\mu}\underline{e}_{\nu}. \tag{35}$$

Examples of some of the operations written in terms of the components are

$$\begin{aligned}
(\mathsf{A}\cdot\underline{x})^{\mu} &= \underline{e}^{\mu}\cdot(\mathsf{A}\cdot\underline{x}) = \underline{e}^{\mu}\cdot\mathsf{A}\cdot(\underline{e}^{\nu}x_{\nu}) = A^{\mu\nu}x_{\nu},\\
(\mathsf{A}\cdot\underline{x})^{\mu} &= A^{\mu\nu}g_{\nu\lambda}x^{\lambda},\\
(\mathsf{A}\cdot\mathsf{B})^{\mu\nu} &= A^{\mu\lambda}B_{\lambda}{}^{\nu} = A^{\mu\lambda}g_{\lambda\gamma}B^{\gamma\nu},\\
\mathsf{A}:\mathsf{B} &= A^{\mu\nu}B_{\mu\nu} = A_{\mu\nu}B^{\mu\nu},\\
\mathsf{A}_s &= A_{\mu}{}^{\mu} = g_{\mu\nu}A^{\mu\nu},\\
(\tilde{A})^{\mu\nu} &= A^{\nu\mu}.
\end{aligned} \tag{36}$$

By the last equation it is evident that a dyadic A is symmetric if

$$A_{\nu\mu} = A_{\mu\nu}, \tag{37a}$$

and antisymmetric if

$$A_{\nu\mu} = -A_{\mu\nu}. \tag{37b}$$

The unit dyadic (or identity dyadic) I_4 is a unique dyadic satisfying

$$\underline{x} = \mathsf{I}_4\cdot\underline{x} = \underline{x}\cdot\mathsf{I}_4 \tag{38}$$

for all four-vectors $\underline{x}$. It may be expressed as

$$\mathsf{I}_4 = \underline{e}^{\mu}\underline{e}_{\mu} = \underline{e}_{\mu}\underline{e}^{\mu} = \underline{e}_k\underline{e}_k - \underline{e}_0\underline{e}_0 = \mathsf{I}_3 - \underline{e}_0\underline{e}_0, \tag{39a}$$

where I_3 is the unit dyadic in $\mathcal{E}_3$:

$$\mathsf{I}_3 = \underline{e}^k\underline{e}_k = \underline{e}_k\underline{e}_k, \tag{39b}$$

which satisfies

$$\underline{e}_0\cdot\mathsf{I}_3 = \mathsf{I}_3\cdot\underline{e}_0 = 0 \tag{39c}$$

and

$$\mathbf{r} = \mathsf{I}_3\cdot\mathbf{r} = \mathbf{r}\cdot\mathsf{I}_3 \tag{39d}$$

for an arbitrary vector $\mathbf{r}\in\mathcal{E}_3$. The $\underline{e}_0$ component of a four-vector is annihilated by I_3:

$$\underline{x}\cdot\mathsf{I}_3 = \mathsf{I}_3\cdot\underline{x} = \mathsf{I}_3\cdot(\mathbf{r} + x^0\underline{e}_0) = \mathsf{I}_3\cdot\mathbf{r} = \mathbf{r}. \tag{40a}$$

Thus I_3 is a projection operator onto $\mathcal{E}_3$. Similarly I_3 annihilates the $\underline{e}_0$ component of the $\Box$ operator:

$$\Box\cdot\mathsf{I}_3 = \mathsf{I}_3\cdot\Box = \mathsf{I}_3\cdot(\boldsymbol{\nabla} + \underline{e}^0\partial_0) = \mathsf{I}_3\cdot\boldsymbol{\nabla} = \boldsymbol{\nabla}. \tag{40b}$$

Note that the unit dyadics are symmetric:

$$\tilde{\mathsf{I}}_4 = \mathsf{I}_4, \quad \tilde{\mathsf{I}}_3 = \mathsf{I}_3, \tag{41}$$

and that

$$(\mathsf{I}_4)_s = 4, \quad (\mathsf{I}_3)_s = 3. \tag{42}$$

We now discuss some of the properties of the dyadic $\mathbf{a}\times\mathsf{I}_3$ which will be used extensively in the sequel. First we show that the dyadics $\mathbf{a}\times\mathsf{I}_3$ and $\mathsf{I}_3\times\mathbf{a}$ are equal. A compact proof is

$$\mathbf{a}\times\mathsf{I}_3 = \mathsf{I}_3\cdot(\mathbf{a}\times\mathsf{I}_3) = (\mathsf{I}_3\times\mathbf{a})\cdot\mathsf{I}_3 = \mathsf{I}_3\times\mathbf{a}. \tag{43}$$

In order to make this calculation more transparent, we write Eq. (43) with $\mathsf{I}_3 = \underline{e}_k\underline{e}_k$. Hence

$$\begin{aligned}
\mathbf{a}\times\underline{e}_k\underline{e}_k &= \underline{e}_l\underline{e}_l\cdot(\mathbf{a}\times\underline{e}_k\underline{e}_k)\\
&= (\underline{e}_l\underline{e}_l\times\mathbf{a})\cdot\underline{e}_k\underline{e}_k = \underline{e}_l\underline{e}_l\times\mathbf{a}.
\end{aligned}$$

This shows explicitly that the second step is simply an interchange of the dot and cross in the vector triple product $\underline{e}_l\cdot\mathbf{a}\times\underline{e}_k$.

The "dot" products of $\mathbf{a}\times\mathsf{I}_3$ with a vector $\mathbf{b}$ are obtained as follows:

$$\begin{aligned}
\mathbf{b}\cdot(\mathbf{a}\times\mathsf{I}_3) &= (\mathbf{b}\times\mathbf{a})\cdot\mathsf{I}_3 = \mathbf{b}\times\mathbf{a},\\
(\mathbf{a}\times\mathsf{I}_3)\cdot\mathbf{b} &= \mathbf{a}\times(\mathsf{I}_3\cdot\mathbf{b}) = \mathbf{a}\times\mathbf{b}.
\end{aligned} \tag{44}$$

Also, we have

$$(\mathbf{a}\times\mathsf{I}_3)\cdot(\mathbf{b}\times\mathsf{I}_3) = \mathbf{a}\times(\mathbf{b}\times\mathsf{I}_3) = \mathbf{ba} - \mathbf{a}\cdot\mathbf{b}\mathsf{I}_3. \tag{45}$$

The dyadic $\mathbf{a}\times\mathsf{I}_3$ is antisymmetric, as can be seen from

$$\begin{aligned}
(\mathbf{a}\times\mathsf{I}_3)_T &= (\mathbf{a}\times\underline{e}_k\underline{e}_k)_T = \underline{e}_k\mathbf{a}\times\underline{e}_k,\\
&= -\underline{e}_k\underline{e}_k\times\mathbf{a} = -\mathsf{I}_3\times\mathbf{a} = -\mathbf{a}\times\mathsf{I}_3.
\end{aligned} \tag{46}$$

We now make use of the above results to derive the most general form for an antisymmetric dyadic G in $\mathcal{M}_4$. To this end note that

$$\mathsf{G} = (\mathsf{I}_3 - \underline{e}_0\underline{e}_0)\cdot\mathsf{G}\cdot(\mathsf{I}_3 - \underline{e}_0\underline{e}_0) = \mathsf{C} + \underline{e}_0\wedge\mathbf{b},$$

where we have used the antisymmetry of G to write

$$\begin{gathered}
\underline{e}_0\cdot\mathsf{G}\cdot\underline{e}_0 = 0,\\
\mathbf{b} = \mathsf{I}_3\cdot\mathsf{G}\cdot\underline{e}_0 = \mathsf{G}\cdot\underline{e}_0 = -\underline{e}_0\cdot\mathsf{G},\\
\mathsf{C} = \mathsf{I}_3\cdot\mathsf{G}\cdot\mathsf{I}_3 = -\tilde{\mathsf{C}}.
\end{gathered}$$

Since I_3 is the projection operator into $\mathcal{E}_3$, it is evident that $\mathbf{b}\in\mathcal{E}_3$ and the dyadic C is composed of vectors in $\mathcal{E}_3$. Moreover, from the antisymmetry of C we have that

$$\begin{aligned}
\mathsf{C} = \tfrac{1}{2}(\mathsf{C} - \tilde{\mathsf{C}}) &= \tfrac{1}{2}(\mathsf{C} - \tilde{\mathsf{C}})\cdot\underline{e}_k\underline{e}_k\\
&= \tfrac{1}{2}(\mathsf{C}\cdot\underline{e}_k - \underline{e}_k\cdot\mathsf{C})\underline{e}_k\\
&= -\tfrac{1}{2}\mathsf{C}_{\times}\times\underline{e}_k\underline{e}_k = -\tfrac{1}{2}\mathsf{C}_{\times}\times\mathsf{I}_3,
\end{aligned}$$

where use has been made of the operation defined by Eq. (22). Introducing the vector in $\mathcal{E}_3$ dual to C by

$$\mathbf{a} = \tfrac{1}{2}\mathsf{C}_{\times},$$

we finally have

$$\mathsf{G} = -\mathbf{a}\times\mathsf{I}_3 + \underline{e}_0\wedge\mathbf{b}. \tag{47a}$$

It is interesting to observe that the above result written in terms of components yields

$$\mathsf{G} = \tfrac{1}{2}\mathbf{a} \cdot (\underline{e}_k \times \underline{e}_l)\mathsf{M}_{(kl)} + b_k \mathsf{M}_{(0k)}, \tag{47b}$$

where

$$\mathsf{M}_{(\mu\nu)} = \underline{e}_\mu \wedge \underline{e}_\nu \tag{48}$$

are the six generators of the infinitesimal Lorentz transformations.[12]

Associated to any antisymmetric dyadic G there is another useful dyadic *G, called the dual of G. For the purpose of defining *G we digress briefly and introduce the concept of polyadics.

A polyad of order n is defined as a tensor product of n four-vectors. A polyadic of order n is then defined as a sum of polyads of order n. For example, a polyadic of order 4 can be written as follows:

$$\mathbf{\Phi} = \sum_{k=1}^{k'} \underline{c}_k \underline{d}_k \underline{f}_k \underline{g}_k .$$

The "double dot" product of $\mathbf{\Phi}$ with a dyadic $\mathsf{A} = \sum_{l=1}^{l'} \underline{a}_l \underline{b}_l$ is defined as

$$\begin{aligned} \mathbf{\Phi}:\mathsf{A} &= \Big(\sum_k \underline{c}_k \underline{d}_k \underline{f}_k \underline{g}_k\Big) : \Big(\sum_l \underline{a}_l \underline{b}_l\Big) \\ &= \sum_k \sum_l \underline{c}_k \underline{d}_k (\underline{f}_k \cdot \underline{a}_l)(\underline{g}_k \cdot \underline{b}_l). \end{aligned} \tag{49}$$

As a generalization of Eq. (27), we define the "exterior product" of four vectors $\underline{c}$, $\underline{d}$, $\underline{f}$, and $\underline{g}$ to be their antisymmetrized tensor product, and it is denoted by

$$\underline{c} \wedge \underline{d} \wedge \underline{f} \wedge \underline{g}.$$

In particular, the exterior product of the basic vectors $\underline{e}_\mu$ is

$$\mathbf{\Gamma} = \underline{e}_1 \wedge \underline{e}_2 \wedge \underline{e}_3 \wedge \underline{e}_0 \equiv \epsilon^{\mu\nu\lambda\delta} \underline{e}_\mu \underline{e}_\nu \underline{e}_\lambda \underline{e}_\delta = -\epsilon_{\mu\nu\lambda\delta} \underline{e}^\mu \underline{e}^\nu \underline{e}^\lambda \underline{e}^\delta, \tag{50}$$

where $\epsilon^{\mu\nu\lambda\delta}$ is the conventional Levi-Civita symbol defined by

$$\epsilon^{\mu\nu\lambda\delta} = \epsilon_{\mu\nu\lambda\delta} = \begin{cases} +1 & \text{if } \mu\nu\lambda\delta \text{ is an even permutation of 1230,} \\ -1 & \text{if } \mu\nu\lambda\delta \text{ is an odd permutation of 1230,} \\ 0 & \text{if two or more of the indices } \mu\nu\lambda\delta \text{ are equal.} \end{cases} \tag{51}$$

$\mathbf{\Gamma}$ is the only independent totally antisymmetric tetradic; thus any other one is proportional to it, e.g.,

$$\begin{aligned} \underline{c} \wedge \underline{d} \wedge \underline{f} \wedge \underline{g} &= c^\mu d^\nu f^\lambda g^\delta \underline{e}_\mu \wedge \underline{e}_\nu \wedge \underline{e}_\lambda \wedge \underline{e}_\delta, \\ &= c^\mu d^\nu f^\lambda g^\delta \epsilon_{\mu\nu\lambda\delta} \underline{e}_1 \wedge \underline{e}_2 \wedge \underline{e}_3 \wedge \underline{e}_0, \\ &= c^\mu d^\nu f^\lambda g^\delta \epsilon_{\mu\nu\lambda\delta} \mathbf{\Gamma}. \end{aligned} \tag{52}$$

[12] S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row and Peterson, Evanston, Illinois, 1961), p. 40.

Making use of $\mathbf{\Gamma}$, we now define the dual *G by

$$^\star\mathsf{G} = \tfrac{1}{2}\mathbf{\Gamma}:\mathsf{G}. \tag{53}$$

Furthermore, in Appendix A we show that

$$^\star(\underline{e}_0 \wedge \mathbf{b}) = -\mathbf{b} \times \mathsf{I}_3$$

and

$$^\star(\mathbf{a} \times \mathsf{I}_3) = \underline{e}_0 \wedge \mathbf{a}.$$

Hence, if

$$\mathsf{G} = -\mathbf{a} \times \mathsf{I}_3 + \underline{e}_0 \wedge \mathbf{b},$$

then

$$^\star\mathsf{G} = -\mathbf{b} \times \mathsf{I}_3 - \underline{e}_0 \wedge \mathbf{a} \tag{54}$$

and

$$^{\star\star}\mathsf{G} = {}^\star({}^\star\mathsf{G}) = \mathbf{a} \times \mathsf{I}_3 - \underline{e}_0 \wedge \mathbf{b} = -\mathsf{G}. \tag{55}$$

2.3. Rules of Differentiation

To conclude this section, we review briefly some useful differential operations in $\mathcal{M}_4$ involving dyadics.

(1) The four-gradient of a four-vector function $\underline{u}(\underline{x})$ is the dyadic:

$$\Box\underline{u} = \underline{e}^\mu(\partial_\mu \underline{u}). \tag{56}$$

(2) The four-gradient of a dyadic function $\mathsf{A}(\underline{x}) = \sum_l \underline{a}_l(\underline{x})\underline{b}_l(\underline{x})$ is the triadic:

$$\Box\mathsf{A} = \underline{e}^\mu \sum_l (\partial_\mu \underline{a}_l)\underline{b}_l + \underline{e}^\mu \sum_l \underline{a}_l(\partial_\mu \underline{b}_l). \tag{57}$$

(3) The four-divergence of a dyadic function $\mathsf{A}(\underline{x})$, obtained by contraction of the previous result, is the four-vector:

$$\Box \cdot \mathsf{A} = \sum_l (\Box \cdot \underline{a}_l)\underline{b}_l + \sum_l (\underline{a}_l \cdot \Box)\underline{b}_l. \tag{58}$$

(4) The partial four-gradient with respect to $\underline{u}$ of a function $f(\underline{u}, \underline{v}, \cdots,)$ of the four-vector $\underline{u}$ and other quantities can be defined as a simple generalization of Eq. (14), and will be equivalently denoted by $\Box_{\underline{u}} f$ or $\partial f/\partial \underline{u}$.

(5) The partial derivative with respect to a dyadic $\partial f/\partial\mathsf{A}$, where $f(\mathsf{A}, \cdots,)$ is a function of the dyadic A and other quantities, is a dyadic defined by the infinitesimal relation

$$df = f(\mathsf{A} + d\mathsf{A}, \cdots,) - f(\mathsf{A}, \cdots,) = d\mathsf{A}:\partial f/\partial\mathsf{A} \tag{59a}$$

with all other independent variables kept fixed.

Observing that f can be regarded as a function of the components $A^{\mu\nu}, \cdots$, we can alternatively write

$$df = dA^{\mu\nu}\frac{\partial f}{\partial A^{\mu\nu}} = (d\mathsf{A}:\underline{e}^\mu \underline{e}^\nu)\frac{\partial f}{\partial A^{\mu\nu}} = d\mathsf{A}:\Big(\underline{e}^\mu \underline{e}^\nu \frac{\partial f}{\partial A^{\mu\nu}}\Big), \tag{59b}$$

which, by comparison with Eq. (59a), yields

$$\frac{\partial f}{\partial \mathsf{A}} = \underline{e}^\mu \underline{e}^\nu \frac{\partial f}{\partial A^{\mu\nu}}. \tag{60}$$

Making use of the above defined operations, we obtain the identities

$$\square \cdot (\mathsf{A} \cdot \mathsf{B}) = (\square \cdot \mathsf{A}) \cdot \mathsf{B} + (\tilde{\mathsf{A}} \cdot \square) \cdot \mathsf{B} = (\square \cdot \mathsf{A}) \cdot \mathsf{B} + \mathsf{A} : (\square \mathsf{B}), \tag{61}$$

$$\square(\mathsf{A} : \mathsf{B}) = (\square \mathsf{A}) : \mathsf{B} + (\square \mathsf{B}) : \mathsf{A}, \tag{62}$$

$$\square f(\underline{\mathbf{u}}(\underline{\mathbf{x}})) = \frac{\partial \underline{\mathbf{u}}}{\partial \underline{\mathbf{x}}} \cdot \frac{\partial f}{\partial \underline{\mathbf{u}}} \equiv (\square \underline{\mathbf{u}}) \cdot \frac{\partial f}{\partial \underline{\mathbf{u}}}, \tag{63}$$

$$\square f(\mathsf{A}(\underline{\mathbf{x}})) = \frac{\partial \mathsf{A}}{\partial \underline{\mathbf{x}}} : \frac{\partial f}{\partial \mathsf{A}} \equiv (\square \mathsf{A}) : \frac{\partial f}{\partial \mathsf{A}}. \tag{64}$$

3. APPLICATIONS TO SPECIAL RELATIVITY

As an illustration of the operations developed in the preceding section, we first derive the Euler–Lagrange equations and the canonical energy–momentum tensor in dyadic form for an arbitrary four-vector field. The advantages of the four-vector and dyadic techniques over the usual tensor formalism will be further demonstrated by considering applications to some specific topics regarding electromagnetic fields and Lorentz transformations.

A. Euler–Lagrange Equations and Canonical Energy–Momentum Tensor

Consider a Lagrangian density of the form $\mathfrak{L}(\boldsymbol{\eta}, \square\boldsymbol{\eta}, \underline{\mathbf{x}})$, where $\boldsymbol{\eta}(\underline{\mathbf{x}})$ is an arbitrary four-vector field. The Euler–Lagrange equations are then obtained by the usual requirement that the action integral

$$I = \int_\Omega \mathfrak{L} d^4x$$

be stationary for arbitrary variations $\delta\boldsymbol{\eta}$ such that $\delta\boldsymbol{\eta} = 0$ on the surface bounding the four-dimensional region of integration Ω. Thus

$$0 = \delta I = \delta \int_\Omega \mathfrak{L} d^4x = \int_\Omega \left[\delta\boldsymbol{\eta} \cdot \frac{\partial \mathfrak{L}}{\partial \boldsymbol{\eta}} + \delta(\square\boldsymbol{\eta}) : \frac{\partial \mathfrak{L}}{\partial \square\boldsymbol{\eta}} \right] d^4x,$$

where $\delta(\square\boldsymbol{\eta}) = \square\delta\boldsymbol{\eta}$. Integration by parts yields

$$0 = \delta I = \int_\Omega d^4x \delta\boldsymbol{\eta} \cdot \left[\frac{\partial \mathfrak{L}}{\partial \boldsymbol{\eta}} - \square \cdot \left(\frac{\partial \mathfrak{L}}{\partial(\square\boldsymbol{\eta})} \right) \right],$$

and since $\delta\boldsymbol{\eta}$ is arbitrary, the above equation implies the Euler–Lagrange equations

$$\frac{\partial \mathfrak{L}}{\partial \boldsymbol{\eta}} - \square \cdot \left(\frac{\partial \mathfrak{L}}{\partial(\square\boldsymbol{\eta})} \right) = 0. \tag{65}$$

To obtain a defining equation for the canonical energy–momentum tensor,[13] we take the total gradient $d\mathfrak{L}/d\underline{\mathbf{x}}$ of the Lagrangian:

$$\frac{d\mathfrak{L}}{d\underline{\mathbf{x}}} \equiv \square\mathfrak{L} = (\square\boldsymbol{\eta}) \cdot \frac{\partial \mathfrak{L}}{\partial \boldsymbol{\eta}} + (\square\square\boldsymbol{\eta}) : \frac{\partial \mathfrak{L}}{\partial(\square\boldsymbol{\eta})} + \frac{\partial \mathfrak{L}}{\partial \underline{\mathbf{x}}}, \tag{66a}$$

which, after substitution from Eq. (65), becomes

$$\square\mathfrak{L} = \square\boldsymbol{\eta} \cdot \left(\square \cdot \frac{\partial \mathfrak{L}}{\partial(\square\boldsymbol{\eta})} \right) + (\square\square\boldsymbol{\eta}) : \frac{\partial \mathfrak{L}}{\partial(\square\boldsymbol{\eta})} + \frac{\partial \mathfrak{L}}{\partial \underline{\mathbf{x}}} = \left(\square \cdot \frac{\partial \mathfrak{L}}{\partial(\square\boldsymbol{\eta})} \right) \cdot \widetilde{\square\boldsymbol{\eta}} + \frac{\partial \mathfrak{L}}{\partial(\square\boldsymbol{\eta})} : (\square\widetilde{\square\boldsymbol{\eta}}) + \frac{\partial \mathfrak{L}}{\partial \underline{\mathbf{x}}}. \tag{66b}$$

Moreover, making use of Eq. (61) gives

$$\square \cdot \mathsf{T}' = -\frac{\partial \mathfrak{L}}{\partial \underline{\mathbf{x}}}, \tag{67}$$

where

$$\mathsf{T}' = \frac{\partial \mathfrak{L}}{\partial(\square\boldsymbol{\eta})} \cdot (\widetilde{\square\boldsymbol{\eta}}) - \mathsf{I}_4 \mathfrak{L} \tag{68}$$

is the canonical energy–momentum tensor.

In particular, the appropriate Lagrangian density for the derivation of the electromagnetic field equations is

$$\mathfrak{L} = -(16\pi)^{-1} \mathsf{F} : \mathsf{F} + (1/c) \underline{\mathbf{J}} \cdot \underline{\mathbf{A}}, \tag{69}$$

where

$$\underline{\mathbf{A}} = \mathbf{A} + \varphi \underline{\mathbf{e}}_0 \tag{70}$$

is the four-vector potential,

$$\underline{\mathbf{J}} = \mathbf{j} + c\rho \underline{\mathbf{e}}_0 \tag{71}$$

is the four-vector current source, and

$$\mathsf{F} = \square \wedge \underline{\mathbf{A}} \equiv \square\underline{\mathbf{A}} - \widetilde{\square\underline{\mathbf{A}}} \tag{72}$$

is the electromagnetic field dyadic. In this case, from Eqs. (60) and (72), we have

$$4\pi \frac{\partial \mathfrak{L}}{\partial(\square\underline{\mathbf{A}})} = -\frac{1}{4} \frac{\partial}{\partial(\square\underline{\mathbf{A}})} (\mathsf{F} : \mathsf{F}) = -\frac{1}{2} \frac{\partial \mathsf{F}}{\partial(\square\underline{\mathbf{A}})} : \mathsf{F} = -\frac{1}{2} \left[\frac{\partial}{\partial(\square\underline{\mathbf{A}})} (\square \wedge \underline{\mathbf{A}}) \right] : \mathsf{F} = -\left[\frac{\partial}{\partial(\square\underline{\mathbf{A}})} \square\underline{\mathbf{A}} \right] : \mathsf{F} = -\mathsf{F}, \tag{73}$$

and

$$\frac{\partial \mathfrak{L}}{\partial \underline{\mathbf{A}}} = \frac{1}{c} \frac{\partial}{\partial \underline{\mathbf{A}}} (\underline{\mathbf{J}} \cdot \underline{\mathbf{A}}) = \frac{1}{c} \underline{\mathbf{J}}. \tag{74}$$

[13] For a comparison of the dyadic forms obtained below with their corresponding component representations see, for example, L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Pergamon Press, Ltd., Oxford, England, 1962), Chap. 4; V. Fock, *The Theory of Space, Time and Gravitation* (Pergamon Press, Ltd., Oxford, England, 1964), Chap. 2. Also see Ref. 12, Sec. 7d.

Inserting these results into the Euler–Lagrange equations

$$\frac{\partial \mathfrak{L}}{\partial \underline{\mathbf{A}}} - \square \cdot \left(\frac{\partial \mathfrak{L}}{\partial (\square \underline{\mathbf{A}})}\right) = 0$$

yields the Maxwell equations

$$\square \cdot \mathsf{F} = -(4\pi/c)\underline{\mathbf{J}} \tag{75}$$

in dyadic form.

Equation (68) for the canonical energy–momentum tensor becomes

$$\begin{aligned} \mathsf{T}' &= \frac{\partial \mathfrak{L}}{\partial (\square \underline{\mathbf{A}})} \cdot \widetilde{\square \underline{\mathbf{A}}} - \mathfrak{L} \mathsf{I}_4 \\ &= -\frac{1}{4\pi} \mathsf{F} \cdot \widetilde{\square \underline{\mathbf{A}}} + \frac{1}{16\pi} \mathsf{F}:\mathsf{F} \mathsf{I}_4 - \frac{1}{c} \underline{\mathbf{J}} \cdot \underline{\mathbf{A}} \mathsf{I}_4, \end{aligned} \tag{76}$$

and substituting into Eq. (67) results in

$$\begin{aligned} &(4\pi)^{-1} \square \cdot (-\mathsf{F} \cdot \widetilde{\square \underline{\mathbf{A}}} + \tfrac{1}{4} \mathsf{F}:\mathsf{F} \mathsf{I}_4) \\ &\qquad = c^{-1}[\square(\underline{\mathbf{J}} \cdot \underline{\mathbf{A}}) - (\square \underline{\mathbf{J}}) \cdot \underline{\mathbf{A}}]. \end{aligned}$$

In order to symmetrize the term on the left-hand side of the above expression, we add to both sides the quantity $(4\pi)^{-1}\square \cdot (\mathsf{F} \cdot \square \underline{\mathbf{A}})$, whereby

$$\begin{aligned} (4\pi)^{-1}\square \cdot (\mathsf{F} \cdot \mathsf{F} + \tfrac{1}{4} \mathsf{F}:\mathsf{F}\mathsf{I}_4) &= c^{-1}(\square \underline{\mathbf{A}} \cdot \underline{\mathbf{J}}) + (4\pi)^{-1}\square \cdot (\mathsf{F} \cdot \square \underline{\mathbf{A}}) \\ &= c^{-1}(\square \underline{\mathbf{A}} \cdot \underline{\mathbf{J}}) + (4\pi)^{-1}[(\square \cdot \mathsf{F}) \cdot \square \underline{\mathbf{A}} \\ &\qquad - (\mathsf{F} \cdot \square) \cdot \square \underline{\mathbf{A}}] \\ &= c^{-1}[\square \underline{\mathbf{A}} \cdot \underline{\mathbf{J}} - \widetilde{\square \underline{\mathbf{A}}} \cdot \underline{\mathbf{J}}], \end{aligned}$$

i.e.,

$$(4\pi)^{-1}\square \cdot (\mathsf{F} \cdot \mathsf{F} + \tfrac{1}{4} \mathsf{F}:\mathsf{F}\mathsf{I}_4) = c^{-1}\mathsf{F} \cdot \underline{\mathbf{J}}. \tag{77}$$

The dyadic

$$\mathsf{T} = (4\pi)^{-1}(\mathsf{F} \cdot \mathsf{F} + \tfrac{1}{4} \mathsf{F}:\mathsf{F}\mathsf{I}_4) \tag{78}$$

is the well-known symmetric energy–momentum tensor for the electromagnetic field. The quantity $(1/c)\mathsf{F} \cdot \underline{\mathbf{J}}$ is the four-force per unit proper volume, also known as the Lorentz force density.

B. Dyadic Form of Maxwell Field Equations

Equation (75) for the dyadic form of the Maxwell source equations was derived by postulating a Lagrangian density and by making use of the Euler–Lagrange equations (65). Here we present an alternate and straightforward procedure for obtaining the field equations in dyadic form. The electromagnetic field dyadic F will be expressed directly in terms of the electric and magnetic field variables $\mathbf{E}$ and $\mathbf{H}$, a form which in some cases is more convenient for calculational purposes.

We consider first the two Maxwell source equations

$$\nabla \times \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j}, \quad \nabla \cdot \mathbf{E} = 4\pi\rho, \tag{79a, b}$$

and write the derivatives in terms of the $\square$ operator:

$$\nabla \times \mathbf{H} = (\square \cdot \mathsf{I}_3) \times \mathbf{H} = \square \cdot (\mathsf{I}_3 \times \mathbf{H}) = \square \cdot (\mathbf{H} \times \mathsf{I}_3), \tag{80}$$

$$-\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = -\frac{\partial \mathbf{E}}{\partial x^0} = -\underline{\mathbf{e}}_0 \cdot \square \mathbf{E} = -\square \cdot (\underline{\mathbf{e}}_0 \mathbf{E}), \tag{81}$$

$$\nabla \cdot \mathbf{E} = (\square \cdot \mathsf{I}_3) \cdot \mathbf{E} = \square \cdot (\mathsf{I}_3 \cdot \mathbf{E}) = \square \cdot \mathbf{E}. \tag{82}$$

Substitution of (80) and (81) into (79a) gives

$$\square \cdot \mathbf{H} \times \mathsf{I}_3 - \square \cdot \underline{\mathbf{e}}_0 \mathbf{E} = (4\pi/c)\mathbf{j}, \tag{83}$$

and multiplication of (79b) on the right-hand side by $\underline{\mathbf{e}}_0$ results in

$$\square \cdot \mathbf{E}\underline{\mathbf{e}}_0 = 4\pi\rho\underline{\mathbf{e}}_0. \tag{84}$$

Addition of Eqs. (83) and (84) and multiplication by (-1) yields, finally,

$$\square \cdot \mathsf{F} = -(4\pi/c)\underline{\mathbf{J}},$$

where

$$\mathsf{F} = -\mathbf{H} \times \mathsf{I}_3 + \underline{\mathbf{e}}_0 \wedge \mathbf{E} \tag{85}$$

is the alternate dyadic form of the electromagnetic field tensor.

By applying the same procedure to the second pair of Maxwell equations

$$\nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} = 0, \quad \nabla \cdot \mathbf{H} = 0,$$

we obtain

$$\square \cdot {}^*\mathsf{F} = 0, \tag{86}$$

where, in accordance with Eqs. (A3) and (A4)

$${}^*\mathsf{F} = -\mathbf{E} \times \mathsf{I}_3 - \underline{\mathbf{e}}_0 \wedge \mathbf{H} \tag{87}$$

is the dual of F.

The usual covariant tensor form of Eq. (75) can be obtained simply by taking the dot product of both sides with $\underline{\mathbf{e}}^\mu$:

$$(\underline{\mathbf{e}}^\nu \partial_\nu) \cdot \mathsf{F} \cdot \underline{\mathbf{e}}^\mu = -(4\pi/c)\underline{\mathbf{J}} \cdot \underline{\mathbf{e}}^\mu,$$

or, from the antisymmetry of F,

$$\partial_\nu F^{\mu\nu} = (4\pi/c)J^\mu. \tag{75'}$$

Similarly, from Eq. (86) it follows that

$$\partial_\nu {}^*F^{\mu\nu} = 0. \tag{86'}$$

The components $F^{\mu\nu}$ are given by

$$\begin{aligned} -F^{k0} = F^{0k} &= \underline{\mathbf{e}}^0 \cdot \mathsf{F} \cdot \underline{\mathbf{e}}^k \\ &= \underline{\mathbf{e}}^0 \cdot (-\mathbf{H} \times \mathsf{I}_3 + \underline{\mathbf{e}}_0 \wedge \mathbf{E}) \cdot \underline{\mathbf{e}}^k \\ &= \underline{\mathbf{e}}^0 \cdot \underline{\mathbf{e}}_0 \mathbf{E} \cdot \underline{\mathbf{e}}^k = E^k, \\ F^{kl} &= \underline{\mathbf{e}}^k \cdot \mathsf{F} \cdot \underline{\mathbf{e}}^l = -\underline{\mathbf{e}}^k \cdot \mathbf{H} \times \underline{\mathbf{e}}^l \\ &= \underline{\mathbf{e}}_k \times \underline{\mathbf{e}}_l \cdot \underline{\mathbf{e}}_m H^m = \epsilon_{klm} H^m, \\ F^{00} &= \underline{\mathbf{e}}^0 \cdot \mathsf{F} \cdot \underline{\mathbf{e}}^0 = 0. \end{aligned}$$

The corresponding components of $^*\mathsf{F}$ are obtained by replacing $\mathbf{E}$ by $-\mathbf{H}$ and $\mathbf{H}$ by $\mathbf{E}$ in the above equations.

The basic simplicity and typographic economy of the dyadic forms, Eqs. (75), (85), (86), and (87), are evident when compared with their corresponding component equations. The situation is analogous to the familiar economy of vector expressions in contrast to their component forms. In addition to its aesthetic appeal, the dyadic formalism entails considerable calculational advantages, as will be demonstrated in the following discussions.

C. Invariants of the Electromagnetic Field

The two invariants of the field under proper Lorentz transformations can be easily derived as special cases of the "double dot" product $\mathsf{G}:\mathsf{G}'$ of the arbitrary antisymmetric dyadics

$$\mathsf{G} = -\mathbf{a} \times \mathsf{I}_3 + \underline{\mathbf{e}}_0 \wedge \mathbf{b}, \tag{88a}$$

$$\mathsf{G}' = -\mathbf{a}' \times \mathsf{I}_3 + \underline{\mathbf{e}}_0 \wedge \mathbf{b}'. \tag{88b}$$

First, recalling Eq. (29) and using the antisymmetry of G, we have

$$\mathsf{G}:\mathsf{G}' = (\tilde{\mathsf{G}} \cdot \mathsf{G}')_s = -(\mathsf{G} \cdot \mathsf{G}')_s. \tag{89a}$$

Furthermore,

$$\begin{aligned} \mathsf{G} \cdot \mathsf{G}' &= (-\mathbf{a} \times \mathsf{I}_3 + \underline{\mathbf{e}}_0 \wedge \mathbf{b}) \cdot (-\mathbf{a}' \times \mathsf{I}_3 + \underline{\mathbf{e}}_0 \wedge \mathbf{b}'), \\ &= \mathbf{a} \times (\mathbf{a}' \times \mathsf{I}_3) + \mathbf{a} \times \mathbf{b}'\underline{\mathbf{e}}_0 - \underline{\mathbf{e}}_0\mathbf{b} \times \mathbf{a}' \\ &\quad + (\underline{\mathbf{e}}_0 \wedge \mathbf{b}) \cdot (\underline{\mathbf{e}}_0 \wedge \mathbf{b}') \\ &= \mathbf{a}'\mathbf{a} \cdot \mathsf{I}_3 - \mathbf{a} \cdot \mathbf{a}'\mathsf{I}_3 + \mathbf{a} \times \mathbf{b}'\underline{\mathbf{e}}_0 - \underline{\mathbf{e}}_0\mathbf{b} \times \mathbf{a}' \\ &\quad - \underline{\mathbf{e}}_0 \cdot \underline{\mathbf{e}}_0\mathbf{b}\mathbf{b}' - \mathbf{b} \cdot \mathbf{b}'\underline{\mathbf{e}}_0\underline{\mathbf{e}}_0 \\ &= \mathbf{a}'\mathbf{a} + \mathbf{b}\mathbf{b}' - \mathbf{a} \cdot \mathbf{a}'\mathsf{I}_3 + \mathbf{a} \times \mathbf{b}'\underline{\mathbf{e}}_0 \\ &\quad - \underline{\mathbf{e}}_0\mathbf{b} \times \mathbf{a}' - \mathbf{b} \cdot \mathbf{b}'\underline{\mathbf{e}}_0\underline{\mathbf{e}}_0. \end{aligned} \tag{90}$$

Consequently,

$$\begin{aligned} \mathsf{G}:\mathsf{G}' &= -\mathbf{a}' \cdot \mathbf{a} - \mathbf{b} \cdot \mathbf{b}' + \mathbf{a} \cdot \mathbf{a}'(\mathsf{I}_3)_s - \mathbf{a} \times \mathbf{b}' \cdot \underline{\mathbf{e}}_0 \\ &\qquad + \underline{\mathbf{e}}_0 \cdot \mathbf{b} \times \mathbf{a}' + \mathbf{b} \cdot \mathbf{b}'\underline{\mathbf{e}}_0 \cdot \underline{\mathbf{e}}_0 \\ &= -\mathbf{a}' \cdot \mathbf{a} - \mathbf{b} \cdot \mathbf{b}' + 3\mathbf{a} \cdot \mathbf{a}' - \mathbf{b} \cdot \mathbf{b}' \\ &= 2(\mathbf{a} \cdot \mathbf{a}' - \mathbf{b} \cdot \mathbf{b}'). \end{aligned} \tag{89b}$$

In particular, setting $\mathsf{G} = \mathsf{G}' = \mathsf{F}$ (i.e., $\mathbf{b} = \mathbf{b}' = \mathbf{E}$ and $\mathbf{a} = \mathbf{a}' = \mathbf{H}$), gives the first invariant quantity

$$\mathsf{F}:\mathsf{F} = 2(H^2 - E^2). \tag{91}$$

The second invariant is obtained by setting $\mathsf{G} = \mathsf{F}$ and $\mathsf{G}' = {}^*\mathsf{F}$ (i.e., $\mathbf{b} = \mathbf{E}$, $\mathbf{a} = \mathbf{H}$, $\mathbf{b}' = -\mathbf{H}$, and $\mathbf{a}' = \mathbf{E}$).

Thus

$$\mathsf{F}:{}^*\mathsf{F} = 2(\mathbf{H} \cdot \mathbf{E} + \mathbf{E} \cdot \mathbf{H}) = 4\mathbf{E} \cdot \mathbf{H}. \tag{92}$$

Note that the substitution $\mathsf{G} = \mathsf{G}' = {}^*\mathsf{F}$ results in

$${}^*\mathsf{F}:{}^*\mathsf{F} = -\mathsf{F}:\mathsf{F}.$$

Hence there are only two independent invariants.

From the symmetric energy–momentum dyadic T given in Eq. (78) we can obtain another interesting quantity[14] $\mathsf{T}:\mathsf{T}$ which is a linear combination of the two above derived invariants. To this end note that

$$\mathsf{T}:\mathsf{T} = (\mathsf{T} \cdot \mathsf{T})_s \tag{93}$$

and

$$\begin{aligned} \mathsf{T} \cdot \mathsf{T} = (4\pi)^{-2}[&(\tfrac{1}{16})(\mathsf{F}:\mathsf{F})^2\mathsf{I}_4 \\ &+ \tfrac{1}{2}(\mathsf{F}:\mathsf{F})\mathsf{F} \cdot \mathsf{F} + \mathsf{F} \cdot \mathsf{F} \cdot \mathsf{F} \cdot \mathsf{F}]. \end{aligned} \tag{94}$$

Furthermore, by specializing the result in Eq. (90), we have

$$\begin{aligned} \mathsf{F} \cdot \mathsf{F} &= \mathbf{H}\mathbf{H} + \mathbf{E}\mathbf{E} - H^2\mathsf{I}_3 - \mathbf{E} \times \mathbf{H}\underline{\mathbf{e}}_0 \\ &\quad - \underline{\mathbf{e}}_0\mathbf{E} \times \mathbf{H} - E^2\underline{\mathbf{e}}_0\underline{\mathbf{e}}_0, \\ {}^*\mathsf{F} \cdot \mathsf{F}^* &= \mathbf{E}\mathbf{E} + \mathbf{H}\mathbf{H} - E^2\mathsf{I}_3 - \mathbf{E} \times \mathbf{H}\underline{\mathbf{e}}_0 \\ &\quad - \underline{\mathbf{e}}_0\mathbf{E} \times \mathbf{H} - H^2\underline{\mathbf{e}}_0\underline{\mathbf{e}}_0, \\ \mathsf{F} \cdot {}^*\mathsf{F} &= -\mathbf{E} \cdot \mathbf{H}\mathsf{I}_4 = -\tfrac{1}{4}\mathsf{F}:{}^*\mathsf{F}\mathsf{I}_4. \end{aligned}$$

Hence

$$\mathsf{F} \cdot \mathsf{F} = {}^*\mathsf{F} \cdot {}^*\mathsf{F} - \tfrac{1}{2}\mathsf{F}:\mathsf{F}\mathsf{I}_4 \tag{95}$$

and

$$\begin{aligned} \mathsf{F} \cdot \mathsf{F} \cdot \mathsf{F} \cdot \mathsf{F} &= \mathsf{F} \cdot \mathsf{F} \cdot ({}^*\mathsf{F} \cdot {}^*\mathsf{F} - \tfrac{1}{2}\mathsf{F}:\mathsf{F}\mathsf{I}_4) \\ &= -\tfrac{1}{4}(\mathsf{F}:{}^*\mathsf{F})\mathsf{F} \cdot {}^*\mathsf{F} - \tfrac{1}{2}(\mathsf{F}:\mathsf{F})\mathsf{F} \cdot \mathsf{F} \\ &= (\tfrac{1}{16})(\mathsf{F}:{}^*\mathsf{F})^2\mathsf{I}_4 - \tfrac{1}{2}(\mathsf{F}:\mathsf{F})\mathsf{F} \cdot \mathsf{F}. \end{aligned} \tag{96}$$

Substituting this last result into Eq. (94) yields

$$\mathsf{T} \cdot \mathsf{T} = (16\pi)^{-2}[(\mathsf{F}:\mathsf{F})^2 + (\mathsf{F}:{}^*\mathsf{F})^2]\mathsf{I}_4, \tag{97}$$

and, by Eq. (93),

$$\mathsf{T}:\mathsf{T} = (8\pi)^{-2}[(\mathsf{F}:\mathsf{F})^2 + (\mathsf{F}:{}^*\mathsf{F})^2]. \tag{98}$$

By combining Eqs. (97) and (98), we have the identity

$$\mathsf{T} \cdot \mathsf{T} = \tfrac{1}{4}(\mathsf{T}:\mathsf{T})\mathsf{I}_4. \tag{99}$$

The invariant $\mathsf{T}:\mathsf{T}$ can be used to obtain a covariant representation of the energy and momentum properties of the electromagnetic field. For this purpose, let $\underline{\mathbf{n}}$ be a unit time-like four-vector, i.e.,

$$\underline{\mathbf{n}} \cdot \underline{\mathbf{n}} = -1.$$

We can then write

$$\begin{aligned} -(2c)^{-2}\mathsf{T}:\mathsf{T} &= (2c)^{-2}(\mathsf{T}:\mathsf{T})\underline{\mathbf{n}} \cdot \underline{\mathbf{n}} = (2c)^{-2}(\mathsf{T}:\mathsf{T})\mathsf{I}_4:\underline{\mathbf{n}}\underline{\mathbf{n}} \\ &= c^{-2}(\mathsf{T} \cdot \mathsf{T}):\underline{\mathbf{n}}\underline{\mathbf{n}} \\ &= c^{-2}\underline{\mathbf{n}} \cdot \mathsf{T} \cdot \mathsf{T} \cdot \underline{\mathbf{n}} = \underline{\mathbf{P}} \cdot \underline{\mathbf{P}}, \end{aligned} \tag{100}$$

where

$$\underline{\mathbf{P}} = (1/c)\mathsf{T} \cdot \underline{\mathbf{n}}. \tag{101}$$

[14] H. S. Ruse, Proc. London Math. Soc. 41, 302 (1936).

In the particular Lorentz frame where $\underline{\mathbf{n}} = \underline{\mathbf{e}}_0$, $\underline{\mathbf{P}}$ reduces to

$$\underline{\mathbf{P}} = (1/c)\mathsf{T} \cdot \underline{\mathbf{e}}_0,$$

which is the usual four-momentum density of the electromagnetic field.

D. Dirac Form of Maxwell's Equations

Here we show how a Dirac-like equation for the electromagnetic field can be derived from the dyadic form of Maxwell's equations (75) and (86). The procedure is more concise and direct than other approaches given in the literature, with the base elements of the Pauli algebra emerging naturally in the form of dyadics.

We begin by multiplying Eq. (86) by i and subtracting from Eq. (75). This results in

$$\Box \cdot \mathsf{G} = -(4\pi/c)\underline{\mathbf{J}}, \tag{102}$$

where

$$\mathsf{G} = \mathsf{F} - i\,{}^*\mathsf{F} = i(\mathbf{Z} \times \mathsf{I}_3 - i\underline{\mathbf{e}}_0 \wedge \mathbf{Z}) \tag{103a}$$

and

$$\mathbf{Z} = \mathbf{E} + i\mathbf{H}. \tag{104}$$

Since $\underline{\mathbf{e}}_0 \cdot \mathbf{Z} = 0$, the dyadic G can be expressed as

$$\begin{aligned}\mathsf{G} &= (i\mathsf{I}_3 \times \mathsf{I}_3 + \underline{\mathbf{e}}_0\mathsf{I}_3 - \underline{\mathbf{e}}_k\underline{\mathbf{e}}_0\underline{\mathbf{e}}_k) \cdot \mathbf{Z} \\ &= [i\mathsf{I}_3 \times \mathsf{I}_3 + \underline{\mathbf{e}}_0\mathsf{I}_4 - \underline{\mathbf{e}}_k(\underline{\mathbf{e}}_0 \wedge \underline{\mathbf{e}}_k)] \cdot \mathbf{Z},\end{aligned} \tag{103b}$$

and Eq. (102) becomes

$$\Box \cdot \mathbf{\Sigma} \cdot \mathbf{Z} = -(4\pi/c)\underline{\mathbf{J}}, \tag{105}$$

where

$$\mathbf{\Sigma} = i\mathsf{I}_3 \times \mathsf{I}_3 + \underline{\mathbf{e}}_0\mathsf{I}_4 - \underline{\mathbf{e}}_k(\underline{\mathbf{e}}_0 \wedge \underline{\mathbf{e}}_k) \tag{106}$$

is a triadic. Note that

$$\begin{aligned}&\underline{\mathbf{e}}^0 \cdot \mathbf{\Sigma} = \mathsf{I}_4, \\ &\underline{\mathbf{e}}^k \cdot \mathbf{\Sigma} = i\underline{\mathbf{e}}_k \times \mathsf{I}_3 - \underline{\mathbf{e}}_0 \wedge \underline{\mathbf{e}}_k = -\frac{i}{2}\epsilon_{klm}\mathsf{M}_{(lm)} - \mathsf{M}_{(0k)},\end{aligned} \tag{107}$$

where the $\mathsf{M}_{(\mu\nu)}$ were introduced in Eq. (48) as the generators of the infinitesimal Lorentz transformations. The four dyadics $\underline{\mathbf{e}}^\mu \cdot \mathbf{\Sigma}$ follow the Pauli algebra, with $\underline{\mathbf{e}}^0 \cdot \mathbf{\Sigma}$ being the identity and $\underline{\mathbf{e}}^k \cdot \mathbf{\Sigma}$ satisfying the relations

$$\begin{aligned}(\underline{\mathbf{e}}^k \cdot \mathbf{\Sigma}) \cdot (\underline{\mathbf{e}}^l \cdot \mathbf{\Sigma}) &= -\underline{\mathbf{e}}_l\underline{\mathbf{e}}_k + \underline{\mathbf{e}}_k \cdot \underline{\mathbf{e}}_l\mathsf{I}_3 + i\underline{\mathbf{e}}_k \times \underline{\mathbf{e}}_l\underline{\mathbf{e}}_0 \\ &\quad - i\underline{\mathbf{e}}_0\underline{\mathbf{e}}_k \times \underline{\mathbf{e}}_l - \underline{\mathbf{e}}_k \cdot \underline{\mathbf{e}}_l\underline{\mathbf{e}}_0\underline{\mathbf{e}}_0 + \underline{\mathbf{e}}_k\underline{\mathbf{e}}_l \\ &= i[i(\underline{\mathbf{e}}_k \times \underline{\mathbf{e}}_l) \times \mathsf{I}_3 - \underline{\mathbf{e}}_0 \wedge (\underline{\mathbf{e}}_k \times \underline{\mathbf{e}}_l)] \\ &\quad + \underline{\mathbf{e}}_k \cdot \underline{\mathbf{e}}_l\mathsf{I}_4 \\ &= i\epsilon_{klm}\underline{\mathbf{e}}^m \cdot \mathbf{\Sigma} + \delta_{kl}\mathsf{I}_4.\end{aligned} \tag{108}$$

E. Lorentz Transformations

We now apply the techniques of Sec. 2 to the study of some properties of the restricted homogeneous Lorentz transformations

$$\underline{\mathbf{x}}' = \mathsf{L} \cdot \underline{\mathbf{x}} \; (= L^\mu{}_\nu\underline{\mathbf{e}}_\mu\underline{\mathbf{e}}^\nu \cdot \underline{\mathbf{x}}), \tag{109}$$

which, by definition, satisfy the conditions:

$$\underline{\mathbf{x}}' \cdot \underline{\mathbf{x}}' = \underline{\mathbf{x}} \cdot \underline{\mathbf{x}}, \quad \text{(invariance)} \tag{110a}$$

$$L^0{}_0 \geq 1, \quad \text{(orthochronicity)} \tag{110b}$$

$$\det(L^\mu{}_\nu) = +1. \tag{110c}$$

From Eq. (110a) it immediately follows that

$$\tilde{\mathsf{L}} \cdot \mathsf{L} = \mathsf{L} \cdot \tilde{\mathsf{L}} = \mathsf{I}_4. \tag{111}$$

To derive explicitly the pure Lorentz transformation L_0 connecting two inertial frames S and S', where S' is moving with velocity $\mathbf{v}$ relative to S, consider a point P at rest in S', which, therefore, moves along the world line in the direction of $\underline{\mathbf{e}}_0$. This point is seen by an observer in S as moving with velocity $\mathbf{v}$, and consequently its world line is along the direction

$$\underline{\mathbf{f}} = \gamma(\underline{\mathbf{e}}_0 + \mathbf{v}/c), \tag{112}$$

where

$$\gamma = (1 - v^2/c^2)^{-\frac{1}{2}},$$

and $\underline{\mathbf{f}}$ has been normalized to make

$$\underline{\mathbf{f}} \cdot \underline{\mathbf{f}} = -1.$$

The four-vectors $\underline{\mathbf{e}}_0$ and $\underline{\mathbf{f}}$ must be related by the Lorentz transformation L_0 according to

$$\mathsf{L}_0 \cdot \underline{\mathbf{f}} = \underline{\mathbf{e}}_0, \tag{113}$$

where the sign on the right-hand side has been selected so that the orthochronicity condition be satisfied, i.e.,

$$\operatorname{sign}(\underline{\mathbf{e}}_0 \cdot \mathsf{L}_0 \cdot \underline{\mathbf{f}}) = \operatorname{sign}(\underline{\mathbf{e}}_0 \cdot \underline{\mathbf{f}}).$$

Moreover, since in this case the $\underline{\mathbf{e}}_0, \mathbf{v}$ plane remains invariant, there is a unit four-vector $\underline{\mathbf{w}}$ in the plane such that

$$\mathsf{L}_0 \cdot \underline{\mathbf{w}} = \hat{\mathbf{v}}. \tag{114}$$

To find $\underline{\mathbf{w}}$ use is made of the invariance of the inner product

$$0 = \underline{\mathbf{e}}_0 \cdot \hat{\mathbf{v}} = \underline{\mathbf{f}} \cdot \tilde{\mathsf{L}}_0 \cdot \mathsf{L}_0 \cdot \underline{\mathbf{w}} = \underline{\mathbf{f}} \cdot \underline{\mathbf{w}}, \tag{115}$$

i.e., $\underline{\mathbf{w}}$ will be orthogonal to $\underline{\mathbf{f}}$ and is found, by inspection, to be

$$\underline{\mathbf{w}} = \gamma[\hat{\mathbf{v}} + (v/c)\underline{\mathbf{e}}_0]. \tag{116}$$

The sign in the above equation has been chosen so that

$$\lim_{v\to 0} \underline{\mathbf{w}}(v) = \hat{\mathbf{v}}$$

in order to satisfy the determinantal condition Eq. (110c). Also, since L_0 will act as the identity in the plane perpendicular to the $\underline{\mathbf{e}}_0, \mathbf{v}$ plane, two unit and mutually orthogonal vectors $\hat{\mathbf{l}}_1$ and $\hat{\mathbf{l}}_2$ in this plane will remain invariant, i.e.,

$$\mathsf{L}_0 \cdot \hat{\mathbf{l}}_1 = \hat{\mathbf{l}}_1, \qquad \mathsf{L}_0 \cdot \hat{\mathbf{l}}_2 = \hat{\mathbf{l}}_2. \tag{117}$$

Thus we have four linearly independent vectors which, by making use of Eqs. (113) and (114), can be used to express L_0 as

$$\mathsf{L}_0 = \hat{\mathbf{l}}_1\hat{\mathbf{l}}_1 + \hat{\mathbf{l}}_2\hat{\mathbf{l}}_2 + \hat{\mathbf{v}}\underline{\mathbf{w}} - \underline{\mathbf{e}}_0\underline{\mathbf{f}} = \mathsf{I}_3 - \hat{\mathbf{v}}\hat{\mathbf{v}} - \underline{\mathbf{e}}_0\underline{\mathbf{f}} + \hat{\mathbf{v}}\underline{\mathbf{w}}. \tag{118a}$$

Alternatively, by substituting Eqs. (112) and (116) into the preceding expression yields

$$\begin{aligned}\mathsf{L}_0 &= \mathsf{I}_3 - \hat{\mathbf{v}}\hat{\mathbf{v}} - \gamma[\underline{\mathbf{e}}_0\underline{\mathbf{e}}_0 + (v/c)\underline{\mathbf{e}}_0\hat{\mathbf{v}}] + \gamma[\hat{\mathbf{v}}\hat{\mathbf{v}} + (v/c)\hat{\mathbf{v}}\underline{\mathbf{e}}_0] \\ &= (\mathsf{I}_4 - \hat{\mathbf{v}}\hat{\mathbf{v}} + \underline{\mathbf{e}}_0\underline{\mathbf{e}}_0) + (\hat{\mathbf{v}}\hat{\mathbf{v}} - \underline{\mathbf{e}}_0\underline{\mathbf{e}}) \cosh\varphi \\ &\qquad + \hat{\mathbf{v}} \wedge \underline{\mathbf{e}}_0 \sinh\varphi,\end{aligned} \tag{118b}$$

where the identification

$$\cosh\varphi = \gamma, \quad \sinh\varphi = \gamma(v/c)$$

has been made. Furthermore, since[15]

$$(\hat{\mathbf{v}} \wedge \underline{\mathbf{e}}_0)^{2k} = \hat{\mathbf{v}}\hat{\mathbf{v}} - \underline{\mathbf{e}}_0\underline{\mathbf{e}}_0, \quad k = 1, 2, \cdots,$$

and

$$(\hat{\mathbf{v}} \wedge \underline{\mathbf{e}}_0)^{2k+1} = \hat{\mathbf{v}} \wedge \underline{\mathbf{e}}_0, \qquad k = 0, 1, 2, \cdots,$$

we can rewrite Eq. (118b) in the exponential form

$$\begin{aligned}\mathsf{L}_0 &= \mathsf{I}_4 - (\hat{\mathbf{v}}\hat{\mathbf{v}} - \underline{\mathbf{e}}_0\underline{\mathbf{e}}_0)(1 - \cosh\varphi) + \hat{\mathbf{v}} \wedge \underline{\mathbf{e}}_0 \sinh\varphi \\ &= \cosh(\varphi\hat{\mathbf{v}} \wedge \underline{\mathbf{e}}_0) + \sinh(\varphi\hat{\mathbf{v}} \wedge \underline{\mathbf{e}}_0) \\ &= \exp(-\varphi\underline{\mathbf{e}}_0 \wedge \hat{\mathbf{v}}).\end{aligned} \tag{118c}$$

Similarly, for a pure rotation it can be easily verified that

$$\mathsf{L}_R = \hat{\mathbf{n}}\hat{\mathbf{n}} - \underline{\mathbf{e}}_0\underline{\mathbf{e}}_0 + (\mathsf{I}_3 - \hat{\mathbf{n}}\hat{\mathbf{n}}) \cos n - (\hat{\mathbf{n}} \times \mathsf{I}_3) \sin n, \tag{119a}$$

where $\hat{\mathbf{n}}$ is a unit vector along the direction of rotation of the reference frame and n is the angle of rotation. By observing that

$$(\hat{\mathbf{n}} \times \mathsf{I}_3)^{2k} = (-1)^k(\mathsf{I}_3 - \hat{\mathbf{n}}\hat{\mathbf{n}}), \quad k = 1, 2, \cdots,$$

and

$$(\hat{\mathbf{n}} \times \mathsf{I}_3)^{2k+1} = (-1)^k\hat{\mathbf{n}} \times \mathsf{I}_3, \qquad k = 0, 1, 2, \cdots,$$

Eq. (119a) can be written, in analogy to Eq. (118c), in the exponential form

$$\mathsf{L}_R = \exp(-\mathbf{n} \times \mathsf{I}_3). \tag{119b}$$

Equations (118c) and (119b) are particular cases of Lorentz transformations expressed as exponentials of an antisymmetric dyadic of a special form. Clearly the more general form

$$\mathsf{L} = \exp(\mathbf{a} \times \mathsf{I}_3 + \underline{\mathbf{e}}_0 \wedge \mathbf{b}), \tag{120}$$

where the exponent is an arbitrary antisymmetric dyadic, is also a Lorentz transformation. It will be shown later that any restricted homogeneous Lorentz transformation can be expressed in this way. Presently we shall consider the problem of expanding Eq. (120) into an explicit dyadic. For this purpose first note that, for arbitrary $\mathbf{q}$ and $\mathbf{q}'$,

$$\begin{aligned}(\mathbf{q} \times \mathsf{I}_3 - i\underline{\mathbf{e}}_0 \wedge \mathbf{q}) \cdot (\mathbf{q}' \times \mathsf{I}_3 + i\underline{\mathbf{e}}_0 \wedge \mathbf{q}') &= \mathbf{q}\mathbf{q}' + \mathbf{q}'\mathbf{q} - \mathbf{q} \cdot \mathbf{q}'\mathsf{I}_3 - i\mathbf{q} \times \mathbf{q}'\underline{\mathbf{e}}_0 \\ &\quad - i\underline{\mathbf{e}}_0\mathbf{q} \times \mathbf{q}' - \mathbf{q} \cdot \mathbf{q}'\underline{\mathbf{e}}_0\underline{\mathbf{e}}_0 \\ &= (\mathbf{q}' \times \mathsf{I}_3 + i\underline{\mathbf{e}}_0 \wedge \mathbf{q}') \cdot (\mathbf{q} \times \mathsf{I}_3 - i\underline{\mathbf{e}}_0 \wedge \mathbf{q}),\end{aligned}$$

i.e.,[16]

$$[\mathbf{q} \times \mathsf{I}_3 - i\underline{\mathbf{e}}_0 \wedge \mathbf{q}\,, \mathbf{q}' \times \mathsf{I}_3 + i\underline{\mathbf{e}}_0 \wedge \mathbf{q}'] = 0. \tag{121}$$

Therefore, by writing

$$\mathbf{a} \times \mathsf{I}_3 + \underline{\mathbf{e}}_0 \wedge \mathbf{b} = \mathsf{Q} + \mathsf{Q}^*, \tag{122}$$

where Q and Q^* commute and are defined by

$$\mathsf{Q} = \mathbf{p} \times \mathsf{I}_3 - i\underline{\mathbf{e}}_0 \wedge \mathbf{p} \tag{123}$$

and

$$\mathbf{p} = \tfrac{1}{2}(\mathbf{a} + i\mathbf{b}), \tag{124}$$

the exponential in Eq. (120) can be separated as

$$\mathsf{L} = \exp(\mathsf{Q} + \mathsf{Q}^*) = \exp \mathsf{Q} \cdot \exp \mathsf{Q}^*. \tag{125}$$

The preceding step has the particular advantage that the resulting exponentials can be easily expanded by making use of the relations

$$\mathsf{Q}^{2k} = (-1)^k p^{2k}\mathsf{I}_4, \quad k = 0, 1, \cdots,$$

and

$$\mathsf{Q}^{2k+1} = (-1)^k p^{2k}\mathsf{Q}, \quad k = 0, 1, \cdots,$$

where

$$\begin{aligned}p &= (\mathbf{p} \cdot \mathbf{p})^{\frac{1}{2}} = \tfrac{1}{2}(a^2 - b^2 + 2i\mathbf{a} \cdot \mathbf{b})^{\frac{1}{2}} = \rho e^{i\theta}, \\ \rho &= \tfrac{1}{2}[(a^2 - b^2)^2 + 4(\mathbf{a} \cdot \mathbf{b})^2]^{\frac{1}{4}}, \\ \theta &= \tfrac{1}{2} \tan^{-1}\left(\frac{2\mathbf{a} \cdot \mathbf{b}}{a^2 - b^2}\right).\end{aligned}$$

Thus

$$\exp \mathsf{Q} = \mathsf{I}_4 \cos p + \mathsf{Q}p^{-1} \sin p, \tag{126a}$$

and similarly,

$$\exp \mathsf{Q}^* = \mathsf{I}_4 \cos p^* + \mathsf{Q}^*(p^*)^{-1} \sin p^*. \tag{126b}$$

Note that in the special case $p = 0$, Eqs. (126a) and

---

[15] The kth power of a dyadic will be defined by

$$(\mathsf{A})^k \equiv \mathsf{A} \cdot (\mathsf{A})^{k-1}, \quad \text{and} \quad (\mathsf{A})^0 = \mathsf{I}_4.$$

[16] The commutator of two dyadics is defined by

$$[\mathsf{A}, \mathsf{B}] \equiv \mathsf{A} \cdot \mathsf{B} - \mathsf{B} \cdot \mathsf{A}.$$

(126b) are to be understood in the limiting sense as $p \to 0$, i.e.,

$$\exp \mathsf{Q}|_{p=0} = \lim_{p\to 0} \exp \mathsf{Q} = \mathsf{I}_4 + \mathsf{Q},$$
$$\exp \mathsf{Q}^*|_{p=0} = \lim_{p^*\to 0} \exp \mathsf{Q}^* = \mathsf{I}_4 + \mathsf{Q}^*. \quad (127)$$

Substitution of Eqs. (126a) and (126b) into (125) finally results in the following explicit dyadic form for L:

$$\begin{aligned}\mathsf{L} &= \mathsf{I}_4 \cos p \cos p^* + \mathsf{Q}p^{-1} \sin p \cos p^* \\ &\quad + \mathsf{Q}^*(p^*)^{-1} \sin p^* \cos p \\ &\quad + \mathsf{Q}\cdot\mathsf{Q}^*(pp^*)^{-1} \sin p \sin p^* \\ &= \mathsf{I}_4 \cos p \cos p^* + [\mathbf{p}p^{-1} \sin p \cos p^* \\ &\quad + \mathbf{p}^*(p^*)^{-1} \sin p^* \cos p] \times \mathsf{I}_3 \\ &\quad - i\underline{e}_0 \wedge [\mathbf{p}p^{-1} \sin p \cos p^* - \mathbf{p}^*(p^*)^{-1} \sin p^* \cos p] \\ &\quad + [\mathbf{p}^*\mathbf{p} + \mathbf{p}\mathbf{p}^* - \mathbf{p}\cdot\mathbf{p}^*(\mathsf{I}_3 + \underline{e}_0\underline{e}_0) \\ &\quad - i\mathbf{p}\times\mathbf{p}^*\underline{e}_0 - i\underline{e}_0\mathbf{p}\times\mathbf{p}^*](pp^*)^{-1} \sin p \sin p^*. \end{aligned} \quad (128)$$

Although the above expression is exhibited in terms of complex quantities, it is evident, from its form, that it is real.

Making use of some of the previous results it becomes rather straightforward to prove that an arbitrary restricted homogeneous Lorentz transformation can be written in the form of Eq. (120). As a first step we observe that L can be expressed as

$$\mathsf{L} = \mathsf{L}_0 \cdot \mathsf{L}_R = \exp(-\varphi \underline{e}_0 \wedge \hat{\mathbf{v}}) \cdot \exp(-\mathbf{n} \times \mathsf{I}_3). \quad (129)$$

(A simple proof of this statement is given in Appendix B.) Moreover, since Eq. (129) is a special case of the more general form

$$\mathsf{L} = \exp(\mathbf{a}_1 \times \mathsf{I}_3 + \underline{e}_0 \wedge \mathbf{b}_1) \cdot \exp(\mathbf{a}_2 \times \mathsf{I}_3 + \underline{e}_0 \wedge \mathbf{b}_2),$$

we consider the latter. Following Eq. (125) we write

$$\begin{aligned} L &= \exp(\mathsf{Q}_1 + \mathsf{Q}_1^*) \cdot \exp(\mathsf{Q}_2 + \mathsf{Q}_2^*) \\ &= \exp \mathsf{Q}_1 \cdot \exp \mathsf{Q}_2 \cdot \exp \mathsf{Q}_1^* \cdot \exp \mathsf{Q}_2^*, \end{aligned} \quad (130)$$

where Q_1, Q_2, $\mathbf{p}_1$, and $\mathbf{p}_2$ are defined in analogy to Eqs. (123) and (124) as

$$\mathsf{Q}_1 = \mathbf{p}_1 \times \mathsf{I}_3 - i\underline{e}_0 \wedge \mathbf{p}_1,$$

and

$$\mathsf{Q}_2 = \mathbf{p}_2 \times \mathsf{I}_3 - i\underline{e}_0 \wedge \mathbf{p}_2,$$
$$\mathbf{p}_1 = \tfrac{1}{2}(\mathbf{a}_1 + i\mathbf{b}_1), \quad \text{and} \quad \mathbf{p}_2 = \tfrac{1}{2}(\mathbf{a}_2 + i\mathbf{b}_2).$$

To complete the proof it is sufficient to show that

$$\exp \mathsf{Q}_1 \cdot \exp \mathsf{Q}_2$$

can be written in the form $\tau \exp \mathsf{Q}$ with

$$\mathsf{Q} = \mathbf{p} \times \mathsf{I}_3 - i\underline{e}_0 \wedge \mathbf{p},$$

and τ equal to either 1 or -1. Equation (126a) is used to expand $\exp \mathsf{Q}_1 \cdot \exp \mathsf{Q}_2$ as

$$\begin{aligned}\exp \mathsf{Q}_1 \cdot \exp \mathsf{Q}_2 &= (\mathsf{I}_4 \cos p_1 + \mathsf{Q}_1 p_1^{-1} \sin p_1) \\ &\quad \cdot (\mathsf{I}_4 \cos p_2 + \mathsf{Q}_2 p_2^{-1} \sin p_2) \\ &= \mathsf{I}_4 s + \mathbf{t} \times \mathsf{I}_3 - i\underline{e}_0 \wedge \mathbf{t}, \end{aligned} \quad (131)$$

where

$$s = \cos p_1 \cos p_2 - \mathbf{p}_1 \cdot \mathbf{p}_2 (p_1 p_2)^{-1} \sin p_1 \sin p_2 \quad (132)$$

and

$$\begin{aligned}\mathbf{t} &= \mathbf{p}_1(p_1)^{-1} \sin p_1 \cos p_2 + \mathbf{p}_2(p_2)^{-1} \sin p_2 \cos p_1 \\ &\quad + \mathbf{p}_1 \times \mathbf{p}_2 (p_1p_2)^{-1} \sin p_1 \sin p_2. \end{aligned} \quad (133)$$

We now look for a vector $\mathbf{p}$ such that the equations

$$\mathbf{p}p^{-1} \sin p = \tau \mathbf{t} \quad (134)$$

and

$$\cos p = \tau s \quad (135a)$$

are simultaneously fulfilled for one of the possible values of τ. Solving Eq. (134) for $\mathbf{p}$ yields

$$\mathbf{p} = \tau p (\sin p)^{-1} \mathbf{t}$$

with

$$p = \begin{cases} \sin^{-1}[(\mathbf{t}\cdot\mathbf{t})^{\frac{1}{2}}] & \text{for} \quad \mathbf{t}\cdot\mathbf{t} \neq 0, \\ 0 & \text{for} \quad \mathbf{t}\cdot\mathbf{t} = 0. \end{cases}$$

It should be noted that in the case $\mathbf{t}\cdot\mathbf{t} \neq 0$ any one of the multiple values of p is permissible. In either case, in order to determine τ, observe that a straightforward calculation from Eqs. (132) and (133) yields

$$1 - \mathbf{t}\cdot\mathbf{t} = s^2.$$

Consequently, by Eq. (134) we have

$$\cos^2 p = 1 - \mathbf{t}\cdot\mathbf{t} = s^2,$$

that is,

$$\cos p = \pm s \quad (135b)$$

or

$$\cos p = \tau s.$$

Thus τ is defined by the sign on the right-hand side of Eq. (135b). Substitution for s and $\mathbf{t}$ from Eqs. (134) and (135a) into (131) results in

$$\begin{aligned}\exp \mathsf{Q}_1 \cdot \exp \mathsf{Q}_2 &= \mathsf{I}_4 \tau \cos p \\ &\quad + \tau(\mathbf{p} \times \mathsf{I}_3 - i\underline{e}_0 \wedge \mathbf{p})p^{-1} \sin p \\ &= \tau(\mathsf{I}_4 \cos p + \mathsf{Q}p^{-1} \sin p) \\ &= \tau \exp \mathsf{Q}. \end{aligned} \quad (136a)$$

By complex conjugation we also have

$$\exp \mathsf{Q}_1^* \cdot \exp \mathsf{Q}_2^* = \tau \exp \mathsf{Q}^*. \quad (136b)$$

Finally, inserting these results into Eq. (130) gives

$$\mathsf{L} = (\tau \exp \mathsf{Q}) \cdot (\tau \exp \mathsf{Q}^*) = \exp(\mathsf{Q} + \mathsf{Q}^*). \quad \text{Q.E.D.}$$

F. Eigenvalues and Eigenvectors of Restricted Homogeneous Lorentz Transformations

We use the decomposition in Eq. (125):

$$\mathsf{L} = \exp(\mathbf{a} \times \mathsf{I}_3 + \underline{\mathbf{e}}_0 \wedge \mathbf{b}) = \exp \mathsf{Q} \cdot \exp \mathsf{Q}^*,$$

where Q and Q* are given by Eq. (123). First there is the trivial case

$$\mathbf{a} = \mathbf{b} = 0$$

in which every vector is an eigenvector of L with eigenvalue $\lambda = 1$. The other two cases follow.

Case I: $p \neq 0$. The fact that Q and Q* commute suggests looking for simultaneous eigenvectors of Q and Q*, which, of course, will also be eigenvectors of L. First let us define the real unit vector $\hat{\mathbf{m}}$ as follows:

(a) If $\hat{\mathbf{p}} \times \hat{\mathbf{p}}^* \neq 0$ (i.e., $\mathbf{a} \times \mathbf{b} \neq 0$) where $\hat{\mathbf{p}} \equiv p^{-1}\mathbf{p}$, then

$$\hat{\mathbf{m}} = i[(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^*)^2 - 1]^{-\frac{1}{2}}(\hat{\mathbf{p}} \times \hat{\mathbf{p}}^*) = [a^2b^2 - (\mathbf{a} \cdot \mathbf{b})^2]^{-\frac{1}{2}}(\mathbf{a} \times \mathbf{b}). \quad (137)$$

Clearly $\hat{\mathbf{m}}$ satisfies the orthogonality conditions

$$\hat{\mathbf{m}} \cdot \hat{\mathbf{p}} = \hat{\mathbf{m}} \cdot \hat{\mathbf{p}}^* = \hat{\mathbf{m}} \cdot \underline{\mathbf{e}}_0 = 0. \quad (138)$$

(b) If $\hat{\mathbf{p}} \times \hat{\mathbf{p}}^* = 0$ (i.e., $\mathbf{a} \times \mathbf{b} = 0$), then let $\hat{\mathbf{m}}$ be any real unit vector satisfying Eq. (138). We now show that the vectors

$$\underline{\mathbf{w}}_{1,s} = \underline{\mathbf{e}}_0 + s\hat{\mathbf{p}}, \quad (139)$$

$$\underline{\mathbf{w}}_{2,s} = \hat{\mathbf{m}} + is\hat{\mathbf{p}} \times \hat{\mathbf{m}}, \quad (140)$$

where $s = \pm 1$, are eigenvectors of Q with eigenvalues $\mu_s = -isp$:

$$\begin{aligned}\mathsf{Q} \cdot \underline{\mathbf{w}}_{1,s} &= p(\hat{\mathbf{p}} \times \mathsf{I}_3 - i\underline{\mathbf{e}}_0 \wedge \hat{\mathbf{p}}) \cdot (\underline{\mathbf{e}}_0 + s\hat{\mathbf{p}}) \\ &= p(i\hat{\mathbf{p}}\underline{\mathbf{e}}_0 \cdot \underline{\mathbf{e}}_0 - i\underline{\mathbf{e}}_0 s\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}) = -isp(\underline{\mathbf{e}}_0 + s\hat{\mathbf{p}}) \\ &= \mu_s \underline{\mathbf{w}}_{1,s},\end{aligned}$$

$$\begin{aligned}\mathsf{Q} \cdot \underline{\mathbf{w}}_{2,s} &= p(\hat{\mathbf{p}} \times \mathsf{I}_3 - i\underline{\mathbf{e}}_0 \wedge \hat{\mathbf{p}}) \cdot (\hat{\mathbf{m}} + is\hat{\mathbf{p}} \times \hat{\mathbf{m}}) \\ &= p(\hat{\mathbf{p}} \times \hat{\mathbf{m}} - is\hat{\mathbf{m}}) = -isp(\hat{\mathbf{m}} + is\hat{\mathbf{p}} \times \hat{\mathbf{m}}) \\ &= \mu_s \underline{\mathbf{w}}_{2,s}.\end{aligned}$$

Taking complex conjugates gives

$$\mathsf{Q}^* \cdot \underline{\mathbf{w}}_{1,s}^* = isp^*\underline{\mathbf{w}}_{1,s}^* = \mu_s^*\underline{\mathbf{w}}_{1,s}^*,$$
$$\mathsf{Q}^* \cdot \underline{\mathbf{w}}_{2,s}^* = isp^*\underline{\mathbf{w}}_{2,s}^* = \mu_s^*\underline{\mathbf{w}}_{2,s}^*,$$

i.e., $\underline{\mathbf{w}}_{1,s}^*$ and $\underline{\mathbf{w}}_{2,s}^*$ are eigenvectors of Q* with eigenvalues $\mu_s^* = isp^*$.

For given values of s and s', we write the equation

$$\alpha_1\underline{\mathbf{w}}_{1,s} + \alpha_2\underline{\mathbf{w}}_{2,s} = \beta_1\underline{\mathbf{w}}_{1,s'}^* + \beta_2\underline{\mathbf{w}}_{2,s'}^*,$$

which is to be solved for α_1, α_2, β_1, and β_2. If there exists a nontrivial solution ($\alpha_1, \alpha_2, \beta_1, \beta_2$ not all zero), then the vector

$$\underline{\mathbf{u}}_{s,s'} = \alpha_1\underline{\mathbf{w}}_{1,s} + \alpha_2\underline{\mathbf{w}}_{2,s} = \beta_1\underline{\mathbf{w}}_{1,s'}^* + \beta_2\underline{\mathbf{w}}_{2,s'}^* \quad (141)$$

is a simultaneous eigenvector of Q and Q* with eigenvalues μ_s and $\mu_{s'}^*$, respectively, i.e.,

$$\begin{aligned}\mathsf{Q} \cdot \underline{\mathbf{u}}_{s,s'} &= \mu_s\underline{\mathbf{u}}_{s,s'}, \\ \mathsf{Q}^* \cdot \underline{\mathbf{u}}_{s,s'} &= \mu_{s'}^*\underline{\mathbf{u}}_{s,s'}.\end{aligned} \quad (142)$$

Substitution from Eqs. (139) and (140) into (141) gives

$$\alpha_1(\underline{\mathbf{e}}_0 + s\hat{\mathbf{p}}) + \alpha_2(\hat{\mathbf{m}} + is\hat{\mathbf{p}} \times \hat{\mathbf{m}}) = \beta_1(\underline{\mathbf{e}}_0 + s'\hat{\mathbf{p}}^*) + \beta_2(\hat{\mathbf{m}} - is'\hat{\mathbf{p}}^* \times \hat{\mathbf{m}}). \quad (143a)$$

Taking the "dot" product of Eq. (143a) with $(-1)\underline{\mathbf{e}}_0$ gives

$$\alpha_1 = \beta_1,$$

and the "dot" product with $\hat{\mathbf{m}}$ gives

$$\alpha_2 = \beta_2.$$

Substitution of these last results into (143a) gives

$$\alpha_1(s\hat{\mathbf{p}} - s'\hat{\mathbf{p}}^*) + \alpha_2 i(s\hat{\mathbf{p}} + s'\hat{\mathbf{p}}^*) \times \hat{\mathbf{m}} = 0. \quad (143b)$$

The vectors $s\hat{\mathbf{p}} - s'\hat{\mathbf{p}}^*$ and $(s\hat{\mathbf{p}} + s'\hat{\mathbf{p}}^*) \times \hat{\mathbf{m}}$ in the above equation are parallel, as can be seen from the following:

$$\begin{aligned}(s\hat{\mathbf{p}} - s'\hat{\mathbf{p}}^*) \times [(s\hat{\mathbf{p}} + s'\hat{\mathbf{p}}^*) \times \hat{\mathbf{m}}] &= (s\hat{\mathbf{p}} + s'\hat{\mathbf{p}}^*)(s\hat{\mathbf{p}} - s'\hat{\mathbf{p}}^*) \cdot \hat{\mathbf{m}} \\ &\quad - (s\hat{\mathbf{p}} - s'\hat{\mathbf{p}}^*) \cdot (s\hat{\mathbf{p}} + s'\hat{\mathbf{p}}^*)\hat{\mathbf{m}}, \\ &= -(s^2 - s'^2)\hat{\mathbf{m}} = 0.\end{aligned}$$

Thus there will be a solution of Eq. (143b) where α_1 and α_2 are not both zero. To obtain the solution, first take the "dot" product of Eq. (143b) with $s\hat{\mathbf{p}}^*$ to give

$$\alpha_1(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* - ss') - \alpha_2 i\hat{\mathbf{p}} \times \hat{\mathbf{p}}^* \cdot \hat{\mathbf{m}} = 0, \quad (144a)$$

and the "dot" product with $s\hat{\mathbf{p}}^* \times \hat{\mathbf{m}}$ to give

$$\alpha_1\hat{\mathbf{p}} \times \hat{\mathbf{p}}^* \cdot \hat{\mathbf{m}} + \alpha_2 i(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* + ss') = 0. \quad (145a)$$

Note that

$$\hat{\mathbf{p}} \times \hat{\mathbf{p}}^* \cdot \hat{\mathbf{m}} = -i[(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^*)^2 - 1]^{\frac{1}{2}} \quad (146)$$

is true whether $\hat{\mathbf{p}}$ is parallel to $\hat{\mathbf{p}}^*$ or not. Substitution from Eq. (146) into (144a) and (145a) gives

$$\alpha_1(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* - ss') - \alpha_2[(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^*)^2 - 1]^{\frac{1}{2}} = 0 \quad (144b)$$

and

$$-\alpha_1[(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^*)^2 - 1]^{\frac{1}{2}} + \alpha_2(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* + ss') = 0, \quad (145b)$$

respectively. If $ss' = -1$, then $\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* - ss' \neq 0$, and division of Eq. (144b) by $(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* - ss')^{\frac{1}{2}}$ results in

$$\alpha_1(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* - ss')^{\frac{1}{2}} - \alpha_2(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* + ss')^{\frac{1}{2}} = 0. \quad (147)$$

If $ss' = 1$, then $\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* + ss' \neq 0$, and division of Eq. (145b) by $(-1)(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* + ss')^{\frac{1}{2}}$ again yields Eq. (147). From (147) we obtain

$$\alpha_1 = \alpha(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* + ss')^{\frac{1}{2}},$$
$$\alpha_2 = \alpha(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* - ss')^{\frac{1}{2}}, \tag{148}$$

where α is arbitrary. Substitution for α_1 and α_2 in Eq. (141) gives

$$\begin{aligned}\underline{\mathbf{u}}_{s,s'} &= \alpha[(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* + ss')^{\frac{1}{2}}\underline{\mathbf{w}}_{1,s} + (\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* - ss')^{\frac{1}{2}}\underline{\mathbf{w}}_{2,s}] \\ &= \tfrac{1}{2}(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* + ss')^{\frac{1}{2}}(\underline{\mathbf{e}}_0 + s\hat{\mathbf{p}}) \\ &\quad + \tfrac{1}{2}(\hat{\mathbf{p}} \cdot \hat{\mathbf{p}}^* - ss')^{\frac{1}{2}}(\hat{\mathbf{m}} + is\hat{\mathbf{p}} \times \hat{\mathbf{m}}),\end{aligned} \tag{149}$$

where we have put $\alpha = \frac{1}{2}$ to normalize $\underline{\mathbf{u}}_{s,s'}$ according to

$$\underline{\mathbf{u}}_{1,1} \cdot \underline{\mathbf{u}}_{-1,-1} = -1,$$
$$\underline{\mathbf{u}}_{1,-1} \cdot \underline{\mathbf{u}}_{-1,1} = 1.$$

Note that, by Eqs. (142), we have

$$(\mathsf{Q} + \mathsf{Q}^*) \cdot \underline{\mathbf{u}}_{s,s'} = (-isp + is'p^*)\underline{\mathbf{u}}_{s,s'}.$$

Consequently,

$$\mathsf{L} \cdot \underline{\mathbf{u}}_{s,s'} = (\exp \mathsf{Q} \cdot \exp \mathsf{Q}^*) \cdot \underline{\mathbf{u}}_{s,s'} = \lambda_{s,s'}\underline{\mathbf{u}}_{s,s'}, \tag{150}$$

where

$$\lambda_{s,s'} = \exp[-i(sp - s'p^*)]. \tag{151}$$

The vectors $\underline{\mathbf{u}}_{s,s'}$, which are null and satisfy the orthonormality properties

$$\underline{\mathbf{u}}_{s,s'} \cdot \underline{\mathbf{u}}_{s'',s'''} = -\tfrac{1}{4}(s - s'')(s' - s'''), \tag{152}$$

can be used to construct the unit dyadic

$$\mathsf{I}_4 = -\sum_{s,s'} ss'\underline{\mathbf{u}}_{s,s'}\underline{\mathbf{u}}_{-s,-s'}, \tag{153}$$

from which it follows immediately that

$$\mathsf{L} = \mathsf{L} \cdot \mathsf{I}_4 = -\sum_{s,s'} ss'\lambda_{s,s'}\underline{\mathbf{u}}_{s,s'}\underline{\mathbf{u}}_{-s,-s'}. \tag{154}$$

Thus we have an alternate explicit expression for the restricted homogeneous Lorentz transformation.[17]

Case II: $p = 0$, $\mathbf{a}$ and $\mathbf{b}$ not both zero. In this case, all eigenvalues of $\mathsf{Q} + \mathsf{Q}^*$ are zero. There will be only two possible eigenvectors[18] $\underline{\mathbf{v}}_e'$ and $\underline{\mathbf{v}}_1'$:

$$(\mathsf{Q} + \mathsf{Q}^*) \cdot \underline{\mathbf{v}}_e' = 0, \tag{155a}$$
$$(\mathsf{Q} + \mathsf{Q}^*) \cdot \underline{\mathbf{v}}_1' = 0. \tag{155b}$$

One eigenvector $\underline{\mathbf{v}}_e'$ is null and may be obtained from

$$\begin{aligned}\underline{\mathbf{v}}_e' &= \lim_{p\to 0}\left(\frac{4pp^*}{\mathbf{p}\cdot\mathbf{p}^*}\right)^{\frac{1}{2}}\underline{\mathbf{u}}_{s,s'} \\ &= \lim_{p\to 0}(\mathbf{p}\cdot\mathbf{p}^*)^{-\frac{1}{2}}[(\mathbf{p}\cdot\mathbf{p}^* + ss'pp^*)^{\frac{1}{2}}(\underline{\mathbf{e}}_0 + sp^{-1}\mathbf{p}) \\ &\quad + (\mathbf{p}\cdot\mathbf{p}^* - ss'pp^*)^{\frac{1}{2}}(\hat{\mathbf{m}} + isp^{-1}\mathbf{p}\times\hat{\mathbf{m}})] \\ &= \underline{\mathbf{e}}_0 + i(\mathbf{p}\cdot\mathbf{p}^*)^{-1}(\mathbf{p}\times\mathbf{p}^*).\end{aligned} \tag{156}$$

In the calculation of this limit, use was made of the fact that $\mathbf{p} \times \mathbf{p}^* \neq 0$ in this case and that

$$\lim_{p\to 0}\hat{\mathbf{m}} = \frac{i\mathbf{p}\times\mathbf{p}^*}{[(\mathbf{p}\cdot\mathbf{p}^*)^2 - p^2p^{*2}]^{\frac{1}{2}}}\bigg|_{p=0} = \frac{i\mathbf{p}\times\mathbf{p}^*}{\mathbf{p}\cdot\mathbf{p}^*},$$

$$\begin{aligned}\lim_{p\to 0}\mathbf{p}\times\hat{\mathbf{m}} &= \mathbf{p}\times\left(\frac{i\mathbf{p}\times\mathbf{p}^*}{\mathbf{p}\cdot\mathbf{p}^*}\right)\bigg|_{p=0} \\ &= \frac{i(\mathbf{p}\cdot\mathbf{p}^*)\mathbf{p} - ip^2\mathbf{p}^*}{\mathbf{p}\cdot\mathbf{p}^*}\bigg|_{p=0} = i\mathbf{p},\end{aligned}$$

$$\lim_{p\to 0}\frac{1}{p}[(\mathbf{p}\cdot\mathbf{p}^* + ss'pp^*)^{\frac{1}{2}} - (\mathbf{p}\cdot\mathbf{p}^* - ss'pp^*)^{\frac{1}{2}}] = 0.$$

The other eigenvector $\underline{\mathbf{v}}_1'$ is found by inspection and is

$$\underline{\mathbf{v}}_1' = \mathbf{p} + \mathbf{p}^*. \tag{157}$$

It is spacelike and orthogonal to $\underline{\mathbf{v}}_e$.

In order to construct a basis we look for two generalized eigenvectors[19] $\underline{\mathbf{w}}_e$ and $\underline{\mathbf{z}}_e$, which, together with $\underline{\mathbf{v}}_e'$ and $\underline{\mathbf{v}}_1'$, form a linearly independent set, i.e., we require that

$$(\mathsf{Q} + \mathsf{Q}^*)^n \cdot \underline{\mathbf{w}}_e' = 0 \tag{158}$$

and

$$(\mathsf{Q} + \mathsf{Q}^*)^m \cdot \underline{\mathbf{z}}_e' = 0 \tag{159}$$

(where m and $n = 2$ or 3). By inspection it is found that, for the choice

$$\underline{\mathbf{w}}_e' = -i(2\mathbf{p}\cdot\mathbf{p}^*)^{-1}(\mathbf{p} - \mathbf{p}^*),$$
$$\underline{\mathbf{z}}_e' = (2\mathbf{p}\cdot\mathbf{p}^*)^{-1}\underline{\mathbf{e}}_0, \tag{160}$$

one has

$$(\mathsf{Q} + \mathsf{Q}^*) \cdot \underline{\mathbf{w}}_e' = \underline{\mathbf{v}}_e' \tag{161}$$

and

$$(\mathsf{Q} + \mathsf{Q}^*) \cdot \underline{\mathbf{z}}_e' = \underline{\mathbf{w}}_e'. \tag{162}$$

Consequently, it follows that

$$(\mathsf{Q} + \mathsf{Q}^*)^2 \cdot \underline{\mathbf{w}}_e' = 0$$

and

$$(\mathsf{Q} + \mathsf{Q}^*)^3 \cdot \underline{\mathbf{z}}_e' = 0.$$

[17] Other polynomial forms for finite restricted homogeneous Lorentz transformations are discussed by S. L. Bazanski, J. Math. Phys. 6, 1201 (1965).

[18] E. Wigner, Ann. Math. 40, 149 (1939).

[19] B. Friedman, *Principles and Techniques of Applied Mathematics* (John Wiley & Sons, Inc., New York, 1956).

Hence $\underline{w}_e'$ and $\underline{z}_e'$ are the desired generalized eigenvectors. Note that if we replace $\underline{z}_e'$, $\underline{w}_e'$, $\underline{v}_1'$, $\underline{v}_e'$ (respectively) by

$$\underline{z}_e = \rho_1 \underline{z}_e' + \gamma_1 \underline{w}_e' + \gamma_2 \underline{v}_1' + \gamma_3 \underline{v}_e', \tag{163a}$$

$$\underline{w}_e = \rho_1 \underline{w}_e' + \gamma_1 \underline{v}_e', \tag{163b}$$

$$\underline{v}_1 = \rho_2 \underline{v}_1' + \gamma_4 \underline{v}_e' \tag{163c}$$

$$\underline{v}_e = \rho_1 \underline{v}_e', \tag{163d}$$

for $\rho_1 \neq 0$, $\rho_2 \neq 0$, then Eqs. (155a), (155b), (161), and (162) remain invariant. In particular, by selecting

$$\rho_1 = (2\mathbf{p} \cdot \mathbf{p}^*)^{\frac{1}{2}}, \quad \rho_2 = 1/\rho_1, \quad \gamma_3 = -(2\rho_1)^{-1},$$

and

$$\gamma_1 = \gamma_2 = \gamma_4 = 0,$$

the above vectors will be normalized according to

$$\begin{aligned} &\underline{v}_1 \cdot \underline{v}_1 = \underline{w}_e \cdot \underline{w}_e = 1, \\ &\underline{v}_e \cdot \underline{z}_e = -1, \\ &\underline{z}_e \cdot \underline{z}_e = \underline{v}_e \cdot \underline{v}_e = 0, \\ &\underline{v}_e \cdot \underline{v}_1 = \underline{v}_e \cdot \underline{w}_e = \underline{v}_1 \cdot \underline{w}_e = \underline{v}_1 \cdot \underline{z}_e = \underline{w}_e \cdot \underline{z}_e = 0. \end{aligned} \tag{164}$$

Also observe that, by exponentiation,

$$\begin{aligned} \mathsf{L} \cdot \underline{v}_1 &= \exp(\mathsf{Q} + \mathsf{Q}^*) \cdot \underline{v}_1 = \underline{v}_1, \\ \mathsf{L} \cdot \underline{v}_e &= \exp(\mathsf{Q} + \mathsf{Q}^*) \cdot \underline{v}_e = \underline{v}_e, \\ \mathsf{L} \cdot \underline{w}_e &= \exp(\mathsf{Q} + \mathsf{Q}^*) \cdot \underline{w}_e = [\mathsf{I}_4 + (\mathsf{Q} + \mathsf{Q}^*)] \cdot \underline{w}_e \\ &= \underline{w}_e + \underline{v}_e, \qquad (165) \\ \mathsf{L} \cdot \underline{z}_e &= \exp(\mathsf{Q} + \mathsf{Q}^*) \cdot \underline{z}_e = [\mathsf{I}_4 + (\mathsf{Q} + \mathsf{Q}^*) \\ &\quad + \tfrac{1}{2}(\mathsf{Q} + \mathsf{Q}^*)^2] \cdot \underline{z}_e \\ &= \underline{z}_e + \underline{w}_e + \tfrac{1}{2}\underline{v}_e. \end{aligned}$$

The unit dyadic in this case is

$$\mathsf{I}_4 = \underline{v}_1\underline{v}_1 + \underline{w}_e\underline{w}_e - \underline{v}_e\underline{z}_e - \underline{z}_e\underline{v}_e, \tag{166}$$

and

$$\begin{aligned} \mathsf{L} = \mathsf{L} \cdot \mathsf{I}_4 &= \underline{v}_1\underline{v}_1 + (\underline{w}_e + \underline{v}_e)\underline{w}_e - \underline{v}_e\underline{z}_e \\ &\quad - (\underline{z}_e + \underline{w}_e + \tfrac{1}{2}\underline{v}_e)\underline{v}_e \\ &= \underline{v}_1\underline{v}_1 + \underline{w}_e\underline{w}_e - \tfrac{1}{2}\underline{v}_e\underline{v}_e + \underline{v}_e \wedge \underline{w}_e \\ &\quad - \underline{v}_e\underline{z}_e - \underline{z}_e\underline{v}_e. \end{aligned} \tag{167}$$

G. Three-Dimensional Complex Orthogonal Representation of the Restricted Homogeneous Lorentz Group

Making use of the transformation properties of the dyadic G defined by Eq. (103a) together with Eqs. (121) and (125), we show how an isomorphism between the three-dimensional complex orthogonal group and the restricted homogeneous Lorentz group can be established by a simple and direct procedure. Let L be an arbitrary restricted homogeneous Lorentz transformation. The dyadic G transforms under L according to

$$\mathsf{G}' = \mathsf{L} \cdot \mathsf{G} \cdot \tilde{\mathsf{L}}, \tag{168}$$

which by Eq. (125), can be expressed as

$$\mathsf{G}' = (\exp \mathsf{Q} \cdot \exp \mathsf{Q}^*) \cdot \mathsf{G} \cdot \exp(-\mathsf{Q}^*) \cdot \exp(-\mathsf{Q}). \tag{169}$$

Recalling the defining equations (103a) and (123) and the commutation relation (121), we see that

$$[\mathsf{Q}^*, \mathsf{G}] = 0.$$

and therefore

$$\mathsf{G}' = (\exp \mathsf{Q}) \cdot \mathsf{G} \cdot \exp(-\mathsf{Q}). \tag{170}$$

We note also that the dyadic

$$\mathsf{S} = \mathbf{p} \times \mathsf{I}_3 + i\underline{e}_0 \wedge \mathbf{p}$$

satisfies the commutation relations

$$[\mathsf{G}, \mathsf{S}] = [\mathsf{Q}, \mathsf{S}] = 0,$$

and we write

$$\begin{aligned} \mathsf{G}' &= (\exp \mathsf{Q} \cdot \exp \mathsf{S}) \cdot \mathsf{G} \cdot \exp(-\mathsf{S}) \cdot \exp(-\mathsf{Q}) \\ &= \exp(\mathsf{Q} + \mathsf{S}) \cdot \mathsf{G} \cdot \exp[-(\mathsf{S} + \mathsf{Q})] \\ &= \exp(2\mathbf{p} \times \mathsf{I}_3) \cdot \mathsf{G} \cdot \exp(-2\mathbf{p} \times \mathsf{I}_3), \\ &= \exp(2\mathbf{p} \times \mathsf{I}_3) \cdot (i\mathbf{Z} \times \mathsf{I}_3 + \underline{e}_0 \wedge \mathbf{Z}) \\ &\quad \cdot \exp(-2\mathbf{p} \times \mathsf{I}_3). \end{aligned} \tag{171}$$

Now observing that

$$\exp(2\mathbf{p} \times \mathsf{I}_3) \cdot \underline{e}_0 = \underline{e}_0 \cdot \exp(-2\mathbf{p} \times \mathsf{I}_3) = \underline{e}_0$$

and

$$\begin{aligned} &\exp(2\mathbf{p} \times \mathsf{I}_3) \cdot \mathbf{Z} \times \mathsf{I}_3 \cdot \exp(-2\mathbf{p} \times \mathsf{I}_3) \\ &\quad = \exp(2\mathbf{p} \times \mathsf{I}_3) \cdot (\mathbf{Z} \times \underline{e}_k)\underline{e}_k \cdot \exp(-2\mathbf{p} \times \mathsf{I}_3) \\ &\quad = [\exp(2\mathbf{p} \times \mathsf{I}_3) \cdot \mathbf{Z}] \times [\exp(2\mathbf{p} \times \mathsf{I}_3) \cdot \underline{e}_k]\underline{e}_k \\ &\qquad \cdot \exp(-2\mathbf{p} \times \mathsf{I}_3) \\ &\quad = \mathbf{Z}' \times [\exp(2\mathbf{p} \times \mathsf{I}_3) \cdot \mathsf{I}_3 \cdot \exp(-2\mathbf{p} \times \mathsf{I}_3)] \\ &\quad = \mathbf{Z}' \times \mathsf{I}_3 \end{aligned}$$

where

$$\mathbf{Z}' = \exp(2\mathbf{p} \times \mathsf{I}_3) \cdot \mathbf{Z}, \tag{172}$$

Eq. (171) becomes

$$\mathsf{G}' = i\mathbf{Z}' \times \mathsf{I}_3 + \underline{e}_0 \wedge \mathbf{Z}'. \tag{173}$$

In summary, we have

$$\mathsf{G}' = \mathsf{L} \cdot \mathsf{G} \cdot \tilde{\mathsf{L}} = \mathsf{L} \cdot (\mathbf{\Sigma} \cdot \mathbf{Z}) \cdot \tilde{\mathsf{L}} = \mathbf{\Sigma} \cdot \mathbf{Z}', \tag{174}$$

where $\mathbf{\Sigma}$ was defined in Eq. (106) and

$$\mathsf{L} = \exp(\mathbf{a} \times \mathsf{I}_3 + \underline{e}_0 \wedge \mathbf{b}) \leftrightarrow \exp(2\mathbf{p} \times \mathsf{I}_3)$$
$$= \exp[(\mathbf{a} + i\mathbf{b}) \times \mathsf{I}_3] \tag{175}$$

is the desired isomorphism.

ACKNOWLEDGMENT

The authors wish to express their appreciation to General Electric Company, and in particular to Dr. Roy W. Hendrick, Jr., for their interest and support of this work.

APPENDIX A

In this Appendix we derive explicitly the dual of the antisymmetric dyadics $\underline{e}_0 \wedge \mathbf{u}$ and $\mathbf{u} \times \mathsf{I}_3$. For this purpose we first prove that $\mathbf{\Gamma}$, as defined in Eq. (50), is invariant under rotation of the orthonormal basis vectors $\underline{e}_k$ in $\mathcal{E}_3$. Thus, if $\underline{e}_k$ is rotated into $\underline{e}_k'$, we have

$$\mathbf{\Gamma}' = \underline{e}_1' \wedge \underline{e}_2' \wedge \underline{e}_3' \wedge \underline{e}_0. \tag{A1}$$

But

$$\underline{e}_1' \wedge \underline{e}_2' \wedge \underline{e}_3' = \epsilon_{klm}\underline{e}_k'\underline{e}_l'\underline{e}_m' = \underline{e}_k' \times \underline{e}_l' \cdot \underline{e}_m'\underline{e}_k'\underline{e}_l'\underline{e}_m'$$
$$= -\underline{e}_k'(\underline{e}_k' \times \underline{e}_m' \cdot \underline{e}_l')\underline{e}_l'\underline{e}_m' = -\underline{e}_k'\underline{e}_k' \times \underline{e}_m'\underline{e}_m'$$
$$= -\mathsf{I}_3 \times \mathsf{I}_3 = \underline{e}_1 \wedge \underline{e}_2 \wedge \underline{e}_3.$$

Hence

$$\mathbf{\Gamma}' = \mathbf{\Gamma}. \tag{A2}$$

Now making use of the definition of the dual in Eq. (53) and choosing a rotated basis with $\underline{e}_3'$ in the direction of $\mathbf{u}$ results in

$$^{\star}(\underline{e}_0 \wedge \mathbf{u}) = \tfrac{1}{2}\mathbf{\Gamma}:(\underline{e}_0 \wedge \mathbf{u})$$
$$= \tfrac{1}{2}u(\underline{e}_1' \wedge \underline{e}_2' \wedge \underline{e}_3' \wedge \underline{e}_0):(\underline{e}_0 \wedge \underline{e}_3')$$
$$= -u(\underline{e}_1' \wedge \underline{e}_2' \wedge \underline{e}_3' \wedge \underline{e}_0):(\underline{e}_3'\underline{e}_0)$$
$$= u\underline{e}_1' \wedge \underline{e}_2' = -u(\underline{e}_1' \times \underline{e}_2') \times \mathsf{I}_3$$
$$= -u\underline{e}_3' \times \mathsf{I}_3 = -\mathbf{u} \times \mathsf{I}_3. \tag{A3}$$

Similarly,

$$^{\star}(\mathbf{u} \times \mathsf{I}_3) = \tfrac{1}{2}\mathbf{\Gamma}:(\mathbf{u} \times \mathsf{I}_3) = \tfrac{1}{2}u\mathbf{\Gamma}:[(\underline{e}_1' \times \underline{e}_2') \times \mathsf{I}_3]$$
$$= -\tfrac{1}{2}u(\underline{e}_1' \wedge \underline{e}_2' \wedge \underline{e}_3' \wedge \underline{e}_0):(\underline{e}_1' \wedge \underline{e}_2')$$
$$= -u(\underline{e}_3' \wedge \underline{e}_0 \wedge \underline{e}_1' \wedge \underline{e}_2'):(\underline{e}_1'\underline{e}_2')$$
$$= -u\underline{e}_3' \wedge \underline{e}_0 = \underline{e}_0 \wedge \mathbf{u}. \tag{A4}$$

APPENDIX B

Here we show that an arbitrary restricted Lorentz transformation can be expressed as a product of a pure Lorentz transformation L_0 and a pure rotation L_R. To this end let

$$\underline{g} = \mathsf{L} \cdot \underline{e}_0. \tag{B1}$$

Since $\underline{g}$ is a unit timelike four-vector, it must be of the form

$$\underline{g} = (1 - u^2/c^2)^{-\frac{1}{2}}(\underline{e}_0 + \mathbf{u}/c), \tag{B2}$$

where $\mathbf{u}$ is defined by

$$\mathbf{u} = -c[(\underline{e}_0 \cdot \underline{g})^{-1}\underline{g} + \underline{e}_0].$$

By introducing the unit vector

$$\underline{y} = (1 - u^2/c^2)^{-\frac{1}{2}}[\hat{\mathbf{u}} + (u/c)\underline{e}_0]$$

which is in the plane of $\underline{e}_0$, $\mathbf{u}$ and is orthogonal to $\underline{g}$, we can construct, as before, the pure Lorentz transformation

$$\mathsf{L}_0 = \mathsf{I}_3 - \hat{\mathbf{u}}\hat{\mathbf{u}} - \underline{e}_0\underline{g} + \hat{\mathbf{u}}\underline{y}. \tag{B3}$$

Clearly

$$\mathsf{L}_0 \cdot \underline{g} = \underline{e}_0. \tag{B4}$$

Substitution of Eq. (B1) into (B4) yields

$$\mathsf{L}_0 \cdot \mathsf{L} \cdot \underline{e}_0 = \underline{e}_0, \tag{B5}$$

i.e., the transformation $\mathsf{L}_R = \mathsf{L}_0 \cdot \mathsf{L}$ (which, by the group property, must also be a Lorentz transformation) is one that leaves the direction $\underline{e}_0$ invariant and is, therefore, a pure rotation. Hence

$$\mathsf{L} = \tilde{\mathsf{L}}_0 \cdot \mathsf{L}_R. \quad \text{Q.E.D.} \tag{B6}$$

Semiclassical and Quantum Descriptions

S. P. MISRA
Regional Engineering College, Rourkela-8, Orissa, India

AND

T. S. SHANKARA
Matscience, The Institute of Mathematical Sciences, Madras-20, India

(Received 12 June 1967)

In the semiclassical descriptions, it is usual to describe a quantum-mechanical system in a classical language with (i) a correspondence between classical functions and operators of quantum mechanics and (ii) with a real, but not necessarily positive, probability density function in phase space corresponding to a particular quantum-mechanical state. The general forms of such semiclassical descriptions is discussed. The conditions for the two descriptions to be equivalent are also examined.

INTRODUCTION

After Sudarshan[1] gave a completely equivalent classical description of Glauber's[2] theory of photon correlations, the phase-space formulation of quantum mechanics has been used much in the study of the coherent properties of the electromagnetic field.[3] Such a formulation allows one to compute the quantum-mechanical expectation values as classical statistical averages over suitable distribution functions on the phase space. Indeed, all these distribution functions turn out to be particular cases of a general form. With a general form for the distribution function on hand, it is the purpose of this paper to compare the semiclassical and quantum descriptions at a universal level instead of restricting the treatment to particular cases (of overcomplete sets of states). Such a comparison in algebraic terms is necessary because it has sometimes been asserted that[4] "the apparent equivalence of these two vastly different types of descriptions, results from the paucity of the set of measurements so far considered for the radiation field and not from any intrinsic unity of the descriptions." For this purpose, we will consider the mapping of semiclassical to quantum description and also the inverse mapping of quantum to semiclassical description. Then we algebraically answer the questions, "When are the two descriptions identical and when does one contain more information than the other?" It turns out that the semiclassical description resulting from Weyl's association of operators to functions is identical with the quantum description and no information need be lost in going from one to the other. On the other hand, antinormal ordering yields a semiclassical description which is more general than the quantum description, provided we allow for arbitrary derivatives of the δ function.

Further, on the basis of these observations we discuss the dynamics of the system by building up the associative algebra in the classical function space which corresponds to the algebra in the quantum-mechanical operator space.

I. GENERAL THEORY

The linear mapping of classical functions to operators is written as

$$\hat{G} = \int \hat{E}(q, p) G(q, p)\, dq\, dp, \tag{1}$$

where $G(q, p)$ is the classical dynamical variable. The cap over every quantity denotes the corresponding operator. When $\hat{G}$ is a bounded operator, we get its expectation value for any normalized vector $|\alpha\rangle$ with $\langle\alpha \mid \alpha\rangle = 1$ as

$$\langle\alpha| \hat{G} |\alpha\rangle = \int \langle\alpha| \hat{E}(q, p) |\alpha\rangle\, G(q, p)\, dq\, dp. \tag{2}$$

Thus, if the distribution function $F_\alpha(q, p)$ is given as

$$F_\alpha(q, p) = \langle\alpha| \hat{E}(q, p) |\alpha\rangle, \tag{3}$$

Eq. (2) reduces to

$$\langle\alpha| \hat{G} |\alpha\rangle = \int F_\alpha(q, p) G(q, p)\, dq\, dp. \tag{4}$$

Thus, with the correspondence established by Eq. (1) between operators and functions and by Eq. (3) between state vectors and distribution functions, the expectation values remain the same.

We also note that instead of the state $|\alpha\rangle$ being given if the density matrix $\hat{\rho}$ is prescribed, the expectation value will be

$$\langle\hat{G}\rangle = \mathrm{Tr}\,(\hat{G}\hat{\rho}) = \int \mathrm{Tr}\,(\hat{E}(q, p)\hat{\rho}) G(q, p)\, dq\, dp \tag{5}$$

[1] E. C. G. Sudarshan, Phys. Rev. Letters **10**, 277 (1963).
[2] R. J. Glauber, Phys. Rev. Letters **10**, 84 (1963).
[3] L. Mandel and E. Wolf, Rev. Mod. Phys. **37**, 231 (1965).
[4] Y. Aharanov, D. Falkoff, E. Lerner, and H. Pendelton, Ann. Phys. (N.Y.) **39**, 498 (1966).

such that Eq. (3) for the distribution function reduces to

$$F(q,p) = \mathrm{Tr}\,(\hat{E}(q,p)\hat{\rho}). \tag{6}$$

The mapping (1) is however not power preserving or more generally, product preserving, and the squares and other powers of expectation values do not agree with the expectation values of the squares of these operators unless by accident. In different semiclassical descriptions we choose different forms of $\hat{E}(q,p)$ and we prefer a correspondence (1) such that q^m and p^n map, respectively, to $\hat{q}^m$ and $\hat{p}^n$ for any positive integral m and n.

We now note how the operators are defined for different semiclassical descriptions. For simplicity we are taking one q and p. This can be immediately generalized for any finite number of degrees of freedom.

(a) For Weyl's correspondence,[5] we have

$$\hat{E}(q,p) = \frac{1}{(2\pi)^2}\int \exp(i\hat{q}\theta + i\hat{p}\tau) \times \exp(-iq\theta - ip\tau)\, d\theta\, d\tau. \tag{7}$$

This gives us, with $\langle q \mid \alpha\rangle = \psi_\alpha(q)$ in coordinate representation,

$$\begin{aligned}
F_\alpha(q,p) &= \frac{1}{(2\pi)^2}\int \langle \alpha \mid q'\rangle\langle q'|\exp(i\hat{q}\theta + i\hat{p}\tau)\,|\alpha\rangle\, dq' \\
&\quad\times \exp(-iq\theta - ip\tau)\, d\theta\, d\tau \\
&= \frac{1}{(2\pi)^2}\int \psi_\alpha^*(q')\exp\left(iq'\theta + \frac{i}{2}\theta\tau\right) \\
&\quad\times \exp(-iq\theta - ip\tau)\, dq'\psi_\alpha(q'+\tau)\, d\theta\, d\tau \\
&= \frac{1}{(2\pi)^2}\int \psi_\alpha^*\left(q - \frac{\tau}{2}\right)\psi_\alpha\left(q+\frac{\tau}{2}\right)\exp(-ip\tau)\, d\tau,
\end{aligned} \tag{8}$$

which is Wigner's distribution function.

(b) It may be seen that in the correspondence described by Cohen[6]

$$\hat{E}(q,p) = \frac{1}{(2\pi)^2}\int \exp(i\hat{q}\theta + i\hat{p}\tau)f(\theta,\tau) \times \exp(-iq\theta - ip\tau)\, d\theta\, d\tau. \tag{9}$$

The procedure for the verification of this is identical as above and $f(\theta,\tau)$ is the same function as introduced by Cohen. This is easily seen when we note that in Ref. 6,

$$\begin{aligned}
F_\alpha(q,p) &= \frac{1}{(2\pi)^2}\iiint \psi_\alpha^*\left(u-\frac{\tau}{2}\right)\exp(i\theta u)\psi_\alpha\left(u+\frac{\tau}{2}\right) du \\
&\quad\times f(\theta,\tau)\exp(-iq\theta - ip\tau)\, d\theta\, d\tau \\
&= \frac{1}{(2\pi)^2}\iint \langle\alpha|\exp(\tfrac{1}{2}i\hat{p}\tau)\exp(i\hat{q}\theta)\exp(\tfrac{1}{2}i\hat{p}\tau)\,|\alpha\rangle \\
&\quad\times f(\theta,\tau)\exp(-iq\theta - ip\tau)\, d\theta\, d\tau \\
&= \frac{1}{(2\pi)^2}\iint \langle\alpha\,|\exp(i\hat{q}\theta + i\hat{p}\tau)\,|\alpha\rangle \\
&\quad\times f(\theta,\tau)\exp(-iq\theta - ip\tau)\, d\theta\, d\tau.
\end{aligned}$$

The conditions

$$f(\theta,0) = 1,$$
$$f(0,\tau) = 1$$

are necessary to ensure that

$$\int F_\alpha(q,p)\, dp = |\psi_\alpha(q)|^2$$

and

$$\int F_\alpha(q,p)\, dq = |\psi_\alpha(p)|^2,$$

respectively. However, when we further want that q^m and p^n should map, respectively, to $\hat{q}^m$ and $\hat{p}^n$ by Eq. (1), it is necessary that $(\partial/\partial\theta)^k f(\theta,\tau)$ and $(\partial/\partial\tau)^k f(\theta,\tau)$ both vanish when $\theta = \tau = 0$ for all integral $k > 0$. It is also to be noted that if $f(\theta,\tau)$ depends on $|\alpha\rangle$ the mapping from state vectors to expectation values of operators will not be bilinear but the linear mapping (1) can still be true for diagonal elements of $\hat{G}$. Here we have to take a basis for which $\hat{\rho}$ is diagonal, and if

$$\langle\alpha|\,\hat{\rho}\,|\beta\rangle = \delta_{\alpha\beta}P_\alpha, \tag{10}$$

then

$$\mathrm{Tr}\,(\hat{G}\hat{\rho}) = \sum_\alpha \langle\alpha|\,\hat{G}\,|\alpha\rangle\, P_\alpha, \tag{11}$$

where

$$\langle\alpha|\,\hat{G}\,|\alpha\rangle = \int\langle\alpha|\,\hat{E}(q,p;\alpha)\,|\alpha\rangle G(q,p)\, dq\, dp. \tag{12}$$

In this way we circumvent the impossibility of defining nondiagonal elements of the operator $\hat{G}$ for such a mapping, and it is clear that the definition of the operator is incomplete.

(c) Recently a correspondence between classical systems with a noncommutative Grassmann algebra and quantum systems has been established by one of

[5] C. L. Mehta, J. Math. Phys. 5, 677 (1964)—the earlier references are quoted here.
[6] L. Cohen, J. Math. Phys. 7, 781 (1966).

the authors[7] and the distribution function obtained.[8] We see that this also falls under the same general scheme. We note that there

$$\hat{E}(q, p) = \frac{1}{(2\pi)^2}\int \frac{1}{-i\theta\tau}[\exp(i\hat{q}\theta), \exp(i\hat{p}\tau)] \times \exp(-iq\theta - ip\tau)\, d\theta\, d\tau. \quad (13)$$

To see this directly,

$$\int \hat{E}(q, p) q^m p^n\, dq\, dp = \int \frac{1}{-i\theta\tau}[\exp(i\hat{q}\theta), \exp(i\hat{p}\tau)] \times i^{m+n}\delta^{(m)}(\theta)\delta^{(n)}(\tau)\, d\theta\, d\tau$$
$$= i^3 \frac{[\hat{q}^{m+1}, \hat{p}^{n+1}]}{(m+1)(n+1)}$$
$$= \frac{[\hat{q}^{m+1}, \hat{p}^{n+1}]}{i(m+1)(n+1)}.$$

We note that the above is the mapping in Ref. 7. Any other correspondence may also be similarly verified. We further note that it belongs to the category of distributions introduced by Cohen[6] and here

$$f(\theta, \tau) = \frac{\sin(\theta\tau/2)}{(\theta\tau/2)}. \quad (14)$$

Thus, although obtained in a totally different way, it is the same distribution function of Born and Jordan[6,9]

(d) When we want standard,[5] normal, or antinormal ordering,[10] we take

$$\hat{E}(q, p) = \frac{1}{(2\pi)^2}\int Q(\exp(i\hat{q}\theta + i\hat{p}\tau)) \times \exp(-iq\theta - ip\tau)\, d\theta\, d\tau, \quad (15)$$

where $Q((\hat{q}\theta + \hat{p}\tau)^n)$ equals the nth degree expression written as an nth order polynomial with the operators appropriately ordered. For the distribution function to be real, we need that $Q((\hat{q}\theta + \hat{p}\tau)^n)$ be Hermitian. We now explicitly verify this for the above types of orderings.

Let

$$\hat{q} = x\hat{\xi} + y\hat{\eta},$$
$$\hat{p} = x'\hat{\xi} + y'\hat{\eta}, \quad (16)$$

and

Q(any operator) = the ordering with all $\hat{\xi}$ to the left (and all $\hat{\eta}$ to the right). (17)

Now here

$$\int \hat{E}(q, p) q^m p^n\, dq\, dp$$
$$= \frac{1}{(2\pi)^2}\int Q(\exp(i\hat{q}\theta + i\hat{p}\tau)) \times \exp(-iq\theta - ip\tau)\, d\theta\, d\tau q^m p^n\, dq\, dp$$
$$= \int Q(\exp(i\hat{q}\theta + i\hat{p}\tau)) i^{m+n}\delta^{(m)}(\theta)\delta^{(n)}(\tau)\, d\theta\, d\tau$$
$$= \int (-1)^{m+n}\delta^{(m)}(\theta)\delta^{(n)}(\tau)\, d\theta\, d\tau \frac{Q((\hat{q}\theta + \hat{p}\tau)^{m+n})}{(m+n)!}$$
$$= (-1)^{m+n}\int \delta^{(m)}(\theta)\delta^{(n)}(\tau)\, d\theta\, d\tau \times \sum \frac{Q((x\hat{\xi}\theta)^{s_1}(y\hat{\eta}\theta)^{s_2}(x'\hat{\xi}\tau)^{s_3}(y'\hat{\eta}\tau)^{s_4})}{s_1!\, s_2!\, s_3!\, s_4!},$$

where $s_1 + s_2 + s_3 + s_4 = m + n$ and the s_i are positive integers or zero. Thus we obtain,

$$\int \hat{E}(q, p) q^m p^n\, dq\, dp$$
$$= \sum_{s_1+s_2=m}\sum_{s_3+s_4=n} \frac{m!}{s_1!\, s_2!}\frac{n!}{s_3!\, s_4!}\hat{\xi}^{s_1+s_3}\hat{\eta}^{s_2+s_4}x^{s_1}y^{s_2}x'^{s_3}y'^{s_4}$$
$$= Q((x\hat{\xi} + y\hat{\eta})^m(x'\hat{\xi} + y'\hat{\eta})^n)$$
$$= Q(\hat{q}^m\hat{p}^n). \quad (18)$$

This proves the earlier statement that for any polynomial, the mapping (1) with (15) gives the desired ordering and (3) or (6) gives the corresponding distribution function.

In order to show the equivalence of the above results with the over-complete set of states used by Sudarshan,[10] let us have in particular, corresponding to Eq. (16),

$$\hat{a} = \frac{\hat{q} + i\hat{p}}{\sqrt{2}} = \hat{\xi},$$
$$\hat{a}^\dagger = \frac{\hat{q} - i\hat{p}}{\sqrt{2}} = \hat{\eta}. \quad (19)$$

Thus, the prescribed ordering defined by $Q = Q_A$ is antinormal ordering. Further, we have the states

$$\hat{a}\,|z\rangle = (x + iy)\,|z\rangle. \quad (20)$$

This gives us, with the state $|k\rangle$ being the state with number operator $\hat{a}^\dagger\hat{a}$ diagonal,

$$\langle \hat{a}^m \hat{a}^{\dagger n}\rangle = \int \sum_{s,k}\langle k|\, \hat{a}^m\, |z\rangle \frac{d^2z}{\pi}\langle z|\, \hat{a}^{\dagger n}\, |s\rangle \rho_{sk}$$
$$= \int \rho_A(z) z^m z^{*n}\, d^2z, \quad (21)$$

[7] T. S. Shankara, Nuovo Cimento **47**, 553 (1967).
[8] T. S. Shankara, Progr. Theoret. Phys. (Kyoto) 37, 1335 (1967).
[9] M. Born and P. Jordan, Z. Phys. **34**, 873 (1925).
[10] C. L. Mehta and E. C. G. Sudarshan, Phys. Rev. **138**, B274 (1965).

where[10]

$$\rho_A(z) = \frac{1}{\pi}\sum_{s,k} \rho_{sk}\langle z \mid s\rangle\langle k \mid z\rangle, \\ = \frac{1}{\pi}\langle z|\, \hat{\rho}\, |z\rangle. \quad (22)$$

In the corresponding semiclassical picture defined by Eq. (15),

$$\langle \hat{a}^m \hat{a}^{\dagger n}\rangle = \int dq\, dp F_A(q,p)\left(\frac{q+ip}{\sqrt{2}}\right)^m \left(\frac{q-ip}{\sqrt{2}}\right)^n, \quad (23)$$

where

$$F_A(q,p) = \frac{1}{(2\pi)^2}\int \mathrm{Tr}\,\{Q_A(\exp(i\hat{q}\theta + i\hat{p}\tau))\hat{\rho}\} \cdot \exp(-iq\theta - ip\tau)\, d\theta\, d\tau. \quad (24)$$

The identity of the results (21) and (23) for arbitrary m and n signifies that we should take

$$x + iy = \frac{q+ip}{\sqrt{2}} \quad (25)$$

and with this association we obtain

$$\rho_A(x+iy) = 2F_A(\sqrt{2}x, \sqrt{2}y)$$

or

$$F_A(q,p) = \tfrac{1}{2}\rho_A\left(\frac{q+ip}{\sqrt{2}}\right), \quad (26)$$

where the equality holds for finding moments for any polynomial in q and p.

It may be noticed that the above results can be generalized to any set of states $|z\rangle$ which may be overcomplete[11] and are eigenstates of $\hat{\xi}$ with $\hat{\xi}^\dagger = \hat{\eta}$. The corresponding semiclassical description in terms of $F(q,p)$ can immediately be obtained where we consider the equations corresponding to Eqs. (21), (23), and (25).

II. EQUIVALENCE OF SEMICLASSICAL AND QUANTUM DESCRIPTIONS

In the above, we have considered the mapping of classical variables to quantum operators and have obtained the quasiprobability density functions. We now consider the inverse problem, i.e., mapping of operators to semiclassical variables and see how far they can be identical.

We write the mapping (1) as

$$\langle\alpha|\, \hat{G}\, |\beta\rangle = \int \langle\alpha|\, \hat{E}(q,p)\, |\beta\rangle\, G(q,p)\, dq\, dp. \quad (27)$$

In case an inverse of (27) exists, we have

$$G(q,p) = \sum_{\alpha,\beta} K_{\alpha,\beta}(q,p)\langle\alpha|\, \hat{G}\, |\beta\rangle \\ = \mathrm{Tr}\,(\hat{K}^t(q,p)\hat{G}), \quad (28)$$

where

$$K^t_{\alpha,\beta}(q,p) = \langle\beta|\, \hat{K}(q,p)\, |\alpha\rangle.$$

Symbolically, we have mappings $\hat{E}$ and $\hat{K}$ such that

$$G \xrightarrow{\hat{E}} \hat{G} \quad (29)$$

and

$$\hat{G} \xrightarrow{\hat{K}} G. \quad (30)$$

Now for the mapping $\hat{K}\hat{E}$ to be identity (in the space of classical dynamical variables), we require that

$$G(q,p) = \int \mathrm{Tr}\,(\hat{K}^t(q,p)\hat{E}(q',p'))G(q',p')\, dq'\, dp'$$

for arbitrary $G(q,p)$ so that we must have

$$\mathrm{Tr}\,(\hat{K}^t(q,p)\hat{E}(q',p')) = \delta(q-q')\delta(p-p'). \quad (31)$$

Similarly, for mapping $\hat{E}\hat{K}$ to be identity (in the operator space of quantum mechanics), we require that

$$\langle\alpha|\, \hat{G}\, |\beta\rangle = \sum_{\alpha',\beta'}\int \langle\alpha|\, \hat{E}(q,p)\, |\beta\rangle\langle\alpha'|\, \hat{K}(q,p)\, |\beta'\rangle \times \langle\alpha'|\, \hat{G}\, |\beta'\rangle\, dq\, dp$$

for arbitrary operators and thus

$$\int \langle\alpha|\, \hat{E}(q,p)\, |\beta\rangle\langle\alpha'|\, \hat{K}(q,p)\, |\beta'\rangle\, dq\, dp = \delta_{\alpha\alpha'}\delta_{\beta\beta'}. \quad (32)$$

In case both (31) and (32) are true, the mapping will be bi-unique and no information need be lost in going from one space to the other. On the other hand, if Eq. (31) is true, but not Eq. (32), the mapping $\hat{K}$ does not have a unique inverse and there are more than one distinct operators which map into the same function. This will imply that there can be information in operator space which will be lost in the semiclassical description and there will be nonzero operators which map to zero functions.

The roles will be reversed if Eq. (31) is not true for some functions, but Eq. (32) is true. In this case the semiclassical description would contain information that may be lost in the quantum-mechanical description and some nonzero functions would map to zero operators.

These comments are relevant when we build up the associative algebra in the function space to correspond with the algebra in the operator space (next section).

We also note that we can determine the density matrix when the mapping $\hat{K}$ exists as in Eq. (28) and a distribution function $F(q,p)$ is given. Thus,

$$\int G(q,p)F(q,p)\, dq\, dp \\ = \int \mathrm{Tr}\,(\hat{K}^t(q,p)\hat{G})F(q,p)\, dq\, dp, \\ = \mathrm{Tr}\,(\hat{G}\hat{\rho}),$$

[11] J. R. Klauder, J. Math. Phys. 5, 177 (1964).

which gives the density matrix $\hat{\rho}$ as

$$\hat{\rho} = \int \hat{K}^t(q, p)F(q, p)\, dq\, dp. \tag{33}$$

In Weyl's association, the inverse $\hat{K}$ exists both as a left and right inverse. To be more general, if we take the mapping of Cohen,[6] i.e., Eq. (9) in Sec. I, we can see that

$$\hat{K}^t(q, p) = \frac{1}{(2\pi)^2}\int \exp(-i\hat{q}\theta - i\hat{p}\tau) \times (f(\theta, \tau))^{-1} \exp(iq\theta + ip\tau)\, d\theta\, d\tau \tag{34}$$

satisfies Eqs. (31) and (32).

To see that with Eqs. (9) and (34), Eq. (31) is satisfied; we first note that

$$\begin{aligned}
&\mathrm{Tr}\,\{\exp(-i\hat{q}\theta - i\hat{p}\tau)\exp(i\hat{q}\theta' + i\hat{p}\tau')\} \\
&= \int \langle q'|\exp(-i\hat{q}\theta - i\hat{p}\tau)|p'\rangle\, dp' \times \langle p'|\exp(i\hat{q}\theta' + i\hat{p}\tau')|q'\rangle\, dq' \\
&= \int \exp(iq'(\theta' - \theta) + ip'(\tau' - \tau))\, |\langle q' \mid p'\rangle|^2\, dq'\, dp' \\
&= 2\pi\delta(\theta - \theta')\delta(\tau - \tau').
\end{aligned}$$

Hence Eqs. (9) and (34) give us

$$\begin{aligned}
&\mathrm{Tr}\,(\hat{K}^t(q, p)\hat{E}(q', p')) \\
&= (2\pi)^{-3}\int 2\pi\delta(\theta - \theta')\delta(\tau - \tau')(f(\theta, \tau))^{-1}f(\theta', \tau') \\
&\qquad \exp(iq\theta + ip\tau)\exp(-iq'\theta' - ip'\tau')\, d\theta'\, d\tau'\, d\theta\, d\tau \\
&= \delta(q - q')\delta(p - p').
\end{aligned}$$

In a similar way, Eq. (32) may also be verified. However, it is assumed that the integral on the right-hand side of Eq. (34) should exist in a formal way to justify various changes in the limit. In such a case, the mapping from semiclassical to quantum descriptions is bi-unique.

Also, Sudarshan[1] has given the expression for $\hat{\rho}$ when $F(q, p)$ is given. Proceeding the same way for anti-normal ordering, we have

$$\rho_A(z) = \rho_A(re^{i\theta}) = \frac{1}{\pi}\sum_{n,m} e^{-r^2}\frac{z^{*n}}{(n!)^{\frac{1}{2}}}\frac{z^m}{(m!)^{\frac{1}{2}}}\rho_{nm}. \tag{35}$$

This gives us

$$\rho_{nm} = \int \rho_A(z)K^t_{nm}(z)\, d^2z, \tag{36}$$

where

$$K^t_{nm}(re^{i\theta}) = \tfrac{1}{2}e^{r^2}\frac{(n!)^{\frac{1}{2}}(m!)^{\frac{1}{2}}}{r(n + m)!}e^{i(n-m)\theta}\left(-\frac{d}{dr}\right)^{n+m}\delta(r). \tag{37}$$

In the above, it is understood symbolically that

$$\int F(r)\left(-\frac{d}{dr}\right)^n\delta(r)\, dr = \int \left(\frac{d}{dr}\right)^n F(r)\delta(r)\, dr, = \left.\frac{d^nF(r)}{dr^n}\right|_{r=0}.$$

Although we have written the above only for $\hat{\rho}$, here it is clear that the mapping $\hat{E}\hat{K}$ is an identity in operator space, but $\hat{K}\hat{E}$ is not necessarily so. Thus, the semiclassical description is more general than quantum mechanical description, but we have to allow for distributions containing arbitrary derivatives of the δ function in the semiclassical description.

III. DYNAMICS

Next we consider the dynamics of quantum and semiclassical systems. The dynamics of the quantum system in a Heisenberg picture is given by the quantum Poisson bracket. We shall find the expression for the corresponding bracket in the semiclassical description.

We first discuss the case when we do not have a meaningful inverse mapping $\hat{K}$ of Eq. (28). We observe that

$$[\hat{A}, \hat{B}] = \int [\hat{E}(q, p), \hat{E}(q', p')] \times A(q, p)B(q', p')\, dq\, dp\, dq'\, dp'. \tag{38}$$

Now *if an L exists such that*

$$[\hat{E}(q, p), \hat{E}(q', p')] = \int L(q, p, q', p', q'', p'')\hat{E}(q'', p'')\, dq''\, dp'', \tag{39}$$

then

$$[\hat{A}, \hat{B}]_{\mathrm{Q.P.B}} = -i[\hat{A}, \hat{B}] = \int \hat{E}(q'', p'')C(q'', p'')\, dq''\, dp'', \tag{40}$$

where

$$C(q'', p'') = -i\int L(q, p, q', p', q'', p'') \times A(q, p)B(q', p')\, dq\, dp\, dq'\, dp'. \tag{41}$$

Clearly, Eq. (41) defines the corresponding semiclassical Poisson bracket when Eq. (39) is true.

Thus,

$$\{A(q, p), B(q, p)\}_{\mathrm{Scl.P.B.}} = -i\int L(q, p, q', p', q'', p'') \times A(q', p')B(q'', p'')\, dq'\, dp'\, dq''\, dp''. \tag{42}$$

In all earlier discussions,[5] essentially, the above arguments are put forth to deduce the classical

[18] J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949).

bracket. The Moyal bracket[12] can also be derived in the above manner, though of course, the derivation may be lengthy.

The problem is much simpler *when the inverse* $\hat{K}$ *exists.* Then we define

$$A(q,p) \times B(q,p)$$
$$= \sum_{\alpha,\beta} K_{\alpha\beta}(q,p)\langle\alpha|\,\hat{A}\hat{B}\,|\beta\rangle$$
$$= \int \mathrm{Tr}\,(\hat{K}^t(q,p)\hat{E}(q',p')\hat{E}(q'',p''))$$
$$\cdot A(q',p')B(q'',p'')\,dq'\,dp'\,dq''\,dp''. \quad (43)$$

With this rule of "multiplication" of functions which is necessarily associative, the semiclassical description merely becomes a different representation of the same algebra as that of the quantum mechanical system, and then the expectation values, dispersions, and dynamics of both become identical. The Poisson bracket in semiclassical description is defined as

$$\{A(q,p), B(q,p)\}_{\mathrm{Scl.P.B}} =$$
$$-i(A(q,p) \times B(q,p) - B(q,p) \times A(q,p)).$$

If only Eq. (39) is true, then the dynamics, of any specific operator is the same in semiclassical and quantum mechanics, but not so the dispersion or for that matter the expectation values of products of operators.

ACKNOWLEDGMENT

The authors would like to express their thanks to Professor Alladi Ramakrishnan for his keen interest and for inviting one of the authors (S. P. M.) to Madras where the problem was started.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Analytic Functionals in Quantum Field Theory and the Regularization of Ultraviolet Divergences*

FRANK B. THIESS
University of California at San Diego

(Received 28 March 1967)

The commutator and propagator distributions associated with free-particle fields are represented as boundary values of analytic kernels in complex space–time. A distribution product of such quantities is then defined as a boundary value of the product of the analytic kernels of its various factors, with the proviso that the product kernel be understood in the sense of Hadamard's "finite part" in those cases when it would otherwise be nonintegrable. For the case of quantum electrodynamics, the momentum-space representation of the products of causal propagators are derived. It is demonstrated by explicit calculation that the second-order scattering matrix elements thus obtained are identical to those usually obtained by means of Pauli–Villars regularization and renormalization.

I. INTRODUCTION

The ultraviolet divergence difficulties of quantum field theory are directly related to the problem of multiplying Feynman (causal) propagators. The Feynman propagators belong to a collection of tempered distributions, called "invariant distributions," or "invariant functions" which arise as commutators and propagators of various quantized fields. Since these distributions contain singular terms such as the δ function, their products are not uniquely defined. In order to give meaning to certain of these products, it has been necessary to first represent the factors as limits of sequences of approximating functions (using regularization methods due to Pauli and Villars[1] and Feynman[2]), and before passing to the limit, subtract from the product of approximating functions certain terms which diverge in the limit. In the case when the products represent scattering amplitudes for particle interactions the subtracted quantities may be interpreted in terms of increments in particle mass and charge that result from the interactions. The subtraction of the divergent terms is then justified, in fact necessitated, by the argument that such increments in mass and charge are not physically observable. There are serious disadvantages to defining distribution products in this way: the calculation of scattering amplitudes is extremely complicated; the regularization procedure introduces into the theory extraneous parameters such as fictitious masses or energy cutoffs which have no physical meaning and the infinite subtraction terms often intrude in such a way as to make rigorous arguments all but impossible and greatly detract from an otherwise elegant theory.

In Secs. II and III of this paper we represent the invariant distributions as analytic functionals. By "analytic functional" we mean a linear mapping of a space of test functions into the complex number system $\mathbb{C}$ which is defined by integrating the test function against a given analytic kernel over a given contour in the complex plane or more generally over a given surface in the space $\mathbb{C}^n$ of n complex variables (see Gel'fand–Shilov[3]). With such representations it is possible to remove from the field theory all considerations of jump functions, delta functions, etc., so that the mathematical operations used are well defined within the context of complex analysis. More important, it is possible to form a Lorentz invariant product of several invariant distributions without recourse to Pauli, Villars, Feynman type regularizations and without introducing into the theory any extraneous cutoff parameters. In Sec. IV this product is defined and its Fourier transform is derived. Finally, in Sec. V, we use this definition of multiplication to calculate the scattering amplitudes for the lowest-order vacuum polarization, electron self energy, and vertex part processes and show that the expressions thus obtained are identical to those arrived at through the standard regularization and subtraction procedures but have the advantage of requiring only finite renormalization.

Notation[4]*:* Points in space are labeled with vectors

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix};$$

events in space–time by 4-dimensional vectors $x = \begin{pmatrix} x_0 \\ \mathbf{x} \end{pmatrix}$. Lorentz innerproducts are denoted by

* This work supported in part by NSF Contract GP-4106.

[1] W. Pauli and F. Villars, Rev. Mod. Phys. **21**, 434 (1949).

[2] R. P. Feynman, Phys. Rev. **76**, 939 (1948).

[3] I. Gel'fand and I. Shilov, *Generalized Functions* (Academic Press Inc., New York, 1964).

[4] N. N. Bogoliubov and O. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience Publications, Inc., New York, 1959). Throughout this paper we shall follow the notation and nomenclature of this book.

$x\,k = g_{\mu\nu}x^\mu x^\nu$, where

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

In particular $x\,x = x^2$. An invariant distribution D is also displayed as a generalized function $D(x)$. The result of applying this distribution to a test function $f = f(x)$ is written

$$\int d^4x D(x) f(x) \equiv (D, f),$$

where $\int d^4x$ denotes integration over-all 4-space.

II. THE FREE-MESON FIELD

A. The Meson Kernel and its Associated Distributions

It has been observed by several authors, in particular, Wightman,[5] that the vacuum expectation value of the product of several field operators is a generalized function which is the boundary value of some analytic function in the following sense: a generalized function $T(x)$ defined on the x axis is said to be the boundary value of function $\mathcal{T}(x)$ analytic in, say, the lower half $z = x + iy$ plane, if

$$(T, f) \equiv \int_{-\infty}^{\infty} dx T(x) f(x) = \lim_{\epsilon \to 0^+} \int_{-\infty}^{\infty} dx \mathcal{T}(x - i\epsilon) f(x)$$

for every function f in the test space of T (see Bremermann and Durand[6]).

One of the simplest examples of a quantum-field invariant function which is the boundary value of an analytic function is the negative frequency commutator

$$D^-(x) = i \langle 0|\, \phi(x)\phi^*(0)\, |0\rangle = \frac{i}{(2\pi)^3} \int d^4k \delta(k^2 - m^2)\theta(-k^0) e^{ikx} \quad (1)$$

of the free field ϕ corresponding to a meson of mass m. Indeed, the integral on the right-hand side of (1) does not converge as it stands and is to be understood as a generalized limit in the following sense: for all admissible test functions $f(x)$

$$(D^-, f) \equiv \int d^4x D^-(x) f(x) = \lim_{\epsilon \to 0^+} \int d^4x D^-(x; \epsilon) f(x), \quad (2)$$

where

$$D^-(x;\epsilon) \equiv \frac{i}{(2\pi)^3} \int d^4k \delta(k^2 - m^2)\theta(-k^0) e^{ik^0(x_0 - i\epsilon) - i\mathbf{k}\cdot\mathbf{x}} = \frac{m}{8\pi} [(x_0 - i\epsilon)^2 - R^2]^{-\frac{1}{2}} H_1^{(1)}(m[(x_0 - i\epsilon)^2 - R^2]^{\frac{1}{2}}). \quad (3)$$

$R^2 \equiv \mathbf{x} \cdot \mathbf{x}$ and $H^{(1)}$ is the Hankel function of the first kind. Thus, D^- is a boundary value of the analytic function

$$\mathfrak{D}(z) = \frac{m}{8\pi(z^2)^{\frac{1}{2}}} H_1^{(1)}[m(z^2)^{\frac{1}{2}}] \quad (4)$$

which we call the "meson kernel"; here we have completed x to a complex valued 4-vector

$$z = x + iy = \begin{pmatrix} x^0 \\ \mathbf{x} \end{pmatrix} + i \begin{pmatrix} y^0 \\ \mathbf{y} \end{pmatrix}$$

and have written $z^2 \equiv g_{\mu\nu} z^\mu z^\nu = x^2 - y^2 + 2ixy$. It is a remarkable fact that not only is D^- a boundary value of the meson kernel (4) but so is every other invariant distribution associated with the free meson field ϕ.

Denote by $|q|^2 \equiv q_0^2 + q_1^2 + q_2^2 + q_3^2$ the square of the "Euclidean length" of the real or complex 4-vector q. Define a "ρ strip" in the space $\mathbb{C}^4$ of complex 4-vectors z to be the set: $\{z/x \in \mathbb{R}^4, |y|^2 < \rho^2\}$, where $\rho > 0$. Let $\mathcal{A}$ be the space of all functions $f(z)$ which are analytic in some ρ strip, where ρ can depend on f, and in that strip $f(z)$ vanishes at infinity faster than any negative power of $|x|$; (i.e., for every integer n, $|x|^n\, |f(z)| \to 0$ as $|x| \to \infty$ uniformly for all $|y|^2 < \rho^2$). For example, the space $\mathcal{A}$ includes all functions of the form $\exp\{-c\,|z|^2\} P(z)$, where c is an arbitrary positive number and $P(z)$ is an arbitrary polynomial; in particular $\mathcal{A}$ includes the Hermite functions. It is not difficult to verify that the space $\mathcal{A}$ is invariant under the full real Lorentz group in the sense that if $f(z) \in \mathcal{A}$ as a function of z then $f[L(z')] \in \mathcal{A}$ as a function of z', where $z' = L^{-1}(z)$ and L is any real Lorentz transformation. We shall use $\mathcal{A}$ as the space of test functions for all distributions considered in this paper. We have adopted this space simply for reasons of mathematical convenience; however, it should be pointed out that distributions defined on $\mathcal{A}$ can be easily extended to other more commonly considered spaces by making use of the fact that functions in these other spaces can be approximated arbitrarily closely by linear combinations of Hermite functions.

The meson kernel (4) is analytic in the 4-vector z, except for polar and branch singularities on the light

[5] A. S. Wightmann, Phys. Rev. **101**, 860 (1956).
[6] H. J. Bremermann and L. Durand III, J. Math. Phys. **2**, 240 (1961).

cone $z^2 = 0$ as displayed by the representation:

$$\mathfrak{D}(z) = \frac{1}{4\pi^2 i z^2} + \frac{m^2}{16\pi^2}\left[1 + \frac{i}{\pi}\log\left(\frac{\gamma m}{2} z^2\right)\right] \times \sum_{j=0}^{\infty} \frac{(-\frac{1}{4}m^2 z^2)^j}{j!(j+1)!} + \frac{1}{16\pi^2 i} \times \sum_{j=0}^{\infty} \frac{(-\frac{1}{4}m^2z^2)^j}{j!(j+1)!}\left(1 + \sum_{k=1}^{j+1}\frac{1}{k}\right), \quad (5)$$

where $\log \gamma \equiv 0.5772\cdots$ is the Euler constant. These singularities will play a central role in our definition and analysis of the invariant distributions. Let us temporarily restrict our considerations to vectors

$$z = \begin{pmatrix} z^0 \\ \mathbf{x} \end{pmatrix} \equiv \begin{pmatrix} x^0 \\ \mathbf{x} \end{pmatrix} + i \begin{pmatrix} y^0 \\ \mathbf{0} \end{pmatrix}$$

that have purely real spacelike parts. For such vectors the meson kernel can be expressed in the form

$$\mathfrak{D}(z) = \mathfrak{D}(z^0, \mathbf{x}) = \frac{1}{4\pi^2 i(z_0^2 - R^2)} + A(z^0, R)\log(z_0^2 - R^2) + B(z^0, R),$$

where for fixed R, A and B are entire functions of z^0. For fixed $\mathbf{x}$ let us specify the "principal sheet" of $\mathfrak{D}(z^0, \mathbf{x})$ by constructing branch cuts $|\mathrm{Re}\,(z^0)| \geq R$, $\mathrm{Im}\,(z^0) \equiv 0$ along the real axis of the z^0 plane and stipulating that

$$\mathrm{Im}\,\{\log(z_0^{*2} - R^2)\} \equiv 2\arg(z_0^{*2} - R^2)^{\frac{1}{2}} = 0$$

for a point z_0^* which lies on the upper edge of the right hand branch cut (see Fig. 1).

Let σ^- be a contour on the principal sheet of $\mathfrak{D}(z^0, \mathbf{x})$ which is parallel to the x^0 axis and passes below both right and left hand branch cuts (see Fig. 2). By making use of this contour we can dispense with the limiting process in representation (2) and write equivalently

$$(D^-, f) = \int d^3\mathbf{x} \int_{\sigma^-} dz^0 \mathfrak{D}(z^0, \mathbf{x}) f(z^0, \mathbf{x}) \quad (6)$$

for each test function $f \in \mathcal{A}$ provided σ^- lies within the strip of analyticity of f. Indeed, it follows from Cauchy's Theorem that the value of integral (6) is

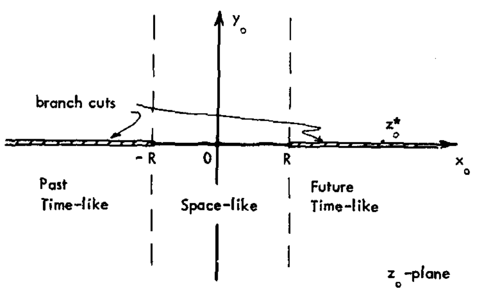


FIG. 1. The principal sheet of the meson kernel $\mathfrak{D}(z^0, \mathbf{x})$.

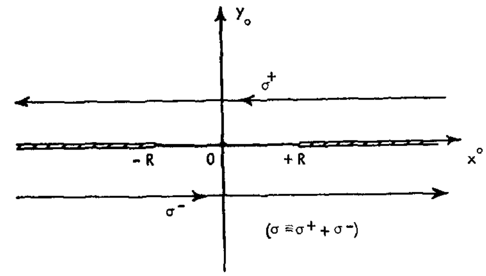


FIG. 2. Commutator contours.

independent of the particular choice of contour σ^- provided that σ^- remains within the strip of analyticity of f and passes below the points of singularity $z^0 = \pm R$. Consequently, σ^- can be pressed upward against the real axis to give the same limit as in Eq. (2). Next we shall show that the other invariant distributions of the meson field are also analytic functionals.

The distributions to be considered are listed below:

(a) The "Negative Frequency Commutator":

$$D^-(x) = i\,\langle 0|\,\phi(x)\phi^*(0)\,|0\rangle.$$

(b) The "Positive Frequency Commutator":

$$D^+(x) = i[\phi(x), \phi^*(0)].$$

(c) The "Full Commutator" (Schwinger function):

$$D(x) = i[\phi(x), \phi^*(0)].$$

(d) The "Causal (Feynman) Propagator":

$$D^c(x) = i\,\langle 0|\,T\phi(x)\phi^*(0)\,|0\rangle.$$

(e) The "Advanced Propagator":

$$D^{\mathrm{adv}}(x) = i\theta(x^0)[\phi(x), \phi^*(0)].$$

(f) The "Retarded Propagator":

$$D^{\mathrm{ret}}(x) = i\theta(-x^0)[\phi(x), \phi^*(0)].$$

T is the time-ordering operator and θ is the Heaviside function.

Let $D^{(\ )}(x)$ be a generic symbol designating any one of the distributions in the above list. We associate with each $D^{(\ )}(x)$ a corresponding contour $\sigma^{(\ )}$ in the principal sheet of the meson kernel $\mathfrak{D}(z^0, \mathbf{x})$. These contours are depicted in Figs. 2 and 3. Using these contours and the meson kernel (4) we define the analytic functionals $D^{(\ )}$ on the space $\mathcal{A}$ as follows

$$(D^{(\ )}, f) \equiv \int d^4x D^{(\ )}(x) f(x) \equiv \int d^3\mathbf{x} \int_{\sigma''} dz^0 \mathfrak{D}(z^0, \mathbf{x}) f(z^0, \mathbf{x}), \quad (7)$$

for each $f \in \mathcal{A}$.

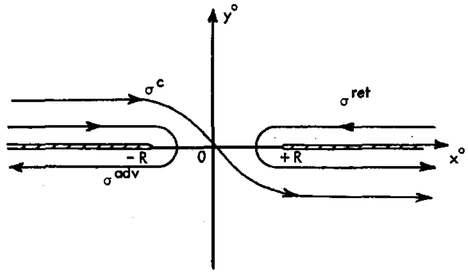


FIG. 3. Propagator contours.

It is not difficult to verify that the definitions of invariant distributions given by (7) are equivalent to the usual definitions. Indeed, the invariant distributions given by the usual definitions are known to satisfy the following relations

$$D^+(x) = -D^-(x), \quad D(x) = D^{ret}(x) - D^{adv}(x)$$
$$D^{ret}(x) = D^c(x) + D^+(x),$$
$$D^c(x) = \theta(x^0)D^-(x) - \theta(-x^0)D^+(x). \qquad (8)$$

On the other hand, the corresponding topological relations which hold for the contours $\sigma^{(\ )}$ (e.g., $\sigma^{ret} = \sigma^c + \sigma^+$, etc.) imply the analytic functionals given by (7) also satisfy the relations (8). This fact together with Eqs. (2), (3), and (6) is sufficient to infer the equivalence of (7) with the usual definitions.

As it stands, integral (7) is the sum (over $\mathbb{R}^3$) of a 3-parameter family of contour integrals. However, it will be helpful to regard it instead as a surface integral over a four-dimensional manifold in the eight-dimensional space $\mathbb{C}^4$ of complex 4-vectors z. Indeed, let $s^{(\ )}$ be a smooth four dimensional surface in $\mathbb{C}^4$ which lies in the sub-space $\mathbf{y} \equiv \mathbf{0}$ and whose intersection with each plane $\mathbf{x}$ = constant is the contour $\sigma^{(\ )}$. Then we may rewrite (7) as the surface integral

$$(D^{(\ )}, f) = \int_{s^{(\ )}} d^4z\mathfrak{D}(z)f(z) \quad \text{for} \quad f \in \mathcal{A}. \qquad (9)$$

Clearly, the surface $s^{(\ )}$ is not unique. A four-dimensional version of Cauchy's Theorem asserts that the value of integral (9) is unchanged under continuous deformations of $s^{(\ )}$ which avoid the complex light-cone $z^2 = 0$. Also, real Lorentz transformations of $s^{(\ )}$ leave the value of (9) unchanged. Thus, there is an equivalence class of surfaces $s^{(\ )}$, any member of which can be used to determine the invariant distribution $D^{(\ )}$ through Eq. (9).

A glance at Fig. 2 reveals that one can always select the surfaces s^-, s^+, and s, associated with the commutator distributions, to be uniformly bounded away from the light cone $z^2 = 0$ which is the only place where the kernel $\mathfrak{D}(z)$ is singular. The integrand of (9) vanishes rapidly as $|x| \to \infty$ [see Eq. (19)]. Consequently, integral (9) is absolutely convergent for the commutator distributions D^-, D^+, and D. The surfaces s^c, s^{ret}, and s^{adv}, associated with the propagator distributions, each touch the complex light cone at the single point $z = 0$. This may be seen from the fact that the contours σ^c, σ^{ret}, and σ^{adv} pass through the gap between the branch cuts and are pinched by the singularities as $R \to 0$. Nevertheless, the singularity of $\mathfrak{D}(z)$ is sufficiently weak [see Eq. (5)] to allow integral (9) to converge absolutely also for the propagator distributions D^c, D^{ret}, and D^{adv}.

B. The Momentum-Space Representation

The Fourier transforms of the meson invariant distributions are themselves analytic functionals. Even though this fact is well known, it will be interesting to review here the momentum space formalism and note how closely it resembles the space–time formalism described above.

Let us extend the momentum vectors k into the complex and write

$$h \equiv \begin{pmatrix} h^0 \\ \mathbf{h} \end{pmatrix} \equiv k + il \equiv \begin{pmatrix} k^0 \\ \mathbf{k} \end{pmatrix} + i\begin{pmatrix} l^0 \\ \mathbf{l} \end{pmatrix},$$

and let us denote by $\hat{\mathcal{A}}$ the space of all functions $g(h)$ such that $g(z)$ belongs to $\mathcal{A}$ when considered as a function of z instead of h. We shall use $\hat{\mathcal{A}}$ as a space of test functions for distributions defined over momentum space. The Fourier transforms $\hat{D}^{(\ )}$ of the distributions $D^{(\ )}$ are defined as analytic functionals on the space $\hat{\mathcal{A}}$ as follows:

$$(\hat{D}^{(\ )}, g) \equiv \int d^4k\hat{D}^{(\ )}(k)g(k)$$
$$\equiv \int d^3\mathbf{k} \int dh^0(m^2 - h_0^2 + \mathbf{k}^2)^{-1} g(h^0, \mathbf{k}), \quad (10)$$

where $\hat{\sigma}^{(\ )}$ are the familiar contours depicted in Figs. 4 and 5.

Denote by

$$\hat{\mathfrak{D}}(h) \equiv (m^2 - h^2)^{-1} \qquad (11)$$

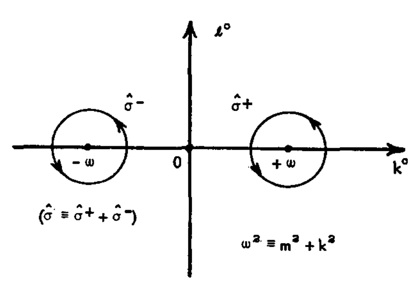


FIG. 4. Commutator contours.

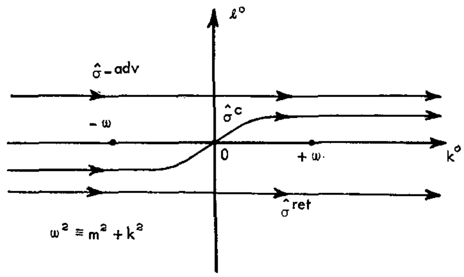


FIG. 5. Propagator contours.

the momentum space representation of the meson kernel, and let $\hat{s}^{(\,)}$ be the four-dimensional surface swept out by the contours $\hat{\sigma}^{(\,)}$ as $\mathbf{k}$ ranges over $\mathbb{R}^3$. Then we can rewrite (10) as the surface integral

$$(\hat{D}^{(\,)}, g) \equiv \int_{\hat{s}^{(\,)}} d^4h \hat{\mathfrak{D}}(h) g(h), \tag{12}$$

which has the same formal structure as the integral (9).

The distributions $D^{(\,)}$ and the Fourier transforms $\hat{D}^{(\,)}$ are coupled by the equations

$$D^{(\,)}(x) = \left[\hat{D}^{(\,)}(k), \frac{e^{ikx}}{(2\pi)^4}\right] = \int_{\hat{s}^{(\,)}} d^4h \hat{\mathfrak{D}}(h) \frac{e^{ihx}}{(2\pi)^4}, \tag{13}$$

$$\hat{D}^{(\,)}(k) = [D^{(\,)}(x), e^{-ikx}] = \int_{s^{(\,)}} d^4z \mathfrak{D}(z) e^{-ikz}. \tag{14}$$

It should be emphasized that integrals (13) and (14) are not convergent and must therefore be evaluated by means of some limiting procedure such as (2) and (3). When this is done one obtains as generalized limits the invariant functions $D^{(\,)}(x)$ and $\hat{D}^{(\,)}(k)$. Comparing Eqs. (9) and (12) one is motivated to say that in a certain sense the analytic kernel $\hat{\mathfrak{D}}(h)$ is the "Fourier transform" of the analytic kernel $\mathfrak{D}(z)$, and that with respect to these kernels the surfaces $\hat{s}^{(\,)}$ are "Fourier transforms" of the corresponding surface $s^{(\,)}$.

III. FREE-SPINOR FIELDS

The analytic formalism described above generalizes in a straight forward way to include all free-spinor fields of Kemmer type. Let $\chi(x) = (\chi_1(x), \cdots, \chi_n(x))$ be an arbitrary Kemmer-type field operator which belongs to a spinor space of dimension n. Associated with this field are the various invariant functions $K^{(\,)}(x) = ([K^{(\,)}(x)_{\xi\eta}])$, which are $n \times n$ matrices whose components $K^{(\,)}(x)$ are distributions defined on $\mathcal{A}$. For example, the causal propagator is $K^c(x)_{\xi\eta} = i\langle 0|\, T\chi_\xi(x)\chi_\eta^*(0)\,|0\rangle$. As may be recalled, Kemmer-type fields are characterized by the property that their invariant functions can be derived from those of the scalar field of same mass by means of a linear first-order differential operation

$$K^{(\,)}(x)_{\xi\eta} = (i\Gamma^\mu_{\xi\eta}\partial_\mu + M_{\xi\eta})D^{(\,)}(x),$$

where $\Gamma^\mu_{\xi\eta}$, $\mu = 0, 1, 2, 3$, and $M_{\xi\eta}$ are $n \times n$ matrices of constant coefficients indexed by spinor numbers ξ, η; we shall use the symbol ∂_μ as shorthand for $\partial/\partial x^\mu$ or $\partial/\partial z^\mu$ depending on context.

Let us associate with the field χ the analytic "particle kernel" $\mathcal{K}(z) \equiv [(\mathcal{K}(z)_{\xi\eta})]$ which is the $n \times n$ matrix whose components are given by

$$\mathcal{K}(z)_{\xi\eta} \equiv (i\Gamma^\mu_{\xi\eta}\partial_\mu + M_{\xi\eta})\mathfrak{D}(z), \tag{15}$$

then for every spinor $f(z) = [f_1(z), \cdots, f_n(z)]$ whose components $f_\eta(z)$ belong to the space $\mathcal{A}$, we have

$$(K^{(\,)}, f)_\xi \equiv \left(\int d^4x K^{(\,)}(x) f(x)\right)_\xi = \left(\int_{s^{(\,)}} d^4z \mathcal{K}(z) f(z)\right)_\xi$$
$$= \sum_\eta \int_{s^{(\,)}} d^4z \mathcal{K}_{\xi\eta}(z) f_\eta(z), \tag{16}$$

where the surfaces of integration $s^{(\,)}$ are the same as those used in the preceding section. Thus, the invariant distributions of free particle fields are all analytic functionals, being simply various "boundary values" of their respective analytic kernels.

To illustrate these ideas let us consider the fields of quantum-electrodynamics. The "electron–positron kernel" associated with the Dirac field ψ is the 4×4 spinor matrix

$$\mathcal{S}(z) \equiv (i\gamma^\mu\partial_\mu + M)\mathfrak{D}(z)$$
$$= \frac{M^2}{8\pi(z^2)^{\frac{1}{2}}} H_1^{(1)}[M(z^2)^{\frac{1}{2}}] - \frac{i\gamma^\mu z_\mu}{8\pi z^2} H_2^{(1)}[M(z^2)^{\frac{1}{2}}],$$

where M is the mass of the electron and γ^μ, $\mu = 0, 1, 2, 3$ are the 4×4 Dirac matrices. The invariant functions of this field are defined by

$$(S^{(\,)}, f) = \int d^4x S^{(\,)}(x) f(x) = \int_{s^{(\,)}} d^4z \mathcal{S}(z) f(z).$$

The "photon kernel" associated with the electromagnetic potential A (in a preselected gauge) is the second rank tensor

$$\mathcal{P}(z)_{\mu\nu} = g_{\mu\nu}\mathfrak{D}(z)\big|_{m=0} = \frac{g_{\mu\nu}}{4\pi^2 i z^2},$$

and the invariant functions of the electromagnetic field are defined by

$$(P^{(\,)}, f) = \int d^4x P^{(\,)}(x) f(x) = \int_{s^{(\,)}} d^4z \mathcal{P}(z) f(z).$$

When the invariant distributions are represented as analytic functions the problem of multiplying then is considerably simplified. We treat this problem in the next section.

IV. MULTIPLICATION OF INVARIANT DISTRIBUTIONS

A. Products in Configuration Space

Let $J^{(\,)}$ and $K^{(\,)}$ be two invariant distributions of the same type, that is, both causal propagators, or both negative frequency commutators, etc. [In quantum field theory we encounter only products whose factors are of the same type. This is because the time-ordering operator T, Heaviside function $\theta(x^0)$, etc., which multiply products of fields affect equally *all* of the fields which occur in the product.] We wish to define the product distribution $J^{(\,)} \cdot K^{(\,)}$ as a linear functional on the space $\mathcal{A}$

$$(J^{(\,)} \cdot K^{(\,)}, f) \equiv \int d^4x J^{(\,)}(x) K^{(\,)}(x) f(x).$$

Clearly, two properties that must be required of this product are: (i) Lorentz invariance, and (ii) away from the light cone $x^2 = 0$ where the invariant functions behave like ordinary functions, this product should behave like the usual function product.

Since $J^{(\,)}$ and $K^{(\,)}$ are "boundary values" of their respective analytic kernels $\mathfrak{J}(z)$ and $\mathfrak{K}(z)$, it would seem natural to take for their product the corresponding "boundary value" of the product of their "kernels"

$$(J^{(\,)} \cdot K^{(\,)}, f) \equiv \int_{s^{(\,)}} d^4z \mathfrak{J}(z) \mathfrak{K}(z) f(z), \quad \text{for } f \in \mathcal{A} \tag{17}$$

This definition certainly fulfills the above-mentioned requirements, but unfortunately integral (17) converges only for the cases where $J^{(\,)}$ and $K^{(\,)}$ are commutator distributions. Indeed, when $J^{(\,)}$ and $K^{(\,)}$ are propagators the surface of integrations $s^{(\,)}$ touches the light-cone at the origin and the singularities of the kernels $\mathfrak{J}(z)$ and $\mathfrak{K}(z)$ combine to make $\mathfrak{J}(z)\mathfrak{K}(z)$ nonintegrable in the neighborhood of this point. (It is this nonintegrability that is responsible for the divergence of the second-order self-energy processes.) We shall therefore temporarily reject definition (17) and seek some modification which will preserve its essential properties but will eliminate the divergence difficulties.

From Eqs. (5) and (15) it is clear that the strongest singularity of the product $\mathfrak{J}(z)\mathfrak{K}(z)$ (or more generally the product of any finite number of kernels) is a term of the form

$$(\text{const}) \frac{\Gamma_1^{\nu_1} z_{\nu_1} \cdots \Gamma_a^{\nu_a} z_{\nu_a}}{(z^2)^b},$$

where a and b are nonnegative integers and Γ^ν are constant matrix coefficients. Let us call

$$\theta = 2b - a - 4 \tag{18}$$

the "index of divergence" of $\mathfrak{J}(z)\mathfrak{K}(z)$. Since $\theta \geq 0$ for any product, $\mathfrak{J}(z)\mathfrak{K}(z)$ is nonintegrable through the origin. However integral (17) would converge there if $f(z)$ had a zero at the origin of order $\theta + 1$. Let us therefore adopt the following definition for the product of two invariant functions of the same type:

$$(J^{(\,)} \cdot K^{(\,)}, f) \equiv \int_{s^{(\,)}} d^4z \mathfrak{J}(z) \mathfrak{K}(z) \times \left\{ f(z) - \sum_{j=0}^{\theta} \frac{1}{j!} (\partial^{\mu_1} \cdots \partial^{\mu_j} f(0)) z_{\mu_1} \cdots z_{\mu_j} \right\}, \tag{17'}$$

for each $f \in \mathcal{A}$. This definition is clearly Lorentz invariant and it makes no use of cutoff parameters or limiting procedures. [It is interesting to note that while θ given by (18) is the smallest upper limit of summation for which integral (17′) converges at the origin, it is also the largest upper limit for which (17′) converges at infinity provided J and K correspond to fields with nonzero mass. This surprising fact follows in a straightforward way from the fact that analytic kernels (15) have the asymptotic behavior

$$\mathfrak{K}(z) \approx (i\Gamma^\mu \partial_\mu + M) \frac{e^{-(3\pi i)/4}}{8\pi} \left(\frac{2m}{\pi}\right)^{\frac{1}{2}} \frac{e^{im(z^2)^{\frac{1}{2}}}}{(z^2)^{\frac{3}{4}}} \quad \text{as} \quad z^2 \to \infty. \tag{19}$$

Thus, the value of index θ is uniquely determined by the requirement that integral (17′) converge both at the origin and at infinity.]

The product $J^{(\,)}(x)K^{(\,)}(x)$ defined by (17) and the product defined by (17′) are identical except possibly at the origin. Indeed, the formal difference between these products is the distribution

$$\sum_{j=0}^{\theta} a^{(\,)}_{\mu_1 \cdots \mu_j} \partial^{\mu_1} \cdots \partial^{\mu_j} \delta(x),$$

where

$$a^{(\,)}_{\mu_1 \cdots \mu_j} \equiv \int_{s^{(\,)}} d^4z \mathfrak{J}(z) \mathfrak{K}(z) z_{\mu_1} \cdots z_{\mu_j}.$$

In the case of products of commutator distributions (17) and (17′) are even identical at the origin, since $a^{(\,)}_{\mu_1 \cdots \mu_j} = 0$ by virtue of the fact that the contours σ^+ and σ^- can be closed in the upper and lower half-planes, respectively. In the case of products of propagator distributions, integral (17′) represents a "Hadamard-type"[7] regularization ("finite part") of integral (17). Such regularizations are often used in the theory of distributions because they have the

[7] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Dover Publications, Inc., New York, 1952).

important property of preserving derivatives, in the sense that the derivative of the Hadamard regularization of a function is equal to the Hadamard regularization of the derivative of the function. In light of the foregoing remarks, it is clear that equation (17) can be regarded as the definition of multiplication provided that the integral is understood in the sense of Hadamard's finite part (17′) whenever it would otherwise diverge.

The scattering matrix terms which correspond to second-order self-energy processes have the following structure

$$\int d^4x_1 \int d^4x_2 J^c(x_1 - x_2)K^c(x_1 - x_2)\phi(x_1)\psi(x_2),$$

where the functions ϕ and ψ are determined by the incoming and outgoing states. The above definition of multiplication enables us to calculate these quantities without encountering ultraviolet divergences.

(To simplify the exposition we shall now restrict our considerations to products of causal propagators.)

Next in order of complication among the scattering-matrix terms that usually diverge is the third-order "vertex part" which has the following structure

$$\int d^4x_1 \int d^4x_2 \int d^4x_3 J^c(x_2 - x_1)K^c(x_3 - x_1)L^c(x_3 - x_2) \times \phi(x_1)\psi(x_2)\Omega(x_3).$$

In order to evaluate such quantities we shall define the product $J^c(x)K^c(x + u)L^c(u)$ as a distribution

$$(J^c(x)K^c(x + u)L^c(u), f(x, u)) = \int d^4x \int d^4u J^c(x)K^c(x + u)L^c(u)f(x, u) \quad (20)$$

on the space $\mathcal{A}^2$ of all functions $f(z, w)$ which are analytic in the two complex 4-vectors $z = x + iy$ and $w = u + iv$ in some "strip": $|y|^2 + |v|^2 < \rho^2$ and within this strip vanish at infinity faster than any negative power of $|x|^2 + |u|^2$. Again applying the rule that the product of "boundary values" of various analytic kernels should be the "boundary value" of the product of these kernels, we define (20) as the two-fold surface integral

$$(J^c(x)K^c(x + u)L^c(u), f(x, u)) \equiv \int_{s^c} d^4z \int_{s^c(-z)} d^4w \mathfrak{J}(z)\mathfrak{K}(z + w)\mathfrak{L}(w)f(z, w), \quad (21)$$

which is to be understood in the sense of Hadamard's finite part as will be explained below.

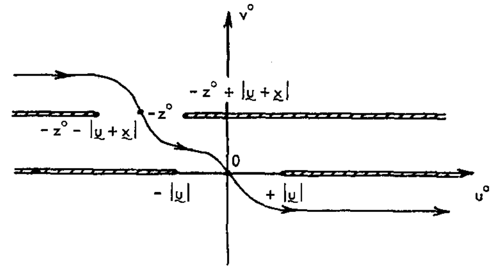


FIG. 6. The contour $\sigma^c(-z)$.

First let us describe the surfaces of integration: the z integration, which we have elected to perform last, is to be taken over the causal surface s^c. Therefore, for purposes of describing the w integration we can suppose that $\mathbf{y} \equiv \mathbf{0}$, and that $-z^0$ lies in the interiors of the 2nd or 4th quadrants except when it passes through the origin. The kernel of the w integration has singularities on the complex light-cones $w^2 = 0$ and $(w + z)^2 = 0$. For fixed $z = \begin{pmatrix} z^0 \\ \mathbf{x} \end{pmatrix}$, and fixed $\mathbf{w} = \mathbf{u} + i\mathbf{0}$ these singularities are located in the $w^0 = u^0 + iv^0$ plane at the points $w^0 = \pm|\mathbf{u}|$ and $w^0 = -z^0 \pm |\mathbf{u} + \mathbf{x}|$ (see Fig. 6). Let us construct branch-cuts from infinity parallel to the real axis which terminate at these four points. Let $\sigma^c(-z)$ be defined as that contour which threads through the gaps between these branch cuts as depicted in Fig. 6. Then integration over $s^c(-z)$ is defined by

$$\int_{s^c(-z)} d^4w \cdots \equiv \int d^3\mathbf{u} \int_{\sigma^c(-z)} dw^0 \cdots$$

Clearly, one can make the contours $\sigma^c(-z)$ depend on $\mathbf{u}$ in a suitable continuous way so that $s^c(-z)$ becomes a smooth four-dimensional surface in w space. Also, by collecting s^c and $s^c(-z)$ together, we can re-express the iterated integral (21) as a surface integral over a smooth eight-dimensional manifold, call it s_2^c, in the sixteen-dimensional space of the two complex four-vectors z and w:

$$[J^c(x)K^c(x + u)L^c(u), f(x, u)] \equiv \iint_{s_2^c} d^4z \wedge d^4w \mathfrak{J}(z)\mathfrak{K}(z + w)\mathfrak{L}(w)f(z, w). \quad (22)$$

The singularities of the integrand of (22) lie on the three 14-dimensional surfaces $z^2 = 0$, $(z + w)^2 = 0$, and $w^2 = 0$. The surface of integration s_2^c touches these singular surfaces when $z = 0$, $z = -w$, and $w = 0$, respectively. However, s_2^c touches these surfaces one at a time except for the case $z = w = 0$. Therefore, we shall assign a finite value to (22) by taking its Hadamard finite part at the point $z = w = 0$:

$$[J^c(x)K^c(x+u)L^c(u), f(x,u)] \equiv \iint_{s_2{}^c} d^4z \wedge d^4w \mathfrak{J}(z)\mathfrak{K}(z+w)\mathfrak{L}(w) \times \left\{ f(z,w) - \sum_{j+r=0}^{\theta} \frac{1}{(j+r)!} \partial^{\mu_1} \cdots \partial^{\mu_j}\partial^{\nu_1} \cdots \partial^{\nu_r} \times f(0,0) z_{\mu_1} \cdots z_{\mu_j} w_{\nu_1} \cdots w_{\nu_r} \right\}. \quad (22')$$

Here θ is the smallest upper limit of summation for which (22′) converges.

B. Momentum-Space Representation of Products

The usefulness of the analytic functional representation for invariant distributions and their products stems from the fact that these quantities are defined completely within the context of complex analysis. As a consequence, the causes of ultraviolet divergences can be traced to bona fide singularities of certain analytic functions, and methods can be devised to "regularize" these singularities without violating the conditions of causality and Lorentz invariance. However, it is certainly not intended that scattering matrix elements be evaluated by computing complicated surface integrals in complex space–time. For this purpose, we shall employ the usual momentum-space representation for causal propagators, viz.,

$$\hat{K}^c(k) = \frac{\Gamma^\mu k_\mu - M}{k^2 - m^2 + i\epsilon} \quad (\text{where } \epsilon \to 0^+),$$

and re-express products such as (17′) and (22′) in terms of these comparatively simple expressions.

For reasons of brevity the derivations presented in this section are purely formal. However, the same results can be obtained in a rigorous way by applying essentially the same argument using sequences of functions which approach the causal propagators, but whose members form convergent convolution integrals.

The Fourier transform $\widehat{J^c \cdot K^c}$ of the product of causal propagators $J^c \cdot K^c$ is defined by the identity $(\widehat{J^c \cdot K^c}, \hat{f}) \equiv (J^c \cdot K^c, f)$, where $\hat{f}$ is the Fourier transform of f. Thus

$$\int d^4k [\widehat{J^c \cdot K^c}](k)\hat{f}(k) = \int d^4x [J^c \cdot K^c](x) f(x) = \int d^4x [\text{unreg } J^c \cdot K^c](x) \times \left\{ f(x) - \sum_{j=0}^{\theta} \frac{1}{j!} \partial^{\mu_1} \cdots \partial^{\mu_j} f(0) x_{\mu_1} \cdots x_{\mu_j} \right\}, \quad (23)$$

where [unreg $J^c \cdot K^c$] stands for the "unregularized" distribution product which is defined on the subspace of test functions that have $\theta + 1$-order zeros at the origin. Taking the Fourier transform of (23) and using the fact that the transform of the unregularized product is the (formal) convolution integral

$$[\widehat{\text{unreg } J^c K^c}](k) \equiv \int \frac{d^4p}{(2\pi)^4} \hat{J}^c(p)\hat{K}^c(k-p).$$

We obtain

$$\int d^4k [\widehat{J^c K^c}](k)\hat{f}(k) = \int d^4k [\widehat{\text{unreg } J^c K^c}](k)[\widehat{f - \textstyle\sum}](k) = \int d^4k \int \frac{d^4p}{(2\pi)^4} \hat{J}^c(p)\hat{K}^c(k-p)[\widehat{f - \textstyle\sum}](k), \quad (24)$$

where

$$[\widehat{f - \textstyle\sum}](k) \equiv \hat{f}(k) - \sum_{j=0}^{\theta} \frac{1}{j!} \partial^{\mu_1} \cdots \partial^{\mu_j} f(0) [\widehat{x_{\mu_1} \cdots x_{\mu_j}}](k).$$

But the Fourier transform of $x_{\mu_1} \cdots x_{\mu_j}$ is

$$[\widehat{x_{\mu_1} \cdots x_{\mu_j}}](k) = (i)^j (2\pi)^4 \partial_{\mu_1} \cdots \partial_{\mu_j} \delta(k).$$

Substituting this expression and the identity

$$(i)^j (2\pi)^4 \partial^{\mu_1} \cdots \partial^{\mu_j} f(0) = (-)^j \int d^4k' k'^{\mu_1} \cdots k'^{\mu_j} \hat{f}(k')$$

into (24), we find

$$\int d^4k [\widehat{J^c \cdot K^c}](k) f(k) = \int d^4k \int \frac{d^4p}{(2\pi)^4} \hat{J}^c(p)\hat{K}^c(k-p)\hat{f}(k) - \int d^4k \int d^4p \int d^4k' \hat{J}^c(p)\hat{K}^c(k-p) \times \sum_{j=0}^{\theta} \frac{(-)^j}{j!} (\partial_{\mu_1} \cdots \partial_{\mu_j} \delta(k)) k'^{\mu^1} \cdots k'^{\mu^j} \hat{f}(k'). \quad (25)$$

In the second term on the right-hand side of (25) we integrate out the k variable and then drop the primes to obtain

$$\int d^4k [\widehat{J^c K^c}](k)\hat{f}(k) = \int d^4k \int \frac{d^4p}{(2\pi)^4} \hat{J}^c(p) \left\{ \hat{K}^c(k-p) - \sum_{j=0}^{\theta} \frac{(-)^j}{j!} \partial_{\mu_1} \cdots \partial_{\mu_j} \hat{K}^c(-p) k^{\mu_1} \cdots k^{\mu_j} \right\} \hat{f}(k),$$

where

$$(-)^j \partial_{\mu_1} \cdots \partial_{\mu_j} \hat{K}^c(-p) \equiv \frac{\partial}{\partial k^{\mu_1}} \cdots \frac{\partial}{\partial k^{\mu_1}} \hat{K}^c(k-p) \bigg|_{k=0}$$

Therefore

$$[\widehat{J^c \cdot K^c}](k) = \int d^4p \tilde{J}^c(p)\Big\{\tilde{K}^c(p-k) - \sum_{j=0}^{\theta} \frac{(-)^j}{j!}\,\partial_{\mu_1}\cdots\partial_{\mu_j}\tilde{K}^c(-p)k^{\mu_1}\cdots k^{\mu_j}\Big\} \quad (26)$$

It is evident from the symmetry of the convolution operation, viz.,

$$\int d^4p\tilde{J}(p)\tilde{K}(k-p) = \int d^4p\tilde{K}(p)\tilde{J}(k-p),$$

that the rôles of $\tilde{J}^c$ and $\tilde{K}^c$ must be interchangeable in representation (26), i.e.,

$$[\widehat{J^c \cdot K^c}](k) = \int d^4p \tilde{K}^c(p)\Big\{\tilde{J}^c(k-p) - \sum_{j=0}^{\theta} \frac{(-)^j}{j!}\,\partial_{\mu_1}\cdots\partial_{\mu_j}\tilde{J}^c(-p)k^{\mu_1}\cdots k^{\mu_j}\Big\}, \quad (26')$$

in order that (26) be the Fourier transform of a uniquely defined product. This is indeed the case, as may be verified by a somewhat intricate calculation which consists of showing that the coefficients of various combinations of the $k^{\mu_1}\cdots k^{\mu_j}$ terms in the difference between (26) and (26′) are integrals of a differential of a three form which vanishes at infinity.

The Fourier transform of the distribution product (22′) is defined by the equation

$$\int d^4k \int d^4q[\widehat{J^c(x)K^c(x+u)L^c(u)}](k,q)\tilde{f}(k,q) \equiv \int d^4x \int d^4u J^c(x)K^c(x+u)L^c(u)f(x,u).$$

A calculation similar to the one given above yields

$$\begin{aligned}&[\widehat{J^c(x)K^c(x+u)L^c(u)}](k,q)\\ &= \int d^4p\tilde{K}^c(p)\Big\{\tilde{J}^c(k-p)\tilde{L}^c(q-p)\\ &\quad - \sum_{j=0}^{\theta}\frac{(-)^{j+r}}{(j+r)!}(\partial_{\mu_1}\cdots\partial_{\mu_j}\tilde{J}^c(-p))\\ &\quad \times (\partial_{\nu_1}\cdots\partial_{\nu_r}\tilde{L}^c(-p))k^{\mu_1}\cdots k^{\mu_j}q^{\nu_1}\cdots q^{\nu_r}\Big\}. \quad (27)\end{aligned}$$

Thus, the effect in momentum space of taking the Hadamard finite part of a product in configuration space is to subtract from the integrand of the divergent convolution integral obtained by formal Fourier transformation of the product, the first few terms in its Maclaurin expansion.

V. EXAMPLES FROM QUANTUM ELECTRODYNAMICS

A. Vacuum Polarization

The scattering amplitude corresponding to second-order vacuum polarization (Fig. 7) is given, before

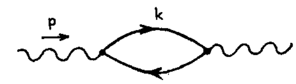


FIG. 7. Scattering amplitude correspondence to second-order vacuum polarization.

regularization and renormalization, by the divergent convolution integral

$$\text{unreg}\prod_{\mu\nu}(p) = \frac{e^e}{(2\pi)^4 i}\,\mathrm{Tr}\int d^4k\gamma_\mu \frac{1}{\not{k}-M+i\epsilon}\gamma_\nu\frac{1}{\not{k}-\not{p}-M+i\epsilon},$$

where $\not{a} \equiv \gamma^\mu a_\mu$. According to the definition of distribution products given in the preceding section, we should instead describe this process by the convergent integral

$$\begin{aligned}\prod_{\mu\nu}(p) &= \frac{e^2}{(2\pi)^4 i}\,\mathrm{Tr}\int d^4k\gamma_\mu\frac{1}{\not{k}-M+i\epsilon}\gamma_\nu\\ &\times\Big\{\frac{1}{\not{k}-\not{p}-M+i\epsilon} - \sum_{j=0}^{2}\frac{1}{j!}\Big[\frac{\partial}{\partial p^{\mu_1}}\cdots\frac{\partial}{\partial p^{\mu_j}}\\ &\times\frac{1}{\not{k}-\not{p}-M+i\epsilon}\Big]_{k=0}p^{\mu_1}\cdots p^{\mu_j}\Big\},\end{aligned}$$

the upper limit of summation $\theta = 2$ being the index of divergence of the kernel product $\mathcal{S}(z)\mathcal{S}(-z)$ that occurs in the space–time description of this process. Carrying out the indicated differentiations and simplifying we find

$$\begin{aligned}\prod_{\mu\nu}(p) &= \frac{e^2}{(2\pi)^4 i}\,\mathrm{Tr}\int d^4k\gamma_\mu\frac{1}{\not{k}-M+i\epsilon}\\ &\times\gamma_\nu\frac{1}{\not{k}-\not{p}-M+i\epsilon}\not{p}\frac{1}{\not{k}-M+i\epsilon}\\ &\times\not{p}\frac{1}{\not{k}-M+i\epsilon}\not{p}\frac{1}{\not{k}-M+i\epsilon}.\end{aligned}$$

By the standard techniques, this integral can be evaluated in a straightforward way to yield

$$\begin{aligned}\prod_{\mu\nu}(p) &= \frac{-e^2}{(2\pi)^2}(p_\mu p_\nu - g_{\mu\nu}p^2)\\ &\times\int_0^1 d\xi(1-\xi)\xi\log\Big(\frac{M^2}{M^2-\xi(1-\xi)p^2}\Big). \quad (28)\end{aligned}$$

B. Electron Self-Energy

Before regularization and renormalization, the second-order electron self-energy process (Fig. 8) is

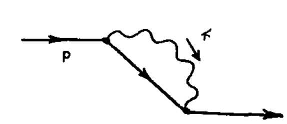


FIG. 8. The second-order electron self-energy process

described by the divergent convolution integral

unreg $\sum(p)$

$$= \frac{e^2}{(2\pi)^4 i} \int d^4k \frac{1}{k^2 + i\epsilon} \gamma^\mu \frac{1}{\not p - \not k - M + i\epsilon} \gamma_\mu .$$

Applying the definitions of Sec. IV, we replace this expression by the integral

$$\begin{aligned} \sum(p) &= \frac{e^2}{(2\pi)^4 i} \int d^4k \frac{1}{k^2 + i\epsilon} \\ &\quad \times \gamma^\mu \left\{ \frac{1}{\not p - \not k - M + i\epsilon} - \frac{1}{-\not k - M + i\epsilon} \right. \\ &\quad \left. + \frac{1}{-\not k - M + i\epsilon} \, p \, \frac{1}{-\not k - M + i\epsilon} \right\} \gamma_\mu \\ &= \frac{e^2}{(2\pi)^4 i} \int d^4k \frac{1}{k^2 + i\epsilon} \gamma^\mu \frac{1}{\not p - \not k - M + i\epsilon} \\ &\quad \times \not p \frac{1}{-\not k - M + i\epsilon} \not p \frac{1}{-\not k - M + i\epsilon} \gamma_\mu , \end{aligned}$$

where we have subtracted from $(\not p - \not k - M + i\epsilon)^{-1}$ the first two terms in its Maclaurin expansion in accordance with the fact that the kernel product $\mathcal{S}(z)\mathcal{T}(z)$, which occurs in the space–time description of electron self-energy, has index of divergence $\theta = 1$. A comparatively simple calculation gives

$$\sum(p) = \frac{e^2}{8\pi^2} \int_0^1 d\xi (2M - \xi \not p) \log \left(\frac{M^2}{M^2 - \xi p^2} \right) . \quad (29)$$

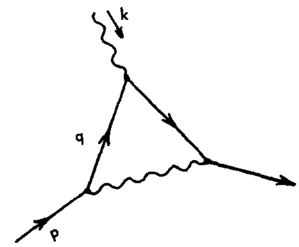


FIG. 9. Vertex part.

C. The Vertex Part

The vertex part (Fig. 9) is described by the divergent convolution integral

$$\begin{aligned} \text{unreg } \Gamma^\nu(p, k) &= \frac{e^2}{(2\pi)^4 i} \int d^4q \frac{1}{(p - q)^2 + i\epsilon} \\ &\quad \times \gamma^\mu \frac{1}{\not q - \not k - M + i\epsilon} \gamma^\nu \frac{1}{\not q - M + i\epsilon} \gamma_\mu . \end{aligned}$$

When this amplitude is modified according to (27) (with $\theta = 0$), we obtain the expression

$\Gamma^\nu(p, k)$

$$\begin{aligned} &= \frac{e^2}{(2\pi)^4 i} \int d^4q \left\{ \frac{1}{(p - \not q)^2 + i\epsilon} \gamma^\mu \frac{1}{\not q - \not k - M + i} \right. \\ &\quad \left. - \frac{1}{q^2 + i\epsilon} \gamma^\mu \frac{1}{\not q - M + i\epsilon} \right\} \gamma^\nu \frac{1}{\not q - M + i\epsilon} \gamma_\mu \end{aligned}$$

A standard-type calculation yields

$$\begin{aligned} \Gamma^\nu(p, k) = \frac{e^2}{8\pi^2} \Biggl\{ &-\gamma^\nu + \gamma^\nu \int_0^1 d\xi \int_0^{1-\xi} d\eta \log \left(\frac{(1 - \xi)m^2}{(1 - \xi)m^2 - \xi(1 - \xi)p^2 - \eta(1 - \eta)k^2 - 2\xi\eta p \cdot k} \right) \\ &+ \int_0^1 d\xi \int_0^1 d\eta \frac{(1 - \xi)m^2 - \xi(1 - \xi)p^2 - \eta(1 - \eta)k^2 - 2\xi\eta p \cdot k}{\gamma^\nu m^2 - 2mk^\nu + (\not p\xi - \not k\eta) \times \gamma^\nu(\not p\xi + \not k(1 - \eta)) - 4m(p^\nu\xi \cdot k^\nu\eta)} \Biggr\} . \quad (30) \end{aligned}$$

It may be easily verified by direct computation that amplitudes (29) and (30) are related by the Ward identity $\partial_\nu \sum(p) = \Gamma_\nu(p, 0)$.

D. Finite Renormalization

In this section we shall use an asterisk to indicate quantities that are renormalized to second order.

The vacuum-polarization amplitude $\prod_{\mu\nu}$ given by (28) agrees precisely with the renormalized amplitude $\prod_{\mu\nu}^*$ that is obtained for this process by means of the standard regularization and subtraction procedures

$$\prod_{\mu\nu}^* (p) = \prod_{\mu\nu} (p), \quad (31)$$

and consequently it requires no renormalization. Thus to second order we have $Z_3 = 1$.

The renormalized second-order electron self-energy amplitude $\sum^*$ is characterized by the (formal) properties

$$\sum^* (p) \big|_{p=M} = 0 \quad \text{and} \quad \frac{\partial}{\partial \not p} \sum (p) \big|_{\not p=M} = 0.$$

This amplitude differs from (29) by a first-degree polynomial in $\not p$

$$\sum^* (p) = \sum (p) - \delta M - (1 - z_2)(\not p - M), \quad (32)$$

where

$$\delta M = \sum (p) \big|_{\not p=M} = \frac{5}{32} \frac{e^2}{\pi^2} M,$$

and

$$(1 - z_2) = \frac{\partial}{\partial \not p} \sum (p) \big|_{p=M} . \quad (33)$$

[The quantities (29), (30), and (33) contain an infrared divergence; therefore during intermediate calculations we shall assume that the photon has an infinitesimal mass.]

The renormalized vertex part amplitude differs from expression (30) by a constant

$$\Gamma^*(p, k) = \Gamma_\nu(p, k) + (1 - z_1)\gamma_\nu. \qquad (34)$$

Taking $z_2 = z_1$, Γ^*_ν will satisfy the Ward identity with respect to $\sum^*$.

The effect of subtractions (32), (34), and trivially (31) is the finite renormalization

$$M \to M^* = M + \delta M \approx (1 + \tfrac{1}{689})M,$$
$$e \to e^* = \frac{z_1}{z_2(z_3)^{\frac{1}{2}}} e = e.$$

Thus, when second-order scattering amplitudes are calculated in accordance with the definitions of distribution products given in Sec. IV, no ultraviolet divergences arise and the only renormalization required is that corresponding to a small increment in electron mass.

The definition of distribution products given in this paper constitutes a type of regularization which has been anticipated by a number of authors; in particular, see the extensive work by Caianiello.[8] Indeed, the "finite part" integrals (17′) and (22′) are special realizations of the Caianiello "$\mathcal{J}$ integrals." However, it should be noted that (17′) and (22′) are not precisely "finite parts" in the strict sense of Hadamard, since we do not first split off the polar and logarithmic singularities of a nonintegrable product of particle kernels and apply the regularization only to these singular terms. Besides simplicity, the advantage of applying the regularization to the entire product lies in the fact that the exponential decay at infinity of kernels corresponding to particles of nonzero mass allows us to avoid the introduction of regularization cut-offs and the nonuniqueness that such cutoffs entail. Unfortunately, when particles of zero mass are involved, infrared divergences are encountered. However, this is a well-understood problem which can be handled by the introduction of infinitesimal masses in the intermediate calculations.

It is possible to generalize the definition of distribution multiplication given in Sec. IV.A to include all such products of causal propagators as arise in the perturbation description of scattering. This generalization will be described in a subsequent paper soon to be published.

The author wishes to express his sincere appreciation to Professor M. M. Schiffer whose questions began the work contained in this paper and whose criticism and suggestions (particularly the one concerning the use of Hadamard's finite part for regularization) have been invaluable.

[8] E. R. Caianello, Nuovo Cimento **13**, 637 (1959); **14**, 185 (1959); **18**, 505 (1960).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Coherent Soft-Photon States and Infrared Divergences. I. Classical Currents*

T. W. B. Kibble
Department of Physics, Imperial College, London, England

(Received 1 May 1967)

As a first step toward a treatment of soft-photon processes which is free of infrared divergences and avoids the necessity of introducing a fictitious photon mass, the specification of asymptotic photon states belonging to non-Fock representations is discussed. As in the work of Chung, a basis consisting of generalized coherent states is used, but in contrast to his work, these states are rigorously defined in terms of von Neumann's infinite tensor product. It is shown that the states must be given an additional label which serves to distinguish various "weakly equivalent" vectors, and which corresponds formally to an infinite phase factor. A nonseparable Hilbert space $\mathcal{H}_{\mathrm{IR}}$ is defined (as a subspace of the infinite tensor-product space) which may be regarded as the space of all possible asymptotic photon states. The interaction of the electromagnetic field with a prescribed classical current distribution is discussed, and it is shown that a unitary S operator, all of whose matrix elements are finite, may be defined on $\mathcal{H}_{\mathrm{IR}}$.

1. INTRODUCTION

The basic principles underlying the infrared divergence problem of quantum electrodynamics have been well understood since the classic paper of Bloch and Nordsieck.[1] The comprehensive treatments given by Yennie, Frautschi, and Suura,[2] Eriksson,[3] and others[4] provide a consistent method of calculating infrared divergence-free transition probabilities to all

* The research reported in this document has been sponsored in part by the Air Force Office of Scientific Research OAR through the European Office Aerospace Research, U.S. Air Force.

[1] F. Bloch and A. Nordsieck, Phys. Rev. **52**, 54 (1937).

[2] D. R. Yennie, S. C. Frautschi, and H. Suura, Ann. Phys. (N.Y.) **13**, 379 (1961).

[3] K. E. Eriksson, Nuovo Cimento **19**, 1010 (1961).

[4] J. M. Jauch and F. Rohrlich, Helv. Phys. Acta **27**, 613 (1954); R. J. Glauber, Phys. Rev. **84**, 395 (1951); E. L. Lomon, Nucl. Phys. **1**, 101 (1956); Phys. Rev. **113**, 726 (1959); S. Weinberg, Phys. Rev. **140**, B516 (1965).

The renormalized vertex part amplitude differs from expression (30) by a constant

$$\Gamma^*(p, k) = \Gamma_\nu(p, k) + (1 - z_1)\gamma_\nu. \qquad (34)$$

Taking $z_2 = z_1$, Γ^*_ν will satisfy the Ward identity with respect to $\sum^*$.

The effect of subtractions (32), (34), and trivially (31) is the finite renormalization

$$M \to M^* = M + \delta M \approx (1 + \tfrac{1}{689})M,$$
$$e \to e^* = \frac{z_1}{z_2(z_3)^{\frac{1}{2}}}\, e = e.$$

Thus, when second-order scattering amplitudes are calculated in accordance with the definitions of distribution products given in Sec. IV, no ultraviolet divergences arise and the only renormalization required is that corresponding to a small increment in electron mass.

The definition of distribution products given in this paper constitutes a type of regularization which has been anticipated by a number of authors; in particular, see the extensive work by Caianiello.[8] Indeed, the "finite part" integrals (17′) and (22′) are special realizations of the Caianiello "$\mathcal{J}$ integrals." However, it should be noted that (17′) and (22′) are not precisely "finite parts" in the strict sense of Hadamard, since we do not first split off the polar and logarithmic singularities of a nonintegrable product of particle kernels and apply the regularization only to these singular terms. Besides simplicity, the advantage of applying the regularization to the entire product lies in the fact that the exponential decay at infinity of kernels corresponding to particles of nonzero mass allows us to avoid the introduction of regularization cut-offs and the nonuniqueness that such cutoffs entail. Unfortunately, when particles of zero mass are involved, infrared divergences are encountered. However, this is a well-understood problem which can be handled by the introduction of infinitesimal masses in the intermediate calculations.

It is possible to generalize the definition of distribution multiplication given in Sec. IV.A to include all such products of causal propagators as arise in the perturbation description of scattering. This generalization will be described in a subsequent paper soon to be published.

The author wishes to express his sincere appreciation to Professor M. M. Schiffer whose questions began the work contained in this paper and whose criticism and suggestions (particularly the one concerning the use of Hadamard's finite part for regularization) have been invaluable.

[8] E. R. Caianello, Nuovo Cimento **13**, 637 (1959); **14**, 185 (1959); **18**, 505 (1960).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Coherent Soft-Photon States and Infrared Divergences. I. Classical Currents*

T. W. B. Kibble
Department of Physics, Imperial College, London, England

(Received 1 May 1967)

As a first step toward a treatment of soft-photon processes which is free of infrared divergences and avoids the necessity of introducing a fictitious photon mass, the specification of asymptotic photon states belonging to non-Fock representations is discussed. As in the work of Chung, a basis consisting of generalized coherent states is used, but in contrast to his work, these states are rigorously defined in terms of von Neumann's infinite tensor product. It is shown that the states must be given an additional label which serves to distinguish various "weakly equivalent" vectors, and which corresponds formally to an infinite phase factor. A nonseparable Hilbert space $\mathcal{H}_{IR}$ is defined (as a subspace of the infinite tensor-product space) which may be regarded as the space of all possible asymptotic photon states. The interaction of the electromagnetic field with a prescribed classical current distribution is discussed, and it is shown that a unitary S operator, all of whose matrix elements are finite, may be defined on $\mathcal{H}_{IR}$.

1. INTRODUCTION

The basic principles underlying the infrared divergence problem of quantum electrodynamics have been well understood since the classic paper of Bloch and Nordsieck.[1] The comprehensive treatments given by Yennie, Frautschi, and Suura,[2] Eriksson,[3] and others[4] provide a consistent method of calculating infrared divergence-free transition probabilities to all

* The research reported in this document has been sponsored in part by the Air Force Office of Scientific Research OAR through the European Office Aerospace Research, U.S. Air Force.

[1] F. Bloch and A. Nordsieck, Phys. Rev. **52**, 54 (1937).

[2] D. R. Yennie, S. C. Frautschi, and H. Suura, Ann. Phys. (N.Y.) **13**, 379 (1961).

[3] K. E. Eriksson, Nuovo Cimento **19**, 1010 (1961).

[4] J. M. Jauch and F. Rohrlich, Helv. Phys. Acta **27**, 613 (1954); R. J. Glauber, Phys. Rev. **84**, 395 (1951); E. L. Lomon, Nucl. Phys. **1**, 101 (1956); Phys. Rev. **113**, 726 (1959); S. Weinberg, Phys. Rev. **140**, B516 (1965).

orders in the fine-structure constant. They are nevertheless not entirely logically satisfactory. For to obtain a finite result it is necessary to give the photon a nonzero mass λ (or to cut off the integrals in the low-energy region), to compute the sum of an infinite number of terms, representing the probabilities for the emission of n soft photons, and only later to take the limit $\lambda \to 0$. In this limit each individual term in the sum vanishes (when computed to all orders in α) although the sum remains finite.

It has long been realized[5] that the fundamental reason for this unfortunate circumstance lies in the assumption, implicit in the conventional perturbation calculation, that the asymptotic states contain finite numbers of photons and so belong to the familiar Fock representation of the canonical commutation relations. In reality, if the initial state contains a finite number of photons, the final state will in general contain infinitely many soft photons, though with finite total energy. A correct description of this process must therefore involve the use of representations unitarily inequivalent to the Fock representation. If we are to allow the initial and final state to belong to any one of the possible representations, then we are forced to use a nonseparable Hilbert space, which may be decomposed into the direct sum of an uncountably infinite number of separable Hilbert spaces.

A significant step in this direction was provided by the paper of Chung[6] which aimed to show that one can find certain states, outside the Hilbert space of the usual Fock representation, between which the S-matrix elements are finite. Two criticisms of this approach can be made, however. One is that it still retains, as a calculational tool in the intermediate steps, the device of giving the photon a finite mass. This is inevitable in any method which relies on summing Feynman diagrams. The second point is perhaps rather more technical, but also more important if one wishes to place the theory on a satisfactory mathematical foundation. The asymptotic states are taken to be coherent states, suitably generalized to allow an infinite expectation value for the total photon number. However these states are only rather loosely defined, and as a result the calculated matrix elements contain infinite phase factors which are simply discarded. Of course an over-all phase factor in the S matrix is unobservable and it might not be unreasonable to discard such a factor even if it were divergent. However the relative phases of different S-matrix elements are in general observable, and should be retained.

It is one of the main aims of the present paper to remedy this defect, and to show that generalized coherent states can be defined in a rigorous way, and that the S-matrix elements between them are finite. This provides a more rigorous justification for the approach adopted by Chung. It is natural to define these states in terms of the infinite tensor product of Hilbert spaces introduced by von Neumann.[7] In fact we shall define a nonseparable subspace $\mathcal{H}_{\text{IR}}$ of this tensor-product space which contains all the states belonging to representations of interest in the discussion of the infrared problem. The space decomposes into the direct sum of uncountably many separable Hilbert spaces in each of which the representation of the canonical commutation relations is irreducible. The representations so obtained are among those which have been discussed by Klauder, McKenna, and Woods,[8] and our work at this point may be regarded as a specialization of theirs.

An important feature of von Neumann's treatment of the infinite tensor product is the distinction he draws between equivalence and weak equivalence of product vectors. Equivalent vectors belong to the same irreducible representation, and have well-defined finite scalar products. Weakly equivalent (but non-equivalent) vectors on the other hand belong to irreducible representations which are unitarily equivalent[8] but defined on distinct, mutually orthogonal, separable subspaces of the tensor-product space. Their scalar products are represented formally by infinite products with divergent phase factors to which the value zero is conventionally assigned. A particular physical state may thus be represented by any one of an uncountably infinite class of weakly equivalent vectors. It follows that the conventional labeling of states is not adequate to select a particular vector (or even ray) in the tensor-product space, and must be supplemented by an additional label which serves to distinguish the various equivalence classes within each weak equivalence class. (One can, however, define a generalization of the concept of a ray such that the generalized rays are in one-to-one correspondence with the physical states of the system.)

We begin in Sec. 2 with a brief discussion of the photon states in the familiar Fock representation. This discussion serves mainly to introduce the notation, and to summarize some of the properties of the coherent states. The explicit construction of

[5] See, for example, K. O. Friedrichs, *Mathematical Aspects of the Quantum Theory of Fields* (Interscience Publishers, Inc., New York, 1953), Secs. 14 and 19.

[6] V. Chung, Phys. Rev. **140**, B1110 (1965).

[7] J. von Neumann, Composito Math. **6**, 1 (1938).

[8] J. R. Klauder and J. McKenna, J. Math. Phys. **6**, 68 (1965); J. R. Klauder, J. McKenna, and E. J. Woods, J. Math. Phys. **7**, 822 (1966).

generalized coherent states belonging to the various infrared representations in terms of the infinite tensor-product space is carried out in Sec. 3. An important role in this discussion is played by certain unitary operators defined as tensor products of unitary operators on the Hilbert spaces from which the tensor-product space is formed. These operators essentially provide an extended representation of the canonical commutation relations on the entire tensor-product space, and by restriction on its subspace $\mathcal{H}_{\mathrm{IR}}$.

In Sec. 4 we consider the interaction of the electromagnetic field with a given classical current distribution. We compute the scattering matrix elements between generalized coherent states, and show that a unitary S operator can be defined on the space $\mathcal{H}_{\mathrm{IR}}$. Indeed we show that the S operator may be identified with one of the unitary operators already introduced. The conclusions are summarized in Sec. 5.

The interaction of the electromagnetic field with other quantized fields may be handled in a similar way. However, the problem is somewhat complicated in this case by renormalization problems and by the need to consider wavefunctions for the charged particles with the characteristic Coulomb asymptotic behavior. This problem will therefore be deferred to a subsequent paper.

2. PHOTON STATES IN THE FOCK REPRESENTATION

Mainly in order to fix our notation, we begin by discussing the states of a free electromagnetic field in the usual Fock representation.

The field may be described by a vector potential of the form

$$A_\mu(x) = \int \frac{d\mathbf{k}}{(2\pi)^3 2k^0} \, [a_\mu(\mathbf{k}) e^{ik\cdot x} + a_\mu^*(\mathbf{k}) e^{-ik\cdot x}], \quad (1)$$

with $k^0 = |\mathbf{k}|$ and $k \cdot x = \mathbf{k} \cdot \mathbf{x} - k^0 x^0$. The creation and annihilation operators satisfy the commutation relations

$$[a_\mu(\mathbf{k}), a_\nu^*(\mathbf{k}')] = \gamma_{\mu\nu}(\mathbf{k})(2\pi)^3 2k^0 \delta_3(\mathbf{k} - \mathbf{k}'), \quad (2)$$

where $\gamma_{\mu\nu}(\mathbf{k})$ is a function depending on the choice of gauge. Since we wish to use a Hilbert space with positive–definite metric containing only physical states, we must choose $\gamma_{\mu\nu}(\mathbf{k})$ to be a projection matrix of rank 2, orthogonal to k_μ; that is, we require

$$\begin{aligned} \gamma_\mu^\rho(\mathbf{k})\gamma_{\rho\nu}(\mathbf{k}) &= \gamma_{\mu\nu}(\mathbf{k}), \\ \gamma_\mu^\mu(\mathbf{k}) &= 2, \\ k^\mu \gamma_{\mu\nu}(\mathbf{k}) &= 0. \end{aligned} \quad (3)$$

Its most general form is

$$\gamma_{\mu\nu}(\mathbf{k}) = g_{\mu\nu} - k_\mu l_\nu^*(\mathbf{k}) - l_\mu(\mathbf{k}) k_\nu, \quad (4)$$

where $l^* \cdot l = 0$ and $k \cdot l = 0$. For example, in the radiation gauge, the only nonvanishing components are

$$\gamma_{ij}(\mathbf{k}) = \delta_{ij} - k_i k_j / \mathbf{k}^2, \quad (5)$$

corresponding to the choice

$$l = (-k^0, \mathbf{k})/2\mathbf{k}^2. \quad (6)$$

Now let us consider the labeling of single-photon states. We introduce first the space $\mathcal{K}_P$ of photon wavefunctions $f^\mu(\mathbf{k})$ which satisfy the conditions:

$$k^\mu f_\mu(\mathbf{k}) = 0, \quad (7)$$

and

$$\int \frac{d\mathbf{k}}{(2\pi)^3 2k^0} f_\mu^*(\mathbf{k}) f^\mu(\mathbf{k}) < \infty. \quad (8)$$

We denote the scalar product in this space by

$$(f^* g) = \int \frac{d\mathbf{k}}{(2\pi)^3 2k^0} f_\mu^*(\mathbf{k}) g^\mu(\mathbf{k}). \quad (9)$$

This scalar product is positive–semidefinite, but not definite, since $\mathcal{K}_P$ contains a subspace $\mathcal{K}_Z$ of vectors of zero norm, namely those of the form

$$f^\mu(\mathbf{k}) = k^\mu g(\mathbf{k}), \quad (10)$$

which by virtue of (7) are orthogonal to all vectors in $\mathcal{K}_P$. Two wavefunctions which differ by an element of $\mathcal{K}_Z$ describe the same physical one-photon state. In other words, the physical one-photon Hilbert space $\mathcal{K}$ may be identified with the factor space $\mathcal{K}_P/\mathcal{K}_Z$. In $\mathcal{K}$ the scalar product defined by (8) is of course positive–definite. The choice of a gauge amounts to choosing a particular representative from each of these equivalence classes of wavefunctions, namely, the one which satisfies

$$\gamma_{\mu\nu}(\mathbf{k}) f^\nu(\mathbf{k}) = f_\mu(\mathbf{k}). \quad (11)$$

Now let us consider the Hilbert space $\mathcal{H}$ of the Fock representation. Corresponding to each $f \in \mathcal{K}_P$ we may introduce creation and annihilation operators denoted, by an obvious extension of the scalar-product notation (9) by $(a^* f)$ and $(f^* a)$. These operators are gauge invariant; that is to say they are unchanged either by making a gauge transformation of the operators

$$a_\mu(\mathbf{k}) \to a_\mu(\mathbf{k}) + k_\mu \lambda(\mathbf{k})$$

or by adding to f an element of $\mathcal{K}_Z$. They have therefore a well-defined meaning as operators on $\mathcal{H}$, independently of any choice of gauge, and may be regarded as being labeled by elements of $\mathcal{K}$ rather than $\mathcal{K}_P$. From the commutation relations (2) together with (4) and (7) one derives the commutators

$$[(f^* a), (a^* g)] = (f^* g). \quad (12)$$

It is often more convenient to re-express these relations in terms of the unitary operators obtained by exponentiation,

$$U(f) = e^{(a^*f)-(f^*a)} = e^{(a^*f)}e^{-(f^*a)}e^{-\frac{1}{2}(f^*f)}, \tag{13}$$

in the form

$$(Uf)U(g) = U(f+g)e^{\frac{1}{2}[(g^*f)-(f^*g)]}. \tag{14}$$

The coherent states[9] in $\mathcal{H}$ are labeled by the elements of $\mathcal{K}$, and are defined by

$$|f\rangle = U(f)\,|0\rangle, \tag{15}$$

where $|0\rangle$ is the vacuum state defined by the condition

$$(g^*a)\,|0\rangle = 0, \tag{16}$$

for all $g \in \mathcal{K}$. These states are all eigenstates of the annihilation operators (g^*a):

$$(g^*a)\,|f\rangle = |f\rangle(g^*f). \tag{17}$$

It follows that these states are in one-to-one correspondence with classical solutions of the wave equation. In fact

$$\langle f|\, A_\mu(x)\,|f\rangle = F_\mu(x) = \int \frac{d\mathbf{k}}{(2\pi)^3 2k^0}\,[f_\mu(\mathbf{k})e^{ik\cdot x} + f_\mu^*(\mathbf{k})e^{-ik\cdot x}]. \tag{18}$$

They form an overcomplete family of states in $\mathcal{H}$ with the scalar products

$$\langle g\,|\,f\rangle = e^{(g^*f)-\frac{1}{2}(g^*g)-\frac{1}{2}(f^*f)}, \tag{19}$$

and the formal completeness relation

$$\int |f\rangle[df]\langle f| = 1. \tag{20}$$

The measure $[df]$ may most simply be defined in terms of an orthonormal basis $\{e_n\}$ in $\mathcal{K}$. If we set $f_n = (e_n^*f)$ then

$$[df] = \prod_n \frac{d^2f_n}{\pi} = \prod_n \frac{d(\operatorname{Re} f_n)\,d(\operatorname{Im} f_n)}{\pi}.$$

(This definition can be made rigorous, but we shall not make it so here.)

Finally we note the effect of translation operators $e^{iP\cdot a}$ on the coherent states. One has

$$e^{iP\cdot a}\,|f\rangle = |f_a\rangle, \tag{21}$$

where

$$f_a^\mu(\mathbf{k}) = e^{ik\cdot a}f^\mu(\mathbf{k}).$$

In the following section we consider the problem of defining generalized coherent states corresponding to wavefunctions $f^\mu(\mathbf{k})$ for which $(f^*f) = \infty$.

3. INFRARED REPRESENTATIONS

This section is devoted to an explicit construction of a certain class of "infrared" representations of the canonical commutation relations. These representations are among those which have been discussed by Klauder, McKenna, and Woods,[8] using the infinite tensor-product spaces introduced by von Neumann.[7] Our discussion differs from theirs mainly in the emphasis we place on the role of the coherent states. We also define and discuss a certain nonseparable subspace of the tensor-product space, which will play an important role in our later discussions.

Although we are of course primarily interested in the case of the electromagnetic field, our discussion in this section will be more general. We assume that we are given a one-particle Hilbert space $\mathcal{K}$ in which a scalar product (f^*g) is defined, and seek representations of unitary operators $U(f)$ satisfying (14). In fact, when we come to look for representations other than the Fock representation, it is too strong a requirement to demand that every operator $U(f)$ be represented. We shall require this not for all $f \in \mathcal{K}$ but only for those in some dense linear subspace $\mathcal{A}$ of $\mathcal{K}$. It will not be necessary at this stage to specify $\mathcal{A}$ more precisely. In fact, however, for application to the infrared problem, the essential feature of the functions $f \in \mathcal{A}$ is that they should vanish appropriately at $\mathbf{k} = \mathbf{0}$.

Now let $\{e_n\}$ be an orthonormal basis in $\mathcal{K}$ such that each $e_n \in \mathcal{A}$. For each $f \in \mathcal{K}$ we denote by $\{f_n\}$ the sequence of complex numbers $f_n = (e_n^*f)$. For later convenience we also define $\mathcal{A}^*$ to be the dual vector space to $\mathcal{A}$. Thus for each $f \in \mathcal{A}$ and $g \in \mathcal{A}^*$, the "scalar products" (f^*g) and (g^*f) are well defined. In particular, since $e_n \in \mathcal{A}$ we can extend the representation by sequences from $\mathcal{K}$ to $\mathcal{A}^*$, and for each $g \in \mathcal{A}^*$ define $g_n = (e_n^*g)$, although of course the sequence $\{g_n\}$ need not be square-summable or even bounded. Clearly also the scalar product may be represented by the absolutely convergent sum $(f^*g) = \sum_n f_n^*g_n$. Note that $\mathcal{A} \subset \mathcal{K} \subset \mathcal{A}^*$. In our case $\mathcal{A}^*$ will represent the extended space of possible photon wavefunctions.

For each n, let $\mathcal{H}_n$ be a Hilbert space on which an irreducible representation of the commutation relations for a single degree of freedom

$$[a_n, a_n^*] = 1 \tag{22}$$

is defined. We denote the corresponding unitary

[9] J. Schwinger, Phys. Rev. **91**, 728 (1953); R. J. Glauber, Phys. Rev. **131**, 2766 (1963); S. S. Schweber, J. Math. Phys. **3**, 831 (1962); V. Bargmann, Proc. Natl. Acad. Sci. U.S. **48**, 199 (1962).

operators defined for all complex numbers z by

$$U_n(z) = \exp(a_n^* z - z^* a_n). \tag{23}$$

The vacuum state $\psi_{0n} \in \mathcal{H}_n$ is defined by

$$a_n \psi_{0n} = 0. \tag{24}$$

We now recall the definition of the infinite tensor-product space

$$\mathcal{H}_\otimes = \prod_n \otimes \mathcal{H}_n. \tag{25}$$

To every sequence $\psi = \{\psi_n\}$ of normalized vectors $\psi_n \in \mathcal{H}_n$, $\|\psi_n\| = 1$, we assign a product vector

$$|\psi\rangle = \prod_n \otimes \psi_n.$$

(The map from sequences to product vectors is not required to be one-to-one. The conditions for the product vectors defined by two sequences to be equal will be specified below.) Two such product vectors $|\phi\rangle$ and $|\psi\rangle$ are equivalent, $|\phi\rangle \sim |\psi\rangle$, if and only if

$$\sum_n |1 - (\phi_n, \psi_n)| < \infty. \tag{26}$$

In that case, their scalar product is defined to be the convergent product

$$\langle \phi \mid \psi \rangle = \prod_n (\phi_n, \psi_n). \tag{27}$$

Scalar products of inequivalent product vectors, represented formally by divergent products, are conventionally assigned the value 0.

Now consider the vector space $\mathcal{H}'_\otimes$ consisting of all finite linear combinations of product vectors, with scalar product defined by linearity from (27). In order to ensure that the scalar product so defined is positive definite, two such linear combinations are defined to be equal if their scalar products with all product vectors are equal. Then the complete direct-product space $\mathcal{H}_\otimes$ is the completion of $\mathcal{H}'_\otimes$ with respect to the norm so defined. It is a nonseparable Hilbert space. However, the closed linear subspace $\mathcal{H}_\otimes(\psi)$ spanned by product vectors equivalent to a given product vector $|\psi\rangle$ is separable. (This space is called the "incomplete direct-product space determined by $|\psi\rangle$," although the name is rather unfortunate since the space is complete!)

Let us now consider a sequence $\{U_n\}$ of unitary operators, with U_n an operator on $\mathcal{H}_n$. Then we show that one can define a tensor-product operator

$$U_\otimes = \prod_n \otimes U_n \tag{28}$$

as an operator on $\mathcal{H}_\otimes$ by its action on product vectors,

$$U_\otimes \prod_n \otimes \psi_n = \prod_n \otimes U_n \psi_n. \tag{29}$$

Clearly the right-hand side is again a product vector. To show that $U_\otimes$ defines an operator on $\mathcal{H}'_\otimes$, it is enough to show that if

$$|\psi\rangle = \sum_{j=1}^{r} c_j \prod_n \otimes \psi_{jn} = 0,$$

then also $U_\otimes |\psi\rangle = 0$. But

$$\begin{aligned} 0 = \langle \psi \mid \psi \rangle &= \sum_{j=1}^{r} \sum_{k=1}^{r} c_j^* c_k \prod_n (\psi_{jn}, \psi_{kn}) \\ &= \sum_{j=1}^{r} \sum_{k=1}^{r} c_j^* c_k \prod_n (U_n \psi_{jn}, U_n \psi_{kn}) \\ &= \|U_\otimes |\psi\rangle\|^2. \end{aligned}$$

(By convention, terms involving divergent products are zero.) This shows also that on $\mathcal{H}'_\otimes$, $U_\otimes$ is unitary. It can therefore be extended by continuity to a unitary operator on $\mathcal{H}_\otimes$.

Note that if $|\phi\rangle \sim |\psi\rangle$, then also $U_\otimes |\phi\rangle \sim U_\otimes |\psi\rangle$. Hence $U_\otimes$ maps each incomplete direct product space $\mathcal{H}_\otimes(\psi)$ isometrically onto $\mathcal{H}_\otimes(\psi')$ with $|\psi'\rangle = U_\otimes |\psi\rangle$. Only in special cases is it true however that $U_\otimes |\psi\rangle \sim |\psi\rangle$. When this is the case, then $U_\otimes$ defines a unitary operator on $\mathcal{H}_\otimes(\psi)$.

We note also that if $\{U_n\}$ and $\{V_n\}$ are two sequences of unitary operators, then

$$U_\otimes V_\otimes = \prod_n \otimes U_n \cdot \prod_n \otimes V_n = \prod_n \otimes (U_n V_n). \tag{30}$$

We now define certain unitary tensor-product operators. Let $\lambda = \{\lambda_n\}$ be any sequence of real numbers. Then we define[7]

$$V_\otimes(\lambda) = \prod_n \otimes e^{i\lambda_n}. \tag{31}$$

If the series $\sum_n \lambda_n$ converges absolutely, we denote its sum by

$$(\lambda) = \sum_n \lambda_n. \tag{32}$$

(This notation is consistent with our earlier notation for scalar products.) In that case,

$$V_\otimes(\lambda) = e^{i(\lambda)}. \tag{33}$$

However if the series $\sum_n |\lambda_n|$ diverges, then $V_\otimes(\lambda)$ is *not* a multiple of the identity operator on $\mathcal{H}_\otimes$. Even in that case however, two sequences define the same operator if their difference yields a series which converges absolutely to zero or a multiple of 2π.

Just as in the familiar case of a separable Hilbert space, two vectors differing only by a phase factor describe the same physical state, so here two vectors related by one of these generalized phase transformations $V_\otimes(\lambda)$ are physically equivalent. For $V_\otimes(\lambda)$ evidently commutes with all operators which can be constructed out of the operators a_n and a_n^*. Thus it would be reasonable to define the *generalized ray* in $\mathcal{H}_\otimes$ determined by a vector $|\psi\rangle$ as the set of all vectors of the form $V_\otimes(\lambda)\,|\psi\rangle$. Then each physical state would be represented by a unique generalized ray.

Two product vectors $|\phi\rangle$ and $|\psi\rangle$ are said to be *weakly equivalent*, $|\phi\rangle \sim_w |\psi\rangle$, if and only if

$$V_\otimes(\lambda)\,|\phi\rangle \sim |\psi\rangle,$$

for some λ. The condition for this is that

$$\sum_n (1 - |(\phi_n, \psi_n)|) < \infty. \tag{34}$$

In that case, $V_\otimes(\lambda)$ maps $\mathcal{H}_\otimes(\phi)$ isometrically onto $\mathcal{H}_\otimes(\psi)$.

Next, we define a unitary operator $U_\otimes(f)$ for each $f \in \mathcal{A}^*$. (Note that we do not restrict f to belong to $\mathcal{K}$.) Let $\{f_n\}$ be the corresponding sequence, $f_n = (e_n^* f)$. Then we set

$$U_\otimes(f) = \prod_n \otimes\, U_n(f_n) \tag{35}$$

with $U_n(f_n)$ defined by (23).

The multiplication law for the operators U_n together with the formula (30) yields for these tensor-product operators the relations

$$V_\otimes(\lambda)V_\otimes(\mu) = V_\otimes(\lambda + \mu), \tag{36}$$

$$V_\otimes(\lambda)U_\otimes(f) = U_\otimes(f)V_\otimes(\lambda), \tag{37}$$

$$U_\otimes(f)U_\otimes(g) = U_\otimes(f+g)V_\otimes(\tfrac{1}{2}i[f^*g - g^*f]), \tag{38}$$

where of course $\frac{1}{2}i[f^*g - g^*f]$ denotes the sequence $\{\frac{1}{2}i[f_n^*g_n - g_n^*f_n]\}$. In the case where the scalar product (f^*g) is defined, that is when f or g belongs to $\mathcal{A}$, or when both belong to $\mathcal{K}$, the corresponding series converges absolutely, and

$$V_\otimes(\tfrac{1}{2}i[f^*g - g^*f]) = \exp \tfrac{1}{2}[(g^*f) - (f^*g)].$$

For vectors $f \in \mathcal{A}$ which are finite linear combinations of the basis vectors e_n, $U_\otimes(f)\,|\psi\rangle \sim |\psi\rangle$ always, and so $U_\otimes(f)$ defines a unitary operator, $U(f)$ say, on each of the incomplete direct-product spaces $\mathcal{H}_\otimes(\psi)$. Thus if we were interested only in such finite linear combinations we could obtain a representation of the canonical commutation relations in each such space, since (38) clearly reduces to (14). For our purposes however we need only a much more restricted class of representations, which may conveniently be described in terms of coherent states.

The vacuum state $|0\rangle \in \mathcal{H}_\otimes$ is the product vector

$$|0\rangle = \prod_n \otimes\, \psi_{0n}. \tag{39}$$

For each $f \in \mathcal{A}^*$ we then define a *coherent state*, as in (15), by

$$|f\rangle = U_\otimes(f)\,|0\rangle. \tag{40}$$

It is easy to verify that for two coherent states, $|f\rangle \sim |g\rangle$ if and only if the sum

$$(f^*g - \tfrac{1}{2}f^*f - \tfrac{1}{2}g^*g) = \sum_n (f_n^*g_n - \tfrac{1}{2}|f_n|^2 - \tfrac{1}{2}|g_n|^2) \tag{41}$$

is absolutely convergent. Their scalar product is then given by

$$\langle f\,|\,g\rangle = \exp(f^*g - \tfrac{1}{2}f^*f - \tfrac{1}{2}g^*g). \tag{42}$$

In particular, of course, $|f\rangle \in \mathcal{H}_\otimes(0) = \mathcal{H}$, the carrier space of the Fock representation, if and only if $f \in \mathcal{K}$. On the other hand $|f\rangle \sim_w |g\rangle$ if and only if the real part of (41) converges, that is if and only if $f - g \in \mathcal{K}$.

It will also be convenient to define the generalized coherent states

$$|f, \lambda\rangle = U_\otimes(f)V_\otimes(\lambda)\,|0\rangle \tag{43}$$

for all $f \in \mathcal{A}^*$ and all sequences λ of real numbers. (Actually it is sufficient for our purposes to consider only those sequences λ which can be constructed as linear combinations of sequences $\{f_n^*g_n\}$ with $f, g \in \mathcal{A}^*$.) The scalar product of two such states is

$$\langle f, \lambda\,|\,g, \mu\rangle = \exp(f^*g - \tfrac{1}{2}f^*f - \tfrac{1}{2}g^*g - i\lambda + i\mu), \tag{44}$$

if the sum formally denoted by the exponent converges absolutely, and is zero otherwise. It is easy to see that $|f, \lambda\rangle \sim_w |g, \mu\rangle$ if and only if $f - g \in \mathcal{K}$, and that $|f, \lambda\rangle \sim |g, \mu\rangle$ if and only if, in addition, the sum

$$(\tfrac{1}{2}i[f^*g - g^*f] + \lambda - \mu)$$

converges absolutely.

Let us denote the closed linear subspace of $\mathcal{H}_\otimes$ spanned by the generalized coherent states by $\mathcal{H}_{\mathrm{IR}}$. This space depends of course on the choice of $\mathcal{A}^*$ or $\mathcal{A}$. It is clear that the operators $U_\otimes(f)$ and $V_\otimes(\lambda)$ map $\mathcal{H}_{\mathrm{IR}}$ onto itself, and therefore define unitary operators (which we again denote by the same symbols) on this subspace. It is interesting to note that if we had made the trivial choice $\mathcal{A} = \mathcal{K} = \mathcal{A}^*$, we would have obtained for $\mathcal{H}_{\mathrm{IR}}$ the subspace spanned by vectors weakly equivalent to those of the Fock representation.

It is easy to see that we have obtained a representation of the canonical commutation relations, that is of unitary operators $U(f)$ for all $f \in \mathcal{A}$ on each of the incomplete direct-product spaces $\mathcal{H}_\otimes(g, \lambda)$ contained in $\mathcal{H}_{\mathrm{IR}}$. For to prove this, we have only to show that $U_\otimes(f)\,|g, \lambda\rangle \sim |g, \lambda\rangle$, since we already know that $U_\otimes(f)$ maps $\mathcal{H}_\otimes(g, \lambda)$ isometrically either onto itself or onto some other incomplete direct product space.

But by (38)

$$\langle g, \lambda | U_{\otimes}(f) | g, \lambda\rangle = \langle 0 | U_{\otimes}(-g) U_{\otimes}(f) U_{\otimes}(g) | 0\rangle = \exp[-\tfrac{1}{2}(f^*f) + (g^*f - f^*g)], \quad (45)$$

and each term in the exponent represents a finite scalar product since $f \in \mathcal{A}$. It follows that $U_{\otimes}(f)$ defines a unitary operator $U(f)$ on $\mathcal{H}_{\otimes}(g, \lambda)$, and that the operators so defined satisfy the relation (14). [They can moreover easily be shown to satisfy the requisite weak continuity condition which allows the generators (a^*f) and (f^*a) to be recovered via Stone's theorem[10].]

The operators $U(f)$ define a cyclic representation of the relations (14) on $\mathcal{H}_{\otimes}(g, \lambda)$, which is completely characterized by the expectation functional (45).[11] In fact, since $U_{\otimes}(g)$ acts as a translation operator on a and a^* according to the formal relations

$$U_{\otimes}(-g) a U_{\otimes}(g) = a + g,$$
$$U_{\otimes}(-g) a^* U_{\otimes}(g) = a^* + g^*,$$

we may identify the representation on $\mathcal{H}_{\otimes}(g, \lambda)$ with the Fock representation of the translated operators $a - g$ and $a^* - g^*$, with $|g, \lambda\rangle$ as the corresponding cyclic "vacuum" state.

It is obvious from (45) that the representations on $\mathcal{H}_{\otimes}(g, \lambda)$ for given g are unitarily equivalent for all λ. Indeed, $V_{\otimes}(\lambda)$ maps $\mathcal{H}_{\otimes}(g)$ isometrically onto $\mathcal{H}_{\otimes}(g, \lambda)$ and commutes with $U_{\otimes}(f)$. Moreover it is easy to verify that the representations on $\mathcal{H}_{\otimes}(f, \lambda)$ and $\mathcal{H}_{\otimes}(g, \mu)$ are unitarily equivalent if and only if $|f, \lambda\rangle \sim_w |g, \mu\rangle$, that is if and only if $f - g \in \mathcal{K}$. For then $U_{\otimes}(-g)$ maps $\mathcal{H}_{\otimes}(g)$ onto $\mathcal{H}_{\otimes}(0)$, $U_{\otimes}(g - f)$ maps $\mathcal{H}_{\otimes}(0)$ onto itself, and $U_{\otimes}(f)$ maps $\mathcal{H}_{\otimes}(0)$ onto $\mathcal{H}_{\otimes}(f)$. Thus

$$U_{\otimes}(f) U_{\otimes}(g - f) U_{\otimes}(-g) = V_{\otimes}(\tfrac{1}{2}i[f^*g - g^*f])$$

maps $\mathcal{H}_{\otimes}(g)$ isometrically onto $\mathcal{H}_{\otimes}(f)$, and induces the unitary equivalence. (The unitary inequivalence for weakly inequivalent $|f\rangle$ and $|g\rangle$ has been established by Klauder, McKenna and Woods.[8])

An interesting consequence of the interpretation of the representation on $\mathcal{H}_{\otimes}(g)$ as a Fock representation of translated operators is the following. We know that in the Fock representation on $\mathcal{H}_{\otimes}(0)$ the operators $U(f)$ are defined for all $f \in \mathcal{K}$, not merely for $f \in \mathcal{A}$, and that the corresponding coherent states (15) form an overcomplete basis with the scalar products (19) and the formal completeness relation (20). It follows that for all $f \in \mathcal{K}$, the operators

$$U_{\otimes}(g) U_{\otimes}(f) U_{\otimes}(-g) = U_{\otimes}(f) V_{\otimes}(i[g^*f - f^*g])$$

define unitary operators on $\mathcal{H}_{\otimes}(g)$, and that the corresponding states

$$U_{\otimes}(g) U_{\otimes}(f) |0\rangle = |g + f, \tfrac{1}{2}i(g^*f - f^*g)\rangle \quad (46)$$

form an overcomplete basis in $\mathcal{H}_{\otimes}(g)$ with the same scalar product and completeness relation as (19) and (20). Thus although the operators (a^*f) and (f^*a) can be defined in all of these representations only if $f \in \mathcal{A}$, we can nevertheless define operators which play the same role for all $f \in \mathcal{K}$, namely $([a^* - g^*]f)$ and $(f^*[a - g])$.

We have so far defined these representations in terms of the infinite tensor-product space $\mathcal{H}_{\otimes}$, whose definition rests on choosing an orthonormal basis in $\mathcal{K}$. However, this is not the only way of defining $\mathcal{H}_{\mathrm{IR}}$. Instead we can work directly in terms of the integral which defines the scalar product in $\mathcal{K}$. We now sketch this definition.

It will be convenient at this point to specialize our discussion by introducing a particular choice for $\mathcal{A}$. Let $\mathcal{A}$ consist of those functions $f \in \mathcal{K}$ for which

$$\int \frac{d\mathbf{k}}{(2\pi)^3 2k^0} (k^0)^{-1} f_\mu^*(\mathbf{k}) f^\mu(\mathbf{k}) < \infty. \quad (47)$$

Then $\mathcal{A}^*$ may be identified with a set of equivalence classes of functions $f_\mu(\mathbf{k})$ for which

$$\int \frac{d\mathbf{k}}{(2\pi)^3 2k^0} \frac{k^0}{k^0 + 1} f_\mu^*(\mathbf{k}) f^\mu(\mathbf{k}) < \infty. \quad (48)$$

[The denominator here is needed to avoid imposing extra conditions on $f_\mu(\mathbf{k})$ for large $\mathbf{k}$.] Of course, as in $\mathcal{K}$, two functions define the same element of $\mathcal{A}^*$ if their difference is proportional to k^μ.

Instead of sequences λ of real numbers, we now consider real measurable functions λ. As before [compare the remark following Eq. (43)], it will be sufficient to consider a restricted class of functions, namely, those which satisfy the condition

$$\int \frac{d\mathbf{k}}{(2\pi)^3 2k^0} \frac{k^0}{k^0 + 1} |\lambda(\mathbf{k})| < \infty, \quad (49)$$

although it is not really essential to impose this restriction. Then for each such function λ and each $f \in \mathcal{A}^*$, we introduce a formal vector $|f, \lambda\rangle$. We say that two such formal vectors are weakly equivalent, $|f, \lambda\rangle \sim_w |g, \mu\rangle$, if and only if $f - g \in \mathcal{K}$, and equivalent, $|f, \lambda\rangle \sim |g, \mu\rangle$, if and only if, in addition, the integral

$$(\tfrac{1}{2}i[f^*g - g^*f] + \lambda - \mu) = \int \frac{d\mathbf{k}}{(2\pi)^3 2k^0} \{\tfrac{1}{2}i[f_\mu^*(\mathbf{k}) g^\mu(\mathbf{k}) - g_\mu^*(\mathbf{k}) f^\mu(\mathbf{k})] + \lambda(\mathbf{k}) - \mu(\mathbf{k})\} \quad (50)$$

[10] See for example, F. Riesz and B. Sz-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1955), p. 380.
[11] H. Araki, J. Math. Phys. 1, 492 (1960).

converges. If they are equivalent, their scalar product is given by (44), where the formal exponent now denotes an integral, while if they are inequivalent it is zero.

Next we consider the vector space $\mathcal{H}'_{\mathrm{IR}}$ comprising all finite linear combinations of these formal vectors, with scalar product defined by linearity, and finally we define $\mathcal{H}_{\mathrm{IR}}$ to be the completion of this vector space. The only problem in carrying through this program is that of proving that the scalar product so defined on $\mathcal{H}_{\mathrm{IR}}$ is positive semidefinite. (That it is positive-definite is then a matter of definition. For, all vectors of zero norm have zero scalar products with all other vectors and are by definition the zero vector.) In fact, however, we can reduce the problem to a triviality. Since the scalar product between inequivalent vectors vanishes, it is sufficient to consider linear combinations of vectors chosen from one equivalence class, say that of $|g, \mu\rangle$. But there is a transformation between the vectors in the equivalence class of $|0\rangle$ and in that of $|g, \mu\rangle$ under which scalar products are preserved, namely,

$$|f, \lambda\rangle \to |g + f, \lambda + \mu + \tfrac{1}{2}i[g^*f - f^*g]\rangle. \quad (51)$$

Thus it is sufficient to prove the result for linear combinations of the basis vectors $|g, 0\rangle$ with $g \in \mathcal{K}$. But this amounts simply to a verification of the positive definiteness of the scalar product in the usual Fock representation.

Having defined $\mathcal{H}_{\mathrm{IR}}$ in this way we can of course define the unitary operators $U_{\otimes}(f)$ and $V_{\otimes}(\lambda)$ by (43) and (36)–(38). In particular this defines a representation of the operators $U(f)$ for $f \in \mathcal{A}$ in each of the subspaces $\mathcal{H}_{\otimes}(g, \mu)$ spanned by vectors equivalent to a given $|g, \mu\rangle$. The properties of these operators may be established as before.

4. INTERACTION WITH A CLASSICAL CURRENT

We now wish to discuss the problem of the electromagnetic field interacting with a prescribed c-number current distribution $J^\mu(x)$ according to the equation

$$\partial_\mu \partial^\nu A_\mu(x) - \partial^2 A_\mu(x) = J_\mu(x), \quad (52)$$

where of course the current must be conserved,

$$\partial_\mu J^\mu(x) = 0. \quad (53)$$

If $J^\mu(x)$ were confined to a finite region of space-time, then we could solve (52) by standard methods and obtain a unitary S operator defined on the Hilbert space of the Fock representation. However if it does not vanish for large times, the S-matrix elements so calculated would be infrared divergent, which means that the interaction is capable of taking us out of this space (in the sense that the in-field and out-field representations are unitarily inequivalent), and we must work instead with the nonseparable Hilbert space $\mathcal{H}_{\mathrm{IR}}$.

We define the Fourier transform of the current to be

$$J^\mu(k) = \int dx e^{-ik\cdot x} J^\mu(x) \quad (54)$$

and its mass-shell restriction as

$$j^\mu(\mathbf{k}) = J^\mu(|\mathbf{k}|, \mathbf{k}). \quad (55)$$

We suppose that j does not belong to $\mathcal{K}$, but that it is restricted by (48) and so belongs to $\mathcal{A}^*$. (We also later require a restriction on the off-mass-shell part of J. Both conditions will be fulfilled for physically reasonable currents.)

Let us now consider the S-matrix element between two generalized coherent states,

$$\langle f, \lambda \text{ out} \mid g, \mu, \text{in}\rangle_J = \left\langle f, \lambda, \begin{Bmatrix}\text{out}\\ \text{in}\end{Bmatrix}\right| S(J) \left| g, \mu, \begin{Bmatrix}\text{out}\\ \text{in}\end{Bmatrix}\right\rangle, \quad (56)$$

where the subscript J denotes the presence of the external current.

The simplest way to compute this matrix element is to use the variational derivative technique developed by Schwinger.[12] The variational derivative with respect to the external current is given by

$$\frac{\delta}{\delta J^\mu(x)} \langle f, \lambda, \text{out} \mid g, \mu, \text{in}\rangle_J = i\langle f, \lambda, \text{out}|\, A_\mu(x)\, |g, \mu, \text{in}\rangle_J. \quad (57)$$

The solution of (52) with appropriate boundary conditions (in the radiation gauge) is

$$A_\mu(x) = A_\mu^{\mathrm{in}(+)}(x) + A_\mu^{\mathrm{out}(-)}(x) + \int dy D^{RG}_{\mu\nu}(x - y) J^\nu(y), \quad (58)$$

with

$$D^{RG}_{\mu\nu}(x - y) = \int \frac{dk}{(2\pi)^4} \gamma_{\mu\nu}(k) \frac{e^{ik\cdot(x-y)}}{k^2 - i\epsilon},$$

where $\gamma_{\mu\nu}(k)$ is again defined by (4) and (6), but for $k^2 \neq 0$. [Of course it no longer satisfies (3).] Inserting this expression in (57) and integrating, we obtain a formal solution

$$\langle f, \lambda, \text{out} \mid g, \mu, \text{in}\rangle_J = \langle f, \lambda, \text{out} \mid g, \mu, \text{in}\rangle_0 \times \exp i\left\{\int dx J^\mu(x) G^{(+)}_\mu(x) + \int dx F^{(-)}_\mu(x) J^\mu(x) + \frac{1}{2}\int dx\, dy J^\mu(x) D^{RG}_{\mu\nu}(x - y) J^\nu(y)\right\}, \quad (59)$$

[12] J. Schwinger, Proc. Natl. Acad. Sci. U.S. 37, 452 (1951).

where $F_\mu(x)$ and $G_\mu(x)$ are the classical fields defined in (18), and the constant of integration has been identified with the value of the matrix element for $J = 0$, given by (44).

The formula (59) has only a formal significance because of course the various factors may all diverge. However we can give it a meaning by adopting a suitable convention. Any expression of this form is to be written as the exponential of an integral over x or k, and all terms are to be summed before doing the final integration. Then if this integral converges (whether or not the integrals of individual terms do so) it yields the required result. If it diverges, the result is zero. This convention is obviously consistent with our treatment of scalar products of coherent states in the preceding section. Moreover it is also consistent with the intention of finding a solution to (57) and (52); for one can verify these equations *a posteriori*.

Proceeding with this program, we now transform the exponent of (59) to momentum space. In virtue of the conservation equation (53), we may replace $D^{\mathrm{RG}}_{\mu\nu}(x-y)$ in the last term of the component by $g_{\mu\nu}D_F(x-y)$, with

$$D_F(x-y) = \int \frac{dk}{(2\pi)^4} \frac{e^{ik\cdot(x-y)}}{k^2 - i\epsilon}.$$

Then we can separate the real and imaginary parts of this term in the form

$$\begin{aligned}&\frac{i}{2}\int \frac{dk}{(2\pi)^4} J^\mu(-k)\frac{1}{k^2 - i\epsilon} J_\mu(k)\\ &= \frac{i}{2}\int \frac{dk}{(2\pi)^4}\frac{J^\mu(-k)J_\mu(k)}{k^2} - \frac{1}{2}\int \frac{d\mathbf{k}}{(2\pi)^3 2\,|\mathbf{k}|} j^*_\mu(\mathbf{k})j^\mu(\mathbf{k})\\ &= i(\sigma) - \tfrac{1}{2}(j^*j), \qquad (60)\end{aligned}$$

say, where (σ) is a principal value integral and therefore real.

It follows that the complete expression for the matrix element (56) is

$$\langle f, \lambda, \text{out} \,|\, g, \mu, \text{in}\rangle_J = \exp\big(f^*g - \tfrac{1}{2}f^*f - \tfrac{1}{2}g^*g + if^*j + ij^*g - \tfrac{1}{2}j^*j - i\lambda + i\mu + i\sigma\big). \quad (61)$$

This is very similar in form to the results obtained by Chung.[6] The matrix element is nonzero if the integral represented formally by the expression in parentheses converges, and zero if it diverges. One remark about the notation here is perhaps needed. The parentheses signify of course an integration over all 3-momenta $\mathbf{k}$. Thus in order to accommodate the quantity σ within this scheme, we have to assume that the integration over k^0 is performed first, leaving as a term in the final integrand the function

$$\frac{\sigma(\mathbf{k})}{2\,|\mathbf{k}|} = \int \frac{dk^0}{2\pi}\frac{J^\mu(-k)J_\mu(k)}{k^2}. \qquad (62)$$

We assume that J^μ is so restricted that σ belongs to the class of functions defined by (49). Separating the real and imaginary parts in the exponent of (61) we obtain

$$\begin{aligned}&\langle f, \lambda, \text{out}\,|\, g, \mu, \text{in}\rangle_J\\ &\quad= \exp\big(-\tfrac{1}{2}[f^* - g^* + ij^*][f - g - ij]\big)\\ &\quad\times \exp i\big(\tfrac{1}{2}i[g^*f - f^*g] + \tfrac{1}{2}[j^*f + f^*j] + \tfrac{1}{2}[j^*g + g^*j] - \lambda + \mu + \sigma\big). \quad (63)\end{aligned}$$

Thus we see that for a given initial state the possible final states belong to the weak equivalence class specified by $g + ij$, and that within this class one can always find one and only one equivalence class for which the value of λ is such as to make the imaginary part of the exponent converge.

We can express the result (61) or (63) in a much more transparent form. For comparison with (44) and (56) shows that

$$\begin{aligned}|g, \mu, \text{in}\rangle &= S(J)\,|g, \mu, \text{out}\rangle\\ &= |g + ij, \tfrac{1}{2}[j^*g + g^*j] + \mu + \sigma, \text{out}\rangle. \qquad (64)\end{aligned}$$

Thus we see that the scattering operator $S(J)$ defined on the nonseparable Hilbert space $\mathcal{H}_{\mathrm{IR}}$ maps coherent states onto coherent states. Indeed it can be identified with one of the operators we have already introduced. It is easy to verify, using (38), that the same transformation is induced by the operator

$$S(J) = U_\otimes(ij)V_\otimes(\sigma). \qquad (65)$$

The physically important factor here is the first one, which depends only on the mass-shell part of the current. The second factor represents a generalized over-all phase factor and may in fact be dropped in this case. (However, in the case of interaction with a quantized field, the corresponding factor is a function of the external momenta, and cannot be dropped.)

It is interesting to note that we can rewrite (65) formally in a more familiar form in terms of the in or out fields,

$$S(J) = \exp i\int \frac{d\mathbf{k}}{(2\pi)^3\,|\mathbf{k}|}\,[a^{*\text{in}}_\mu(\mathbf{k})j^\mu(\mathbf{k}) + j^{\mu *}(\mathbf{k})a^{\text{in}}_\mu(\mathbf{k})] \times \exp \tfrac{1}{2}i\int \frac{dk}{(2\pi)^4}\frac{J^\mu(-k)J_\mu(k)}{k^2},$$

or, returning to coördinate space,

$$S(J) = \exp i \int dx J^{\mu}(x) A_{\mu}^{\text{in}}(x) \times \exp \tfrac{1}{2} i \int dx\, dy J^{\mu}(x) \bar{D}(x-y) J_{\mu}(y), \quad (66)$$

where of course $\bar{D}$ is the Fourier transform of the principal-value function $1/k^2$.

This is just the result which would be obtained by a naive and straightforward calculation. However it has now been given a precise significance even in the case where the integrals involved are divergent.

5. DISCUSSION

We have shown that one can define a nonseparable Hilbert space $\mathcal{H}_{\text{IR}}$ which contains all the photon states which are relevant to the infrared problem, and that for any classical current satisfying physically reasonable restrictions, a unitary S operator may be defined on this space. The space $\mathcal{H}_{\text{IR}}$ decomposes into the direct sum of uncountably many separable Hilbert spaces, in each of which the representations of the canonical commutation relations for both the in and the out fields are irreducible, although in general unitarily inequivalent (since S maps one such space onto a different one).

We recall that the vectors in $\mathcal{H}_{\text{IR}}$ contain a physically irrelevant label λ which serves to distinguish between different equivalence classes within a weak equivalence class. Any physical state may be represented by any one of a large class of different vectors, namely those belonging to a "generalized ray" in $\mathcal{H}_{\text{IR}}$. For consistency we must of course require that the physical results for such quantities as transition probabilities be independent of the choice of a vector from its generalized ray. That this is indeed the case is ensured by the fact that the S operator commutes with the operators $V_{\otimes}(\lambda)$ which induce generalized phase transformations. Thus if we alter the choice of the vector representing the initial state, the vector representing the final state also changes, but in such a way as to correspond to the same physical state. If we wish to compute the transition probability between given initial and final physical states we may choose any vector representing the initial state. Then, if the transition probability is nonzero there will be some vector representing the final state which has a nonvanishing matrix element with the chosen initial state vector, although there will of course be other such vectors with zero matrix elements. The transition probability is then the squared modulus of this particular matrix element. We can rephrase this statement as follows. If we choose arbitrary vectors $|i, \text{in}\rangle$ and $|f, \text{out}\rangle$ representing the initial and final states, then we have to find a value of λ for which

$$\langle f, \text{out}|\, V_{\otimes}(\lambda)\, |i, \text{in}\rangle$$

is nonvanishing. When we have done this, the transition probability is given by the square of the absolute value of this matrix element. More briefly, what we have shown is that one can drop infinite-phase factors represented by $V_{\otimes}(\lambda)$ just as one can drop finite phase factors. (Of course, if we want to consider states which are linear combinations of the basis states, we can only drop an over-all phase factor in the complete matrix element, retaining the relative phases). For transition probabilities between coherent states, this observation provides a justification for the methods used by Chung.[6]

In a subsequent publication we intend to discuss the problem of the interaction of the electromagnetic field with other quantized fields, along similar lines, choosing asymptotic states from the Hilbert space $\mathcal{H}_{\text{IR}} \otimes \mathcal{H}_0$, where $\mathcal{H}_0$ describes the charged-particle states.

Evolution of the Probability Distributions and of the Correlation Functions in a Bogoliubov Plasma*

ARNOLD H. KRITZ
Aeronautical Research Associates of Princeton, Inc., Princeton, New Jersey
AND
S. RADIN†
Lehigh University, Bethlehem, Pennsylvania
AND
G. SANDRI
Aeronautical Research Associates of Princeton, Inc., Princeton, New Jersey

(Received 1 May 1967)

Recently we developed a matrix formulation of nonequilibrium statistical mechanics which we applied to dilute gases ($nr_0^3 \ll 1$) and to small momentum transfer interactions ($\phi_0/kT \ll 1$). We now show that the matrix formulation of the nonequilibrium equations can be extended to the Bogoliubov gas ($\phi_0/kT = 1/nr_0^3 \ll 1$). The asymptotic behavior is calculated for times long compared with the plasma frequency for: (i) the probability distribution functions, and (ii) the Mayer correlation functions. The collision integrals that determine the higher-order kinetic equations for the Bogoliubov gas are thereby constructed. The calculations performed directly with the nonlinearly coupled Mayer functions are shown to be equivalent to those performed with the linearly coupled probability distribution functions. In lowest order our theory coincides with a result obtained previously by Bogoliubov.

1. INTRODUCTION

The BBGKY hierarchy of equations of nonequilibrium statistical mechanics has been formulated in matrix notation.[1] This technique makes it possible to investigate in detail the solution of these equations to all orders in a perturbation parameter. In a neutral short-range gas the natural small parameter is the number of particles in an interaction sphere. The short range of the interactions means that binary collisions will dominate over multiple collisions in determining the time evolution of the single-particle distribution function. The lowest-order result is then the Boltzmann equation. Density corrections should involve multiple-particle interactions. We have shown previously[2] that, in the perturbation solution for the dilute gas, the asymptotic behavior of the s-particle distribution function is a polynomial in time. The lowest-order term in the expansion (i.e., the linear secularity for the single-particle distribution function) is the Boltzmann collision integral. Higher-order terms identify the contributions from multiple collisions needed for density corrections. Similar considerations for the Landau gas[1] (fairly dense, but dominated by small-momentum transfer binary collisions) have yielded similar results, with the linear secularity yielding the Landau collision integral.

In this paper, we extend our technique to the Bogoliubov regime, i.e., the regime appropriate to a gas of electrons in a neutralizing positive background. The interaction is that of Debye-shielded collisions with small momentum transfer. Since the Debye shielding is the result of collective effects, it is clear that s-electron distribution functions must be considered. We thus consider the perturbation expansion of the s-electron distribution function in powers of the plasma parameter. However, since the BBGKY hierarchy of equations for the distribution functions does not allow a systematic decoupling, it has been useful to use a Mayer expansion to express the distribution functions in terms of the correlation functions. This yields another hierarchy for the correlation functions. The perturbation expansion for the correlation hierarchy is also considered.

The perturbation expansion in the Bogoliubov regime is considerably more complicated than for the Boltzmann or Landau regimes because of collective effects. We are able, however, to carry out the perturbation solution, as in the previous cases, and again obtain asymptotic behavior (for both the distribution functions and the correlation functions). The coefficients of the secularities are simply characterized. The lowest-order term coincides with the Lenard–Guernsey–Balescu collision integral.[3] Our results depend on: (i) an expansion theorem for the

* This work was sponsored by the Aerospace Research Laboratories, Office of Aerospace Research, United States Air Force.

† Present address: Aeronautical Research Associates of Princeton, Incorporated, Princeton, New Jersey.

[1] P. Goldberg and G. Sandri, Phys. Rev. **154**, 188 (1967); **154**, 199 (1967).

[2] A. Kritz, S. Radin, and G. Sandri, Phys. Rev. **155**, 101 (1967).

[3] N. Bogoliubov, in *Studies in Statistical Mechanics* I (North-Holland Publishing Company, Amsterdam 1962); A. Lenard, Ann. Phys. (N.Y.) **10**, 390 (1960); R. L. Guernsey, Ph.D. thesis, University of Michigan (1960); R. Balescu, Phys. Fluids **3**, 52 (1960); for a discussion of these and other derivations, see T. Y. Wu, *Kinetic Equations of Gases and Plasmas* (Addison-Wesley Publishing Company, Reading, Mass., 1966).

propagation function that includes the polarization of the medium (Sec. 2), and (ii) a general equivalence theorem that allows us to express the higher-order collision integrals through *either* the distribution functions or the correlation functions.

The secular behaviors of the distribution functions are obtained in Sec. 2. In Sec. 3 we show that the molecular chaos initial condition yields directly the zero values usually assumed for certain low-order correlation functions. The secular behavior of the correlation functions is obtained in Sec. 4. Our equivalence theorem implies that calculations for the Bogoliubov plasma may be carried out through the linearly coupled distribution function hierarchy rather than through the more usually used nonlinearly coupled correlation function hierarchy.

2. SECULAR BEHAVIOR OF ELECTRON DISTRIBUTION FUNCTIONS

The BBGKY equations for the distribution functions can be rewritten in the form of a single matrix equation.[2] Since this formulation allows us to treat the distribution function for s electrons on equal footing with the distribution function for any other number of electrons, we refer to the matrix equation as the cluster equation. If the cluster equation is made dimensionless by employing for units of length, time, and potential energy those units which characterize, respectively, the range of interaction (r_0), the duration of the collision (τ_p), and the mean potential during the interaction (ϕ_0), the resulting equation contains two nondimensional parameters. These are ϕ_0/kT and nr_0^3, where T is the temperature, and n is the number density. The quantity ϕ_0/kT is a measure of the strength of the interaction, and nr_0^3 is a measure of the diluteness of the electron gas.

For a plasma, the range of interaction is the Debye length, $\lambda_D \equiv (kT/4\pi ne^2)^{\frac{1}{2}}$, and the duration of the interaction is, on the average, the Debye length divided by the thermal velocity, i.e., $\tau_p = \omega_p^{-1}$, where ω_p is the plasma frequency. Since the electrons interact over a distance of the order of the Debye length, the unit of interaction energy ϕ_0 is taken as e^2/λ_D. The two dimensionless parameters are thus related in a plasma, yielding a single small parameter

$$\epsilon = \phi_0/kT = (n\lambda_D^3)^{-1} \ll 1. \tag{2.1}$$

We use this parameter to establish a perturbation solution of the cluster equation. Using molecular chaos as the initial condition, we investigate the behavior of distribution functions for times long compared to the duration of a collision. We are able to deduce the asymptotic form of the expanded distribution functions by extending to a plasma the previously derived necessary and sufficient conditions for secular behavior.[2]

The cluster equation is given by

$$\frac{\partial F}{\partial t} + (K - L)F = \epsilon IF, \tag{2.2}$$

where F is a column matrix whose s element is the s-body distribution function F^s. The operator matrices K and I are diagonal and of infinite order when we allow the number of particles to become infinite. The (s, s) element of the K and I matrices are given by

$$K^s = \sum_{i=1}^{s} \mathbf{v}_i \cdot \nabla_i \tag{2.3}$$

and

$$I^s = \sum_{1 \le i < j \le s} I_{ij}, \tag{2.4}$$

where

$$I_{ij} = \nabla_i \phi_{ij} \cdot \nabla_{\mathbf{v}_i} + \nabla_j \phi_{ij} \cdot \nabla_{\mathbf{v}_j}. \tag{2.5}$$

The dimensionless plasma potential $\phi_{ij}(|\mathbf{x}_i - \mathbf{x}_j|)$ has been normalized by $\phi_0 = e^2/\lambda_D$. The L matrix is also of infinite order, but has the nonzero elements L^s only at the $(s, s+1)$ locations:

$$L^s = \sum_{i=1}^{s} L_{i,s+1} = \sum_{i=1}^{s} \int d\mathbf{x}_{s+1}\, d\mathbf{v}_{s+1} I_{i,s+1}. \tag{2.6}$$

The expansion in ϵ of the matrix distribution function is given by

$$F = \sum_{\nu=0}^{\infty} \epsilon^\nu F^\nu. \tag{2.7}$$

When Eq. (2.7) is substituted into Eq. (2.1) and terms proportional to ϵ^ν are equated, the following matrix recursion relation results:

$$\frac{\partial F^\nu}{\partial t} + (K - L)F^\nu = IF^{\nu-1}. \tag{2.8}$$

We investigate the long-time evolution of the column matrix F. The long-time behavior of F^ν will be denoted by $F^{*\nu}$, and the symbol $\sim$ will be used to mean "is asymptotic to." Therefore,

$$F^\nu \sim F^{*\nu}. \tag{2.9}$$

In obtaining the asymptotic behavior of F^ν we assume that the expansion of F is uniform. Consequently, we require that for large times (measured in units of the duration of a collision), F^ν is time independent; i.e.,

$$\partial F^\nu/\partial t \sim 0. \tag{2.10}$$

Equation (2.10) contains the physical condition that the system approach equilibrium for large time. However, when $F^{*\nu}$ is determined by direct perturbation theory, we find that Eq. (2.10) is violated. A main purpose of this paper is to understand the exact nature of the violation so that a method more sophisticated than direct perturbation (for example, the method of extension[4,5]) can be applied to solve Eq. (2.1) in such a manner that the physical conditions for equilibrium are, in fact, satisfied.

In obtaining the steady state solution to Eq. (2.18), we utilize the following operator identities:

$$\zeta^*[i(K - L)]I = \sum_{A=0}^{\infty} \mathbf{L}^A \mathbf{I} = \int_0^{\infty} dt e^{-(K-L)t} I, \quad (2.11)$$

where

$$\mathbf{I} \equiv \zeta^*(iK)I; \quad \mathbf{L} \equiv \zeta^*(iK)L \quad (2.12)$$

and for $\nu \geq 1$,

$$\prod_{i=1}^{\nu} \sum_{A_i=0}^{\infty} \mathbf{L}^{A_i}\mathbf{I} = \sum_{i=0}^{\infty} \{\mathbf{L}^i \mathbf{I}^{\nu-1}\}\mathbf{I}. \quad (2.13)$$

We have introduced the notation of the curly brackets to represent the sum over all possible distinguishable sequences in which the $\mathbf{L}$ operator appears i times and the $\mathbf{I}$ operator $\nu - 1$ times. The sum does not reduce since the operators do not commute. It is worth emphasizing the significance of the result given by Eq. (2.11). The *integral* operator L in the argument of the propagation function ζ^* represents the collective shielding effects (polarization) in the charged gas, and introduces only minor complications when our matrix approach to the BBGKY hierarchy is applied to a Bogoliubov plasma.

Employing the identities given in Eqs. (2.11) and (2.13), we find that for a spatially homogeneous plasma the steady-state solution of the recursion equation (2.8) is

$$F^\nu(t) \sim \sum_{i=0}^{\infty} \{\mathbf{L}^i \mathbf{I}^{\nu-1}\}\mathbf{I}F(0), \quad \nu \geq 1, \quad (2.14)$$

where we have used the result obtained in Appendix A, $F^0(t) = F(0)$. Molecular chaos has been taken as the initial condition so that the s element of the column matrix $F(0)$ is

$$F^s(0) = F^{s,0}(0) = \prod_{i=1}^{s} f_i^0,$$

where f_i^0 is the zero-order one-particle distribution function for the ith particle. We show in Appendix B that Eq. (2.14) does express the proper result for $F^{1*\nu}$ although not all the terms in Eq. (2.8) are nonzero for the first ($s = 1$) element. ($I^1 \equiv 0$ and $K^1F^{1,\nu} = 0$ since the plasma is homogeneous.)

In our previous analysis of the asymptotic behavior of the expanded distribution functions for a dense gas, the necessary and sufficient conditions for secular behavior were presented.[2] It was shown that the expanded distribution functions grow in time when the propagation operators $\zeta^*(iK)$ act on space-independent functions. The space-independent functions can be classified in terms of "generalized collision integrals" $\mathbf{\Lambda}$. These integrals represent the general class of functions in which L_{ij} acts on a function of $\mathbf{x}_{ij}$, i.e.,

$$\mathbf{\Lambda}F(0) = L_{ij}\zeta^*(iK_{ij}^2)\mathfrak{I}(\mathbf{x}_{ij})F(0), \quad (2.15)$$

where $\mathfrak{I}(\mathbf{x}_{ij})$, the effective interaction operator, is a sequence of operators in which at least one I_{ij} must appear, in which L does not act on a free particle or a space-independent function, and which depends on position only through $\mathbf{x}_{ij}$. The simplest form of $\mathbf{\Lambda}$ is the Landau collision integral which corresponds to the phase mixing of two particles that have undergone a completed collision. In Appendix C we consider the simplest form of $\mathbf{\Lambda}$ which, for the Coulomb potential, is precisely the Landau collision integral. Furthermore, we show that this integral can be expressed in the Fokker–Planck form

In general, $\mathbf{\Lambda}$ corresponds to the phase mixing of two particles that have undergone a number of collisions. The function $\mathbf{\Lambda}$ contains a minimum of one I and one L operator. The term proportional to t^n arises when the function $\zeta^*\mathbf{\Lambda}$ occurs n times. Thus, in order to have a secularity of order n, a minimum of n I's and n L's is required. Utilizing this requirement, we find that Eq. (2.14) can be written as a power series in time,

$$F^\nu \sim \sum_{i=0}^{\nu} C_i t^i. \quad (2.16)$$

Note that the secular behavior is independent of the cluster index. The secular behavior of F^ν exhibited in Eq. (2.16) is directly responsible for the irreversible behavior of the plasma. It is the growing terms that describe the changes which occur during the relaxation of the plasma to equilibrium.

3. EQUIVALENCE OF INITIAL MOLECULAR CHAOS AND STANDARD ORDERING OF CORRELATION FUNCTIONS

In the previous section we investigated the perturbation solution of the hierarchy equations for the s-particle distribution function. In particular, we determined the secular behavior of the individual terms in the plasma parameter expansions of the distribution function. This is an essential step in obtaining corrections to the lowest-order one-particle distribution function. However, the lowest-order plasma kinetic equation (the Lenard–Guernsey–Balescu kinetic equation) is not derived systematically

[4] G. Sandri, in *Proceedings of the Second Symposium on Nonlinear Partial Differential Equations, Delaware, 1965*, W. Ames, Ed. (Academic Press Inc., New York, 1967).

[5] G. Sandri, Ann. Phys. (N.Y.) **24**, 332, 380 (1963).

from the hierarchy for the distribution functions F^s, but instead, from the hierarchy for the correlation functions g^s.[3] The correlation hierarchy is obtained by substituting the Ursell–Mayer relations for F^s in terms of g^s (given in Appendix D) into the hierarchy for F^s. The correlation functions are then expanded in power series in the plasma parameter. In order to be able to systematically decouple lower-order equations, a set of correlation functions is taken to be identically zero. This allows derivation of the plasma kinetic equation in lowest order. The justification for this ordering is based on an analogy with equilibrium results. We demonstrate below that the vanishing of this set of correlation functions is a direct consequence of using molecular chaos as the initial condition.

The hierarchy of equations for the correlation functions is

$$\frac{\partial g^s}{\partial t} + K^s g^s - \epsilon I^s g^s - \epsilon \sum_{1\le i<j\le s} I_{ij} \sum_{k=1}^{s-1} g^k_{\{i\}} g^{s-k}_{\{j\}} = L^s g^{s+1} + \sum_{i=1}^{s} L_{i,s+1} \sum_{k=1}^{s} g^k_{\{i\}} g^{s+1-k}_{\{s+1\}}, \quad (3.1)$$

where $g^k_{\{i\}} g^{s-k}_{\{j\}}$ means all possible products of k-particle and $(s-k)$-particle correlation functions formed from the s particles, and in which particle i is in $g^k_{\{i\}}$ and particle j is in $g^{s-k}_{\{j\}}$.[6] If now the expansion in the small plasma parameter ϵ given by

$$g^s = \sum_{\nu=0}^{\infty} \epsilon^\nu g^{s,\nu} \quad (3.2)$$

is substituted into Eq. (3.1) and coefficients of ϵ^ν are equated, we obtain

$$\left(\frac{\partial}{\partial t} + K^s\right) g^{s,\nu} - L^s g^{s+1,\nu} - \sum_{i=1}^{s} L_{i,s+1} \sum_{k=1}^{s} \sum_{m=0}^{\nu} g^{k,m}_{\{i\}} g^{s+1-k,\nu-m}_{\{s+1\}} = I^s g^{s,\nu-1} + \sum_{1\le i<j\le s} I_{ij} \sum_{k=1}^{s-1} \sum_{m=0}^{\nu-1} g^{k,m}_{\{i\}} g^{s-k,\nu-m-1}_{\{j\}}. \quad (3.3)$$

To solve this set of differential equations, we require the initial values of the correlations. From the expressions for the correlation functions in terms of the distribution functions given in Appendix D, it follows that as a direct consequence of molecular chaos

$$g^{s,\nu}(0) = \begin{cases} f^0(0) & s=1 \text{ and } \nu = 0 \\ 0 & \text{otherwise.} \end{cases} \quad (3.4)$$

We solve Eqs. (3.3) for $g^{s,\nu}$ $(\nu \le s-2)$ using the initial conditions of Eq. (3.4). It is seen that the set of equations for $g^{s,\nu}$ $(\nu \le s-2)$ contains correlation functions of the form $g^{s,\nu}$ $(\nu \le s-2)$. Thus, the set of equations for $g^{s,\nu}$ $(\nu \le s-2)$ is decoupled from the rest of the hierarchy, and may be solved separately. We solve this set of equations for the initial condition of molecular chaos and obtain

$$g^{s,\nu}(t) = 0 \quad \text{for} \quad \nu \le s-2. \quad (3.5)$$

This is precisely the set of correlation functions noted above as being taken identically equal to zero by analogy with the equilibrium results.

To obtain the result expressed in Eq. (3.5) we use an induction method. We first consider the equations for $\nu = 0$, $s \ge 1$ and show that Eq. (3.5) holds for this case. Using this result, we then consider the equations for $\nu = 1$, $s \ge 3$ and show Eq. (3.5) holds for this case also. Finally, we demonstrate that, if Eq. (3.5) is valid for $\nu = n-1$, it is valid for $\nu = n$, thus completing the induction proof of Eq. (3.5).

A. Case $\nu = 0$, $s \ge 1$

From Eq. (3.3) we observe that the zero-order correlation functions are decoupled from the remainder of the hierarchy:

$$\frac{\partial g^{s,0}}{\partial t} + K^s g^{s,0} - L^s g^{s+1,0} - \sum_{i=1}^{s} L_{i,s+1} \sum_{k=1}^{s-1} g^{k,0}_{\{i\}} g^{s+1-k,0}_{\{s+1\}} = 0. \quad (3.6)$$

Applying the initial conditions on the $g^{s,\nu}$ given in Eq. (3.4), we obtain

$$\left.\frac{\partial g^{s,0}}{\partial t}\right|_{t=0} = 0. \quad (3.7)$$

Now, differentiating Eq. (3.6) with respect to time, and applying initial conditions on $g^{s,0}$ and $\partial g^{s,0}/\partial t$ for $s \ge 2$ we find

$$\left.\frac{\partial^2 g^{s,0}}{\partial t^2}\right|_{t=0} = 0. \quad (3.8)$$

Clearly, this procedure can be continued with the result that the initial conditions require that $g^{s,0}$ and all derivatives of $g^{s,0}$ are initially zero for $s \ne 1$. Consequently, $g^{s,0}$ must remain zero for all time, i.e.,

$$g^{s,0}(t) = 0 \quad \text{for} \quad s \ne 1. \quad (3.9)$$

The case $s = 1$ requires special consideration since the initial condition on $g^{1,0}$ is nonzero. It readily follows that

$$g^{1,0}(t) = g^{1,0}(0) = f^0(0). \quad (3.10)$$

[6] M. Green, in *Lectures in Theoretical Physics*, W. E. Britten, B. W. Downs, and J. Downs, Eds. (Interscience Publishers, Inc., New York, 1961), Vol. III.

B. Case $\nu = 1$, $s \geq 3$

Using Eq. (3.9), we find that when $\nu = 1$, Eq. (3.3) reduces to, for $s \geq 3$,

$$\frac{\partial g^{s,1}}{\partial t} + K^s g^{s,1} - L^s g^{s+1,1} - \sum_{i=1}^{s} L_{i,s+1}[g^{1,0}_{\{i\}} g^{s,1}_{\{s+1\}} + g^{s,1}_{\{i\}} g^{1,0}_{\{s+1\}}] = 0. \quad (3.11)$$

We again can calculate all time derivatives of $g^{s,1}$ ($s \geq 3$) at the initial time by successive differentiation of Eq. (3.11) and substitution of the initial conditions. We find that all derivatives and the function are zero initially so that

$$g^{s,1}(t) = 0 \quad \text{if} \quad s \geq 3. \quad (3.12)$$

C. Case $\nu = n$, $s \geq n + 2$

The above procedure may be carried through to higher ν. We assume

$$g^{s,n-1}(t) = 0 \quad \text{if} \quad s \geq n + 1 \quad (3.13)$$

and demonstrate that Eq. (3.13) implies

$$g^{s,n}(t) = 0 \quad \text{if} \quad s \geq n + 2. \quad (3.14)$$

Substituting Eq. (3.13) into Eq. (3.3) for $s \geq n + 2$, we find that the only remaining terms are

$$\frac{\partial g^{s,n}}{\partial t} + K^s g^{s,n} - L^s g^{s+1,n} - \sum_{i=1}^{s} L_{i,s+1}[g^{1,0}_{\{i\}} g^{s,n}_{\{s+1\}} + g^{s,n}_{\{i\}} g^{1,0}_{\{s+1\}}] = 0. \quad (3.15)$$

If we evaluate the derivatives of $\partial g^{s,n}/\partial t$ ($s \geq n + 2$) at $t = 0$ as above, we find that all derivatives are zero. Consequently, Eq. (3.14) is valid.

Clearly, n is entirely arbitrary so that Eq. (3.5) follows for arbitrary ν. This result has been obtained from the correlation function hierarchy using only the initial conditions on $g^{s,\nu}$ [Eq. (3.4)] deduced from the molecular chaos initial condition on the distribution functions [Eq. (A3)]. Thus, we conclude that the standard assumption of setting certain correlation functions identically zero is equivalent to using molecular chaos as the initial value of the distribution functions.

It is worth pointing out that the induction demonstration given in this section is based on the fact that the equations for the Mayer correlation functions are first order in time. Thus, the Cauchy–Kowalewski theorem can be applied even though the equations are nonlinear as far as their g dependence is concerned.[7]

[7] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1962), Vol. 2, pp. 39–56.

4. SECULAR BEHAVIOR OF THE CORRELATION FUNCTIONS

In Sec. 2 we determined the secular behavior of the electron distribution functions employing molecular chaos as the initial value. We then obtained the equivalent initial value statement for the correlation functions in order to demonstrate that the result expressed by Eq. (3.5) is a consequence of the initial condition of molecular chaos. Using Eq. (3.5), we now derive the secular behavior of the correlation functions. We show below that the correlation functions can be expressed as a sum of a special class of terms, that is, terms whose graphical representation contains only "connected" graphs.[1] This knowledge is utilized in determining the secular behavior of the correlation functions. We also demonstrate that the long-time behavior of g^s, obtained from the correlation function hierarchy, is identical to that deduced from the asymptotic solution for F^s. This result is obtained using the expressions for g^s in terms of F^s given in Appendix D.

The hierarchy for the expanded correlation functions, Eq. (3.3) takes the following form when the substitution $m' = m - k + 1$ is made, and the result given in Eq. (3.5) is utilized:

$$\begin{aligned}\left(\frac{\partial}{\partial t} + K^s\right) g^{s,\nu} - \sum_{i=1}^{s} L_{i,s+1} g^{1,0}_{\{i\}} g^{s,\nu}_{\{s+1\}} &= L^s g^{s+1,\nu} + \sum_{i=1}^{s} L_{i,s+1} \sum_{k=2}^{s-1} g^{k,k-1}_{\{i\}} g^{s+1-k,\nu-k+1}_{\{s+1\}} \\ &\quad + \sum_{i=1}^{s} L_{i,s+1} \sum_{k=1}^{s-1} \sum_{m=1}^{\nu-s+1} g^{k,m+k-1}_{\{i\}} g^{s+1-k,\nu-m-k+1}_{\{s+1\}} \\ &\quad + I^s g^{s,\nu-1} + \sum_{1\leq i<j\leq s} I_{ij} \sum_{k=1}^{s-1} \sum_{m=0}^{\nu-s+1} g^{k,m+k-1}_{\{i\}} g^{s-k,\nu-m-k}_{\{j\}},\end{aligned} \quad (4.1)$$

where we have dropped the primes from the m's. Equation (4.1) has been arranged so that all $g^{s,\nu}$ functions are on the left, and all other correlations are on the right. Inspection of Eq. (4.1) indicates that interdependence of the correlation functions can be represented by the array shown in Fig. 1. A given $g^{s,\nu}$ depends on all those $g^{t,\mu}$ which lead to it by arrows. Thus, $g^{1,1}$ depends on $g^{1,0}$ and $g^{2,1}$, and $g^{2,2}$ depends

$$\begin{tikzcd} g^{1,0} \arrow[r] \arrow[dr] & g^{1,1} \arrow[r] \arrow[dr] & g^{1,2} \arrow[r] \arrow[dr] & g^{1,3} \cdots \\ & g^{2,1} \arrow[u] \arrow[r] \arrow[dr] & g^{2,2} \arrow[u] \arrow[r] \arrow[dr] & g^{2,3} \cdots \arrow[u] \\ & & g^{3,2} \arrow[u] \arrow[r] \arrow[dr] & g^{3,3} \cdots \arrow[u] \\ & & & g^{4,3} \cdots \arrow[u] \end{tikzcd}$$

FIG. 1. Dynamical coupling of Mayer correlation functions.

on $g^{3,2}$, $g^{2,1}$, $g^{1,1}$, and $g^{1,0}$. To solve Eq. (4.1), we first fix $\nu - s = -1$ and increase s starting from $s = 1$. Secondly, $\nu - s$ is increased in integer steps, and the process is repeated.

A. Case $\nu - s = -1$

Upon elimination of the index ν, Eq. (4.1) reduces to

$$\left(\frac{\partial}{\partial t} + K^s\right) g^{s,s-1} - \sum_{i=1}^{s} L_{i,s+1} g^{1,0}_{\{i\}} g^{s,s-1}_{\{s+1\}} = \sum_{i=1}^{s} L_{i,s+1} \sum_{k=2}^{s-1} g^{k,k-1}_{\{i\}} g^{s+1-k,s-k}_{\{s+1\}} + \sum_{1\le i<j\le s} I_{ij} \sum_{k=1}^{s-1} g^{k,k-1}_{\{i\}} g^{s-k,s-k-1}_{\{j\}}. \quad (4.2)$$

For $s = 1$, Eq. (4.2) yields

$$\frac{\partial}{\partial t} g_1^{1,0} - L_{12} g_1^{1,0} g_2^{1,0} = 0. \quad (4.3)$$

We have assumed that the plasma is homogeneous so that g^1, which is the one-body distribution function, is space independent. Since L^1 operating on a space-independent function is zero, we observe that $g^{1,0}$ is constant in time,

$$g^{1,0}(t) = g^{1,0}(0). \quad (4.4)$$

For $s = 2$, Eq. (4.2) becomes

$$\left(\frac{\partial}{\partial t} + K^2\right) g^{2,1} - \sum_{i=1}^{2} L_{i,3} g^{1,0}_{\{i\}} g^{2,1}_{\{3\}} = I_{12} g_1^{1,0} g_2^{1,0}. \quad (4.5)$$

The asymptotic solution for $g^{2,1}$ then is

$$g^{2,1} \sim \zeta^*[i(K^2 - \Gamma^2)] I_{12} g_1^{1,0} g_2^{1,0}, \quad (4.6)$$

where

$$\zeta^*[i(K^2 - \Gamma^2)] = \int_0^\infty e^{-(K^2-\Gamma^2)\lambda}\, d\lambda = \sum_{A=0}^{\infty} [\zeta^*(iK^2)\Gamma^2]^A \zeta^*(iK^2) \quad (4.7)$$

and

$$\Gamma^2 A_{12} \equiv L_{13} g_1^{1,0} A_{23} + L_{23} g_2^{1,0} A_{13}. \quad (4.8)$$

A_{12} is any function of particles 1 and 2. Employing the expansion for $\zeta^*[i(K^2 - \Gamma^2)]$ given in Eq. (4.7), we obtain

$$g^{2,1} \sim \zeta^*(iK^2) I_{12} g_1^{1,0} g_2^{1,0} + \zeta^*(iK^2)\Gamma^2\zeta^*(iK^2) I_{12} g_1^{1,0} g_2^{1,0} + \zeta^*(iK^2)\Gamma^2\zeta^*(iK^2)\Gamma^2\zeta^*(iK^2) I_{12} g_1^{1,0} g_2^{1,0} + \cdots. \quad (4.9)$$

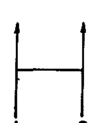


FIG. 2. Zero-order shielding contribution to the two-electron correlation.

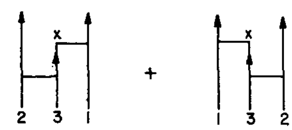


FIG. 3. First-order shielding contribution to the two-electron correlation.

We prove below that this expression, expressed in graph form, contains only connected graphs.

We employ the following graph notation. An undirected vertical line (no arrow) represents a free particle; a directed vertical line represents the propagator operator $\zeta^*(iK)$; a horizontal line, the interaction operator I^2, and a horizontal line with a cross, the phase-mixing operator L^1, the cross indicating the particle which is averaged. By a connected graph we mean that there is a path that connects all vertical lines. The first term on the right side of Eq. (4.9) is represented by the connected graph given in Fig. 2. The second term

$$\zeta^*(iK^2)\Gamma^2\zeta^*(iK^2) I_{12} g_1^{1,0} g_2^{1,0} = \zeta^*(iK^2)[L_{13} g_1^{1,0} \zeta^*(iK^2) I_{23} g_2^{1,0} g_3^{1,0} + L_{23} g_2^{1,0} \zeta^*(iK^2) I_{13} g_1^{1,0} g_3^{1,0}] \quad (4.10)$$

is given by the two connected graphs shown in Fig. 3. The representation of the third term on the right side of Eq. (4.9) contains graphs of the form given in Fig. 4. Again, these are all connected graphs.

Now, consider an arbitrary term in the expansion. The first two particles, reading from right to left, are connected by an I_{ij} interaction term. To this point the graph is connected. The subsequent application of Γ^2 brings in a new particle which interacts with a particle previously present through a phase-mixing (L) operation (with the averaging process over the new particle). This once again produces a connected graph. Further Γ^2 terms just add more particles in a similar manner, always with a phase-mixing interaction between the new particle and one previously present, phase averaging over the new particle. Thus, all the terms in Eq. (4.9) are represented by connected graphs since all the vertical lines are connected by horizontal lines (interactions). That $g^{2,1}$ is represented by only connected graphs can also be seen by a counting argument. There is always one I_{ij} and $(n - 2)$ L_{pr}'s in a term containing n particles. This is a total of $(n - 1)$ interactions, and since by our arguments, no

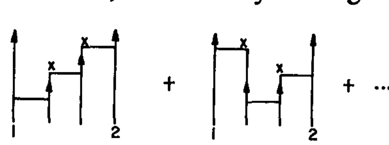


FIG. 4. Second-order shielding contribution to the two-electron correlation.

interaction occurs between the same two particles as any other interaction, these $(n-1)$ interactions exactly connect all the vertical lines.

For $s = 3$, the asymptotic solution of Eq. (4.2) is

$$g^{3,2} = \sum_{A=0}^{\infty} [\zeta^*(iK^3)\Gamma^3]^A \zeta^*(iK^3)\Big[\sum_{i=1}^{3} L_{1,4} g_{\{i\}}^{2,1} g_{\{4\}}^{2,1} + \sum_{1\le i<j\le 3} I_{ij}(g_{\{i\}}^{1,0} g_{\{j\}}^{2,1} + g_{\{i\}}^{2,1} g_{\{j\}}^{1,0})\Big], \quad (4.11)$$

where

$$\Gamma^n A_{123\cdots n} \equiv L_{1,n+1} g_1^{1,0} A_{23\cdots n} + L_{2,n+1} g_2^{1,0} A_{13\cdots n} + L_{3,n+1} g_3^{1,0} A_{124\cdots n} + \cdots + L_{n,n+1} g_n^{1,0} A_{1\cdots n-1}. \quad (4.12)$$

Since $g^{2,1}$ has been shown to contain only connected graphs, the operators L_{i4} and L_{ij} simply connect products of correlation functions, producing connected graphs. By our previous discussion of the effect of the Γ^2 operator, we see that the Γ^3 operator operating on connected graphs yields only new connected graphs. Thus, $g^{3,2}$ is represented by only connected graphs.

Consider the case $s = n$. Clearly, the argument presented above may be continued to include all correlations for which $s - \nu = -1$ since the right-hand side of the equation for $g^{n,n-1}$, by Fig. 1, depends only on terms previously shown to be connected. We show that if $g^{s,s-1}$ $(s \le n-1)$ contains only connected graphs, then $g^{n,n-1}$ also contains only connected graphs.

The steady-state solution of Eq. (4.2) with $s = n$ is

$$g^{n,n-1} = \sum_{A=0}^{\infty} [\zeta^*(iK^n)\Gamma^n]^A \Big[\sum_{i=1}^{n} L_{i,n+1} \sum_{k=2}^{n-1} g_{\{i\}}^{k,k-1} g_{\{n+1\}}^{n-k+1,n-k} + \sum_{1\le i<j\le n} I_{ij} \sum_{k=1}^{n-1} g_{\{i\}}^{k,k-1} g_{\{j\}}^{n-k,n-k-1}\Big]. \quad (4.13)$$

Note that $g_{\{i\}}^{k,k-1}$, $g_{\{n+1\}}^{n-k+1,n-k}$, and $g_{\{j\}}^{n-k,n-k-1}$ all contain only connected graphs by our assumption. Furthermore, $L_{i,n+1} g_{\{i\}}^{k,k-1} g_{\{n+1\}}^{n-k+1,n-k}$ is a connected graph since $L_{i,n+1}$ connects two connected graphs, and $I_{ij} g_{\{i\}}^{k,k-1} g_{\{j\}}^{n-k,n-k-1}$ is a connected graph because I_{ij} connects two connected graphs. The Γ^n terms then act on connected graphs. Since, by the arguments presented in the discussion of $g^{2,1}$, the effect of the Γ^n operator is to bring in another particle and connect it to the particles already present, we see that $g^{n,n-1}$ contains only connected graphs. Clearly, n is arbitrary, and so we have shown that $g^{s,s-1}$ is represented by only connected graphs.

The asymptotic expression for $g^{s,s-1}$ is not secular, i.e., $g^{s,s-1}$ does not contain any time-dependent terms. This is true since to have a secularity we must have an L_{ij} operator acting on a function of $\mathbf{x}_{ij}$ as exhibited in Eq. (2.15). However, this cannot occur because all graphs have been shown to be connected and because there are only exactly the correct number of interaction terms (L_{pr} and I_{mn}) to connect all the particles without any two particles being connected more than once. Thus, I_{pr} cannot precede an L_{pr} acting on the same pair.

B. Case $\nu - s = 0$

Let us now consider Eq. (4.1) for $\nu = s$:

$$\Big(\frac{\partial}{\partial t} + K^s\Big) g^{s,s} - \sum_{i=1}^{s} L_{i,s+1} g_{\{i\}}^{1,0} g_{\{s+1\}}^{s,s} = L^s g^{s+1,s} + \sum_{i=1}^{s} L_{i,s+1} \sum_{k=2}^{s-1} g_{\{i\}}^{k,k-1} g_{\{s+1\}}^{s+1-k,s+1-k} + \sum_{i=1}^{s} L_{i,s+1} \sum_{k=1}^{s-1} g_{\{i\}}^{k,k} g_{\{s+1\}}^{s+1-k,s-k} + I^s g^{s,s-1} + \sum_{1\le i<j\le s} I_{ij} \sum_{k=1}^{s-1} \sum_{m=0}^{1} g_{\{i\}}^{k,m+k-1} g_{\{j\}}^{s-k,s-k-m}. \quad (4.14)$$

For $s = 1$, Eq. (4.14) becomes

$$\frac{\partial g_1^{1,1}}{\partial t} - L_{1,2} g_1^{1,0} g_2^{1,1} = L_{1,2} g_{12}^{2,1}. \quad (4.15)$$

Since g^1 is space independent, $L_{1,2} g_1^{1,0} g_2^{1,1}$ equals zero, and since $L_{1,2} g_{12}^{2,1}$ is independent of t, it follows from Eq. (4.15) that

$$g^{1,1} \sim t L_{1,2} g^{2,1}. \quad (4.16)$$

Thus, $g^{1,1}$ is linearly secular and represented by a connected graph.

For arbitrary s we have the following asymptotic solution:

$$g^{s,s} \sim \sum_{A=0}^{\infty} [\zeta^*(iK^s)\Gamma^s]^A B, \quad (4.17)$$

where B is the right-hand side of Eq. (4.14). Clearly, the $L^s g^{s+1,s}$ and $I^s g^{s,s-1}$ terms of B are connected because of our previous results. But there are also terms including $g^{n,n}$ $(n \le s-1)$. These are connected to terms of the form $g^{k,k-1}$ by L_{pr} or I_{ij} operators. If $g^{n,n}$ $(n \le s-1)$ is a connected graph, then the entire term is connected. Thus, if we assume that $g^{n,n}$ $(n \le s-1)$ is represented by connected graphs, then B contains only connected graphs, and it follows that the function $g^{s,s}$ is also represented by only connected graphs. Clearly, since we have seen that $g^{1,1}$ is represented by a connected graph, then $g^{2,2}$ is also represented by connected graphs. It follows by induction that $g^{s,s}$ will be represented by connected graphs only.

We note that $g^{s,s}$ must have a linear secularity, i.e., a term proportional to t. This occurs because of the

$L^s g^{s+1,s}$ term. The $g^{s+1,s}$ term is represented by a connected graph, such that there is exactly the number of connections needed to just fill the spaces. But L^s operating on $g^{s+1,s}$ produces a phase-averaging interaction (L_{ij}) after a direct interaction (I_{ij}), resulting in a term with an L_{ij} acting on a function of $\mathbf{x}_{ij}$. Such a term yields a secularity. Since $g^{s+1,s}$ itself is not secular, the resulting expression for $g^{s,s}$ has a term proportional to t. Terms proportional to t^0 also occur because of other terms on the right side of Eq. (4.17). For the special case of $s = 1$, where there are no other terms on the right, there is no time-independent term.

C. Arbitrary $\nu - s$

The above arguments for the $g^{s,\nu}$ graphs being all connected can be extended from the cases already considered to the general case by arguments similar to those already presented. The general steady-state solution for $g^{s,\nu}$ is

$$g^{s,\nu} \sim \sum_{A=0}^{\infty} [\zeta^*(iK^s)\Gamma^s]^A \zeta^*(iK^s)C, \tag{4.18}$$

where C is the right-hand side of Eq. (4.1). Figure 1 gives for any s and ν, the $g^{t,\mu}$ terms which appear in C. These $g^{t,\mu}$ all come from lower diagonals or from the same diagonal of which $g^{s,\nu}$ is a member. Consequently, we may easily see that $g^{1,2}$ is represented by only connected graphs since we have already shown that $g^{s,s}$ and $g^{s,s-1}$ are all represented by only connected graphs. Similarly, since $g^{2,3}$ depends on $g^{1,2}$, $g^{s,s}$, and $g^{s,s-1}$, it is also easily seen that it also is represented by only connected graphs. Clearly, the process may be continued to show that all $g^{s,\nu}$ are represented by only connected graphs.

The question of what secularities occur in $g^{s,\nu}$ may now be attacked in either of two ways. Since we know that $g^{s,\nu}$ contains only connected graphs we can express $g^{s,\nu}$ in terms of the distribution functions, keeping only those parts that are represented by connected graphs. Since we already know the secular behavior of the distribution functions, we can obtain the secular behavior of $g^{s,\nu}$. The second method of attack is through the correlation function hierarchy directly. We have already used these to show: that $g^{1,0}$ is nonsecular, that $g^{1,1}$ is directly proportional to t and has no t^0 term, that $g^{s,s-1}$ is nonsecular for all s, and that $g^{s,s}$ has both t^0 and t^1 terms for $s \geq 2$. We shall continue the derivation of the $g^{s,\nu}$ secularities from the correlation function equations, and then consider their derivation through the distribution functions.

In our discussion of the secularities of $g^{s,s}$ we saw that the highest secularity came from the $L^s g^{s+1,s}$ term, and that this term yielded a secularity of order one higher than that of $g^{s+1,s}$. The corresponding term in the general expression for $g^{s,\nu}$ [Eq. (4.18)] is $L^s g^{s+1,\nu}$. By the same argument as presented in the $g^{s,s}$ case, it is clear that the $L^s g^{s+1,\nu}$ term yields a secularity of order one higher than that of $g^{s+1,\nu}$. Referring again to Fig. 1, we see that the order of the highest secularity depends only on what diagonal the term is in, and that counting up from the $g^{s,s-1}$ diagonal to $g^{s,s}$, etc., the order increases by one for each diagonal. Thus, the highest secularity in the expression for $g^{s,\nu}$ is proportional to $t^{\nu-s+1}$.

We might also ask what other orders of secularities occur in $g^{s,\nu}$. Since the only term to appear in C [on the right side of Eq. (4.18)] for $s = 1$ is $L^1 g^{2,\nu}$, the case $s = 1$ is a special case. But $g^{2,\nu}$ contains terms from t^0 up to $t^{\nu-1}$. The L^1 operator acting on $g^{2,\nu}$ then yields secularities of orders from t to t^ν, but not t^0. Since there are no other terms which can contribute, $g^{1,\nu}$ has secularities from t to t^ν, but not t^0. This is not true for $s \geq 2$ since other terms in C can then contribute secularities of orders from t^0 to $t^{\nu-s}$, and the $L^s g^{s+1,\nu}$ term contributes secularities of orders from t^0 to $t^{\nu-s+1}$. Thus, we see that $g^{s,\nu}$ contains secularities of orders from t^0 to $t^{\nu-s+1}$ and $g^{1,\nu}$ contains secularities of orders from t to t^ν.

As noted above, the secularities of the correlation functions $g^{s,\nu}$ may also be obtained from the secularities of the distribution functions $F^{n,\mu}$ by expressing $g^{s,\nu}$ in terms of the $F^{n,\mu}$ functions. We present this as a check on our calculations and as an alternative approach. Since we have shown that $g^{s,\nu}$ contains only connected graphs, we need consider only those terms in the expression for $F^{n,\mu}$ which consist of connected graphs. The only such term is $F^{s,\nu}$ since all other terms are products of $F^{n,\mu}$ functions, and therefore, are represented by disconnected graphs. Furthermore, only connected graphs from $F^{s,\nu}$ contribute to $g^{s,\nu}$. All other terms must cancel. We must thus consider the connected graphs of $F^{s,\nu}$.

The solution for $F^{s,\nu}$ can be written in the form

$$F^{s,\nu} = \mathbf{I}^\nu F^{s,0} + \{\mathbf{L}^1, \mathbf{I}^{\nu-1}\}\mathbf{I}F^{s+1,0} + \{\mathbf{L}^2, \mathbf{I}^{\nu-1}\}\mathbf{I}F^{s+2,0} + \cdots. \tag{4.19}$$

In order to have a connected graph we must have ν at least equal to $(s - 1)$ so as to be able to connect all particles with an I_{ij} or an L_{pr}. If $\nu < s - 1$, all graphs will be disconnected. If $\nu > s - 1$, there will be some spaces with more than one connection; in fact, there will be $(\nu - s + 1)$ spaces with extra connections. In some terms these $(\nu - s + 1)$ extra connections will be arranged so as to yield secularities of order $t^{\nu-s+1}$.

These must occur since all orderings of the L's and $(\nu - 1)I$'s must occur. The lowest-order secularity, except for $s = 1$, is t^0 because of the $\mathbf{I}^\nu F^{s,0}$ term. The $s = 1$ case is special because the $\mathbf{I}^\nu F^{s,0}$ term does not exist. For $\nu \geq 1$, the lowest term is then $\{\mathbf{L}^1, \mathbf{I}^{\nu-1}\}\mathbf{I}F^{2,0}$ which is proportional to t. For the special case $\nu = 0$, we find $F^{1,0}(t) = F^{1,0}(0)$. Thus, we obtain the same results for the secular behavior of the correlation functions deduced from the distribution function behavior as obtained from the correlation function hierarchy.

5. CONCLUSIONS

We have considered the perturbation solutions for the electron distribution and correlation functions for a plasma. We have seen that the perturbation expansion is not uniformly valid, but leads to secular terms for times long compared to the inverse plasma frequency. In order to apply uniformizing techniques designed to eliminate these secularities and obtain a description of a plasma valid for times of the order of its relaxation time to equilibrium, it is necessary to analyze the breakdown of the perturbation expansions. In this paper we have determined the secularities which occur in the plasma expansion. The s-electron distribution function in νth order has secularities of orders from 0 to ν. For the electron correlation functions we have shown that $g^{s,\nu}(t) = 0$, for $\nu \leq s - 2$ follows from the initial condition of molecular chaos, and that the s-electron correlation function in νth order has secularities of orders from 0 to $\nu - s + 1$ (except for the one-particle correlation function which has secularities of orders from 1 to ν). The secular terms determine the behavior of the plasma for times long compared to the duration of a collision, i.e., the inverse plasma frequency. Consequently, it is these terms which must be understood in order to describe the irreversible behavior of a plasma. Our study has classified the secular terms which appear, and thus forms a foundation for corrections to the plasma kinetic equation.

APPENDIX A: SOLUTION FOR $F^0(t)$

The infinite matrix equation is obtained as the limit of matrix equations for N electrons as N is allowed to become infinite. In zero order, we consider first the finite matrix equation which involves N electrons [the elements of which are given in Eqs. (A1) and (A2) below]. After this equation is solved, using as the initial condition molecular chaos, we allow N to go to infinity, and thus obtain the solution for the infinite system.

When the number of electrons is finite, Eq. (2.8) yields the following equations for F^s in zero order:

$$\frac{\partial F^{s,0}}{\partial t} + K^s F^{s,0} = L^s F^{s+1,0} \quad (s < N), \tag{A1}$$

$$\frac{\partial F^{N,0}}{\partial t} + K^N F^{N,0} = 0. \tag{A2}$$

The equation for $F^{N,0}$ has no source term because limiting the number of electrons to N causes distribution functions for more than N electrons to vanish. The solution for $F^{N,0}(t)$ under the simple initial value condition

$$F(0) = \begin{bmatrix} f_1^0 \\ f_1^0 \; f_2^0 \\ \cdot \\ \cdot \\ \cdot \\ \prod_{i=1}^{s} f_i^0 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{A3}$$

is

$$F^{N,0}(t) = e^{-K^N t} \prod_{i=1}^{N} f_i^0(0), \tag{A4}$$

where f_i^0 is the zero-order one-electron distribution function for the ith electron. For a homogeneous plasma, f_i^0 is space independent so that Eq. (A4) reduces to

$$F^{N,0}(t) = \prod_{i=1}^{N} f_i^0(0). \tag{A5}$$

Now, consider Eq. (A1) for $s = N - 1$,

$$\frac{\partial F^{N-1,0}}{\partial t} + K^{N-1} F^{N-1,0} = L^N F^{N,0}. \tag{A6}$$

But, since L acting on a space-independent function yields zero, Eq. (A6) reduces to the same form as Eq. (A2), and has the solution

$$F^{N-1,0}(t) = \prod_{i=1}^{N-1} f_i^0(0). \tag{A7}$$

Clearly, the successive equations for $N - 2$, $N - 3$, etc., will reduce to the form of Eq. (A2) so that the solution for the zeroth order s-body distribution function becomes

$$F^{s,0}(t) = \prod_{i=1}^{s} f_i^0(0) \quad (s \leq N). \tag{A8}$$

The above sequence of logical steps may be applied to a succession of N-electron systems with N increasing to infinity. Clearly, the logic follows the same pattern independent of N, yielding Eq. (A8) as the solution.

Thus, we see that if the initial state of the plasma is molecular chaos, it remains so for all time so that for the infinite-column matrix distribution function

$$F^0(t) = F^0(0) = F(0). \tag{A9}$$

APPENDIX B: SOLUTION FOR F^{1*}

We wish to demonstrate that $F^{1*\nu}$ is given correctly by the general expression for $F^{*\nu}$ in Eq. (2.14). The one-body distribution function in νth order satisfies the equation

$$\frac{\partial F^{1\nu}}{\partial t} = L^1 F^{2\nu}. \tag{B1}$$

From Eq. (B1) it follows that

$$\frac{\partial F^{1\nu}}{\partial t} \sim L^1 F^{2*\nu}. \tag{B2}$$

Therefore, by Whittaker and Watson's theorem which states that the asymptotic expansion of an integral equals the integral of the asymptotic expansion of the integrand,[8] we have

$$F^{1*\nu} = \int_0^\infty L^1 F^{2*\nu}\, d\lambda. \tag{B3}$$

We note that, since $F^{2*\nu}$ depends only on the separation of particles *1* and *2*, $L^1F^{2*\nu}$ is a homogeneous function. Consequently,

$$e^{-K^1\lambda} L^1 F^{2*\nu} = L^1 F^{2*\nu}, \tag{B4}$$

so that Eq. (B3) can be written

$$F^{1*\nu} = \int_0^\infty e^{-K^1\lambda} L^1 F^{2*\nu}\, d\lambda = \mathbf{L}^1 F^{2*\nu}. \tag{B5}$$

When $F^{2*\nu}$, given by Eq. (2.14), is substituted into the right side of Eq. (B5), we obtain

$$F^{1*\nu} = \mathbf{L}^1[\mathbf{I}^2]^\nu f^0 f^0 + \mathbf{L}^1\mathbf{L}^2[\mathbf{I}^3]^\nu f^0 f^0 f^0 + \mathbf{L}^1\mathbf{I}^2\mathbf{L}^2[\mathbf{I}^3]^{\nu-1} f^0 f^0 f^0 + \cdots, \tag{B6}$$

which is the expression for $F^{1*\nu}$ given directly by Eq. (2.14).

APPENDIX C: LANDAU COLLISION INTEGRAL

The simplest form of $\mathbf{\Lambda}$ is given by

$$\mathbf{\Lambda} = L^1\zeta^*(iK^2) I_{12} f_1^0 f_2^0. \tag{C1}$$

We note that Eq. (C1) is in the nondimensional form discussed in Sec. 2. We show that the expression given in Eq. (C1) is just the Landau collision integral.[9] We also put this into the usual Fokker–Planck form. In expanded form, Eq. (C1) is

$$\mathbf{\Lambda} = n\lambda_D^3 \left(\frac{\Phi_0}{kT}\right)^2 \frac{\partial}{\partial v_{1i}} \int d\mathbf{v}_2\, d\mathbf{x}_2 \frac{\partial \Phi(|\mathbf{x}_{12}|)}{\partial x_{12i}} \times \int_0^\infty d\lambda \frac{\partial \Phi(|\mathbf{x}_{12} - \mathbf{v}_{12}\lambda|)}{\partial x_{12j}} D_j f_1^0 f_2^0, \tag{C2}$$

where

$$D_j \equiv \left(\frac{\partial}{\partial v_{1j}} - \frac{\partial}{\partial v_{2j}}\right), \quad \mathbf{v}_{12} \equiv \mathbf{v}_1 - \mathbf{v}_2, \quad \mathbf{x}_{12} = \mathbf{x}_1 - \mathbf{x}_2,$$

and where we have used

$$e^{-K^2\lambda}\Phi(|\mathbf{x}_{12}|) = \Phi(|\mathbf{x}_{12} - \mathbf{v}_{12}\lambda|).$$

This collision integral expression has been previously derived and used in this form in the treatments of the weakly-coupled gas[5] and the plasma in a strong magnetic field.[10]

The $\mathbf{x}_2$ integration may be changed into an integration over $\mathbf{x}_{12}$ and Parseval's theorem used to express (C2) in terms of Fourier transforms. This is

$$\mathbf{\Lambda} = (n\lambda_D^3)\left(\frac{\Phi_0}{kT}\right)^2 \frac{(2\pi)^3}{2} \frac{\partial}{\partial v_{1i}} \int d\mathbf{v}_2 (D_j f_1^0 f_2^0) \times \int d\mathbf{k} k_i k_j \tilde{\Phi}^2(|\mathbf{k}|) \int_{-\infty}^\infty e^{-i\mathbf{k}\cdot\mathbf{v}_{12}\lambda}\, d\lambda, \tag{C3}$$

where $\tilde{\Phi}(|\mathbf{k}|)$ is the Fourier transform of $\Phi(|\mathbf{x}_{12}|)$. The λ integration is easily performed, yielding a delta function,

$$\int_{-\infty}^\infty d\lambda e^{-i\mathbf{k}\cdot\mathbf{v}_{12}\lambda} = \frac{2\pi}{|\mathbf{v}_{12}|}\delta(k_\parallel), \tag{C4}$$

where $k_\parallel$ is the component of $\mathbf{k}$ parallel to $\mathbf{v}_{12}$. The $\mathbf{k}$ integration is then expressed as

$$\int d\mathbf{k} = \int d\mathbf{k}_\perp \int dk_\parallel. \tag{C5}$$

Upon performing the $k_\parallel$ integration and part of the $\mathbf{k}_\perp$ integration, we obtain

$$\mathbf{\Lambda} = 8\pi^5 (n\lambda_D^3)\left(\frac{\Phi_0}{kT}\right)^2 \frac{\partial}{\partial v_{1i}} \int d\mathbf{v}_2 \times \frac{1}{|\mathbf{v}_{12}|}\left(\delta_{ij} - \frac{v_{12i}v_{12j}}{v_{12}^2}\right) D_j f_1^0 f_2^0 \int_0^\infty dk_\perp k_\perp^3 \tilde{\Phi}^2(k_\perp), \tag{C6}$$

where $k_\perp = |\mathbf{k}_\perp|$. With the insertion of the Coulomb potential $[\Phi(|\mathbf{x}_{12}|) = 1/|\mathbf{x}_{12}|$ or $\tilde{\Phi}(k) = (2\pi^2 k^2)^{-1}]$, and with $\Phi_0 = e^2/\lambda_D$, Eq. (C6) takes on precisely the

[8] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University Press, Cambridge, England, 1958), 4th ed., p. 153.

[9] L. Landau, Zh. Eksp. Teor. Fiz. 7, 203 (1937).

[10] S. Radin, A. Kritz, and G. Sandri, Phys. Rev. 157, 150 (1967).

form of the Landau collision integral[9] (in nondimensional notation):

$$\Lambda = 2\pi(n\lambda_D^3)\frac{(e^2/\lambda_D)^2}{(kT)^2}\frac{\partial}{\partial v_{1i}}\int d\mathbf{v}_2\frac{1}{|\mathbf{v}_{12}|}\left(\delta_{ij}-\frac{v_{12i}v_{12j}}{v_{12}^2}\right)\times\left[\left(\frac{\partial}{\partial v_{1j}}-\frac{\partial}{\partial v_{2j}}\right)f_1^0f_2^0\right]\int_0^\infty\frac{dk_\perp}{k_\perp}. \quad (C7)$$

The $k_\perp$ integration yields the logarithmic divergence, and requires cutoffs at $k_\perp$'s corresponding to the Debye distance and to the distance of closest approach.

A slight rearrangement of Eq. (C7) puts it into the Fokker–Planck form, as follows:

$$\Lambda = -\frac{\partial}{\partial v_{1i}}(A_i f_1^0)+\frac{1}{2}\frac{\partial^2}{\partial v_{1i}\partial v_{1j}}(B_{ij}f_1^0), \quad (C8)$$

where

$$A_i = 2\pi n\lambda_D^3\frac{e^4/\lambda_D^2}{(kT)^2}\int_0^\infty\frac{dk_\perp}{k_\perp}\int d\mathbf{v}_2 f_2^0\left(\frac{\partial}{\partial v_{1j}}-\frac{\partial}{\partial v_{2j}}\right)\times\frac{\delta_{ij}-v_{12i}v_{12j}/|\mathbf{v}_{12}|^2}{|\mathbf{v}_{12}|} \quad (C9)$$

or

$$A_i = -8\pi n\lambda_D^3\frac{e^4/\lambda_D^2}{(kT)^2}\int_0^\infty\frac{dk_\perp}{k_\perp}\int d\mathbf{v}_2 f_2^0\frac{v_{12i}}{|\mathbf{v}_{12}|^3}, \quad (C10)$$

and

$$B_{ij} = 4\pi n\lambda_D^3\frac{e^4/\lambda_D^2}{(kT)^2}\int_0^\infty\frac{dk_\perp}{k_\perp}\times\int d\mathbf{v}_2 f_2^0\frac{\delta_{ij}-v_{12i}v_{12j}/|\mathbf{v}_{12}|^2}{|\mathbf{v}_{12}|}. \quad (C11)$$

The Landau kinetic equation is, in nondimensional form,

$$\partial f/\partial t = \Lambda. \quad (C12)$$

If we write Eq. (C12) in dimensional form, we have

$$\frac{\partial f_1^0}{\partial t} = \frac{(kT/m)^{\frac{1}{2}}}{\lambda_D}\Lambda, \quad (C13)$$

where now t has the dimensions of time, and Λ is still in the nondimensional form. In completely dimensional form, Eq. (C13) becomes

$$\frac{\partial f_1^0}{\partial t} = 2\pi\frac{ne^4}{(kT)^2}\left(\frac{kT}{m}\right)^{\frac{1}{2}}\frac{\partial}{\partial v_{1i}}\int d\mathbf{v}_2\frac{1}{|\mathbf{v}_{12}|}\left(\delta_{ij}-\frac{v_{12i}v_{12j}}{v_{12}^2}\right)\times\left[\left(\frac{\partial}{\partial v_{1j}}-\frac{\partial}{\partial v_{2j}}\right)f_1^0f_2^0\right]\int_0^\infty\frac{dk_\perp}{k_\perp}, \quad (C14)$$

where we have used the normalization

$$\int d\mathbf{v}_1 f_1^0 = 1. \quad (C15)$$

APPENDIX D: RELATIONSHIP BETWEEN CORRELATION FUNCTIONS AND DISTRIBUTION FUNCTIONS

The s-particle distribution function may be expressed in terms of a sum over products of correlation functions. Each term in the expansion must involve precisely s particles, and all possible arrangements of products of correlation functions involving a total of s particles must appear. Furthermore, all distinct permutations of numbered particles must appear. Thus, for example, the first few are:

$$F^1 = g^1,$$
$$F_{12}^2 = g_1^1g_2^1 + g_{12}^2,$$
$$F_{123}^3 = g_1^1g_2^1g_3^1 + g_1^1g_{23}^2 + g_2^1g_{13}^2 + g_3^1g_{12}^2 + g_{123}^3.$$

In general then,

$$F^s = \sum(g^1)^p(g^2)^q(g^3)^r\cdots, \quad (D1)$$

where there are p single-particle correlation functions, q two-particle correlation functions, etc., and

$$p + 2q + 3r + \cdots = s. \quad (D2)$$

The sum is over all $p, q, r, \cdots$ which satisfy Eq. (D2) and over all distinct permutations of the numbered particles. This set of equations actually serves to define the correlation functions.

Equations (D1) may also be inverted to obtain expressions for the correlation functions in terms of the distribution functions.[6] Examples are:

$$g^1 = F^1,$$
$$g_{12}^2 = F_{12}^2 - F_1^1F_2^1,$$
$$g_{123}^3 = F_{123}^3 - F_1^1F_{23}^2 - F_2^1F_{13}^2 - F_3^1F_{23}^2 + 2F_1^1F_2^1F_3^1.$$

In general,

$$g^s = \sum(-1)^{p+q+r+\cdots-1}(p+q+r+\cdots-1)!\,(F^1)^p(F^2)^q(F^3)^r\cdots, \quad (D3)$$

where

$$p + 2q + 3r + \cdots = s. \quad (D4)$$

Here the sum is over all $p, q, r, \cdots$ which satisfy Eq. (D4) and over all distinct permutations of the numbered particles.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 2 FEBRUARY 1968

Another Interpretation of the Optical Scalars*

R. KANTOWSKI†
The University of Texas, Austin, Texas

(Received 29 June 1967)

In this paper an interpretation of the optical scalars θ and σ (the expansion and shear of an irrotational null congruence) is given in terms of the principal curvatures of a two-dimensional subspace of the instantaneous rest frame of an arbitrary observer. The two space is defined as the intersection of the observer's rest frame and the particular null hypersurface the observer is intersecting.

I. INTRODUCTION

The geometrical-optics approximation for a test electromagnetic field in a given curved space-time introduces a sequence of null hypersurfaces $S(x^\alpha) =$ const.[1] These null hypersurfaces have normals $k_a \equiv S_{,a}$ which are null and tangent to an irrotational geodesic congruence. The optical scalars θ and σ are then given by

$$\theta \equiv \tfrac{1}{2} k^a_{;a} \tag{1}$$

$$\sigma \equiv [\tfrac{1}{2} k_{(a;b)} k^{a;b} - \theta^2]^{\frac{1}{2}}. \tag{2}$$

Sachs[2] has interpreted θ and σ, respectively, as the rate of expansion and shear of a shadow cast by some opaque object placed orthogonally in the beam of light. The purpose of this paper is to give another geometrical interpretation of θ and σ. Before proceeding with the interpretation a short review of two-surfaces is in order.[3]

Consider an open set of a three-dimensional positive-definite Riemannian manifold which is covered by a single coordinate patch. Let the coordinates be x^α ($\alpha = 1, 2, 3$) and the metric tensor have components in the natural basis $g_{\alpha\beta}$. Let $f(x^\alpha) = 0$ or $x^\alpha = x^\alpha(y^A)$, ($A = 1, 2$) be a two-surface in this three space. There are two fundamental structures defined on this surface, they are the first and second fundamental forms. The first fundamental form is the metric and it gives the intrinsic structure of the two surface. The second fundamental form gives the relation (at least locally) of the two-space to its embedding three space. Let η^α be a unit normal to the surface. The first and second fundamental forms are given, respectively, by[4]

$$g_{AB} = g_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^A} \frac{\partial x^\beta}{\partial y^B}, \tag{3}$$

$$N_{AB} = \frac{\partial x^\alpha}{\partial y^A} \frac{\delta \eta_\alpha}{\delta y^B}. \tag{4}$$

Because N_{AB} is symmetric it has two orthogonal eigendirections called the principal curvature directions and two eigenvalues $K_\pm$ called the principal curvatures.

The eigenvalue equation is

$$N_{AB} E^B_\pm = K_\pm g_{AB} E^B_\pm, \tag{5}$$

where $E^B_\pm$ are unit eigenvectors. The eigenvalue equation (5) can easily be written in the following form:

$$\frac{\delta \eta^\alpha}{\delta s_\pm} = K_\pm e^\alpha_\pm, \tag{6}$$

where $e^\alpha_\pm = (\partial x^\alpha / \partial y^A) E^A_\pm$ and the derivative is taken along a curve in the surface whose tangent is $e^\alpha_\pm$. Equation (6) follows from Eq. (5) when you recall that $(\delta/\delta s_\pm)(\eta^\alpha \eta_\alpha) = 0$.

Now let $x^\alpha(s) = x^\alpha[y^A(s)]$ be a curve in the surface through the point p. The curvature of this curve K is defined by

$$\frac{\delta^2 x^\alpha}{\delta s^2} = K N^\alpha, \tag{7}$$

where $N^\alpha N_\alpha = 1$ and $K > 0$.

The curve $x^\alpha(s)$ is called normal at p if and only if $N^\alpha = \pm \eta^\alpha$, i.e., the normal to the curve and the normal to the surface are parallel at p. By expanding $(\delta/\delta s)[(dx^\alpha/ds)n_\alpha] = 0$ for a normal curve $x^\alpha(s)$ in a principal direction it follows that $K_\pm$ (the principal curvatures) are equal in magnitude to the normal curvatures in the principal curvature directions. We are now ready to proceed with the main theorem.

II. CONSTRUCTION

Let u^a ($a = 0, 1, 2, 3$) be the unit tangent to the world line of some observer traveling in a given patch of space-time and suppose the observer intercepts the null hypersurface $S(x^a) = 0$ (electromagnetic wave front) at some point p. At p we now construct the instantaneous rest frame of the observer. It is defined to be the locus of all geodesics through p which are orthogonal to u^a. The intersection of the null hypersurface $S = 0$ with the instantaneous rest frame of the observer forms a two-dimensional subspace Σ. This

* Research supported by Aerospace Research Labs, OAR, AF-33(615)-1029.

† Present address: Southwest Center for Advanced Studies, Dallas, Texas.

[1] A. Trautman, *Brandeis Lectures on General Relativity* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964), Vol. 1, p. 134.

[2] R. K. Sachs, Proc. Roy. Soc. (London) **264**, 309 (1961).

[3] J. A. Schouten, *Ricci-Calculus* (Springer-Verlag, Berlin, 1954).

[4] $\delta/\delta v$ is the usual invariant differentiation operator.

two-space is a surface in the instantaneous rest frame and therefore at p has two principal curvatures $k_\pm$. We now relate these principal curvatures to the optical scalars by the following.

Theorem: $$(\theta \pm \sigma) + (k^a u_a) k_\pm = 0. \tag{8}$$

The theorem[5] shows how simply the optical scalars are related to the curvature structure of the electromagnetic wave fronts.

Before presenting the proof it may be instructive to consider a simple example. Consider an observer in Minkowski space a distance d away from a source of light; everything is at rest in the observers rest frame. Suppose the light source emits a single pulse of light. The null hypersurface is then a spherical shell traveling out from the source. When it strikes the observer the intersection of his instantaneous rest frame and the null hypersurface is a sphere of radius d. The two principal curvatures for a sphere are both equal to $1/d$. The expansion and shear are given by

$$\theta = 2\pi\nu/d,$$
$$\sigma = 0, \tag{9}$$

where

$$k^a u_a = -2\pi\nu.$$

All of which is in agreement with the above theorem which we now prove. Let u^a be the unit normal to the instantaneous rest frame of the observer. $u^a u_a = -1$ and u^a is equal to the four velocity of the observer at p. Let η^a be the unit normal to Σ which is orthogonal to u^a. This is the normal drawn in the instantaneous rest frame of the observer. Because the two surface is contained in $S = 0$, it follows that η^a is not linearly independent of k^a and u^a. In fact

$$\eta^a = -\left(u^a + \frac{k^a}{k^b u_b}\right). \tag{10}$$

An important observation to make at this point is that the second fundamental form of the rest frame vanishes at p. To see this we go to a set of geodesic normal coordinates adapted to the rest frame. The metric becomes

$$ds^2 = -(dx^0)^2 + g_{\alpha\beta}\, dx^\alpha\, dx^\beta \tag{11}$$

and $u^a = \delta_0^a$.

The second fundamental form is given by

$$N_{\alpha\beta} = -\delta_\alpha^a \frac{\delta(\delta_a^0)}{\delta x^\beta} = \tfrac{1}{2} g_{\alpha\beta,0}. \tag{12}$$

[5] The signature is taken to be $(-, +, +, +)$ and $-k^a u_a$ the angular frequency of the electromagnetic radiation seen by the observer.

Because the hypersurface is defined as the locus of all geodesics normal to u^a at p we will see that $g_{\alpha\beta,0} = 0$ at p. Let $x^a(s)$ be one of the geodesics. Then

$$\frac{d^2x^a}{ds^2} + \left\{{a \atop bc}\right\} \frac{dx^b}{ds}\frac{dx^c}{ds} = 0. \tag{13}$$

But $x^0(s) = 0$ and Eq. (13) implies

$$\left\{{0 \atop bc}\right\} \frac{dx^b}{ds}\frac{dx^c}{ds} = 0. \tag{14}$$

Because Eq. (14) must be satisfied at p for all dx^a/ds orthogonal to u^a we have

$$\left\{{0 \atop \alpha\beta}\right\}\Big|_p = 0 \rightarrow g_{\alpha\beta,0}\big|_p = 0 \tag{15}$$

and we conclude that the second fundamental form at the point p vanishes. This is equivalent to saying that

$$\left.\frac{\delta u^a}{\delta v}\right|_p = 0 \tag{16}$$

for all derivatives taken tangent to the instantaneous rest frame.

Now let $e_\pm^a$ be unit vectors which are tangent to the two-surface and point in the principal directions and let $k_\pm$ be the two principal curvatures. If $x_\pm^a(s_\pm)$ are two curves through p in the two principal directions then it follows that at p,

$$\frac{\delta\eta^a}{\delta s_\pm} = k_\pm e_\pm^a. \tag{17}$$

The invariant derivative is computed using the full metric, however, at p the second fundamental form of the rest frame vanishes and the derivatives in Eq. (17) involve only the metric of the rest frame. Equation (17) is easily established by using the geodesic normal coordinates and recalling the definitions of principal curvatures and principal curvature directions given in the Introduction.

Now let us compute Eq. (17) using the representation of η^a in terms of the u^a and k^a, Eq. (10):

$$\frac{\delta\eta^a}{\delta s_\pm} = \frac{-\delta}{\delta s_\pm}\left(u^a + \frac{k^a}{k^b u_b}\right). \tag{18}$$

Recalling Eq. (16) we have

$$-\frac{\delta k^a}{\delta s_\pm}\frac{1}{k^b u_b} + \frac{k^a[(\delta k^b/\delta s_\pm) a_b]}{(k^c u_c)^2} = +k_\pm e_\pm^a; \tag{19}$$

$e_\pm^a$ are not only orthonormal but also normal to k^a because they are tangent to the null hypersurface of which k^a is the normal.

Transvecting Eq. (19) with $e_\pm^a$ we obtain

$$e_\pm^a \frac{\delta k_a}{\delta s_\pm} + (k^b u_b) k_\pm = 0. \tag{20}$$

If we are dealing with a congruence of null rays we can write

$$e^a_\pm \frac{\delta k_a}{\delta s_\pm} = e^a_\pm k_{a;b} e^b_\pm. \tag{21}$$

By writing the metric in the following form:

$$g^{ab} = 2k^{(a}m^{b)} + e^a_+ e^b_+ + e^a_- e^b_-, \tag{22}$$

where $m^a k_a = +1$, $m_a e^a_\pm = m_a m^a = 0$, we observe that

$$\theta = -k^b u_b (k_+ + k-)/2,$$
$$\sigma = -k^b u_b (k_+ - k-)/2$$

or

$$(\theta \pm \sigma) + (k^b u_b) k_\pm = 0 \tag{23}$$

and the theorem is complete.