JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 5, NUMBER 10 October 1964

Lorentz Invariant Multichannel Scattering Formalism

THOMAS F. JORDAN*
Institut für theoretische Physik, Universität Bern, Bern, Switzerland
(Received 25 February 1964)

A manifestly Lorentz invariant Hamiltonian formalism for multichannel scattering and production processes is developed by making two simple and natural extensions of the ordinary quantum mechanical formalism. The first is the asymptotic covariance postulate of Fong and Sucher which is essentially a necessary and sufficient condition for Lorentz invariance of the scattering amplitudes. The second is a Lorentz invariant extension of the asymptotic condition. It is shown that the latter is in fact no extension at all in a case where the total momentum operators of the asymptotic (unperturbed) systems are the same as the total momentum operators of the interacting system. In such a case the ordinary multichannel scattering formalism is completely Lorentz invariant whenever asymptotic covariance is satisfied.

I. INTRODUCTION

THERE is quite a difference between the concepts and techniques of ordinary nonrelativistic quantum mechanics and those of relativistic quantum field theory or dispersion theory. One would like to know if this reflects a fundamental inability of Hamiltonian theories to provide a suitable Lorentz invariant quantum mechanical description of scattering and production processes or if a Hamiltonian formalism might yet be found to be useful in understanding such phenomena. There is reason to doubt that a Lorentz invariant Hamiltonian formalism can be used to describe nontrivial particle motion in terms of covariant world-lines in space–time; for two or three particles in classical mechanics it has been shown that only nonaccelerated motion can be so described.[1-3] But it has been emphasized by several authors that one can construct a Hamiltonian theory of particle dynamics which is symmetric under all special relativistic transformations of reference frame[1,4-7] and it has been shown by the construction of a specific two-particle model that such a theory can produce a nontrivial Lorentz invariant scattering amplitude according to the rules of ordinary "nonrelativistic" scattering theory.[8]

General conditions for the Lorentz invariance of the S matrix in a Hamiltonian theory of the scattering of a fixed number of particles have been developed by Fong and Sucher.[9] They postulate a condition which they call "asymptotic covariance" (because of its natural physical interpretation) and show that this condition is essentially equivalent to the Lorentz invariance of the S matrix. On the other hand, the multichannel scattering formalism of ordinary Hamiltonian quantum mechanics is capable, in principle, of describing the scattering

* Present address: Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania.

[1] D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963).

[2] D. G. Currie, J. Math. Phys. 4, 1470 (1963).

[3] J. T. Cannon and T. F. Jordan, J. Math. Phys. **5**, 299 (1964).

[4] P. A. M. Dirac, Rev. Mod. Phys. **21**, 392 (1949).

[5] L. H. Thomas, Phys. Rev. **85**, 868 (1952); B. Bakamjian and L. H. Thomas, Phys. Rev. **92**, 1300 (1953); B. Bakamjian, Phys. Rev. **121**, 1849 (1961).

[6] L. L. Foldy, Phys. Rev. **122**, 275 (1961).

[7] E. C. G. Sudarshan, *1961 Brandeis Summer Institute Lectures in Theoretical Physics* (W. A. Benjamin Company, New York, 1962).

[8] T. F. Jordan, A. J. Macfarlane, and E. C. G. Sudarshan, Phys. Rev. **133**, B487 (1964).

[9] R. Fong and J. Sucher, J. Math. Phys. **5**, 456 (1964); R. Fong, thesis, University of Maryland, 1963.

and production of any number of particles. This paper is based on the observation that the inclusion of asymptotic covariance can make this "nonrelativistic" multichannel scattering formalism completely and manifestly Lorentz invariant.

We use a mathematical formulation of multichannel scattering in terms of wave operators or Møller matrices.[10,11] Some of the essential elements of this formalism are outlined in Sec. II. In Sec. III we state the definition and consequences of asymptotic covariance. It should be recognized that the definition and theorems of this section are mostly a straightforward extension to the multichannel case of the work of Fong and Sucher.[9] In particular, we show that asymptotic covariance is still essentially equivalent to the Lorentz invariance of the scattering amplitudes. In Sec. IV, we review Fong and Sucher's interpretation of asymptotic covariance, develop an alternate interpretation in terms of operators analogous to those occurring in quantum field theory, and discuss the situation that results from the incorporation of asymptotic covariance into the multichannel scattering formalism. The situation is essentially that the wave operators now have an intertwining property between a unitary representation of the Poincaré (inhomogeneous Lorentz) group for the interacting system and the representation for each of the asymptotic or unperturbed channel systems. This is a Lorentz invariant extension of the intertwining property of the wave operators between the Hamiltonian of the interacting system and the "free" Hamiltonian of each of the channel systems.

In Sec. V we state a formal definition of the ordinary "nonrelativistic" multichannel scattering problem in terms of three conditions to be satisfied by the wave operators: the wave operators are partially isometric between subspaces which satisfy certain conditions; they have the intertwining property between the Hamiltonian and each of the channel Hamiltonians; they satisfy a boundary or asymptotic condition. Asymptotic covariance is a Lorentz invariant extension of the second condition. The first condition can be made Lorentz invariant rather trivially because the subspaces involved are each invariant under the relevant representation of the Poincaré group. A natural Lorentz invariant extension of the third condition, the asymptotic condition, gives a multichannel scattering formalism which is completely Lorentz invariant. This is displayed in Sec. VI as a formal definition of the Lorentz invariant multichannel scattering problem. The Lorentz invariance is manifest in that only Lorentz invariant sets of operators and Lorentz invariant subspaces are involved in the definition.

In our formulation of the ordinary "nonrelativistic" scattering problem we explain our choice of the boundary condition satisfied by the wave operators by showing that the ordinary "nonrelativistic" asymptotic limits are the unique solution. We show also that similar Lorentz invariant asymptotic limits are the unique solution of the Lorentz invariant scattering problem. But we want to emphasize that the choice of this particular asymptotic condition is probably not essential for a Lorentz invariant extension of the formalism. The incorporation of asymptotic covariance, which is the foundation of our extension, is completely independent of the asymptotic condition.

In Sec. VII we raise the question: what restrictions has our Lorentz invariant extension of the scattering problem put on its possible solutions? Since the restrictions imposed by asymptotic covariance are evidently just what are needed to obtain Lorentz invariant scattering amplitudes, we will be really concerned only with the restrictions imposed by the Lorentz invariant extension of the asymptotic condition. It may happen that there are actually no restrictions from the latter source because it may happen that the Lorentz invariant asymptotic condition contains no actual extension of the ordinary "nonrelativistic" asymptotic condition. In Sec. VIII it is shown that this is the case whenever the total momentum operators (generators of space translations in the representation of the Poincaré group) for each of the asymptotic or unperturbed channel systems are essentially the same as the total momentum operators for the interacting system. This property has been assumed as a matter of convenience in the construction of various models of interacting systems.[5,6,8,9] We have here a more fundamental motivation. If a system has this property and satisfies asymptotic covariance, then the ordinary quantum mechanical scattering formalism with the "nonrelativistic" asymptotic condition is completely Lorentz invariant.

In Sec. IX we outline briefly the problem of an analogous Galilei invariant scattering formalism. The asymptotic or unperturbed channel systems have unitary representations up to a factor of the inhomogeneous Galilei group with different factors for different channels. Asymptotic covariance and

[10] J. M. Jauch, Helv. Phys. Acta **31**, 127, 661 (1958).

[11] T. F. Jordan, J. Math. Phys. **3**, 414, 429 (1962). Further references to the literature of quantum mechanical scattering formalism can be found in these papers.

Galilei invariant scattering amplitudes are not always possible and one must develop a generalized formalism in which the scattering amplitudes are invariant only up to phase factors which may depend on the channels and on the transformation in the inhomogeneous Galilei group. Besides asymptotic covariance, the conditions involved in the Galilei invariant formalism are at least as strong as, if not actually stronger than, those of the Lorentz invariant formalism. This suggests again that the Lorentz invariant asymptotic condition does not restrict too severely the possible solutions of the scattering problem.

To make a casual reading possible, the proofs of all the theorems have been collected in the final Sec. X. Simplicity takes precedence over mathematical rigor in certain familiar situations.

II. MULTICHANNEL SCATTERING FORMALISM—BASIC EQUATIONS

We use a quantum mechanical multichannel scattering formalism[10,11] which may be described as follows in terms of linear operators on a separable Hilbert space $\mathcal{H}$. The different channels are labeled by an index a (or b) which may take on a finite or countable number of values. For each channel a there are ten linear operators H_a, $\mathbf{P}_a$, $\mathbf{J}_a$, $\mathbf{K}_a$ which are respectively the generators for time translations, space translations, space rotations, and rotation-free Lorentz transformations in a description of the "free" asymptotic or unperturbed motion in that channel. These operators are assumed to generate a (continuous) unitary representation of the Poincaré (proper orthochronous inhomogeneous Lorentz) group on some subspace of $\mathcal{H}$. (This subspace may be different for each a.) In other words, H_a, $\mathbf{P}_a$, $\mathbf{J}_a$, $\mathbf{K}_a$ are self-adjoint operators satisfying the "commutation relations" of the Poincaré group on a subspace of $\mathcal{H}$.

The dynamics of the actual interacting system is described by a Hamiltonian H which also is assumed to be a self-adjoint linear operator on a subspace of $\mathcal{H}$. Scattering is defined by a comparison between the interacting dynamics determined by H and the "free" asymptotic or unperturbed dynamics determined by each of the channel Hamiltonians H_a. This comparison is effected by a family of linear operators $\Omega_\pm^a$, two for each channel. For each channel a the wave operators $\Omega_\pm^a$ are assumed to be partially isometric operators[12] from a subspace $\mathfrak{D}_a$ in the continuum subspace[13] of H_a (subspace of continuous spectrum "eigenstates" of H_a or, in other words, the orthogonal complement of the subspace of true eigenvectors or "bound states" of H_a) to subspaces $\mathcal{R}_a^{(\pm)}$ in the continuum subspace of H. In other words, if E_a and $F_\pm^a$ are the projection operators whose ranges are, respectively, $\mathfrak{D}_a$ and $\mathcal{R}_a^{(\pm)}$ ($E_a\mathcal{H}=\mathfrak{D}_a$ and $F_\pm^a\mathcal{H} = \mathcal{R}_a^{(\pm)}$), it is assumed that[14]

$$\Omega_\pm^{a\dagger}\Omega_\pm^a = E_a \tag{2.1}$$

$$\Omega_\pm^a\Omega_\pm^{a\dagger} = F_\pm^a. \tag{2.2}$$

It is assumed, furthermore, that $\mathcal{R}_a^{(+)}$ is orthogonal to $\mathcal{R}_b^{(+)}$ and that $\mathcal{R}_a^{(-)}$ is orthogonal to $\mathcal{R}_b^{(-)}$ for $a \neq b$, and that the direct sum over a of the subspaces $\mathcal{R}_a^{(+)}$ is the same as the direct sum over a of the subspaces $\mathcal{R}_a^{(-)}$. Let $\mathcal{R}$ denote this direct sum subspace and let F be the projection operator whose range is $\mathcal{R}$ ($F\mathcal{H} = \mathcal{R}$). In terms of projection operators we have that

$$F_+^aF_+^b = 0 = F_-^aF_-^b \tag{2.3}$$

if $a \neq b$, and

$$\sum_a F_+^a = \sum_a F_-^a = F. \tag{2.4}$$

From these assumptions it follows that

$$\Omega_\pm^{a\dagger}\Omega_\pm^b = \delta_{ab}E_a, \tag{2.5}$$

$$\Omega_\pm^aE_a = \Omega_\pm^a, \tag{2.6}$$

$$F_\pm^a\Omega_\pm^b = \delta_{ab}\Omega_\pm^a. \tag{2.7}$$

It is assumed that for each channel a

$$H\Omega_\pm^a = \Omega_\pm^aH_a \tag{2.8}$$

or, more precisely, that

$$e^{iHt}\Omega_\pm^a = \Omega_\pm^ae^{iH_at} \tag{2.9}$$

for all real t. From this assumption it follows that $\mathcal{R}_a^{(\pm)}$ reduce[15] H and $\mathfrak{D}_a$ reduces H_a or, in other words, that $F_\pm^a$ commute with H and E_a commutes with H_a for each a. From Eq. (2.4) it then follows that H is reduced also by $\mathcal{R} = F\mathcal{H}$ or, in other words, that H commutes with F.

Let ϕ^a be a vector in $\mathfrak{D}_a$ representing a certain state of the "free" asymptotic or unperturbed motion in channel a (e.g., a packet of "plane-wave momen-

[12] F. J. Murray and J. von Neumann, Ann. Math. 37, 116 (1936).

[13] M. H. Stone, *Linear Transformations in Hilbert Space* (American Mathematical Society Colloquium Publications, New York, 1932), Vol. 15, Theorem 5.13.

[14] Throughout this paper when we write equations involving operators with both + and − indices we mean that they should be read as two separate equations, one with the upper and the other with the lower set of indices.

[15] F. Riesz and B. Sz. Nagy, *Functional Analysis* (Ungar Publishing Company, New York, 1955), p. 302.

tum eigenstates") and let ϕ^b be a similar state for channel b. The probability amplitude for scattering from ϕ^a to ϕ^b is defined to be

$$(\Omega_-^b\phi^b,\ \Omega_+^a\phi^a) = (\phi^b,\ \Omega_-^{b\dagger}\Omega_+^a\phi^a). \qquad (2.10)$$

In terms of the S operator defined by[10]

$$S = \sum_a \Omega_+^a \Omega_-^{a\dagger} \qquad (2.11)$$

we have, from Eqs. (2.5) and (2.6), that the scattering amplitude (2.10) is equal to the matrix elements

$$(\psi_b^{(-)},\ S\psi_a^{(-)}) = (\psi_b^{(+)},\ S\psi_a^{(+)}),$$

where

$$\psi_a^{(\pm)} = \Omega_\pm^a \phi^a \qquad (2.12)$$

are the states of the interacting system corresponding to the state ϕ^a of the channel system. From Eqs. (2.2), (2.4), (2.5), and (2.6) we find that the operator S defined by Eq. (2.11) satisfies the equations

$$S^\dagger S = SS^\dagger = F. \qquad (2.13)$$

Thus S is a unitary operator on the subspace $\mathfrak{R} = F\mathfrak{K}$.

This scattering formalism has so far involved only the Hamiltonians. We turn now to our main business which is to incorporate the whole Poincaré group.

III. DEFINITION AND CONSEQUENCES OF ASYMPTOTIC COVARIANCE

Let L denote an element of the Poincaré group, and for each channel a let $U_a(L)$ be the unitary representation of the Poincaré group generated by H_a, $\mathbf{P}_a$, $\mathbf{J}_a$, $\mathbf{K}_a$. We assume that $\mathfrak{D}_a = E_a\mathfrak{K}$ is contained in the subspace of $\mathfrak{K}$ which supports the unitary representation $U_a(L)$.

Definition: A scattering system has the property of *asymptotic covariance* if for each element L of the Poincaré group there exists a (bounded) linear operator $W(L)$ on $\mathfrak{K}$ such that for each channel a

$$W(L)\Omega_\pm^a = \Omega_\pm^a U_a(L). \qquad (3.1)$$

Theorem 1. For each element L of the Poincaré group let $W(L)$ be a (bounded) linear operator on $\mathfrak{K}$ which satisfies the defining relations (3.1) for asymptotic covariance. Then $W(L)$ leaves invariant each of the subspaces $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathfrak{K}$ and $\mathfrak{R} = F\mathfrak{K}$. The operation of $W(L)$ on the subspace $\mathfrak{R} = F\mathfrak{K}$ is uniquely determined by the relation

$$W(L)F = \sum_a \Omega_+^a U_a(L)\Omega_+^{a\dagger} = \sum_a \Omega_-^a U_a(L)\Omega_-^{a\dagger}. \qquad (3.2)$$

The operators $W(L)F$ are unitary on $\mathfrak{R} = F\mathfrak{K}$ and form on $\mathfrak{R}$ a (continuous) unitary representation of the Poincaré group. The one-parameter subgroup of operators $W(L)F$ representing time translations L is the same as the one-parameter subgroup $e^{iHt}F$ (t real).

Note that the relation (3.1) is just an extension to the whole Poincaré group of the relation (2.9) which is assumed in traditional scattering theory to hold for the one-parameter subgroup of time translations.

Note also that from Eqs. (3.2) and the adjoint of Eqs. (2.7) it follows immediately that

$$W(L)F_\pm^a = \Omega_\pm^a U_a(L)\Omega_\pm^{a\dagger}. \qquad (3.3)$$

From this we see that the spectra of the generators of the part of the unitary representation $W(L)$ on $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathfrak{K}$ are the same as the spectra of the corresponding generators H_aE_a, $\mathbf{P}_aE_a$, $\mathbf{J}_aE_a$, $\mathbf{K}_aE_a$ of the part of the unitary representation $U_a(L)$ on $\mathfrak{D}_a = E_a\mathfrak{K}$.

Theorem 2. For a scattering system to have the property of asymptotic covariance it is necessary that $U_a(L)$ commutes with E_a for each element L of the Poincaré group and for each channel a. When this necessary condition is satisfied, each of the following properties of a scattering system is equivalent to asymptotic covariance:

(i) For each element L of the Poincaré group

$$\sum_a \Omega_+^a U_a(L)\Omega_+^{a\dagger} = \sum_a \Omega_-^a U_a(L)\Omega_-^{a\dagger}; \qquad (3.4)$$

(ii)
$$\sum_a \Omega_+^a \mathbf{K}_a \Omega_+^{a\dagger} = \sum_a \Omega_-^a \mathbf{K}_a \Omega_-^{a\dagger}; \qquad (3.5)$$

(iii) For each element L of the Poincaré group and for each pair of channels a, b

$$U_b^\dagger(L)\Omega_-^{b\dagger}\Omega_+^a U_a(L) = \Omega_-^{b\dagger}\Omega_+^a; \qquad (3.6)$$

(iv) For each pair of channels a, b

$$\Omega_-^{b\dagger}\Omega_+^a \mathbf{K}_a = \mathbf{K}_b \Omega_-^{b\dagger}\Omega_+^a; \qquad (3.7)$$

(v) There exists a (continuous) unitary representation $W(L)$ of the Poincaré group defined on the subspace $\mathfrak{R} = F\mathfrak{K}$ such that for each channel a and each element L of the Poincaré group

$$U_a^\dagger(L)\Omega_\pm^{a\dagger} A \Omega_\pm^a U_a(L) = \Omega_\pm^{a\dagger} W^\dagger(L) A W(L) \Omega_\pm^a, \qquad (3.8)$$

respectively, for every bounded operator $A = AF = FA$ which has the property, respectively, that $AF_\pm^b = F_\pm^b A$ for each b.

(vi) There exists a (continuous) unitary representation $W(L)$ of the Poincaré group defined on the subspace $\mathfrak{R} = F\mathfrak{K}$ such that for each channel a and each element L of the Poincaré group

$$\sum_a \Omega_\pm^a U_a(L) A_a U_a^\dagger(L) \Omega_\pm^{a\dagger} = W(L) \sum_a \Omega_\pm^a A_a \Omega_\pm^{a\dagger} W^\dagger(L) \tag{3.9}$$

for every family of bounded operators A_a with the property that $A_a = A_a E_a = E_a A_a$ for each a.

When one of the conditions (v) or (vi) is satisfied [and when $U_a(L)$ commutes with E_a so that asymptotic covariance is also satisfied], the unitary representation $W(L)$ of the Poincaré group is uniquely determined by that condition and is the same as the unitary representation which satisfies the asymptotic covariance condition.

IV. INTERPRETATION AND REMARKS

The interpretation that Fong and Sucher[9] have given to the asymptotic covariance condition has an immediate generalization for the multichannel case. Let ϕ^a represent a state of the "free" asymptotic or unperturbed system for channel a. Then $\psi^{(\pm)} = \Omega_\pm^a \phi^a$ represent corresponding states of the interacting system. A transformation L of the Poincaré group gives us the state $U_a(L)\phi^a$ of the channel system and the corresponding states $\Omega_\pm^a U_a(L)\phi^a$ of the interacting system. The asymptotic covariance condition states that the transformation of the interacting states from $\Omega_\pm^a \phi^a$ to $\Omega_\pm^a U_a(L)\phi^a$ should be effected by a linear operator $W(L)$ which is the same for the $+$ and $-$ cases. In other words, there must be a linear operator $W(L)$ which transforms the interacting states just as $U_a(L)$ transforms the "free" asymptotic or unperturbed states.

To be more specific about this, we may suppose that a complete set of states of the channel system can be labeled by some parameters k (for example the momenta of free particles) such that the transform of each state is the state labeled by the transformed parameter $L(k)$,

$$U_a(L)\phi_k^a = \phi_{L(k)}^a.$$

The asymptotic covariance postulate is then the statement that the corresponding interacting states $\psi_k^{(\pm)} = \Omega_\pm^a \phi_k^a$ transform as

$$\psi_{L(k)}^{(\pm)} = W(L)\psi_k^{(\pm)}$$

with $W(L)$ a linear operator which is the same for the $+$ and $-$ cases.

Asymptotic covariance can be interpreted also in terms of the following operators. For each channel a let A_a be a (bounded) linear operator on the subspace $\mathfrak{D}_a = E_a\mathcal{H}$. For each such family of operators A_a we define two operators

$$(A_a)_\pm = \sum_a \Omega_\pm^a A_a \Omega_\pm^{a\dagger}. \tag{4.1}$$

The operators $(A_a)_\pm$ have the properties that

$$(A_a)_\pm = (A_a)_\pm F = F(A_a)_\pm, \tag{4.2}$$

$$(A_a)_\pm F_\pm^b = F_\pm^b (A_a)_\pm, \tag{4.3}$$

for each b. We have several examples available. If $A_a = E_a$ for each a, then $(A_a)_\pm = F$; if A_a is nonzero only for $a = b$ and if $A_b = E_b$, then $(A_a)_\pm = F_\pm^b$; if $A_a = U_a(L)$ for each a, then $(A_a)_\pm = W(L)F$; etc. There is a kind of inverse to this operation. For each (bounded) linear operator A on the subspace $\mathfrak{R} = F\mathcal{H}$ we define, for each a,

$$A_{\substack{\text{in}\\\text{out}}}^a = \Omega_\pm^{a\dagger} A \Omega_\pm^a. \tag{4.4}$$

These operators have the properties that

$$A_{\substack{\text{in}\\\text{out}}}^a = A_{\substack{\text{in}\\\text{out}}}^a E_a = E_a A_{\substack{\text{in}\\\text{out}}}^a. \tag{4.5}$$

Again several examples are available. If $A = F$, then $A_{\substack{\text{in}\\\text{out}}}^a = E_a$ for each a; if $A = W(L)$, then $A_{\substack{\text{in}\\\text{out}}}^a = U_a(L)E_a$ for each a; etc. If we start with an operator A which has the properties that

$$A = AF = FA, \tag{4.6}$$

$$AF_+^a = F_+^a A, \tag{4.7}$$

for each a, we find that the "in" operations (4.4) followed by the "+" operation (4.1) bring us back to the same operator A,

$$(A_{\text{in}}^a)_+ = A \tag{4.8}$$

while if we start with an operator A which satisfies Eq. (4.6) and

$$AF_-^a = F_-^a A \tag{4.9}$$

for each a, we find that the "out" operations (4.4) followed by the "$-$" operation (4.1) bring us back to the same operator A,

$$(A_{\text{out}}^a)_- = A. \tag{4.10}$$

Conversely, if we start with a family of operators A_a which have the property that

$$A_a = A_a E_a = E_a A_a \tag{4.11}$$

for each a, then we find that the "+" or "$-$" operations (4.1) followed respectively by the "in" or "out" operations (4.4) bring us back to the same family of operators A_a,

$$((A_a)_\pm)_{\substack{\text{in}\\\text{out}}}^b = A_b \tag{4.12}$$

for each b. Thus we see that the set of all operators $(A_a)_+$ is just the same as the set of all operators A which satisfy Eqs. (4.6) and (4.7), the set of all operators $(A_a)_-$ is just the same as the set of all operators A which satisfy Eqs. (4.6) and (4.9), and the set of all operators A^a_{in} is the same as the set of all operators A^a_{out} is just the same as the set of all operators A_a which satisfy Eqs. (4.11).

The operators (4.1) and the S operator (2.11) satisfy the relation

$$(A_a)_+S = S(A_a)_- \tag{4.13}$$

for every family of operators A_a. This relation serves to define the S operator (and hence the scattering amplitudes) to within phase factors which may be different for the different channels. More precisely, if S' is an operator which satisfies the same equations (2.13) and (4.13) as S, then

$$S' = \sum_a C_a F^a_+ S = \sum_a C_a S F^a_-,$$

where C_a are complex numbers with $|C_a| = 1$. This follows from the fact that the sets of operators $(A_a)_\pm$ include all operators satisfying Eqs. (4.6) and (4.7) or (4.9), respectively. (Since the statements of this section are included for interpretation and not as an integral part of the development of our formalism, their proofs, which are straightforward in every case, are left to the reader.)

The interpretation of asymptotic covariance in terms of the operators (4.1) and (4.4) is provided by conditions (v) and (vi) of Theorem 2. Condition (v) requires that there exists a unitary representation $W(L)$ of the Poincaré group such that for each channel a and each element L of the Poincaré group

$$U^\dagger_a(L)A^a_{\substack{\text{in}\\ \text{out}}}U_a(L) = (W^\dagger(L)AW(L))^a_{\substack{\text{in}\\ \text{out}}},$$

respectively, for every (bounded) operator A which belongs to the set of operators $(A_b)_+$ or the set of operators $(A_b)_-$, respectively. Condition (vi) requires that there exists a unitary representation $W(L)$ of the Poincaré group such that for each channel a and each element L of the Poincaré group

$$(U_a(L)A_aU^\dagger_a(L))_\pm = W(L)(A_a)_\pm W^\dagger(L)$$

for every family of (bounded) operators A_a which belong to the sets of operators A^a_{in} or A^a_{out}. Both of these conditions say essentially that there must be a unitary representation $W(L)$ of the Poincaré group which transforms the operators A, $(A_a)_+$ or $(A_a)_-$, respectively, of the interacting system just as the unitary representations $U_a(L)$ transform the operators A^a_{in}, A^a_{out}, or A_a of the "free" asymptotic or unperturbed channel systems. The essential point again is that $W(L)$ is the same for the $+$ and $-$ cases.

The condition (iii) of Theorem 2 is just the requirement of invariance of the S matrix or scattering amplitudes. We can see this more explicitly as follows. The probability amplitude for scattering from a state ϕ^a of the channel system a to a state ϕ^b of the channel system b is given by the matrix elements (2.10). The analogous probability amplitude for scattering from the transformed state $U_a(L)\phi^a$ to the corresponding transformed state $U_b(L)\phi^b$ is

$$(U_b(L)\phi^b,\ \Omega^{b\dagger}_-\Omega^a_+U_a(L)\phi^a) = (\phi^b,\ U^\dagger_b(L)\Omega^{b\dagger}_-\Omega^a_+U_a(L)\phi^a).$$

Condition (iii) is the statement that these scattering amplitudes are the same for all transformations L of the Poincaré group. Condition (iv) is the statement that they are the same for infinitesimal pure Lorentz transformations.

That asymptotic covariance implies invariance of the scattering amplitudes can be seen also from the fact that the unitary representation $W(L)F$ of the Poincaré group given by Eqs. (3.2) commutes with the S operator (2.11),

$$W(L)FS = SW(L)F$$

from which it follows that

$$(\psi^{(\pm)}_b,\ S\psi^{(\pm)}_a) = (W(L)\psi^{(\pm)}_b,\ SW(L)\psi^{(\pm)}_a)$$

which displays the invariance of the scattering amplitudes as matrix elements with respect to the states (2.12) of the interacting system.

Thus we see that asymptotic covariance can be motivated not only by a natural interpretation but also by the fact that asymptotic covariance is exactly what is needed to obtain invariant scattering amplitudes. The suggestion is compelling that asymptotic covariance should be satisfied in any Lorentz invariant scattering theory.

If asymptotic covariance is satisfied, we have a uniquely determined unitary representation $W(L)F$ of the Poincaré group defined on the subspace $\mathcal{R} = F\mathcal{H}$. This is an extension of the representation $e^{iHt}F$ of the one-parameter subgroup of time translations, and the defining Eq. (3.1) of asymptotic covariance is an extension to the whole Poincaré group of the basic intertwining relation (2.9) of the usual scattering formalism. This suggests that any Lorentz invariant scattering formalism should involve a unitary representation $W(L)$ of the whole Poincaré group in which the Hamiltonian H appears as the generator of the time translations and that the wave operators $\Omega^a_\pm$ should have the extended intertwining property (3.1) between the unitary representation $W(L)$ for the interacting system and the unitary representation $U_a(L)$ for each of the channel systems.

For a given family of unitary representations $U_a(L)$ and for a given Hamiltonian H, asymptotic covariance provides a program for completing the formalism outlined above. One solves the scattering problem determined by the Hamiltonian H and the channel Hamiltonians H_a to find the wave operators $\Omega_\pm^a$ and then constructs the unitary representation $W(L)$ according to Eq. (3.2). Of course this is not always possible. The two summations in Eq. (3.2) may turn out to be different. In fact Theorem 2 [condition (i)] states that this program will produce a scattering formalism in which asymptotic covariance is satisfied just when the formula (3.2) for $W(L)$ in terms of the operators Ω_+^a is the same as that in terms of the operators Ω_-^a. From condition (ii) of Theorem 2 we see that it is not necessary to consider all transformations L of the Poincaré group; infinitesimal pure Lorentz transformations are enough.

V. MULTICHANNEL SCATTERING FORMALISM—BOUNDARY CONDITIONS

The basic operators $\Omega_\pm^a$ of our scattering formalism have been assumed so far to satisfy only the partial isometry requirements (2.1) and (2.2) [with ranges satisfying Eqs. (2.3) and (2.4)] and the intertwining requirement (2.9). There can be a large number of operators having these properties. A complete scattering theory must contain some boundary condition which selects just two operators Ω_+^a and Ω_-^a for each channel. From the time-dependent point of view this appears as the asymptotic condition which demands that in the distant past or future the motion of the interacting system coincides with the "free" asymptotic motion in some channel. From the time-independent or perturbation theory point of view it appears as the boundary condition which selects stationary states of the interacting system which have the form of stationary states of the "free" or unperturbed channel system plus "incoming or outgoing scattered waves". Complete with boundary condition, the usual quantum mechanical multichannel scattering formalism can be cast in the form of the following.[16,17]

Definition: The linear operators $\Omega_\pm^a$ are solutions of the scattering problem defined by the self-adjoint operators H and H_a if they satisfy the following conditions:

(α) The operators $\Omega_\pm^a$ satisfy Eqs. (2.1) and (2.2) where E_a is a projection operator onto a subspace in the continuum subspace of H_a and $F_\pm^a$ are projection operators which satisfy Eqs. (2.3) and (2.4);

(β) $e^{iHt}\Omega_\pm^a = \Omega_\pm^a e^{iH_a t}$ for all real t; (2.9)

(γ) $\lim_{t\to\mp\infty} e^{iH_a t}\Omega_\pm^{a\dagger} e^{-iH_a t}E_a = E_a.$

Theorem 3. The conditions (α), (β), and (γ) may be replaced by the equivalent conditions (α), (β), and:

(γ') $\lim_{t\to\mp\infty} e^{iHt}\Omega_\pm^{a\dagger} e^{-iHt} = F_\pm^a.$

From (α) and (β) we can deduce Eqs. (2.5), (2.6), and (2.7) as before. It also follows that the projections $F_\pm^a$ commute with H and that the projection E_a commutes with H_a. By multiplying condition (β) on the right by $\Omega_\pm^{a\dagger}$ and using condition (α) we find that

$$e^{iHt}F_\pm^a = \Omega_\pm^a e^{iH_a t}\Omega_\pm^{a\dagger}$$

for all real t which shows that the parts of the operator H in the subspaces $\mathcal{R}_a^{(\pm)} = F_\pm^a\mathcal{H}$ are unitarily equivalent to the part of the operator H_a in the subspace $\mathcal{D}_a = E_a\mathcal{H}$. From this and the condition (α) that $\mathcal{D}_a = E_a\mathcal{H}$ is in the continuum subspace of H_a it follows that the subspaces $\mathcal{R}_a^{(\pm)} = F_\pm^a\mathcal{H}$ are in the continuum subspace of H. The connection of the boundary conditions (γ) or (γ') with the asymptotic limits is established by the following.

Theorem 4. If a solution of the conditions (α), (β), and (γ) or (γ') exists, it is the set of operators

$$\Omega_\pm^a = \lim_{t\to\mp\infty} e^{iHt}e^{-iH_a t}E_a. \tag{5.1}$$

These operators are a solution whenever the limits (5.1) converge on a subspace $\mathcal{D}_a = E_a\mathcal{H}$ which reduces H_a and is in the continuum subspace of H_a for each a and have ranges $\mathcal{R}_a^{(\pm)} = F_\pm^a\mathcal{H}$ satisfying Eqs. (2.3) and (2.4). This solution is unique in the sense that there is only one set of operators $\Omega_\pm^a$ which satisfy conditions (α), (β), and (γ) or (γ') for given domains $\mathcal{D}_a = E_a\mathcal{H}$ and ranges $\mathcal{R}_a^{(\pm)} = F_\pm^a\mathcal{H}$.

VI. LORENTZ INVARIANT BOUNDARY CONDITIONS

Now if we examine the conditions (α), (β), and (γ) or (γ') with an eye for finding a Lorentz invariant extension, we see immediately that a natural Lorentz

[16] We assume that the domains of definition of the various operators are such that all of the equations are meaningful. It is clear that the asymptotic conditions require assumptions of this kind beyond what are needed for the other parts of the scattering formalism. This is a feature of the quantum mechanical scattering formalism; it is not particularly relevant for the possibility of a Lorentz invariant extension.

[17] All operator limits are in the strong operator topology [see M. A. Naimark, *Normed Rings*, translated by L. F. Boron (P. Noordhoff Ltd., Groningen, The Netherlands, 1959), p. 441].

invariant extension of condition (β) is the equation (3.1) of asymptotic covariance. Let us assume, therefore, that we have not only a Hamiltonian H but a set of self-adjoint operators H, $\mathbf{P}$, $\mathbf{J}$, $\mathbf{K}$ which satisfy the "commutation relations" of the Poincaré group and generate respectively the time translations, space translations, space rotations, and pure Lorentz transformations in our description of the interacting system. Let $W(L)$ be the continuous unitary representation of the Poincaré group for which H, $\mathbf{P}$, $\mathbf{J}$, $\mathbf{K}$ are the generators. We now require that the operators $\Omega_\pm^a$ satisfy the asymptotic covariance equation (3.1). This gives us a manifestly Lorentz invariant extension of condition (β). In addition it quite nearly establishes (α) as a Lorentz invariant condition since we can now prove that the subspaces $\mathfrak{D}_a = E_a\mathcal{H}$ are invariant under the respective unitary representation $U_a(L)$ and that the subspaces $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathcal{H}$ are invariant under the unitary representation $W(L)$. We just need to make a Lorentz invariant extension of the requirement that the subspace $\mathfrak{D}_a = E_a\mathcal{H}$ is contained in the continuum subspace of H_a.

It remains also to formulate a Lorentz invariant extension of condition (γ) or (γ') and a Lorentz invariant restriction of the possible solutions (5.1). This can be accomplished in a very natural way by replacing Ht by $P_\mu t^\mu = Ht_0 - \mathbf{P}\cdot\mathbf{t}$ and $H_a t$ by $P_{a\mu}t^\mu = H_a t_0 - \mathbf{P}_a\cdot\mathbf{t}$, where $t_\mu = (t_0, \mathbf{t})$ is an arbitrary timelike real four-vector. Thus we propose a Lorentz invariant multichannel scattering formalism in the form of the following.

Definition: The linear operators $\Omega_\pm^a$ are solutions of the Lorentz invariant scattering problem defined by the (continuous) unitary representations $W(L)$ and $U_a(L)$ of the Poincaré group if they satisfy the following conditions:

(A) $$\Omega_\pm^{a\dagger}\Omega_\pm^a = E_a \qquad (2.1)$$

$$\Omega_\pm^a\Omega_\pm^{a\dagger} = F_\pm^a, \qquad (2.2)$$

where E_a is a projection operator onto a subspace which is in the continuum subspace of the operator $P_{a\mu}t^\mu = H_a t_0 - \mathbf{P}_a\cdot\mathbf{t}$ for every real four-vector $t_\mu = (t_0, \mathbf{t})$ such that $t_\mu t^\mu = t_0^2 - \mathbf{t}\cdot\mathbf{t} > 0$ and where $F_\pm^a$ are projection operators which satisfy the conditions

$$F_+^aF_+^b = F_-^aF_-^b = 0 \quad \text{for} \quad a \neq b \qquad (2.3)$$

$$\sum_a F_+^a = \sum_a F_-^a; \qquad (2.4)$$

(B) $$W(L)\Omega_\pm^a = \Omega_\pm^a U_a(L)$$

for all elements L of the Poincaré group;

(C) $$\lim_{s\to\mp\infty} U_a^\dagger(st_\mu)\Omega_\pm^{a\dagger}U_a(st_\mu)E_a = E_a$$

for all real four-vectors $t_\mu = (t_0, \mathbf{t})$ such that $t_\mu t^\mu = t_0^2 - \mathbf{t}\cdot\mathbf{t} > 0$ and $t_0 > 0$, where $U_a(st_\mu) = e^{-iP_{a\mu}t^\mu s}$ is the unitary representative $U_a(L)$ of the element L of the Poincaré group corresponding to space–time translation by st_μ.

Theorem 5. The conditions (A), (B), and (C) may be replaced by the equivalent conditions (A), (B), and:

(C′) $$\lim_{s\to\mp\infty} W^\dagger(st_\mu)\Omega_\pm^{a\dagger}W(st_\mu) = F_\pm^a$$

for all real four-vectors $t_\mu = (t_0, \mathbf{t})$ such that $t_\mu t^\mu = t_0^2 - \mathbf{t}\cdot\mathbf{t} > 0$ and $t_0 > 0$, where $W(st_\mu) = e^{-iP_\mu t^\mu s}$ is the unitary representative $W(L)$ of the element L of the Poincaré group corresponding to space–time translation by st_μ.

Theorem 6. If a solution of the conditions (A), (B), and (C) or (C′) exists, it is the set of operators

$$\Omega_\pm^a = \lim_{s\to\mp\infty} W^\dagger(st_\mu)U_a(st_\mu)E_a \qquad (6.1)$$

in which the limits converge to the same operators for all real four-vectors $t_\mu = (t_0, \mathbf{t})$ such that $t_\mu t^\mu = t_0^2 - \mathbf{t}\cdot\mathbf{t} > 0$ and $t_0 > 0$ [with $W(st_\mu)$ and $U_a(st_\mu)$ being the operators defined in (C) and (C′)]. These operators are a solution whenever the limits (6.1):(1) converge to the same operators for all such four-vectors t_μ on subspaces $\mathfrak{D}_a = E_a\mathcal{H}$ which are in the continuum subspaces of the respective operators $P_{a\mu}t^\mu$ for all real four-vectors t_μ such that $t_\mu t^\mu > 0$; (2) have ranges $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathcal{H}$ satisfying Eqs. (2.3) and (2.4); and (3) either satisfy condition (B) or satisfy the weaker equations

$$e^{i\mathbf{K}\cdot\mathbf{v}}\Omega_\pm^a = \Omega_\pm^a e^{i\mathbf{K}_a\cdot\mathbf{v}} \qquad (6.2)$$

for all real three-vectors $\mathbf{v}$ and each a and have domains $\mathfrak{D}_a = E_a\mathcal{H}$ which reduce the respective operators H_a. This solution is unique in the sense that there is only one set of operators $\Omega_\pm^a$ which satisfy conditions (A), (B), and (C) or (C′) for given domains $\mathfrak{D}_a = E_a\mathcal{H}$ and ranges $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathcal{H}$.

It may be helpful if we analyze briefly the conditions that we have stated for the existence of the solutions (6.1). If the limits (6.1) converge as stated (1), the conditions (2) on the ranges are sufficient to insure that condition (A) is statisfied. If conditions (A) and (B) are both satisfied, then the stated convergence of the limits (6.1) is sufficient for condition (C) also to be satisfied. The weaker alternative (3) to condition (B) rests on the fact (Theorem 4) that, if the domains $\mathfrak{D}_a = E_a\mathfrak{H}$ reduce the respective operators H_a, the limits (6.1) [which define the same operators as the limits (5.1)] satisfy condition (β). But condition (B) is much stronger than condition (β); it implies all of the consequences of asymptotic covariance, for example the invariance of the scattering amplitudes. Equation (6.2) is just the added element that is needed to establish the complete condition (B), as we state formally in the following.

Lemma 1. If (bounded) linear operators $\Omega_\pm^a$ satisfy condition (β) and also satisfy Eq. (6.2) (for all real three-vectors $\mathbf{v}$), then they satisfy condition (B).

Condition (A) implies Eqs. (2.5), (2.6), and (2.7) as before. Conditions (A) and (B) imply that the subspaces $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathfrak{H}$ reduce the unitary representation $W(L)$ and that the subspaces $\mathfrak{D}_a = E_a\mathfrak{H}$ reduce the respective unitary representation $U_a(L)$. This is a Lorentz invariant generalization of the reduction of H by $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathfrak{H}$ and the reduction of H_a by the respective subspace $\mathfrak{D}_a = E_a\mathfrak{H}$. Conditions (A) and (B) imply also a Lorentz invariant extension of the condition that the subspaces $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathfrak{H}$ be contained in the continuum subspace of H. We state these formally in the following.

Lemma 2. In order for conditions (A) and (B) to be satisfied for given projections $F_\pm^a$ and E_a and unitary representations $W(L)$ and $U_a(L)$ of the Poincaré group, it is necessary that $W(L)$ commutes with $F_\pm^a$ and $U_a(L)$ commutes with the respective projection E_a for each a and for each element L of the Poincaré group. It is also necessary that the subspaces $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathfrak{H}$ be contined in the continuum subspace of the operator $P_\mu t^\mu = Ht_0 - \mathbf{P}\cdot\mathbf{t}$ for every real four-vector $t_\mu = (t_0, \mathbf{t})$ such that $t_\mu t^\mu = t_0^2 - \mathbf{t}\cdot\mathbf{t} > 0$.

VII. DISCUSSION

As we have already remarked, conditions (A), (B), and (C) or (C′) are completely and manifestly Lorentz invariant. This is because we deal only with invariant sets of operators and invariant subspaces of the Hilbert space. Condition (B), for example, looks the same with respect to any special relativistic reference frame because with respect to any given frame any given transformation from the Poincaré group will be represented by one of the operators $W(L)$ for the interacting system and by the corresponding operator $U_a(L)$ for each of the channel systems. Condition (A) involves projections E_a onto subspaces of state vectors of the respective channel systems and projections $F_\pm^a$ onto subspaces of state vectors of the interacting system. According to Lemma 2, the subspace $\mathfrak{D}_a = E_a\mathfrak{H}$ must be invariant under the representation $U_a(L)$ of the Poincaré group for the channel system and the subspaces $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathfrak{H}$ must be invariant under the representation $W(L)$ for the interacting system. Hence these subspaces will look the same from any special relativistic reference frame. Also in condition (A) we have the requirement that the subspace $\mathfrak{D}_a = E_a\mathfrak{H}$ be in the continuum subspace of the operator $P_{a\mu}t^\mu$ for every timelike four-vector t_μ. This is again a Lorentz invariant set of operators. It includes the Hamiltonian for the channel system with respect to every special relativistic reference frame. Conditions (C) or (C′) look the same from any frame because they are statements that involve equally the representations of space–time translations in every forward (backward) timelike direction. With respect to any given special relativistic reference frame, translations in a given timelike direction will be represented for some choice of t_μ by the operators $W(st_\mu)$ for the interacting system and by the corresponding operators $U_a(st_\mu)$ for the channel system. In particular when condition (C) or (C′) is satisfied the usual "nonrelativistic" asymptotic condition of the type (γ) or (γ') is satisfied with respect to every frame.

We note that the solutions (6.1) of our relativistic scattering problem are also manifestly invariant. We get the same operators $\Omega_\pm^a$ if we evaluate the limits (6.1) in any special relativistic reference frame. In particular these operators are equal to ordinary "nonrelativistic" limits of the form (5.1) in every frame. Here we see quite clearly how the invariance of our formalism depends on the comparison of the interacting and channel systems. A change of reference frame is represented in the interacting system by the operator $W(L)$ and in the channel systems by the operators $U_a(L)$. If we tried to use a single representation throughout, we would not obtain invariance, for example, of the operators $\Omega_\pm^a$ computed according to the limits (6.1).

Our formulation of the relativistic scattering problem includes that of the ordinary "nonrelativistic" scattering problem. If operators $\Omega_\pm^a$ satisfy conditions (A), (B), and (C) or (C′), then they also satisfy conditions (α), (β), and (γ) or (γ'). Our relativistic formalism also includes asymptotic covariance. All of the statements (for example those of Theorem 2) which are consequences of asymptotic covariance hold for any solution of our relativistic scattering problem. In particular the scattering amplitudes are invariant.

We would like to know if the possible solutions of the scattering problem are restricted much more by conditions (A), (B), and (C) or (C′) than by conditions (α), (β), and (γ) or (γ') of the ordinary "nonrelativistic" scattering problem. We want to know the price that we have to pay for the Lorentz invariance of our formalism. The extension of condition (β) to condition (B) involves only the inclusion of asymptotic covariance. Since this is just what is needed to make the scattering amplitudes invariant, it appears that the restrictions introduced here are fundamental for a Lorentz invariant theory. In addition to the requirements of condition (α), condition (A) includes the statement that the subspace $\mathfrak{D}_a = E_a\mathfrak{H}$ must be contained in the continuum subspace of the operator $P_{a\mu}t^\mu$ for every timelike four-vector t_μ. [The corresponding statement for the subspaces $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathfrak{H}$ and the operators $P_\mu t^\mu$ for the interacting system are consequences of conditions (A) and (B); Lemma 2.] The restriction here is only on the channel systems. This seems to be a reasonable requirement; it says only that the relevant asymptotic or unperturbed states of the channel system should be in the continuum subspace of the channel Hamiltonian with respect to any special relativistic reference frame. It may be that there are severe restrictions imposed by conditions (C) or (C′) as compared to conditions (γ) or (γ'). However we shall see in the next section that this need not be true in every case.

VIII. REDUCTION OF THE BOUNDARY CONDITIONS WHEN $\mathbf{P} = \mathbf{P}_a$

In the construction of models of scattering systems it has been customary to assume that the total momentum operator $\mathbf{P}$ for the interacting system and the total momentum operators $\mathbf{P}_a$ for the "free" asymptotic or unperturbed channel systems are esentially the same. In such a case the Lorentz invariant boundary conditions add no new restrictions to the scattering problem; the boundary conditions (γ) and (γ') are already Lorentz invariant.

Theorem 7. In a case where

$$F_\pm^a e^{-i\mathbf{P}\cdot\mathbf{r}} e^{i\mathbf{P}_a\cdot\mathbf{r}} E_a = F_\pm^a E_a \tag{8.1}$$

for all real three-vectors $\mathbf{r}$ and each channel a, the conditions (A), (B), and (C) or (C′) reduce to conditions (A), (B), and (γ) or (γ').

For such a case it follows that whenever the ordinary "nonrelativistic" asymptotic limits (5.1) converge to operators $\Omega_\pm^a$ which satisfy conditions (A) and (B), these operators $\Omega_\pm^a$ satisfy also condition (γ) (Lemma 2; Theorem 4) and hence condition (C) (Theorem 7) and therefore are equal to the Lorentz invariant asymptotic limits (6.1) which converge to $\Omega_\pm^a$ for all real four-vectors $t_\mu = (t_0, \mathbf{t})$ such that $t_\mu t^\mu = t_0^2 - \mathbf{t}\cdot\mathbf{t} > 0$ and $t_0 > 0$ (Theorem 6). In particular, we can see that for such a case the ordinary "nonrelativistic" asymptotic limits (5.1) define solutions of the Lorentz invariant conditions (A), (B), and (C) or (C′) and are equal to the Lorentz invariant asymptotic limits (6.1) for all such real four-vectors t_μ whenever they converge on a subspace $\mathfrak{D}_a = E_a\mathfrak{H}$ which reduces H_a and is in the continuum subspace of H_a to operators $\Omega_\pm^a$ which have ranges $\mathfrak{R}_a^{(\pm)} = F_\pm^a\mathfrak{H}$ satisfying Eqs. (2.3) and (2.4) and satisfy Eq. (6.2) for all real three-vectors $\mathbf{v}$ (Theorem 4; Lemma 1). In other words, if Eq. (8.1) is valid for all real three-vectors $\mathbf{r}$, the Lorentz invariance of the asymptotic limits as well as the Lorentz invariance of the boundary conditions is gotten for free; if wave operators $\Omega_\pm^a$ either satisfy the ordinary "nonrelativistic" boundary conditions (γ) or (γ') or are defined by the ordinary "nonrelativistic" asymptotic limits (5.1), and if they satisfy the other conditions of the ordinary "nonrelativistic" scattering problem plus the condition of asymptotic covariance, then they satisfy also the Lorentz invariant boundary conditions (C) and (C′) and are equal also to the Lorentz invariant asymptotic limits (6.1).

IX. GALILEI INVARIANCE

We can try to modify the preceding to develop a similar multichannel scattering formalism which is Galilei invariant rather than Lorentz invariant. By comparison we can see then exactly how the "relativistic" Lorentz invariant scattering problem differs from the "nonrelativistic" Galilei invariant scattering problem.

We use the same basic multichannel scattering formalism as was outlined in Sec. II. Let us first suppose that unitary representations of the Poincaré group are replaced by unitary representations of the (proper orthochronous) inhomogeneous Galilei group. Asymptotic covariance is defined then in exact analogy to the definition of Sec. III and Theorem 1 has a precise analog. The first statement of Theorem 2 also has a precise analog as does the second statement insofar as it involves conditions (i) and (iii). Conditions (ii) and (iv) must be modified to include the analogous equations in the channel angular momentum operators $\mathbf{J}_a$ in addition to Eqs. (3.5) and (3.7) in the channel Galilei generators $\mathbf{K}_a$. [This is because the commutation relations (10.11) do not hold for the Galilei group.] Our proof of the sufficiency of conditions (v) and (vi) for asymptotic covariance does not hold (because the Galilei group may have one-dimensional unitary representations other than the identity representation) but an alternative proof can be constructed [from the intertwining property (2.9) for the Hamiltonians and the fact that the one-dimensional representations of the Galilei group differ from the identity representation only in the one-parameter subgroup of time translations]. However the last statement of Theorem 2 does not hold; the unitary representations $W(L)$ of the inhomogeneous Galilei group which satisfy condition (v) or (vi) may differ from the representation satisfying asymptotic covariance by phase factors which may be different on different subspaces $\mathcal{R}_a^{(\pm)} = F_\pm^a \mathcal{H}$. For a system with Galilei invariance which has a true unitary representation of the inhomogeneous Galilei group for each of the channel systems, the definition, interpretation, and consequences of asymptotic covariance are thus quite the same as for a system with Lorentz invariance.

Now the inhomogeneous Galilei group, unlike the Poincaré group, has unitary representations up to a factor (unitary ray representations) which can not be reduced to true unitary representations.[18] We must allow such representations for the operators $U_a(L)$ which represent the inhomogeneous Galilei group in the asymptotic or unperturbed channel systems. (To allow only true representations would be to assume that each of the channel systems has zero total mass.) If for these representations up to a factor we adopt the same definition of asymptotic covariance as for the true representations, we find that the operators $W(L)$ form a representation up to a factor of the inhomogeneous Galilei group if and only if the factors for the representations $U_a(L)$ are the same for all channels a. Inclusion of these common factors then results in a situation similar to that outlined above for true representations. We find further that it is possible for asymptotic covariance to be satisfied only if the factors for the representation $U_a(L)$ are the same as the factors for the representation $U_b(L)$ for all pairs of channels a, b between which there is scattering. This will not be true in general. (Channel systems between which there is scattering will not always have the same total mass.) Hence our definition of asymptotic covariance is not suitable for the most general Galilei invariant scattering systems. This does not mean that a Galilei invariant scattering formalism is impossible. Asymptotic covariance implies a statement of the invariance of the scattering amplitudes [condition (iii) of Theorem 2] which is actually stronger than is necessary. The scattering amplitudes need to be invariant only to within phase factors which may be different for different channels and for different transformations in the inhomogeneous Galilei group. We have to modify the formalism to allow the inclusion of these phase factors. This kind of weakening of the formalism is possible also for the Lorentz invariant case. But since the phase factors can be removed more often in a Lorentz invariant system, the generalized formalism might not be needed so much there.

Where condition (A) for a Lorentz invariant system requires that $\mathfrak{D}_a = E_a\mathcal{H}$ be in the continuum subspace of the operators $P_{a\mu}t^\mu = H_a t_0 - \mathbf{P}_a\cdot\mathbf{t}$ for all timelike four-vectors $t_\mu = (t_0, \mathbf{t})$, the analogous condition (A) for a Galilei invariant system requires the same thing for all four-vectors. This is necessary just to include the Hamiltonian operators $H_a - \mathbf{P}_a\cdot\mathbf{v}$ resulting from Galilei transformations to frames moving with all real three-vector velocities $\mathbf{v}$. Galilei transformations do not respect the timelike property. Hence condition (A) appears stronger for a Galilei invariant system than for a Lorentz invariant system. A similar strengthening occurs in the asymptotic condition. Where conditions (C) or (C′) for a Lorentz invariant system must hold for all four-vectors $t_\mu = (t_0, \mathbf{t})$ in the forward timelike cone, the Galilei invariant analog must hold for all four-vectors $t_\mu = (t_0, \mathbf{t})$ for which t_0 is positive. Thus we see that the conditions that we have added to asymptotic covariance to make a Lorentz invariant scattering formalism are no stronger, and perhaps even less strong than are needed for a Galilei invariant formalism.

X. PROOF OF THE THEOREMS

Let $W(L)$ be a bounded linear operator on $\mathcal{H}$ satisfying the relations (3.1). For any vector f in $\mathcal{H}$ it follows from Eqs. (2.2), (3.1), and (2.7) that

$$\begin{aligned} W(L)F_\pm^a f &= W(L)\Omega_\pm^a \Omega_\pm^{a\dagger} f \\ &= \Omega_\pm^a U_a(L)\Omega_\pm^{a\dagger} f \\ &= F_\pm^a \Omega_\pm^a U_a(L)\Omega_\pm^{a\dagger} f, \end{aligned}$$

which shows that the subspaces $\mathcal{R}_a^{(\pm)} = F_\pm^a\mathcal{H}$ are invariant under $W(L)$ or, equivalently, that

$$W(L)F_\pm^a = F_\pm^a W(L) F_\pm^a. \qquad (10.1)$$

From Eq. (2.4) we can conclude that the subspace $\mathcal{R} = F\mathcal{H}$ is also invariant under $W(L)$ or, equiv-

[18] V. Bargmann, Ann. Math. 59, 1 (1954).

alently, that

$$W(L)F = FW(L)F. \tag{10.2}$$

This proves the first statement of Theorem 1.

By multiplying Eq. (3.1) on the right by $\Omega_{\pm}^{a\dagger}$, using Eq. (2.2), we obtain Eq. (3.3). By summing the latter over a, using Eq. (2.4), we obtain Eq. (3.2). By taking the adjoint of Eq. (3.2) we find that

$$(W(L)F)^{\dagger} = \sum_a \Omega_{\pm}^a U_a^{\dagger}(L)\Omega_{\pm}^{a\dagger}$$

from which we can see that $(W(L)F)^{\dagger}$ is a well-defined operator on $\mathcal{R} = F\mathcal{H}$.

From Eqs. (2.1), the adjoints of Eqs. (3.1) and (2.6), and the fact that the operators $U_a(L)$ form a unitary representation of the Poincaré group, it follows that for any vector f in $\mathcal{H}$

$$\begin{aligned} U_a(L)E_a f &= U_a^{\dagger}(L^{-1})\Omega_+^{a\dagger}\Omega_+^a f \\ &= \Omega_+^{a\dagger}W^{\dagger}(L^{-1})\Omega_+^a f \\ &= E_a\Omega_+^{a\dagger}W^{\dagger}(L^{-1})\Omega_+^a f \end{aligned}$$

which shows that the subspace $\mathfrak{D}_a = E_a\mathcal{H}$ is invariant under $U_a(L)$ or, equivalently, that

$$U_a(L)E_a = E_aU_a(L)E_a. \tag{10.3}$$

By changing L to L^{-1} and taking the adjoint of the latter equation we find that

$$E_aU_a(L) = E_aU_a^{\dagger}(L^{-1}) = E_aU_a^{\dagger}(L^{-1})E_a = E_aU_a(L)E_a$$

and by comparing this with Eq. (10.3) we conclude that

$$U_a(L)E_a = E_aU_a(L). \tag{10.4}$$

From Eqs. (3.2), (2.5), (10.4), (2.6), (2.2), (2.4), and the unitarity of $U_a(L)$ on $\mathfrak{D}_a = E_a\mathcal{H}$ it follows that

$$\begin{aligned} (W(L)F)^{\dagger}W(L)F &= \sum_{ab} \Omega_+^a U_a^{\dagger}(L)\Omega_+^{a\dagger}\Omega_+^b U_b(L)\Omega_+^{b\dagger} \\ &= \sum_a \Omega_+^a U_a^{\dagger}(L)E_aU_a(L)\Omega_+^{a\dagger} \\ &= \sum_a \Omega_+^a\Omega_+^{a\dagger} \\ &= \sum_a F_+^a = F. \end{aligned} \tag{10.5}$$

In a similar manner we can show that

$$W(L)F(W(L)F)^{\dagger} = F. \tag{10.6}$$

Thus we can conclude that $W(L)F$ is a unitary operator on $\mathcal{R} = F\mathcal{H}$.

Let L_1 and L_2 be two elements of the Poincaré group and L_1L_2 their group-theoretic product. By using Eqs. (3.2), (2.5), (10.4), (2.6), and the fact that the operators $U_a(L)$ form a representation of the Poincaré group, we find that

$$\begin{aligned} W(L_1)FW(L_2)F &= \sum_a \Omega_+^a U_a(L_1)U_a(L_2)\Omega_+^{a\dagger} \\ &= \sum_a \Omega_+^a U_a(L_1L_2)\Omega_+^{a\dagger} \\ &= W(L_1L_2)F, \end{aligned}$$

which shows that the operators $W(L)$ form a representation of the Poincaré group on $\mathcal{R} = F\mathcal{H}$. [The continuity of this representation in the strong operator topology as a function of the element L of the topological Poincaré group can be shown to follow from the continuity of the representations $U_a(L)$.]

From Eq. (2.9) we see that the operators e^{iHt}, t real, are solutions for $W(L)$ of Eqs. (3.1) in the cases where L is a time translation. But we have found that the operators $W(L)$ satisfying Eqs. (3.1) are unique on $\mathcal{R} = F\mathcal{H}$. Hence the one-parameter group $e^{iHt}F$, t real, must be the same as the one-parameter group $W(L)F$ for time translations L.

We have completed the proof of Theorem 1 and of the first statement of Theorem 2. We have showed also that condition (i) of Theorem 2 is necessary for asymptotic covariance. The necessity of condition (ii) follows from that of (i) by taking $U_a(L) = e^{i\mathbf{K}_a\cdot\mathbf{r}}$ (with $\mathbf{r}$ a real three-vector) to first order in $\mathbf{r}$. [To zero order in $\mathbf{r}$ Eq. (3.4) is just Eq. (2.4).] From Eqs. (3.1), (2.7), (10.5) and their adjoints it follows that

$$\begin{aligned} &U_b^{\dagger}(L)\Omega_-^{b\dagger}\Omega_+^a U_a(L) \\ &\quad = \Omega_-^{b\dagger}W^{\dagger}(L)W(L)\Omega_+^a = \Omega_-^{b\dagger}FW^{\dagger}(L)W(L)F\Omega_+^a \\ &\quad = \Omega_-^{b\dagger}(W(L)F)^{\dagger}W(L)F\Omega_+^a = \Omega_-^{b\dagger}F\Omega_+^a = \Omega_-^{b\dagger}\Omega_+^a \end{aligned}$$

which is just Eq. (3.6). This establishes the necessity for asymptotic covariance of condition (iii). By multiplying on the left by $U_b(L)$ and using the unitarity of $U_a(L)$ on $\mathfrak{D}_a = E_a\mathcal{H}$, we may write Eq. (3.6) in the equivalent form

$$\Omega_-^{b\dagger}\Omega_+^a U_a(L) = U_b(L)\Omega_-^{b\dagger}\Omega_+^a. \tag{10.7}$$

By taking $U_a(L) = e^{i\mathbf{K}_a\cdot\mathbf{r}}$ and $U_b(L) = e^{i\mathbf{K}_b\cdot\mathbf{r}}$ (with $\mathbf{r}$ a real three-vector) to first order in $\mathbf{r}$ we obtain Eq. (3.7). This establishes the necessity for asymptotic covariance of condition (iv). It remains to show that these conditions are sufficient for asymptotic covariance.

Let condition (i) be valid, and define $W(L)$ by

$$W(L) = \sum_a \Omega_{\pm}^a U_a(L)\Omega_{\pm}^{a\dagger}. \tag{10.8}$$

It follows from Eq. (2.7) that

$$W(L)F_{\pm}^{a} = \Omega_{\pm}^{a} U_a(L) \Omega_{\pm}^{a\dagger}. \tag{10.9}$$

By multiplying on the right by $\Omega_{\pm}^{a}$, using Eqs. (2.7), (2.1), (10.4), and (2.6), we obtain Eq. (3.1). This establishes the sufficiency for asymptotic covariance of condition (i).

Suppose that condition (iii) is valid. Then Eq. (10.7) is valid also. By multiplying Eq. (10.7) on the left by Ω_{-}^{b} and on the right by $\Omega_{+}^{a\dagger}$, using Eq. (2.2), we find that

$$F_{-}^{b} \Omega_{+}^{a} U_a(L) \Omega_{+}^{a\dagger} = \Omega_{-}^{b} U_b(L) \Omega_{-}^{b\dagger} F_{+}^{a}.$$

If we sum this equation over a and b, using Eqs. (2.4) and (2.7), we obtain Eq. (3.4). This shows that condition (i) is a consequence of condition (iii) and establishes the sufficiency for asymptotic covariance of condition (iii).

Next suppose that Eq. (3.7) is valid. From the assumptions of the scattering formalism, namely, Eq. (2.8), we have that

$$\Omega_{-}^{b\dagger} \Omega_{+}^{a} H_a = \Omega_{-}^{b\dagger} H \Omega_{+}^{a} = H_b \Omega_{-}^{b\dagger} \Omega_{+}^{a}.$$

From this, Eq. (3.7), and the commutation relations

$$\mathbf{P}_a = i(H_a \mathbf{K}_a - \mathbf{K}_a H_a) \tag{10.10}$$

$$\mathbf{J}_{ai} = i\epsilon_{jkm}(K_{ak}K_{am} - K_{am}K_{ak}) \tag{10.11}$$

of the Poincaré group, it follows that

$$\Omega_{-}^{b\dagger} \Omega_{+}^{a} \mathbf{P}_a = \mathbf{P}_b \Omega_{-}^{b\dagger} \Omega_{+}^{a}$$

$$\Omega_{-}^{b\dagger} \Omega_{+}^{a} \mathbf{J}_a = \mathbf{J}_b \Omega_{-}^{b\dagger} \Omega_{+}^{a}.$$

The same relation clearly holds for any "power series" in any linear combination of the generators H_a, $\mathbf{P}_a$, $\mathbf{J}_a$, $\mathbf{K}_a$. From this we can see that Eq. (10.7) is a consequence of Eq. (3.7). By multiplying Eq. (10.7) on the left by $U_b^{\dagger}(L)$, using the unitarity of $U_b(L)$ on $\mathfrak{D}_b = E_b \mathcal{H}$ and Eq. (2.6), we obtain Eq. (3.6). This shows that condition (iii) is a consequence of condition (iv) and establishes the sufficiency for asymptotic covariance of condition (iv).

Finally, suppose that Eq. (3.5) is valid. From Eqs. (2.8), (2.2), and (2.4) we have that

$$\sum_a \Omega_{+}^{a} H_a \Omega_{+}^{a\dagger} = H \sum_a \Omega_{+}^{a} \Omega_{+}^{a\dagger} = H \sum_a \Omega_{-}^{a} \Omega_{-}^{a\dagger} = \sum_a \Omega_{-}^{a} H_a \Omega_{-}^{a\dagger}.$$

From Eqs. (10.10), (2.6), (2.5), and the fact the E_a commutes with H_a and $\mathbf{K}_a$ it follows that

$$\sum_a \Omega_{+}^{a} \mathbf{P}_a \Omega_{+}^{a\dagger}$$
$$= i \sum_a \Omega_{+}^{a} H_a \mathbf{K}_a \Omega_{+}^{a\dagger} - i \sum_a \Omega_{+}^{a} \mathbf{K}_a H_a \Omega_{+}^{a\dagger}$$
$$= i \sum_{ab} \Omega_{+}^{a} H_a \Omega_{+}^{a\dagger} \Omega_{+}^{b} \mathbf{K}_b \Omega_{+}^{b\dagger} - i \sum_{ab} \Omega_{+}^{a} \mathbf{K}_a \Omega_{+}^{a\dagger} \Omega_{+}^{b} H_b \Omega_{+}^{b\dagger}$$
$$= \sum_a \Omega_{-}^{a} \mathbf{P}_a \Omega_{-}^{a\dagger}.$$

In a similar manner we can show that

$$\sum_a \Omega_{+}^{a} \mathbf{J}_a \Omega_{+}^{a\dagger} = \sum_a \Omega_{-}^{a} \mathbf{J}_a \Omega_{-}^{a\dagger}.$$

By the same technique we can show that the same relation holds for any "power series" in any linear combination of H_a, $\mathbf{P}_a$, $\mathbf{J}_a$, $\mathbf{K}_a$. From this we can see that Eq. (3.4) is a consequence of Eq. (3.5). This shows that condition (ii) implies condition (i), establishes the sufficiency for asymptotic covariance of condition (ii), and completes the proof that conditions (i)–(iv) are necessary and sufficient for asymptotic covariance.

That conditions (v) and (vi) are necessary for asymptotic covariance follows immediately from Theorem 1 and Eq. (3.1). We must show now that these conditions are also sufficient for asymptotic covariance when $U_a(L)$ commutes with E_a.

Let $W(L)$ be a unitary representation of the Poincaré group defined on $\mathfrak{R} = F\mathcal{H}$ and satisfying Eqs. (3.8) for all bounded operators $A = AF = FA$ with the respective property that $AF_{\pm}^{b} = F_{\pm}^{b}A$ for all b. Letting $A = F_{\pm}^{a}$ we find that

$$E_a = \Omega_{\pm}^{a\dagger} W^{\dagger}(L) F_{\pm}^{a} W(L) \Omega_{\pm}^{a}.$$

Multiplying on the left by $\Omega_{\pm}^{a}$ and on the right by $\Omega_{\pm}^{a\dagger}$, using Eqs. (2.6) and (2.2), we deduce that

$$F_{\pm}^{a} = F_{\pm}^{a} W^{\dagger}(L) F_{\pm}^{a} W(L) F_{\pm}^{a}$$

from which it follows that

$$F_{\pm}^{a} W^{\dagger}(L)(1 - F_{\pm}^{a}) W(L) F_{\pm}^{a} = 0.$$

Rewriting this in the form

$$((1 - F_{\pm}^{a}) W(L) F_{\pm}^{a})^{\dagger} (1 - F_{\pm}^{a}) W(L) F_{\pm}^{a} = 0,$$

we conclude that

$$F_{\pm}^{a} W(L) F_{\pm}^{a} = W(L) F_{\pm}^{a}.$$

By changing L to L^{-1}, taking the adjoint and comparing [just as in proceeding from Eq. (10.3) to (10.4)] we have that

$$W(L) F_{\pm}^{a} = F_{\pm}^{a} W(L). \tag{10.12}$$

Now multiplying Eq. (3.8) on the left by $W(L)\Omega_{\pm}^{a}$ and on the right by $U_a^{\dagger}(L)\Omega_{\pm}^{a\dagger}$, using Eqs. (2.2), (10.12) and the unitary property of $W(L)$ and $U_a(L)$, we have for each channel a and each element L of the Poincaré group that

$$W(L) \Omega_{\pm}^{a} U_a^{\dagger}(L) \Omega_{\pm}^{a\dagger} A = A W(L) \Omega_{\pm}^{a} U_a^{\dagger}(L) \Omega_{\pm}^{a}$$

for every bounded operator A such that $A = AF^a_\pm = F^a_\pm A$, respectively. From this it follows that for each channel a and each element L of the Poincaré group

$$W(L)\Omega^a_\pm U^\dagger_a(L)\Omega^{a\dagger}_\pm = C^{(\pm)}_a(L)F^a_\pm$$

or, after multiplying on the right by $\Omega^a_\pm U_a(L)$, using Eqs. (2.1), (2.6), (2.7) and the commutability of $U_a(L)$ and E_a, that

$$W(L)\Omega^a_\pm = C^{(\pm)}_a(L)\Omega^a_\pm U_a(L), \qquad (10.13)$$

where $C^{(\pm)}_a(L)$ are complex numbers. Now from the unitary property of $W(L)$ and $U_a(L)$ and the isometric property of $\Omega^a_\pm$ we have that for any vector f in $\mathcal{H}$

$$\begin{aligned}||E_af|| &= ||W(L)\Omega^a_\pm f||\\ &= ||C^{(\pm)}_a(L)\Omega^a_\pm U_a(L)f|| = |C^{(\pm)}_a(L)|\ ||E_af||\end{aligned}$$

from which we conclude that

$$|C^{(\pm)}_a(L)| = 1.$$

Let L_1 and L_2 be any two elements of the Poincaré group. By a repeated use of Eq. (10.13) we find that

$$\begin{aligned}&C^{(\pm)}_a(L_1L_2)\Omega^a_\pm U_a(L_1)U_a(L_2)\\ &\quad= C^{(\pm)}_a(L_1L_2)\Omega^a_\pm U_a(L_1L_2) = W(L_1L_2)\Omega^a_\pm\\ &\quad= W(L_1)W(L_2)\Omega^a_\pm = W(L_1)C^{(\pm)}_a(L_2)\Omega^a_\pm U_a(L_2)\\ &\quad= C^{(\pm)}_a(L_1)C^{(\pm)}_a(L_2)\Omega^a_\pm U_a(L_1)U_a(L_2)\end{aligned}$$

from which it follows that

$$C^{(\pm)}_a(L_1L_2) = C^{(\pm)}_a(L_1)C^{(\pm)}_a(L_2).$$

The complex numbers $C^{(\pm)}_a(L)$ form a one-dimensional unitary representation of the Poincaré group. The continuity of this representation follows from the continuity of the representations $W(L)$ and $U_a(L)$ as follows. Let L_1 converge to L_2 in the topology of the Poincaré group. Then for any vectors f and g in $\mathcal{H}$ it is a consequence of the continuity of the representation $U_a(L)$ that

$$(\Omega^{a\dagger}_\pm g,\ U_a(L_1)E_af) \to (\Omega^a_\pm g,\ U_a(L_2)E_af).$$

But from the continuity of the representation $W(L)$, Eqs. (10.13), (2.6), and the commutability of $U_a(L)$ and E_a we have also that

$$\begin{aligned}(g,\ W(L_1)\Omega^a_\pm f) &= C^{(\pm)}_a(L_1)(\Omega^{a\dagger}_\pm g,\ U_a(L_1)E_af)\\ &\to (g,\ W(L_2)\Omega^a_\pm f) = C^{(\pm)}_a(L_2)(\Omega^a_\pm g,\ U_a(L_2)E_af).\end{aligned}$$

For this to be true it is necessary that

$$C^{(\pm)}_a(L_1) \to C^{(\pm)}_a(L_2).$$

Now since the Poincaré group has no one-dimensional continuous unitary representations except the identity representation, we can conclude that

$$C^{(\pm)}_a(L) = 1.$$

Equation (10.13) thus becomes the asymptotic covariance relation (3.1). This completes the proof that condition (v) is sufficient for asymptotic covariance. We next prove the sufficiency of condition (vi).

Let $W(L)$ be a unitary representation of the Poincaré group defined on $\mathcal{R} = F\mathcal{H}$ and satisfying Eq. (3.9) for every family of bounded operators A_a such that $A_a = A_aE_a = E_aA_a$ for each a. Choosing $A_a = 0$ for all but a single value of a and $A_a = E_a$ for that value of a, we find, using Eqs. (2.6), (2.2), the unitarity of $U_a(L)$ and the commutability of $U_a(L)$ and E_a, that

$$W(L)F^a_\pm W^\dagger(L) = F^a_\pm$$

or, using also the unitarity of $W(L)$, that

$$W(L)F^a_\pm = F^a_\pm W(L)$$

for each channel a and each element L of the Poincaré group. Now we use this, Eq. (2.1), the unitarity of $W(L)$ and $U_a(L)$, and the commutability of $U_a(L)$ and E_a to multiply Eq. (3.9) on the left by $U^\dagger_a(L)\Omega^{a\dagger}_\pm$ and on the right by $W(L)\Omega^a_\pm$ for a case where A_a is chosen to be nonzero only for a single value of a and conclude that

$$U^\dagger_a(L)\Omega^{a\dagger}_\pm W(L)\Omega^a_\pm A_a = A_aU^\dagger_a(L)\Omega^{a\dagger}_\pm W(L)\Omega^a_\pm$$

for each channel a, each element L of the Poincaré group, and for every bounded operator A_a such that $A_a = A_aE_a = E_aA_a$. From this it follows that for each channel a and each element L of the Poincaré group

$$U^\dagger_a(L)\Omega^{a\dagger}_\pm W(L)\Omega^a_\pm = C^{(\pm)}_a(L)E_a$$

or, after multiplying on the left by $\Omega^a_\pm U_a(L)$, using Eqs. (2.2), (2.6), (2.7), the unitarity of $U_a(L)$, and the commutability of $U_a(L)$ with E_a and of $W(L)$ with $F^a_\pm$, that

$$W(L)\Omega^a_\pm = C^{(\pm)}_a(L)\Omega^a_\pm U_a(L),$$

where $C^{(\pm)}_a(L)$ are complex numbers. The remainder of the proof that condition (vi) is sufficient for asymptotic covariance is identical to the last part [from Eq. (10.13)] of the proof for condition (v). From these proofs it is clear that the operators $W(L)$ occurring in conditions (v) and (vi) are the same as those of the asymptotic covariance condition. Theorem 1 tells us that these are uniquely determined. This completes the proof of Theorem 2.

The proof of Theorem 3 can be obtained as a specialization of the proof of Theorem 5, and the proof of Theorem 4 can be gotten by adapting the proof of Theorem 6 to the relevant special cases. We shall therefore present only the proofs of Theorems 5 and 6. But first we shall prove the lemmas of Sec. VI.

Let $\Omega_\pm^a$ be (bounded) linear operators which satisfy condition (β) and also satisfy Eqs. (6.2) for all real three-vectors $\mathbf{v}$. Working again with the nonrigorous language of commutators and power series, we convert our hypotheses into the form of the equations

$$H\Omega_\pm^a = \Omega_\pm^a H_a$$

$$\mathbf{K}\Omega_\pm^a = \Omega_\pm^a \mathbf{K}_a$$

and use Eq. (10.10) and the analogous commutation relation

$$\mathbf{P} = i(H\mathbf{K} - \mathbf{K}H)$$

to write

$$\mathbf{P}\Omega_\pm^a = i(H\mathbf{K}\Omega_\pm^a - \mathbf{K}H\Omega_\pm^a) = i(H\Omega_\pm^a\mathbf{K}_a - \mathbf{K}\Omega_\pm^a H_a)$$
$$= i(\Omega_\pm^a H_a \mathbf{K}_a - \Omega_\pm^a \mathbf{K}_a H_a) = \Omega_\pm^a \mathbf{P}_a.$$

Using Eq. (10.11) and the analogous commutation relation

$$J_i = i\epsilon_{ijk}(K_j K_k - K_k K_j)$$

we find similarly that

$$\mathbf{J}\Omega_\pm^a = \Omega_\pm^a \mathbf{J}_a.$$

The extension of this relation to power series in linear combinations of the ten generators establishes condition (B). This proves Lemma 1.

Lemma 2 is already almost completely proved. The proof that $U_a(L)$ commutes with E_a is exactly the same as the argument used to establish Eqs. (10.3) and (10.4) in the proof of Theorems 1 and 2. Also from conditions (A) and (B) we can establish Eq. (9.1) by the same argument that we used in the proof of Theorem 1. Then by the same method as was used to move from Eq. (10.3) to (10.4), namely changing L to L^{-1}, taking the adjoint and comparing, we can deduce from Eq. (9.1) that $W(L)$ commutes with $F_\pm^a$. If we consider condition (B) for a case where L corresponds to space–time translation by a four-vector t_μ with $t_\mu t^\mu > 0$, multiply on the right by $\Omega_\pm^a$, and use Eq. (2.2) of condition (A), we find that

$$e^{iP_\mu t^\mu} F_\pm^a = \Omega_\pm^a e^{iP_{a\mu} t^\mu} \Omega_\pm^{a\dagger}.$$

In view of the fact that $\Omega_\pm^a$ are partially isometric by condition (A), this shows that the parts of the operator $P_\mu t^\mu$ in the subspaces $\mathcal{R}_a^{(\pm)} = F_\pm^a \mathcal{H}$ have the same spectrum as the part of the operator $P_{a\mu} t^\mu$ in the subspace $\mathfrak{D}_a = E_a\mathcal{H}$. Since the subspace $\mathfrak{D}_a = E_a\mathcal{H}$ is required by condition (A) to be contained in the continuum subspace of $P_{a\mu}t^\mu$, it follows that the subspaces $\mathcal{R}_a^{(\pm)} = F_\pm^a\mathcal{H}$ are contained in the continuum subspace of $P_\mu t^\mu$. This completes the proof of Lemma 2.

We turn now to the proof of Theorem 6. Let $\Omega_\pm^a$ be linear operators which satisfy conditions (A), (B), and (C). Multiplying Eq. (C) on the left by $\Omega_\pm^a$ and using Eqs. (2.6), (2.2), condition (B), and the facts that $U_a(L)$ are a unitary representation of the Poincaré group and that $F_\pm^a$ commute with $W(L)$ (Lemma 2), we find that

$$\Omega_\pm^a = \lim_{s\to\mp\infty} \Omega_\pm^a U_a(-st_\mu)\Omega_\pm^{a\dagger} U_a(st_\mu) E_a$$
$$= \lim_{s\to\mp\infty} W(-st_\mu)\Omega_\pm^a \Omega_\pm^{a\dagger} U_a(st_\mu) E_a$$
$$= \lim_{s\to\mp\infty} F_\pm^a W_a^\dagger(st_\mu) U_a(st_\mu) E_a$$

with the limits converging to the same operators for all real four-vectors $t_\mu = (t_0, \mathbf{t})$ such that $t_\mu t^\mu = t_0^2 - \mathbf{t}\cdot\mathbf{t} > 0$ and $t_0 > 0$. But for all such four-vectors t_μ and for any vector f in $\mathcal{H}$ it follows from the isometric property (A) of the operators $\Omega_\pm^a$ and the properties of (strong) operator convergence that

$$\lim_{s\to\mp\infty} ||F_\pm^a W^\dagger(st_\mu) U_a(st_\mu) E_a f|| = ||E_a f||$$

and so, using the isometric properties of the operators $W^\dagger(st_\mu)$ and $U_a(st_\mu)$, we have that

$$\lim_{s\to\mp\infty} ||(1 - F_\pm^a) W^\dagger(st_\mu) U_a(st_\mu) E_a f||^2$$
$$= \lim_{s\to\mp\infty} \{||E_a f||^2 - ||F_\pm^a W^\dagger(st_\mu) U_a(st_\mu) E_a f||^2\} = 0$$

from which we can conclude that Eq. (6.1) holds with the limits converging as stated. This proves the first statement of Theorem 6 and also demonstrates the uniqueness. It remains to show that this solution does exist under the conditions (1), (2), and (3).

Suppose that the limits (6.1) converge to the operators $\Omega_\pm^a$ for all real four-vectors t_μ such that $t_\mu t^\mu > 0$ and $t_0 > 0$ on subspaces $\mathfrak{D}_a = E_a\mathcal{H}$ which are in the continuum subspaces of the respective operators $P_{a\mu}t^\mu$ for all real four-vectors t_μ such that $t_\mu t^\mu > 0$. From the properties of (strong) operator convergence and the isometric property of the operators $W^\dagger(st_\mu)$ and $U_a(st_\mu)$ it follows that $\Omega_\pm^a$ are isometric operators on the subspaces $\mathfrak{D}_a = E_a\mathcal{H}$.

Letting $\mathcal{R}_a^{(\pm)} = F_\pm^a \mathcal{H}$ be the ranges of the respective operators $\Omega_\pm^a$, we can express the isometric property of $\Omega_\pm^a$ in the form of Eqs. (2.1) and (2.2). Our hypothesis that the projections $F_\pm^a$ satisfy Eqs. (2.3) and (2.4) completes condition (A).

Letting t_μ take the particular value $t_\mu = (1, \mathbf{0})$, we see that the limits (6.1) define the same operators as the limits (5.1). Using the latter we find that for all real t

$$\begin{aligned} e^{iHt}\Omega_\pm^a &= \lim_{s\to\mp\infty} e^{iH(t+s)}e^{-iH_a(t+s)}E_a e^{iH_a t} \\ &= \Omega_\pm^a e^{iH_a t} \end{aligned}$$

under the hypotheses that E_a commutes with H_a. Under the additional hypotheses that Eq. (6.2) is valid for all real three-vectors $\mathbf{v}$, Lemma 1 gives us condition (B).

Now suppose that the limits (6.1) converge for all real four-vectors t_μ such that $t_\mu t^\mu > 0$ and $t_0 > 0$ to operators $\Omega_\pm^a$ that satisfy conditions (A) and (B). We then have that

$$\begin{aligned} &\lim_{s\to\mp\infty} U_a^\dagger(st_\mu)\Omega_\pm^{a\dagger} U_a(st_\mu)E_a \\ &\qquad = \Omega_\pm^{a\dagger} \lim_{s\to\mp\infty} W^\dagger(st_\mu)U_a(st_\mu)E_a = \Omega_\pm^{a\dagger}\Omega_\pm^a = E_a \end{aligned}$$

for all such four-vectors t_μ, which is just condition (C). This completes the proof of Theorem 6.

To prove Theorem 5 we note that with conditions (A) and (B) condition (C) is equivalent to the statement that

$$\Omega_\pm^a = \lim_{s\to\mp\infty} W^\dagger(st_\mu)U_a(st_\mu)E_a \tag{6.1}$$

with the limits converging to the same operators for all real four-vectors t_μ such that $t_\mu t^\mu > 0$ and $t_0 > 0$. This is evident from the proof of Theorem 6. We can show that in the presence of conditions (A) and (B) condition (C′) is also equivalent to the above statement as follows. Suppose the above statement holds. Multiplying Eq. (6.1) on the right by $\Omega_\pm^{a\dagger}$, using Eq. (2.2), the adjoint of Eq. (2.6), and the adjoint of condition (B), we arrive at condition (C′). The return trip is accomplished by multiplying condition (C′) on the right by $\Omega_\pm^a$ and using Eq. (2.7), condition (B), Eq. (2.1), and the fact that E_a commutes with $U_a(st_\mu)$ (Lemma 2). This completes the proof of Theorem 5.

Finally we prove Theorem 7. That conditions (C) and (C′) imply conditions (γ) and (γ') respectively can be seen by letting t_μ have the particular value $t_\mu = (1, \mathbf{0})$. Let Eq. (8.1) be valid for all real three-vectors $\mathbf{r}$. We must show that conditions (A), (B) and (γ) or (γ') imply conditions (A), (B), and (C) or (C′) respectively. Using condition (B), Eqs. (8.1), the adjoint of Eq. (2.7), and the fact that E_a commutes with H_a (Lemma 2), we have that

$$\begin{aligned} &\lim_{s\to\mp\infty} U_a^\dagger(st_\mu)\Omega_\pm^{a\dagger} U_a(st_\mu)E_a \\ &\qquad = \lim_{s\to\mp\infty} e^{iH_a t_0 s}e^{-i\mathbf{P}_a\cdot\mathbf{t}s}\Omega_\pm^{a\dagger} e^{i\mathbf{P}_a\cdot\mathbf{t}s}e^{-iH_a t_0 s}E_a \\ &\qquad = \lim_{s\to\mp\infty} e^{iH_a t_0 s}\Omega_\pm^{a\dagger} e^{-i\mathbf{P}\cdot\mathbf{t}s}e^{i\mathbf{P}_a\cdot\mathbf{t}s}e^{-iH_a t_0 s}E_a \\ &\qquad = \lim_{s\to\mp\infty} e^{iH_a s}\Omega_\pm^{a\dagger} e^{-iH_a s}E_a \end{aligned}$$

for all real four-vectors $t_\mu = (t_0, \mathbf{t})$ for which $t_0 > 0$. Since in the presence of conditions (A) and (B) condition (γ) is equivalent to (γ') and condition (C) is equivalent to (C′) (Theorem 3 and 5), this completes the proof of Theorem 7.

ACKNOWLEDGMENTS

The author is grateful to Professor J. Sucher for helpful discussions. This work was done while the author was a U. S. National Science Foundation postdoctoral fellow. It is a pleasure to thank Professor A. Mercier and the staff of the Institut für theoretische Physik, Universität Bern for their hospitality.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 10 OCTOBER 1964

The Concept of Nonlocalizable Fields and its Connection with Nonrenormalizable Field Theories*

B. SCHROER†
Institute for Advanced Study, Princeton, New Jersey
(Received 3 April 1964)

From an investigation of the two-point function in nonrenormalizable field theories, it is shown that at least in certain approximations, a nonrenormalizable field is nonlocalizable. This is intimately connected with the occurrence of essential singularities on the light cone in the Wightman functions of the field. Green's functions cannot be defined and the observables, in particular the scattering matrix elements, have to be expressed directly in terms of the unordered expectations values.

I. INTRODUCTION

ANY attempt to understand nonrenormalizable theories within the framework of general quantum field theory is tantamount to interpreting the breakdown of Feynman–Dyson's perturbation theory as a pitfall of computational techniques rather than of physical principles involved in quantum field theory. In all the expositions of general quantum field theory it is assumed that the fields in terms of which the theory is formulated are localizable, i.e., can be averaged with smearing function of finite support in space–time. This rules out any essential light-cone singularity respectively an exponential growth in momentum space for the expectation values of the field. The introduction of nonlocalizable fields in the general framework of quantum field theory permits understanding in what sense nonrenormalizable fields differ from renormalizable fields. In particular, any formalism based on Green's functions (like time-ordered or retarded functions) breaks down. This becomes evident from the discussion in the second section. At first sight, the nonexistence of Green's function seems to be a serious obstacle for the computation of scattering quantities, since the conventional stationary scattering formulas (reduction formulas) express scattering quantities in terms of the interpolating field via time-ordered or retarded functions. Happily, the formulation of collision theory does not depend on the availability of these technical tools. This is discussed in the third section for the special case of elastic scattering.

In the last section the more theoretical problems of the relation between nonlocalizability of fields and Einstein causality for the local observables is discussed.

* This work supported is by the National Science Foundation.

† Present address: Physics Department, Pittsburgh University, Pittsburgh, Pennsylvania.

II. DEFINITION AND PROPERTIES OF NONLOCALIZABLE FIELDS

It is well known that a quantum field $A(x)$ has to be averaged with smearing functions $f(x)$ in order to get bona fide operators. To obtain quantities which belong to a finite space–time region, it is usually assumed that the space of allowed smearing functions $\mathfrak{K}$ contains the class of all smooth and localizable functions $\mathfrak{D}_x$[1] (class of infinitely differentiable functions with compact support in x space),

$$\mathfrak{D}_x \subseteq \mathfrak{K}. \tag{1}$$

A field with this property will be called localizable. Wightman functions[2]

$$W(\xi_1 \cdots \xi_n) = \langle A(x_0) \cdots A(x_n)\rangle_0, \quad \xi_i = x_{i-1} - x_i \tag{2}$$

of a localizable field have the following "short-distance" behavior:

$$W(\xi_1 \cdots \xi_{i-1}, \lambda u, \xi_{i+1} \cdots \xi_n) \underset{\lambda\to 0}{\longrightarrow} \infty, \tag{3}$$

not worse than $1/\lambda^m$ for a certain $m > 0$. Here ξ_j for $j \neq i$ is fixed (or rather smeared with $\mathfrak{K}$ function) and u is a spacelike unit vector.

Definition: A field with a small-distance behavior worse than that for localizable fields (3) is called nonlocalizable. A theory defined in terms of several fields is called nonlocalizable if at least one of its fields is nonlocalizable.

In the next section we show (in a certain approxi-

[1] Here we use the notation of L. Schwartz, *Théorie des distributions* (Hermann & Cie., Paris, 1957). The indices x and p indicate x-space and momentum-space functions.

[2] This short-distance behavior is intimately related to the statement that the Fourier transform of a Wightman function belonging to a localizable field is tempered in one variable if the rest of the variables are fixed. This is a well-known generalization of a statement originally formulated for the two-point function by A. S. Wightman, Phys. Rev. **101**, 864 (1956).

mation) that theories which are classified as nonrenormalizable in Feynman–Dyson's perturbation theory are nonlocalizable in the sense of the previous definition.

For nonlocalizable fields, the class $\mathfrak{K}$ of allowed smearing function does not contain the localizable function $\mathfrak{D}_x$. Since the class of all smooth and rapidly decreasing functions $\mathfrak{S}_x$ contains $\mathfrak{D}_x$ and has the property that under Fourier transformation it goes over into itself, i.e., $\mathfrak{S}_x = \mathfrak{S}_p$, $\mathfrak{K}$ must be smaller than $\mathfrak{S}_p$. On the other hand, $\mathfrak{D}_p$ is the smallest (from a physical viewpoint) acceptable class,[3] and therefore

$$\mathfrak{D}_p \subseteq \mathfrak{K} \subset \mathfrak{S}_p. \tag{4}$$

The specific form of $\mathfrak{K}$ depends (as in the localizable case) on the particular model.

We would like to illustrate these properties in some examples. Any infinite Wick series[4]

$$B(x) = A(x) + \sum_{\nu=2}^{\infty} \frac{C_\nu}{\nu!} :A^\nu(x):, \tag{5}$$

where $A(x)$ is a free scalar field of mass m, and C_ν real constants, gives rise to a nonlocalizable field. [For the more general case, including derivatives of A or vector (spinor) fields this is also true]. This is evident from the following two statements:

(a) $B(x)$ is not localizable;

(b) The smearing of $B(x)$ with $\mathfrak{D}_p$ functions gives rise to Hermitian operators which can be successively applied to the vacuum.

The state space obtained in this way is irreducible with respect to the algebra of smeared-out B fields.

Statement (a) follows from the form of the two-point function,

$$\langle B(x)B(y)\rangle = \langle A(x)A(y)\rangle + \sum_{\nu=2}^{\infty} \frac{C_\nu^2}{\nu!} [i\Delta^{(+)}(x-y)]^\nu. \tag{6}$$

According to Wightman,[5] the measure ρ defined by

$$\langle B(x)B(y)\rangle = i \int_0^\infty \rho(\kappa^2)\, \Delta^{(+)}(x-y, \kappa^2)\, d\kappa^2, \tag{7}$$

has to be of slow increase if B would be localizable. But ρ belonging to (6) has the form

$$\rho(\kappa) = \delta(\kappa^2 - m^2) + \sum_{2}^{\infty} C_\nu^2 \rho_\nu(\kappa^2), \tag{8}$$

where $\rho_\nu(\kappa^2)$ is proportional to the phase-space volume of κ particles which goes for large κ^2 like $(\kappa^2)^{\nu-2}$. Since all the ρ_ν's are positive, there can be no cancellation and hence ρ can not be of slow increase.

Statement (b) is a consequence of the fact that, in the application of $B(f) = \int B(x)\, f(x)\, d^4x$ with $\tilde{f}(p) \in \mathfrak{D}_p$ onto the vacuum state, only a finite number of terms of the series (5) are contributing (this is only true for $m \neq 0$). The state thus obtained contains only momenta below a certain value which is given by the size of the support of $\tilde{f}(p)$. The application of $B(g)$ with $g \in \mathfrak{D}_p$ to this state again receives only contributions from a finite number of terms. In this way the smeared-out field can be applied successively onto the vacuum, and the states $B(f_1) \cdots B(f_n)\, |0\rangle$ stay in the domain of the $B(f)$'s. The irreducibility follows from Ruelle's[6] consideration.

For purpose of later application, we discuss the special case

$$B(x) = :e^{gA(x)}:, \tag{9}$$

when g is a real constant. We want to study the (optimal) class of smearing functions $\mathfrak{K}$, such that $B(f)$ for $f \in \mathfrak{K}$ is applicable onto the vacuum. Let $\mathfrak{H}_M$ be the subspace of the statespace $\mathfrak{H}$ on which the mass operator is bounded by

$$||P^2|| \leq M^2. \tag{10}$$

Let E_M be the projector onto $\mathfrak{H}_M$. We then compute the following norm:

$$\begin{aligned} ||E_M B(f)\mid 0\rangle||^2 &= \sum_{\nu=0}^{\infty} \int f_M^*(x) f_M(x) \frac{g^\nu}{\nu!} [i\, \Delta^{(+)}(x-x')]^\nu\, dx\, dx' \\ &= \int |\tilde{f}_M(p)|^2\, F(p)\, d^4p, \end{aligned} \tag{11}$$

with

$$\tilde{f}_M(p) = \begin{cases} f(p) & \text{for } p^2 \leq M^2, \\ 0 & \text{for } p^2 > M^2, \end{cases}$$

$$F(p) = \frac{1}{(2\pi)^3} \left\{ \frac{1}{(2\pi)}\, \delta(p) + \sum_{\nu=1}^{\infty} \frac{g^{2\nu}}{\nu!}\, \rho_\nu(p^2)\theta(p^2)\theta(p_0) \right\}, \tag{12}$$

where $\rho_\nu(p^2)$ is the already mentioned phase-space volume of ν particles with total momentum p.

Therefore the limit of (11) for $M \to \infty$ exists if and only if $f(p)$ decreases in such a way that

$$H(\kappa^2) = \int |\tilde{f}(p)|^2\, \delta(p^2 - \kappa^2)\theta(p_0)\, d^4p \tag{13}$$

[3] It is necessary in order to approximate plane-wave states.

[4] Local Wick-ordered polynomials of free fields when first used as field theoretical models by A. S. Wightman, Paris lectures, 1956 (manuscript).

[5] See Ref. 2.

[6] The proof can be found in D. Ruelle, Helv. Phys. Acta 35, (1962) Appendix.

compensates for the growth of the phase-space sum in (12). For $m = 0$ this phase-space sum is

$$F(p) = \frac{1}{(2\pi)^3}\left[\frac{1}{2\pi}\,\delta(p) + g^2\,\delta(p^2)\theta(p_0) + \left\{\sum_{\nu=2}\left(\frac{g}{4\pi}\right)^{2\nu}\frac{(p^2)^{\nu-1}}{\nu^2[(\nu-1)!]^3}\right\}\theta(p^2)\theta(p_0)\right]. \quad (14)$$

The phase-space sum converges against an entire function which is of the order of $\frac{1}{3}$ (in the sense of Titchmarsh[7]).

For the special test function

$$f(x) = \delta(\mathbf{x})\,\frac{1}{(2\pi)}\int e^{-\beta E}e^{Et}\,dE \quad (15)$$

(this test function averages in time only), one gets the particularly simple result

$$||B(f)\mid 0\rangle||^2 = e^{g^2 i\Delta^{(+)}(0,2i\beta)}, \quad (16)$$

when $i\Delta^{(+)}(0,\,2i\beta)$ denotes the free-field two-point function for purely imaginary time and vanishing space component. We could have computed the analytic Wightman function directly by realizing that

$$\langle B(x)B(y)\rangle = \sum\frac{g^\nu}{\nu!}\,[i\,\Delta^{(+)}(\xi)]^\nu = e^{g^2 i\Delta^{(+)}(\xi)} = \lim_{\epsilon\to 0} F(\xi^2 - i\epsilon\xi_0) \quad (17)$$

converges in the analyticity domain of $i\Delta^{(+)}(Z)$, which is the cut plane with a cut extending from 0 to ∞. This Wightman function has an essential singularity at $Z = 0$. Despite this singularity, the Wightman boundary value

$$\lim_{\epsilon\to 0} F(\xi^2 - i\epsilon\xi_0) \quad (18)$$

makes sense as a distribution over $\mathfrak{D}_p$. This is evident from the fact that the Fourier transform

$$\int e^{-ip\xi}\lim_{\epsilon\to 0} F(\xi^2 - i\epsilon\xi_0)\,d^4\xi = \int_C\int e^{-ip_0\xi_0 + i\vec{p}\xi}F(\xi^2)\,d\xi_0\,d^3\xi \quad (19)$$

exists, since the integration path which follows from the boundary prescription (18) is a closed path going around the cut as indicated in Fig. 1.

The decreasing property of the allowed smearing functions $\mathfrak{K}$ in p space lead to analyticity properties

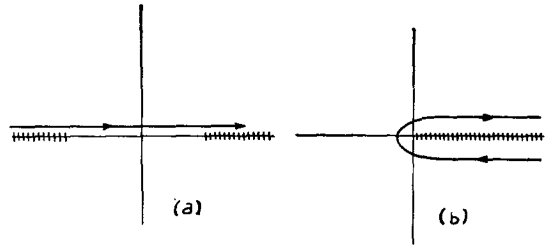

FIG. 1. (a) Path of integration for $\int e^{ipx}\lim_{\epsilon\to\infty} F(x^2 - i\epsilon t)$ in the complex $t = x_0$ plane. F is analytic with the exception of the indicated cuts; (b) Path after transformation into the complex $\xi = t^2 - r^2$ plane.

of the Fourier-transformed functions such that

$$\int f(\xi)\lim_{\epsilon\to 0} F(\xi^2 - i\epsilon\xi_0)\,d^4\xi$$

can be converted into a path integral analog to (19). The Wightman functions in x space are therefore distributions over certain classes of analytic test functions. Such distributions have been studied by Gel'fand.[8] In contradistinction to the Wightman functions, Green's functions, like time-ordered functions or retarded functions, can not be defined for nonlocalizable fields. For the special case of the two-point function, this is evident from the fact that

$$\tilde{\Delta}'_F(p) \sim \int\frac{\rho(\kappa^2)}{p^2 - \kappa^2 + i\epsilon}\,d\kappa^2 \quad (20)$$

cannot be defined by a finite number of subtractions, or alternatively that the boundary prescription in x space does not lead to a closed contour integral [which made it possible to give a distribution theoretical meaning to the singular expression (18)]. In order to show that nonlocalizable fields may lead to a more singular behavior than that of the previous example, we discuss the field

$$B(x) = :A(x)e^{gA^2(x)}:, \quad (21)$$

where g is a real constant having the dimensions of a length. A straightforward computation of the two-point function gives a series

$$\langle B(x)B(y)\rangle = \sum_0^\infty \Delta^{(+)}(\xi)[2gi\,\Delta^{(+)}(\xi)]^{\nu+2}\binom{-\frac{1}{2}}{\nu}(-1)^\nu = i\Delta^{(+)}(\xi)\{1 - [2gi\,\Delta^{(+)}(\xi)]^2\}^{-1}, \quad (22)$$

which converges for

$$4g^2 i\Delta^{(+)^2}(\xi) < 1.$$

The convergence holds for all spacelike ξ with

$$\xi^2 < -a^2,$$

[7] In the sense of E. C. Titchmarsh, *The Theory of Functions*, (Oxford University Press, London, 1939), 2nd ed., p. 248.

[8] I. M. Gel'fand, *Verallgemeinerte Funktionen, Vol. 2* (VEB Deutscher Verlag der Wissenschaften, Berlin, 1960).

where a is a solution of the transcendental equation

$$i\Delta^{(+)}(-a^2) = 1/2g. \tag{23}$$

It is easy to see that the Lehman spectral function of (22) has a much worse growth than the one studied in the previous example. In the specification of how the allowed test functions have to decrease for large p, the length a plays a role. This is intimately connected with the fact that the cut extends slightly into the spacelike region up to the distance a.

For later purposes, it is convenient to introduce the following definition:

Definition: A nonlocalizable field is said to be of the first kind, if the Wightman functions are analytic in the "extended permuted tube" of Bargman–Wightman and Hall.[9] Nonlocalizable fields like (21) which lead to singularities in this region are said to be of the second kind.

III. THE CONNECTION BETWEEN NONRENORMALIZABILITY AND NONLOCALIZABILITY

We want to show that, in a certain approximation (uncrossed string approximation) a theory which is unrenormalizable in the sense of Feynman and Dyson, leads to a "nonlocalizable" two-point function as defined in the previous section.

Consider for example a theory of a boson–spinor coupling,

$$(\gamma^\mu\partial_\mu + M)\psi(x) = ig\mathcal{O}(x)\psi(x), \tag{24}$$

where $\mathcal{O}(x)$ is linear in the boson field (and contains γ matrices). Approximating the boson field by a free field,

$$\mathcal{O}(x) = \mathcal{O}^{(0)}(x), \tag{25}$$

and taking only contraction between this free field into account, we obtain the following equation for the two-point function:

$$(\gamma^\mu_{\alpha\beta}\partial_\mu + M)\langle\psi_\beta(x)\bar{\psi}_\gamma(y)\rangle_0(-\gamma^\mu_{\gamma\delta}\partial_\mu + M) = -g^2\langle\mathcal{O}^{(0)}_{\alpha\beta}(x)\mathcal{O}^{(0)}_{\gamma\delta}(y)\rangle_0\langle\psi_\beta(x)\bar{\psi}_\gamma(y)\rangle. \tag{26}$$

This differential equation is formally equivalent to

$$\begin{aligned}(\gamma^\mu_{\alpha\beta}\partial_\mu + M)S_F^{\beta\gamma}(x-y)(-\gamma^\mu_{\gamma\delta}\partial_\mu + M) \\ = -g^2\langle T\mathcal{O}^{(0)}_{\alpha\beta}(x)\mathcal{O}^{(0)}_{\gamma\delta}(y)\rangle S_F^{\beta\gamma}(x-y) \\ + (\gamma^\mu_{\alpha\delta}\partial_\mu + M)Z_2^{-1}\delta^4(x-y) + Z_2^{-1}\delta M\,\delta^4(x-y),\end{aligned} \tag{27}$$

where

$$S_F(x-y) = \theta(x_0-y_0)\langle\psi(x)\bar{\psi}(y)\rangle + \theta(y_0-x_0)\langle\bar{\psi}(y)\psi(x)\rangle$$

and Z_2 and δM^2 are renormalization terms coming from the differentiation of the step function. (These serve as counter terms to infinities arising in the product of the two time-ordered functions on the right-hand side.)

Equation (27) can be written as an integral equation,

$$\begin{aligned}S_F(x-y) &= Z_2^{-1}S_F^{(0)}(x-y) \\ &+ Z_2^{-1}\,\delta M\int S_F^{(0)}(x-x')S_F^{(0)}(x'-y)\,dx' \\ &+ \int S_F^{(0)}(x-x')\langle T\mathcal{O}^{(0)}(x')\mathcal{O}^{(0)}(y')\rangle \\ &\cdot S_F(x'-y')S_F^{(0)}(y'-y)\,d^4x'\,d^4y'.\end{aligned} \tag{28}$$

This is the linearized version of the renormalized Dyson equation for the propagator. (The equation corresponds to the string approximation of Fig. 2).

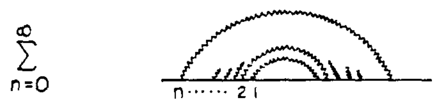

FIG. 2. String approximation.

Starting from (28), Eq. (26) can be viewed as the differential equation for the "absorptive" part of (28). Contrary to (28) there is no renormalization problem in (26). It has to be emphasized that the connection between (26), (27), and (28) is purely formal. As we will see, neither (28) nor (27) can be given a meaning for nonrenormalizable couplings. But, (26) poses a well-defined problem.

Decomposing the two-point function into its invariants (P and C invariance assumed)

$$\langle\psi(x)\bar{\psi}(y)\rangle = \gamma^\mu\partial_\mu F(\xi^2) + G(\xi^2), \tag{29}$$

one obtains a second-order coupled system of differential equations for F and G. For reasons of simplicity, let M be zero. Then the equations for F and G decouple.

For the special case of a derivative coupling

$$\mathcal{O}^{(0)}(x) = \gamma^\mu\partial_\mu A^{(0)}(x), \tag{30}$$

with a zero-rest-mass boson field

$$\langle A^{(0)}(x)A^{(0)}(y)\rangle = \lim_{\epsilon\to 0}\frac{-1}{4\pi(\xi^2 - i\epsilon\xi_0)}, \tag{31}$$

one obtains with $Z = -\xi^2$

$$3F'' + ZF''' = \tfrac{3}{4}(g^2/\pi^2)(1/Z^2)F', \tag{32}$$

$$Z^2G'' + 2G' = 0. \tag{33}$$

Here the fact that F and G are boundary values of

[9] D. Hall and A. S. Wightman, Kgl. Danske Videnskab. Selskab. Mat. Fys. Medd. 31, 1 (1957).

analytic functions has been used to obtain differential equations of the Fuchsian type. (Any differential equation arising from the "string approximation" is of this type.) The Wightman distribution can be recovered at the end of the calculation by taking the $i\epsilon\xi_0$ prescription for the boundary value.

By making the substitution $\zeta = 1/Z^{\frac{1}{2}}$ and $H = \zeta^{-2}F'$, one obtains

$$\frac{d^2}{d\zeta^2}H + \frac{1}{\zeta}\frac{d}{d\zeta}H - \left(\frac{4}{\zeta^2} + \frac{3}{\pi^2}g^2\right)H = 0. \tag{34}$$

The two independent solutions are

$$I_2[(g/\pi)\sqrt{3}\zeta], \qquad K_2[(g/\pi)\sqrt{3}\zeta], \tag{35}$$

or in terms of the original quantities,

$$F'(\xi^2) = \lim_{\epsilon\to 0}\frac{1}{-\xi^2 + i_\epsilon\xi_0}\left\{C_1 I_2\left[\frac{g}{\pi}\sqrt{3}\frac{1}{(-\xi^2 + i\epsilon\xi_0)^{\frac{1}{2}}}\right]\right.$$
$$\left. + C_2 K_2\left(\frac{g}{\pi}\sqrt{3}\frac{1}{(-\xi^2 + i\epsilon\xi_0)^{\frac{1}{2}}}\right)\right\}. \tag{36}$$

The only solution which is in agreement with the fall-off property for large spacelike distances following from the spectrum conditions[10] is

$$F'(\xi^2) = \lim_{\epsilon\to 0}\frac{1}{3g^2(-\xi^2+i\epsilon\xi_0)} I_2\left[\frac{g}{\pi}\sqrt{3}\frac{1}{(-\xi^2+i\epsilon\xi_0)^{\frac{1}{2}}}\right], \tag{37}$$
$$G = 0.$$

It is easy to check that this is also the only solution which leads to a positive-definite spectral function.[11] So the K_2 solution has to be rejected notwithstanding the fact that it leads to a nice exponential damping, and that the function belonging to the K_2 solution is the only solution of the propagator integral equation (28). The string approximation for nonrenormalizable theories therefore leads to an exponential growth for $\rho(\kappa)$ in agreement with the behavior of nonlocalizable fields studied in the previous section. In agreement with the remarks in the previous section, the time-ordered boundary value which is obtained from (37) by replacing the $i\epsilon\xi_0$ by $i\epsilon$ does not define a distribution, i.e., can neither be Fourier transformed nor smeared with test functions.

One is tempted to ask the question if this exponential growth is only a feature of the particular approximation involved, i.e., if it is offset for example by the inclusion of crossed graphs. The feeling of the author is that the growth is a characteristic feature of nonrenormalizable theories. The only evidence for this conjecture comes from the fact that in the case of the considered derivation coupling (31) the graphical structure which would occur if $A(x)$ is a neutral field (occurrence of all crossed structures) can be exactly summed and gives

$$\langle\psi(x)\bar{\psi}(y)\rangle = \langle\psi_0(x)\psi_0(y)\rangle e^{g^2 i\Delta^{(+)}(x-y)}. \tag{38}$$

For the case of zero rest mass, one can show that the Lehmann spectral function ρ_s of the string approximation (37) is a minorization of the Lehmann function ρ belonging to (38),

$$\rho_s(\kappa) \leq \rho(\kappa). \tag{39}$$

But unfortunately we are not able to show, in general, that the inclusion of crossed graphs does not offset the exponential growth.

IV. SCATTERING THEORY WITHOUT THE USE OF GREEN'S FUNCTION

The elastic scattering can be computed, once the matrix element[12]

$$\langle p' \,|A(x)A(y)|\, p\rangle \tag{40}$$

is known. For this matrix element in turn one can derive a differential equation in a certain approximation analogous to (27). The solution of this partial differential equation is subject to boundary conditions which follow from the requirement that (40) should fulfill causality and spectrum condition. The computational problem which this differential equation poses will be discussed elsewhere. From the previous discussion of the two-point function, we expect that for nonrenormalizable theories (40) may have an essential singularity at the light cone $(x - y)^2 = 0$ which could lead to an exponential growth of the Fourier transform and hence Green's function would not be available. Unfortunately, the scattering amplitude has only been expressed in terms of the interpolating fields $A(x)$ with the help of Green's function. The typical stationary scattering formula derived from the asymptotic convergence

$$A(x) \underset{t\to\mp\infty}{\longrightarrow} A_{\substack{\text{in}\\ \text{out}}}(x) \tag{41}$$

expresses the scattering amplitude as the mass-shell

[10] The fall-off property of the two-point function is a result of the mass spectrum of $\rho(\kappa^2)$. In the case of zero-rest mass, the decrease is only like $1/\xi^2$ instead of exponentially. For the case of a spinor field, this leads to (37). In case of a general matrix element, $\langle P\,|A(x/2)A(-x/2)|\,Q\rangle$ locality has to be used in order to obtain similar decreasing properties. See H. Araki, K. Hepp, D. Ruelle, Helv. Phys. Acta **35** (1962).

[11] The constant in front of the I_2 can be fixed by normalizing the decrease in spacelike direction,

$$\langle\psi(x)\bar{\psi}(y)\rangle_{\xi^2 \to -\infty} \longrightarrow \langle\psi^0(x)\bar{\psi}(y)^0\rangle.$$

[12] In order to avoid all unnecessary complications, we write the formulas in terms of scalar fields. The transcription to spinor fields is entirely straightforward.

Fourier transform of a retarded or time-ordered matrix element.[13]

$$\langle p'k' \overset{\text{out}}{|} \overset{\text{in}}{pk}\rangle - \langle p'k' \overset{\text{out}}{|} \overset{\text{in}}{pk}\rangle \equiv -iT$$

$$= \frac{1}{(2\pi)^3} \int e^{-ikx+ik'y} K_x \langle p' \,|\theta(y-x)[A(x)j(y)]|\, p\rangle, \quad (42)$$

where $K_x = \Box_x - m^2$.

For localizable fields the retarded product involved in (42) can only be defined up to an arbitrary number of derivations of $\delta(x-y)$ functions with arbitrary coefficients (which are functions of p and p'). It can easily be seen that, due to the presence of the Klein–Gordon operator K_x, these ambiguities vanish at the mass shell. More generally speaking, one can add any distribution with compact support in $x - y$ to the matrix element $\langle p' \,|\theta(y-x)[A(x)j(y)]|\, p\rangle$ without changing the mass shell quantities.

Since the scattering amplitude does not depend on the ambiguities which are always arising in the definition of Green's function, one suspects that there must be a stationary scattering formula which expresses T in terms of the Wightman functions (unordered matrix elements) directly. We want to derive the following statement:

The matrix element (40) has the following structure (one-particle structure)

$$\langle p' \,|\tilde{A}(q)\tilde{j}(-k)|\, p\rangle$$
$$= i[P/(q^2-m^2)]F - i\pi\delta(q^2-m^2)G + H, \quad (43)$$
$$q^2 \sim m^2, \quad q_0 > 0, \qquad k^2 = m^2, \quad k_0 > 0,$$

where P denotes the Cauchy principle value, and the scattering amplitude (42) can be written as

$$T = [1/(2\pi)^3](G + iF). \quad (44)$$

The first part of the statement follows from the fact that

$$\operatorname{Im} T = \tfrac{1}{2} \lim_{\substack{q\to k'\\ k'^2=m^2}} [(q^2-m^2)\langle p' \,|\tilde{A}(q)j(-h)|\, p\rangle] \quad (45)$$

is a finite quantity. More precisely, the asymptotic condition states that $(q^2-m^2)\langle p' \,|\tilde{A}(q)j(-h)|\, p\rangle$ is free of a singularity in the variable $q^2 - m^2$ near $q^2 - m^2 \sim 0$. Equation (43) with F, G, and H being free of singularities in $q^2 - m^2$ at $q^2 - m^2 \sim 0$ is an immediate consequence of (45).

The second part follows from the following computation:

$$\lim_{t\pm\infty} \langle p' \,|A_{fk'}(t)\tilde{j}(-k)|\, p\rangle$$
$$= \lim_{t\to\pm\infty} \frac{1}{(2\pi)^{5/2}} \int \langle p' \,|A(\mathbf{k}_1'k_0')\tilde{j}(-k)|\, p\rangle$$
$$\cdot[k_0' + (\mathbf{k}'^2 + m')^{\frac{1}{2}}] \exp[(\mathbf{k}'^2+m^2)^{\frac{1}{2}} - k_0']t\, dk_0'$$
$$= \lim_{t\to\pm\infty} \frac{i}{(2\pi)^{5/2}} \int \left\{\frac{P}{q^2-m^2} F - \pi\delta(q^2-m^2)G + H\right\}$$
$$\cdot[k_0' + (k'^2+m^2)^{\frac{1}{2}}] \exp[(k'^2+m^2)^{\frac{1}{2}} - k_0]t\, dk_0'$$
$$= \tfrac{1}{2}(2\pi)^{-\frac{3}{2}}(F \mp iG),$$

and

$$-iT = \frac{1}{(2\pi)^{\frac{3}{2}}} \lim_{t\to+\infty} \langle p' \,|A_{fk'}(t)\tilde{j}(-h)|\, p\rangle$$
$$= \frac{1}{(2\pi)^3}\, 1/2(F - iG). \quad (46)$$

For $t \to -\infty$ we obtain

$$\frac{1}{(2\pi)^{\frac{3}{2}}} \lim_{t\to-\infty} \langle p' \,|A_{fk'}(t)\tilde{j}(-h)|\, p\rangle$$
$$= \frac{1}{(2\pi)^3}\, 1/2(F + iG) = iT^*, \quad (47)$$

which is in accordance with

$$\lim_{t\to-\infty} \langle p' \,|A_f(t)(A_{\text{in}}(-h) - A_{\text{out}}(-h))|\, p\rangle$$
$$= \langle p'k' \overset{\text{in}}{|} \overset{\text{in}}{ph}\rangle - \langle p'h' \overset{\text{out}}{|} ph\rangle^*. \quad (48)$$

During recent years, it has been recognized that the asymptotic condition is not an additional requirement, but rather a consequence of a local quantum field theory. It can be shown that the stationary scattering formula (42) as well as (43), (44) are obtainable from Haag–Ruelle's collision theory, which is valid for any local field theory fullfilling nonzero-rest-mass spectrum conditions.[14] This has the interesting consequence that causality and spectrum conditions are the only boundary conditions which one has for the differential equation of the amplitude (40) mentioned in the beginning of this section. There is no additional boundary condition like an ingoing or outgoing condition in Schrödinger theory.

Nonlocalizable field theories of the first kind as defined in Sec. II lead (like localizable field theories) to Wightman functions which are analytic and symmetric in spacelike points. This property of the Wightman function, together with the spectrum conditions, leads to the asymptotic convergence of

[13] For example H. Lehmann, K. Symansik, and W. Zimmermann, Nuovo Cimento 1, 425 (1955).

[14] A proof of this statement inside the Haag–Ruelle collision theory can be given analog to K. Hepp, Acta Phys. Austriaca 17, 92 (1963).

fields and hence the scattering formulas (43), (44) are valid. Formula (42), however, would break down due to the fact that the multiplication with θ inside the integrand cannot be defined.

V. NONLOCALIZABILITY AND EINSTEIN CAUSALITY

The feature which makes quantum field theory most distinctively different from any other quantum theory is the postulate of Einstein causality, namely, the requirement that local observables[15] belonging to spacelike separated region commute.

For a Hermitian field $A(x)$ which commutes with all the superselecting quantities of the theory,[16] the local observables $\mathcal{O}_{\mathfrak{B}}(A)$ generated by A and belonging to a four-dimensional space–time region $\mathfrak{B}$ are the smeared-out fields:

$$\mathcal{O}_{\mathfrak{B}}(A)=\left\{\int A(x)f(x)\,dx : f(x)\in\mathfrak{D}_{\mathfrak{B}}\right\},$$

when $\mathfrak{D}_{\mathfrak{B}}$ is the set of all smooth functions which have their support in $\mathfrak{B}$. In order that Einstein causality can be formulated, these local observables must exist, which amounts to saying that $A(x)$ must be a localizable field. Einstein causality does not put any (explicit) restriction on nonobservable fields, i.e., fields which do not commute with the superselecting quantities. The best known example for this state of affairs is quantum electrodynamics in the Coulomb gauge. Neither the potential nor the electron spinor field commute (anticommute) for spacelike distances (the renormalized spinor field does not even anticommute for a fixed time[17]). There is, however, no violation of Einstein causality, since all the observable fields like currents, field strength, etc., commute for spacelike distances. In the special context of nonrenormalizable theories, it has to be stressed that even the occurrence of nonlocalizable fields is not in disagreement with Einstein causality, as long as these fields are unobservable. This seems to be true for all realistic models which play a role in weak interactions. Of course we do not know if the observable fields (like currents) are localizable. Hence it seems to be not possible at the moment to decide whether realistic nonrenormalizable models are still inside the usual field-theoretical frame or not.

The models studied in the first section are not of much help in the pursuit of this problem. The violation of Einstein causality for the Hermitian field [Eq. (21)] is a consequence of the occurrence of an elementary length, but this mathematical model unfortunately does not have any physical consequences, since it suffers from the common disease of having a trivial S matrix.

ACKNOWLEDGMENTS

The author is indebted to Professor F. Dyson for critical remarks and helpful suggestions. He wishes to thank Professor J. R. Oppenheimer for his hospitality at the Institute for Advanced Study.

It is also a pleasure to thank Professor A. Pais and Professor T. Regge for interesting conversations.

[15] The notion of local observables has been most concisely introduced by H. Araki, Zurich lecture notes, 1962.

[16] In a theory which admits a gauge group, a field generating local observables must naturally be also gauge-invariant.

[17] For a discussion of quantum electrodynamics in the Coulomb gauge see for example L. Evans, thesis, the Johns Hopkins University (1960).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 10 OCTOBER 1964

The Graviton as a Spin-2 Particle*

OLOF BRULIN AND STIG HJALMARS
Royal Institute of Technology, Stockholm, Sweden
(Received 9 June 1964)

It is proved that the linearized gravitational field can be described by means of a first-order relativistic wave equation with matrix coefficients, obtained in a simple way from the generators of the full linear group in five dimensions.

SINCE the beginning of relativistic quantum theory, many approaches have been made to the problem of casting the field equations of all known elementary particles in the same form as the ordinary Dirac equation, i.e.,

$$(\Gamma_k p_k - \lambda)\psi = 0, \tag{1}$$

where

$$p_k = -i\hbar\, \partial/\partial x_k, \qquad x_4 = ict$$

(Latin indices = 1, 2, 3, 4).

As is well-known, the covariance of (1) requires

$$[q_{kl}, \Gamma_r] = \delta_{lr}\Gamma_k - \delta_{kr}\Gamma_l, \tag{2}$$

$$[q_{kl}, \lambda] = 0. \tag{3}$$

Here

$$q_{kl} = -q_{lk} \tag{4}$$

are the infinitesimal generators of the Lorentz transformation

$$\psi' = Q\psi = (1 + \tfrac{1}{2}\epsilon_{rs}q_{rs})\psi, \tag{5}$$

where

$$p_k' = p_k + \epsilon_{kl}p_l, \qquad \epsilon_{kl} = -\epsilon_{lk}. \tag{6}$$

The q_{kl} satisfy the group condition of the orthogonal group:

$$[q_{kl}, q_{rs}] = \delta_{lr}q_{ks} - \delta_{kr}q_{ls} + \delta_{ls}q_{rk} - \delta_{ks}q_{rl}. \tag{7}$$

For spin $\frac{1}{2}$, 0, and 1 the following special connection between Γ_k and q_{rs} is satisfied:

$$q_{kl} = [\Gamma_k, \Gamma_l]. \tag{8}$$

As pointed out by Klein[1] in 1936, the assumption

$$\Gamma_m = iq_{m5} \tag{9}$$

makes the Eqs. (2), (7), and (8) immediate consequences of the group condition for the 5-dimensional orthogonal group, i.e., (7) with Greek indices, taking the values 1, 2, 3, 4, 5.

Moreover, Klein has devised a procedure (described in Ref. 2 for the case of spin 0 and 1) to derive natural representations for the Γ_k matrices for integer spin particles by means of the irreducible tensor representations of the 5-dimensional orthogonal group. In this method the components of ψ are identified with the independent components of the corresponding irreducible 5-dimensional tensor. The condition that the infinitesimal transformation matrix of ψ has to transform each component of ψ as the assumed tensor component, gives an explicit representation of $q_{\kappa\lambda}$ and thus of q_{kl} and Γ_r from (9).

The idea of Klein to utilize the 5-dimensional orthogonal group is accordingly very adequate to describe spin-0 and -1 particles. Moreover, as shown by Brulin and Hjalmars[3,4] these methods can also be adapted to the case of spin-0 and -1 particles in external gravitational fields.

For higher spin particles, the 5-dimensional aspect gives in general particles with a mass spectrum. Bhabha[5] investigates, e.g., the case when λ in (1) is a constant times the unit matrix. Equation (8) is then shown to have the consequence that all components of ψ do not satisfy the same Klein–Gordon equation, but only a higher-order equation containing a product of Klein–Gordon operators with different mass parameters. However, as shown by Brulin and Hjalmars[6,7] the condition (8) admits higher-spin particles with unique mass, if the term λ in (1) is assumed to be a more general matrix, satisfying (3).

For spin-2 particles, described by (1) and (8), the commutation relations for the Γ_k have been constructed by Madhava Rao.[8] The component differential equations for the case of λ being pro-

* Submitted in honor of the seventieth birthday of Professor O. Klein.

[1] O. Klein, Arkiv Mat. Astr. Fysik **25A**, No. 15 (1936).

[2] S. Hjalmars, Arkiv Fysik **1**, 41 (1949).

[3] O. Brulin and S. Hjalmars, Arkiv Fysik **5**, 163 (1952).

[4] S. Hjalmars, J. Math. Phys. **2**, 663 (1961).

[5] H. J. Bhabha, Rev. Mod. Phys. **17**, 200 (1945).

[6] O. Brulin and S. Hjalmars, Arkiv Fysik **14**, 49 (1958).

[7] O. Brulin and S. Hjalmars, Arkiv Fysik **16**, 19 (1959).

[8] B. S. Madhava Rao, Proc. Ind. Acad. Sci., **A15**, 139 (1942).

portional to the unit matrix have been given by Tsuneto, Hirosige, and Fujiwara.[9,10] The component equations for the most general λ are given by Brulin and Hjalmars in Ref. 6 and as special cases of the still more general formalism, developed in Refs. 7 and 11.

As well-known, the graviton can be considered as a zero-mass limit of a spin-2 particle, just as the photon can be considered as a zero-mass limit of a spin-1 particle. Neither in the spin-1 nor in the spin-2 case this zero-mass limit can be obtained by simply putting $\lambda = 0$ in (1). As shown by Harish-Chandra,[12] the photon case can be obtained from (1) only by considering the most general λ, satisfying (3). This λ turns out to be a linear combination of two projection operators and only one of the two arbitrary constants in this combination has to be put equal to zero in order to obtain the photon case. A first attempt to obtain the graviton in a similar way would be to start with a spin-2 equation (1), satisfying (8) and with the most general λ, satisfying (3), and then to make some of the arbitrary constants of λ equal to zero. As can be seen from the treatment in Refs. 6, 7, and 11 such a scheme is not sufficiently general to contain the graviton case.

In order to describe the graviton by means of a spin-2 equation of the type (1) it is in fact necessary to widen the scheme and make use of the generators of the full linear group in five dimensions instead of the orthogonal group, used by Klein. This generalization was worked out in Refs. 7 and 11. It was shown there, that if the matrix coefficients of (1), constructed from the generators of the full linear group, were to have the closest possible analogy with the coefficients, obtained from the orthogonal group, preserving the condition (8), it was not possible to obtain the graviton case directly. In fact, it was in this case necessary to introduce the rather artificial trick of multiplying the coefficient matrices by suitable projection operators, satisfying (3).

It will be shown below that if the close connection with the generators of the 5-dimensional orthogonal group is abandoned, i.e., the freedom of the full linear group is utilized, it is possible to obtain the equations of the linearized gravitational field in a straightforward way.

According to (5) it is only the antisymmetric part of the generators, which plays a rôle in the orthogonal transformation, whereas the symmetric part is completely undetermined. The simplest assumption as to the symmetric part is of course to put it equal to zero, according to (4). We now generalize the scheme by allowing a symmetric part in the generators, putting

$$Q = 1 + \epsilon_{rs} Q_{rs}, \tag{10}$$

where the antisymmetric part

$$q_{rs} = Q_{rs} - Q_{sr} \tag{11}$$

has to satisfy (2), (3), and (7).

Evidently, we are able to satisfy (2) by the following two alternative assumptions:

$$[Q_{rs}, \eta_l] = -\delta_{lr}\eta_s, \tag{12}$$

or

$$[Q_{rs}, \vartheta_l] = \delta_{ls}\vartheta_r, \tag{13}$$

where η_l and ϑ_l are two possible sets of Γ_l.

Consequently, the condition of covariance is satisfied for a wave equation of the form

$$[(a\eta_k + b\vartheta_k)p_k - \lambda]\psi = 0, \tag{14}$$

where a and b are arbitrary constants.

Now, consider the full linear group in five dimensions, where a vector is transformed according to the formula

$$A^{\mu\prime} = A^\mu + \epsilon^\mu_\nu A^\nu. \tag{15}$$

The infinitesimal group generators $Q_\mu{}^\nu$, defined by

$$\psi' = Q\psi = (1 + \epsilon^\nu_\mu Q_\mu{}^\nu)\psi, \tag{16}$$

satisfy the group condition of the full linear group

$$[Q_\alpha{}^\beta, Q_\delta{}^\gamma] = g_\delta{}^\beta Q_\alpha{}^\gamma - g_\alpha{}^\gamma Q_\delta{}^\beta. \tag{17}$$

We now notice, that with the definition

$$\eta_k = -iQ_5{}^k, \tag{18}$$

$$\vartheta_k = iQ_k{}^5, \tag{19}$$

$$Q_{rs} = Q_r{}^s, \tag{20}$$

Eqs. (12), (13), and (7) with (11) all follow from (17). Moreover,

$$[\eta_k, \eta_l] = [\vartheta_k, \vartheta_l] = 0 \tag{21}$$

and

$$q_{kl} = [\vartheta_k, \eta_l] + [\eta_k, \vartheta_l] = [\eta_k + \vartheta_k, \eta_l + \vartheta_l], \tag{22}$$

which is the counterpart to the special condition (8).

It follows from the preceding arguments, that every irreducible representation of the full linear

[9] T. Tsuneto, T. Hirosige, and I. Fujiwara, Progr. Theoret. Phys. (Kyoto) 14, 267 (1955).
[10] T. Tsuneto and I. Fujiwara, Progr. Theoret. Phys. (Kyoto) 20, 439 (1958).
[11] O. Brulin and S. Hjalmars, Arkiv Fysik 18, 209 (1960).
[12] Harish-Chandra, Proc. Roy. Soc. (London) A186, 502 (1946).

group in five dimensions gives rise to a possible wave equation (1). Now these representations can be obtained as tensors with symmetries in the indices, corresponding to the different primitive idempotents of the Frobenius algebra for the symmetric group of permutations of the indices. When the components of the wavefunction ψ are assumed to be the independent components of such a tensor, the identification of $Q\psi$ with the tensor transformation of ψ gives, just as in the orthogonal case, an explicit representation of the $Q_\kappa^{\ \lambda}$ and thus of η_k, ϑ_k, q_{kl}.

For spin 2 there are three possibilities, i.e., the tensors

$$\psi_{\kappa\lambda} = (\kappa\lambda)T_{\kappa\lambda}, \tag{23}$$

$$\psi_{\kappa\lambda\mu} = \begin{pmatrix}\kappa\lambda\\ \mu\end{pmatrix}T_{\kappa\lambda\mu}, \tag{24}$$

$$\psi_{\kappa\lambda\mu\nu} = \begin{pmatrix}\kappa\lambda\\ \mu\nu\end{pmatrix}T_{\kappa\lambda\mu\nu}, \tag{25}$$

where the brackets denote the standard Young tableaux operating on the indices of the subsequent tensor. It may be noted that the symmetry (25) is the same as that of the Riemann–Christoffel tensor.

As can be seen from Refs. 6, 7, and 11, only the case (25) is able to give the graviton equations from (14). In this case the procedure of constructing the matrix representation gives the following result:

$$(\eta_k\psi)_{\alpha\beta\gamma\delta} = -i(\delta_{5\alpha}\psi_{k\beta\gamma\delta} + \delta_{5\beta}\psi_{\alpha k\gamma\delta} + \delta_{5\gamma}\psi_{\alpha\beta k\delta} + \delta_{5\delta}\psi_{\alpha\beta\gamma k}), \tag{26}$$

$$(\vartheta_k\psi)_{\alpha\beta\gamma\delta} = i(\delta_{k\alpha}\psi_{5\beta\gamma\delta} + \delta_{k\beta}\psi_{\alpha5\gamma\delta} + \delta_{k\gamma}\psi_{\alpha\beta5\delta} + \delta_{k\delta}\psi_{\alpha\beta\gamma5}). \tag{27}$$

The matrix λ consists of projection operators, projecting into the subspaces, spanned by the components of the different types of 4-dimensional tensors, into which $\psi_{\alpha\beta\gamma\delta}$ decomposes.

Using for short the notation $\partial/\partial x_k = \partial_k$ and

$$\psi_{5klm} = \varphi_{klm}, \quad \psi_{5k5l} = \varphi_{kl}, \quad \psi_{rkrl} = \psi_{kl}, \quad \psi_{5rkr} = \varphi_{rkr} = \varphi_k, \quad \psi_{5r5r} = \varphi_{rr} = \varphi, \quad \psi_{rsrs} = \psi, \tag{28}$$

we obtain the following field equations by putting $a = 0$ in (14):

$$\begin{aligned}\partial_k\varphi_{lmn} - \partial_l\varphi_{kmn} + \partial_m\varphi_{nkl} - \partial_n\varphi_{mkl} &= C_1\psi_{klmn} + C_2(\psi_{ln}\delta_{km} - \psi_{lm}\delta_{kn} + \psi_{km}\delta_{ln} - \psi_{kn}\delta_{lm}) + C_3\psi(\delta_{km}\delta_{ln} - \delta_{kn}\delta_{lm})\\ &\quad + C_4(\varphi_{ln}\delta_{km} - \varphi_{lm}\delta_{kn} + \varphi_{km}\delta_{ln} - \varphi_{kn}\delta_{lm}) + C_5\varphi(\delta_{km}\delta_{ln} - \delta_{kn}\delta_{lm}),\end{aligned} \tag{29}$$

$$\partial_l\varphi_{km} - \partial_m\varphi_{kl} = C_6\varphi_{klm} + C_7(\varphi_l\delta_{km} - \varphi_m\delta_{kl}), \tag{30}$$

$$0 = C_8\varphi_{kl} + C_9\varphi\delta_{kl} + C_{10}\psi_{kl} + C_{11}\psi\delta_{kl}, \tag{31}$$

where the constants $C_1 \ldots C_{11}$ are to be suitably chosen.

The simplest identification with the linearized theory of gravitation is

$$\varphi_{kl} = (1/\epsilon)(g_{kl} - \delta_{kl}), \qquad (\epsilon \text{ small}). \tag{32}$$

Putting

$$C_6 = 1, \qquad C_7 = 0, \tag{33}$$

Eq. (30) defines φ_{klm} as a superpotential. With

$$C_1 = 4, \qquad C_2 = C_3 = C_4 = C_5 = 0, \tag{34}$$

Eq. (29) defines ψ_{klmn} as the Riemann–Christoffel curvature tensor R_{klmn}. Thus, we obtain the ordinary definition of the curvature tensor:

$$R_{klmn} = \tfrac{1}{2}(\partial_k\partial_m\varphi_{ln} - \partial_k\partial_n\varphi_{lm} + \partial_l\partial_n\varphi_{km} - \partial_l\partial_m\varphi_{kn}). \tag{35}$$

As is well-known, the gravitational field equations with cosmological term for empty space,

$$R_{kl} - \tfrac{1}{2}g_{kl}R + kg_{kl} = 0, \tag{36}$$

give

$$k = \tfrac{1}{4}R. \tag{37}$$

The field equations can thus be written

$$R_{kl} - \tfrac{1}{4}g_{kl}R = 0. \tag{38}$$

These equations in the linear approximation are in our scheme obtained from (31) by putting

$$C_8 = C_9 = 0, \qquad C_{10} = 1, \qquad C_{11} = -\tfrac{1}{4}. \tag{39}$$

The extra condition

$$k = 0, \tag{40}$$

giving the ordinary Einstein equations

$$R_{kl} - \tfrac{1}{2}g_{kl}R = 0, \tag{41}$$

is obtained in the linear approximation from (31) by changing the choice of constants (39) into

$$C_8 = C_9 = 0, \qquad C_{10} = 1, \qquad C_{11} = -\tfrac{1}{2}. \tag{42}$$

It has thus proved possible to describe the graviton as a massless spin-2 particle by means of a first-order wave equation with matrix coefficients, obtained from the generators of the full linear group in five dimensions.

The Hose Instability Dispersion Relation

STEVEN WEINBERG*
Department of Physics, University of California, Berkeley, California
(Received 6 April 1964)

The dispersion relation is obtained for the "hose" instability in a modulated beam of charged particles traveling through a finite ohmic plasma channel. For a relativistic beam, the low-frequency mode obeys a dispersion relation of the same form as for the unmodulated beam, but it involves a constant which will have to be obtained by machine computation. For large enough modulation frequency the dispersion relation is identical to that for an unmodulated beam.

1. INTRODUCTION AND SUMMARY

IN this article we present a mathematical derivation of the dispersion relation for the "hose" instability in a modulated beam.[1] The unperturbed system consists of a beam of charged particles with mass m and charge e, moving with velocity v in the z direction, through an ohmic plasma channel of radius R and conductivity σ. The density of beam particles is assumed to be of the form

$$n(r) \sum_s f_s \exp\left[isp(z - vt)\right] \qquad (f_0 \equiv 1), \tag{1.1}$$

where pv is the modulation frequency, the f_s are Fourier coefficients with $s = 0, \pm 1, \pm 2$, etc., and $n(r)$ is some smooth cylindrically symmetric time-averaged beam particle number density.

The instability is assumed to have two characteristic features:

(1) The beam moves rigidly from side to side, with a displacement in a fixed lateral direction having z and t dependence of the form

$$e^{-i\Omega t} \sum_s d(\kappa - sp) \exp\left[i(\kappa - sp)(z - vt)\right]. \tag{1.2}$$

We can loosely refer to κ and Ω as the wavenumber and frequency characterizing a particular mode. The assumption of rigid beam displacement is essential, and is made throughout.

(2) The plasma conductivity σ is much larger than $|\Omega|$, pv, or κv. This assumption is less essential, and will not be used until Sec. 7.

Our conclusion is that the dispersion relation must be found by solving a set of coupled linear equations:

$$\begin{aligned}[\Omega^2 - \omega_\beta^2 \mathfrak{F}(\Omega + \kappa v - spv)]d(\kappa - sp) &= \omega_\beta^2 \sum_{s' \neq s} f_{s'-s}[\mathfrak{F}(\Omega + \kappa v - spv) \\ &\quad - \mathfrak{F}_0([s' - s]pv)]d(\kappa - s'p),\end{aligned} \tag{1.3}$$

where ω_β is the betatron frequency determined by the average beam density

$$\omega_\beta^2 \equiv 2\pi \langle n \rangle e^2 v^2 / m\gamma c^2, \tag{1.4}$$

$$\gamma \equiv (1 - v^2/c^2)^{-\frac{1}{2}}, \tag{1.5}$$

$$\langle n \rangle \equiv \int d^2 r\, n(r)^2 \Big/ \int d^2 r\, n(r), \tag{1.6}$$

and $\mathfrak{F}(\omega)$ and $\mathfrak{F}_0(\omega)$ are functions defined by

$$\mathfrak{F}(\omega) \equiv F(q^2) - \frac{i\pi}{2}\left[H_0^{(1)}(qR)/J_0(qR)\right] q^2 \rho(q)^2, \tag{1.7}$$

$$\begin{aligned}\mathfrak{F}_0(\omega) &\equiv F(q^2) - \frac{i\pi}{2}\left[H_0^{(1)\prime}(qR)/J_0'(qR)\right] q^2 \rho(q)^2 \qquad (\gamma \cong 1), \\ &\equiv \mathfrak{F}(\omega) \qquad (\gamma \gg 1),\end{aligned} \tag{1.8}$$

$$q^2(\omega) \equiv 4\pi i\sigma\omega/c^2, \tag{1.9}$$

$$F(q^2) \equiv \frac{-\frac{i\pi}{2} q^2 \int_0^\infty r\,dr \int_0^\infty r'\,dr'\, n(r) H_0^{(1)}(qr_>) J_0(qr_<) n(r')}{\int_0^\infty n(r)^2 r\,dr}, \tag{1.10}$$

$$\rho(q) \equiv \int_0^\infty n(r) J_0(qr) r\,dr \Big/ \left[\int_0^\infty n(r)^2 r\,dr\right]^{\frac{1}{2}}. \tag{1.11}$$

The characteristic beam functions $F(q^2)$ and $\rho(q)$ are discussed in detail in Appendix B.

Sections 2–7 are devoted to a derivation of (1.3)–(1.11). In Sec. 2 we solve Maxwell's equations for the unperturbed electric and magnetic fields. The beam current density associated with a small beam displacement is calculated in Sec. 3, and used in Sec. 4 to find the perturbed electric and magnetic fields. (Details concerning the effect of the finite plasma channel radius are presented in Appendix A.) These fields and currents are used in Sec. 5 to find

---

* Alfred P. Sloan Foundation Fellow.

[1] Earlier work on the hose instability by M. Rosenbluth, Phys. Fluids 3, 932 (1960), dealt with an unmodulated beam.

the total force per unit beam length, and the exact general dispersion relation is derived from Newton's second law in Sec. 6. The high-σ approximation is invoked in Sec. 7, yielding (1.3)–(1.11).

If there is no beam modulation then (1.3) gives the explicit dispersion relation

$$\Omega^2 = \omega_\beta^2 \mathfrak{F}(\Omega + kv). \qquad (1.12)$$

This dispersion relation is discussed in detail in Sec. 8. Our conclusions are similar to those of Rosenbluth,[1] except that we encounter no logarithmic divergences, even if the plasma channel radius R is taken infinite.[2] In particular we find that for $\Omega + kv$ sufficiently small, (1.12) becomes

$$\Omega^2 = -i\omega_0(\Omega + kv), \qquad (1.13)$$

where

$$\omega_0 = \frac{\omega_\beta^2 \pi\sigma a^2}{c^2} \ln\left(\frac{C^2R^2}{4r_0^2}\right) \times \left(\text{if } \frac{4\pi\sigma\ |\Omega + kv|}{c^2} \ll a^{-2} \text{ and } R^{-2}\right), \qquad (1.14)$$

$$\omega_0 = \frac{\omega_\beta^2 \pi\sigma a^2}{c^2} \ln\left(\frac{ic^2}{4\pi\sigma(\Omega + kv)r_0^2}\right) \times \left(\text{if } R^{-2} \ll \frac{4\pi\sigma\ |\Omega + kv|}{c^2} \ll a^{-2}\right). \qquad (1.15)$$

Here $\ln C = 0.577 \cdots$, and a and r_0 are characteristic beam radii:

$$a = \sqrt{2} \int_0^\infty n(r) r\, dr \bigg/ \left[\int_0^\infty n(r)^2 r\, dr\right]^{\frac{1}{2}} \qquad (1.16)$$

$$\ln r_0^2 = \int_0^\infty \ln\left(\frac{C^2 r^2}{4}\right) d \times \left[\int_0^r n(r')r'\, dr' \bigg/ \int_0^\infty n(r) r\, dr\right]^2. \qquad (1.17)$$

Even if the beam is modulated, we still get a dispersion relation of the same form for both Ω and $\Omega + \kappa v$ sufficiently small, provided that $\gamma \gg 1$. We show in Sec. 9 that in this case

$$\Omega^2 = -i\omega_1(\Omega + \kappa v), \qquad (1.18)$$

where

$$\omega_1 = \omega_0 + \frac{4\pi\sigma a^2 \omega_\beta^2}{c^2} \sum_{s\neq 0} f_s u_s \mathfrak{F}(spv), \qquad (1.19)$$

the coefficients u_s being determined for $s \neq 0$ by the linear equations (with $s \neq 0$)

$$-\mathfrak{F}(-spv)u_s = \frac{-ic^2\mathfrak{F}'(-spv)}{4\pi\sigma a^2} f_s^* + \sum_{s'\neq 0, s} f_{s'-s}[\mathfrak{F}(-spv) - \mathfrak{F}([s' - s]pv)]u_{s'}. \qquad (1.20)$$

The nonrelativistic case is somewhat more complicated, and is also treated in Sec. 9.

In Sec. 10 we show that the dispersion relation for a modulated beam is precisely the same as for an unmodulated beam, provided that the chopping frequency pv is large enough so that

$$pv \gg \frac{c^2}{2\pi\sigma ba}\left[\frac{1}{L}\sum_{s\neq 0}\frac{|f_s|^2}{s^2}\right]^{\frac{1}{2}}, \qquad (1.21)$$

where L is the logarithm in (1.14) or (1.15), and b is another effective beam radius

$$b^2 = \int_0^\infty n(r)^2 r\, dr \bigg/ \int_0^\infty \left[\frac{dn(r)}{dr}\right]^2 r\, dr. \qquad (1.22)$$

This conclusion is quite reasonable, since the relaxation rate for the perturbed fields is roughly of order $2\pi\sigma a^2/c^2$.

For definiteness we have derived the boundary conditions on the unperturbed and perturbed fields by assuming a finite uniform plasma channel with nothing outside. However, our work can be adapted very easily to any other boundary, such as a conducting tube outside a finite plasma channel, by a simple modification of the function $\mathfrak{F}(\omega)$ introduced in (1.3).

A program of numerical computation of the frequency ω_1 in the dispersion relation (1.18) has been started at Stanford Research Institute.

2. THE UNPERTURBED BEAM

The beam particles have mass m, charge e, and velocity $\mathbf{v}$ in the z direction. Their number density is of the form

$$n_0(\mathbf{r}, z, t) = n(r)f(z - vt). \qquad (2.1)$$

Here and throughout, $\mathbf{r}$ will be understood as a two-dimensional vector $\{x, y, 0\}$. We are assuming cylindrical symmetry, so n only depends on the distance $r = (x^2 + y^2)^{\frac{1}{2}}$ to the beam axis.

The beam electric current corresponding to (2.1) is

$$\mathbf{J}_{0B}(\mathbf{r}, z, t) = evn(r)f(z - vt). \qquad (2.2)$$

In addition the plasma contributes a current

$$\mathbf{J}_{0P}(\mathbf{r}, z, t) = \sigma\mathbf{E}_0(\mathbf{r}, z, t). \qquad (2.3)$$

The electric field $\mathbf{E}_0$ and magnetic field $\mathbf{B}_0$ are determined by three of Maxwell's equations

[2] G. Ascoli and H. Chang have independently obtained this result for the special case of an uniform beam. In this case, (1.16) is just the radius of the beam.

$$\nabla \times \mathbf{B}_0(\mathbf{r}, z, t) = (1/c)\dot{\mathbf{E}}_0(\mathbf{r}, z, t) + (4\pi/c)[\mathbf{J}_{0B}(\mathbf{r}, z, t) + \mathbf{J}_{0P}(\mathbf{r}, z, t)], \quad (2.4)$$

$$\nabla \times \mathbf{E}_0(\mathbf{r}, z, t) = -(1/c)\dot{\mathbf{B}}_0(\mathbf{r}, z, t), \quad (2.5)$$

$$\nabla \cdot \mathbf{B}_0(\mathbf{r}, z, t) = 0. \quad (2.6)$$

We do not use the fourth Maxwell equation, because the plasma charge density is implicately determined by (2.3) and charge conservation.

To solve these equations we will Fourier-analyze f:

$$f(z - vt) = \int_{-\infty}^{\infty} f(k)e^{ik(z-vt)}\, dk. \quad (2.7)$$

We then find that

$$\mathbf{E}_0(\mathbf{r}, z, t) = \int_{-\infty}^{\infty} \mathbf{E}_0(\mathbf{r}, k)e^{ik(z-vt)}\, dk, \quad (2.8)$$

$$\mathbf{B}_0(\mathbf{r}, z, t) = \int_{-\infty}^{\infty} \mathbf{B}_0(\mathbf{r}, k)e^{ik(z-vt)}\, dk, \quad (2.9)$$

$$\mathbf{E}_0(\mathbf{r}, k) = \frac{4\pi e f(k)}{[1 + 4\pi i\sigma/kv]}[\nabla_\perp - i\hat{\mathbf{v}}q_0^2(k)/k]\phi_0(r, k), \quad (2.10)$$

$$\mathbf{B}_0(\mathbf{r}, k) = [4\pi e f(k)/c][\mathbf{v} \times \nabla_\perp]\phi_0(r, k), \quad (2.11)$$

where

$$q_0^2(k) = -k^2[1 - (v/c)^2] + 4\pi i\sigma kv/c^2, \quad (2.12)$$

$$\nabla_\perp = \left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, 0\right\} = \hat{\mathbf{r}}\frac{\partial}{\partial r},$$

$$\hat{\mathbf{v}} = \mathbf{v}/v = \{0, 0, 1\},$$

and

$$[\nabla_\perp^2 + q_0^2(k)]\phi_0(r, k) = n(r). \quad (2.13)$$

The most general solution for ϕ_0 which stays finite at the beam axis is

$$\phi_0(r, k) = \int_0^\infty G_0(r, r'; k)r'n(r')\, dr', \quad (2.14)$$

where G_0 is a cylindrical Green's function

$$G_0(r, r'; k) = (-\tfrac{1}{2}i\pi)J_0(q_0(k)r_<)H_0^{(1)}(q_0(k)r_>) - \alpha_0(k)J_0(q_0(k)r)J_0(q_0(k)r'). \quad (2.15)$$

Here $q_0(k)$ is the root of (2.12) with

$$\operatorname{Im} q_0(k) > 0, \quad (2.16)$$

and as usual $r_>$ and $r_<$ are the greater and lesser of r and r'. *The parameter $\alpha_0(k)$ is to be determined by imposing a boundary condition on ϕ_0 outside the beam.* [Taking $q_0(k)$ with opposite sign would just be equivalent to a redefinition of $\alpha_0(k)$.]

It is very important at this point to recognize that (2.8)–(2.16) give the correct solution for $\mathbf{E}_0$ and $\mathbf{B}_0$ only where the plasma conductivity σ is constant. It is safe to assume that σ *is* constant inside the beam, which is where we need to know $\mathbf{E}_0$ and $\mathbf{B}_0$. But outside the beam σ may drop to zero sharply or gradually, or may become infinite at the walls of the system. These different possibilities will be reflected in our choice of boundary conditions on ϕ_0, and hence in the value of $\alpha_0(k)$. In order to calculate $\alpha_0(k)$ we will assume that there is a region just outside the beam where n is negligible but σ is still constant. The potential ϕ_0 in this region is given by (2.14) as

$$\phi_0(r, k) = [-\tfrac{1}{2}i\pi H_0^{(1)}(q_0(k)r) - \alpha_0(k)J_0(q_0(k)r)] \times \int_0^\infty J_0(q_0(k)r')r'n(r')\, dr' \quad (2.17)$$

so we must find $\alpha_0(k)$ by using Maxwell's equations to integrate out from this region through the region where σ varies with r, and then impose the correct boundary condition at infinity or at the system walls.

For example, if σ is actually constant everywhere, then (2.17) gives the fields everywhere outside the beam. We must choose the exponentially decaying solution; so in this case

$$\alpha_0(k) = 0. \quad (2.18)$$

On the other hand, suppose that the plasma conductivity σ stays constant only out to some plasma radius R, and then drops sharply to zero. In this case the fields outside the beam are given by (2.10), (2.11), and (2.17) for $r < R$, while in the vacuum $r > R$ they are given by setting $\sigma = 0$ in (2.10)–(2.13):

$$\mathbf{E}_0(\mathbf{r}, k) = 4\pi e f(k)[\nabla_\perp + ik\gamma^{-2}\hat{\mathbf{v}}]\phi_0(r, k) \quad (2.19)$$

$$\mathbf{B}_0(\mathbf{r}, k) = [4\pi e f(k)/c][\mathbf{v} \times \nabla_\perp]\phi_0(r, k) \quad (2.20)$$

$$[\nabla_\perp^2 - k^2\gamma^{-2}]\phi_0(r, k) = 0. \quad (2.21)$$

As usual, $\gamma \equiv [1 - (v/c)^2]^{-\frac{1}{2}}$. Observe that the beam cannot radiate light into the vacuum, because kc is greater than the "frequency" kv. Instead, we must choose the exponentially decaying solution

$$\phi_0(r, k) \propto K_0(rk/\gamma). \quad (2.22)$$

The conditions required for matching (2.17) and (2.22) at $r = R$ are given by the usual discontinuity relations at the surface of a cylinder of finite conductivity:

$$\Delta(E_\theta) = 0, \quad (2.23)$$

$$\Delta(E_z) = 0, \quad (2.24)$$

$$\Delta(\dot{E}_r + 4\pi\sigma E_r) = 0, \quad (2.25)$$

$$\Delta(B_\theta) = 0, \tag{2.26}$$

$$\Delta(B_z) = 0, \tag{2.27}$$

$$\Delta(B_r) = 0. \tag{2.28}$$

Conditions (2.23), (2.27), and (2.28) are satisfied trivially. Condition (2.24) gives

$$q_0^2(k)[1 + 4\pi i\sigma/kv]^{-1}\phi_0(R - \epsilon, k) = -k^2\gamma^{-2}\phi_0(R + \epsilon, k). \tag{2.29}$$

Conditions (2.25) and (2.26) both give

$$\phi_0'(R - \epsilon, k) = \phi_0'(R + \epsilon, k). \tag{2.30}$$

Our boundary condition on (2.17) is therefore

$$\phi_0'(R, k)/\phi_0(R, k) = -q_0(k)\beta_0(k) \tag{2.31}$$

$$\beta_0(k) = q_0(k)\gamma\{[1 + 4\pi i\sigma/kv]k\}^{-1} \times [K_0'(kR/\gamma)/K_0(kR/\gamma)]. \tag{2.32}$$

Using (2.17) and solving for $\alpha_0(k)$ gives finally

$$\alpha_0(k) = \left(\frac{-i\pi}{2}\right)\frac{H_0^{(1)\prime}(q_0(k)R) + \beta_0(k)H_0^{(1)}(q_0(k)R)}{J_0'(q_0(k)R) + \beta_0(k)J_0(q_0(k)R)}. \tag{2.33}$$

We see here again that if R is so large that $|\text{Im } q_0(k)|\, R \gg 1$, then $\alpha_0(k)$ is exponentially small, as in (2.18). On the other hand if $k \to 0$ then

$$\beta_0(k) \to \gamma^2\beta^2[q_0(k)R \ln (CkR/2\gamma)]^{-1}, \tag{2.34}$$

$$\ln C = 0.5772 \cdots ; \qquad \beta \equiv (v/c).$$

Using (2.34) in (2.33) gives the limit

$$\alpha_0(k) \underset{k\to 0}{\to} \frac{\ln (CkR/2\gamma)}{\gamma^2\beta^2} + \ln [-iCq_0(k)R/2] \tag{2.35}$$

or neglecting terms of order unity

$$\alpha_0(k) \to (1/\gamma^2\beta^2 + \tfrac{1}{2}) \ln k. \tag{2.36}$$

If the whole system of plasma plus beam is surrounded by a conducting tube of radius greater than R, then we can still use (2.33) and (2.32), with the proviso that the function K_0 in (2.32) must be replaced by some appropriate linear combination of K_0 and I_0, chosen so that the fields (2.19), (2.20) satisfy the correct boundary conditions at the inner tube surface. We will not pursue this possibility further.

[It may be worth mentioning that the time-averaged fields $\langle \mathbf{E}_0\rangle$, $\langle \mathbf{B}_0\rangle$ are obtained by setting $k = 0$ everywhere above. In this case we have

$$q_0^2(k) \to 4\pi i\sigma kv/c^2 \to 0,$$

and the boundary-value parameter $\alpha_0(k)$ can be calculated from the approximate formula (2.36) as

$$\alpha_0(k) \sim \ln k \to \infty. \tag{2.37}$$

This seems very odd, particularly since we know that $\alpha_0(k) = 0$ if the plasma is infinite. However, the infinity (2.37) is entirely spurious, for (2.14) and (2.15) give $\phi_0(\mathbf{r}, k)$ in the limit $k \to 0$ as

$$\phi_0(\mathbf{r}, k) \to \int_0^\infty r'\langle n_0(r')\rangle \ln r_> \, dr' - \alpha_0(k). \tag{2.38}$$

The constant term does not contribute to the fields, which are given by (2.10) and (2.11) for $k = 0$ as

$$\langle \mathbf{E}_0(\mathbf{r})\rangle = 0, \tag{2.39}$$

$$\langle \mathbf{B}_0(\mathbf{r})\rangle = \frac{4\pi e}{c}[\mathbf{v} \times \nabla_\perp]\phi_0(r, 0) = \frac{4\pi e}{rc}[\mathbf{v} \times \hat{\mathbf{r}}] \int_0^r r'\langle n_0(r')\rangle \, dr'. \tag{2.40}$$

This is of course just the solution we would have obtained if we had started with $f(z - vt)$ constant. Together with Ohm's law (2.3), Eq. (2.39) shows that the time-averaged plasma current is zero. This is not necessarily the most general case, but we will not consider the possibility of a constant plasma current in this article.]

3. THE PERTURBED BEAM CURRENT

We suppose now that the beam axis can move transversely by an amount $\mathbf{d}(z, t)$, in such a way that every beam particle maintains a constant distance from the moving beam axis, and keeps a constant velocity v in the z direction. Then the modified beam density will be

$$\tilde{n}(\mathbf{r}, z, t) = n(|\mathbf{r} - \mathbf{d}(z, t)|)f(z - vt). \tag{3.1}$$

A beam particle at a (fixed) distance $\boldsymbol{\varrho}$ from the beam axis moves in time dt from

$$\{\boldsymbol{\varrho} + \mathbf{d}(z, t), z\}$$

to

$$\{\boldsymbol{\varrho} + \mathbf{d}(z + v\,dt, t + dt), z + vt\},$$

and hence its velocity is

$$\tilde{\mathbf{v}}(z, t) = \mathbf{v} + [\partial/\partial t + v(\partial/\partial z)]\mathbf{d}(z, t). \tag{3.2}$$

The modified beam current is therefore

$$\tilde{\mathbf{J}}_B(\mathbf{r}, z, t) = en(|\mathbf{r} - \mathbf{d}(z, t)|)f(z - vt) \times [\mathbf{v} + [\partial/\partial t + v(\partial/\partial z)]\mathbf{d}(z, t)], \tag{3.3}$$

$$\tilde{\epsilon}_B(\mathbf{r}, z, t) = en(|\mathbf{r} - \mathbf{d}(z, t)|)f(z - vt). \tag{3.4}$$

It is straightforward to check that (3.3) and (3.4) still satisfy the continuity condition

$$\nabla \cdot \tilde{\mathbf{J}}_B + \partial \tilde{\epsilon}_B / \partial t = 0.$$

We will next assume that the beam displacement $\mathbf{d}$ is so small that

$$|\mathbf{d} \cdot \nabla n| \ll n \qquad |\dot{\mathbf{d}}| \ll v \qquad |\partial \mathbf{d}/\partial z| \ll 1. \tag{3.5}$$

We can then expand (3.3) and (3.4) to first order in $\mathbf{d}$:

$$\tilde{\mathbf{J}}_B(\mathbf{r}, z, t) = \mathbf{J}_{0B}(\mathbf{r}, z, t) + \mathbf{J}_{1B}(\mathbf{r}, z, t), \tag{3.6}$$

$$\tilde{\epsilon}_B(\mathbf{r}, z, t) = en_0(\mathbf{r}, z, t) + \epsilon_{1B}(\mathbf{r}, z, t), \tag{3.7}$$

where

$$\mathbf{J}_{1B}(\mathbf{r}, z, t) = -e\mathbf{v}(\mathbf{d}(z, t) \cdot \nabla) n(r) f(z - vt) + en(r) f(z - vt)[\partial/\partial t + v(\partial/\partial z)]\mathbf{d}(z, t), \tag{3.8}$$

$$\epsilon_{1B}(\mathbf{r}, z, t) = -e(\mathbf{d}(z, t) \cdot \nabla) n(r) f(z - vt). \tag{3.9}$$

It is very important at this point to realize that the unperturbed beam is uniform in time from the point of view of a beam particle. Hence $\mathbf{d}(z, t)$, and all quantities of first order in $\mathbf{d}$, can be written as a function of $z - vt$ times an exponential $\exp(-i\Omega t)$. We will also assume that the beam displacement is always in the same direction $\hat{\mathbf{d}}$, so that

$$\mathbf{d}(z, t) = d(z - vt) \exp(-i\Omega t)\hat{\mathbf{d}}. \tag{3.10}$$

Therefore (3.8) and (3.9) give

$$\mathbf{J}_{1B}(\mathbf{r}, z, t) = e\, d(z - vt) f(z - vt) e^{-i\Omega t} \times [-\mathbf{v}(\hat{\mathbf{d}} \cdot \nabla) - i\Omega \hat{\mathbf{d}}] n(r), \tag{3.11}$$

$$\epsilon_{1B}(\mathbf{r}, z, t) = -e\, d(z - vt) f(z - vt) e^{-i\Omega t} (\hat{\mathbf{d}} \cdot \nabla) n(r). \tag{3.12}$$

The instability frequency seen in the *beam* frame is $\gamma\Omega$.

4. THE PERTURBED FIELDS

The first-order perturbations $\mathbf{E}_1$ and $\mathbf{B}_1$ in the electric and magnetic fields are determined by Maxwell's equations

$$\nabla \times \mathbf{B}_1(\mathbf{r}, z, t) = (1/c)\dot{\mathbf{E}}_1(\mathbf{r}, z, t) + (4\pi/c)[\mathbf{J}_{1B}(\mathbf{r}, z, t) + \mathbf{J}_{1P}(\mathbf{r}, z, t)], \tag{4.1}$$

$$\nabla \times \mathbf{E}_1(\mathbf{r}, z, t) = -(1/c)\dot{\mathbf{B}}_1(\mathbf{r}, z, t), \tag{4.2}$$

$$\nabla \cdot \mathbf{B}_1(\mathbf{r}, z, t) = 0, \tag{4.3}$$

where the perturbed plasma current is

$$\mathbf{J}_{1P}(\mathbf{r}, z, t) = \sigma \mathbf{E}_1(\mathbf{r}, z, t). \tag{4.4}$$

To solve these equations we Fourier-analyze the beam displacement,

$$d(z - vt) = \int_{-\infty}^{\infty} d(k) e^{ik(z - vt)}\, dk, \tag{4.5}$$

so that

$$d(z - vt) f(z - vt) = \int_{-\infty}^{\infty} g(k) e^{ik(z - vt)}\, dk, \tag{4.6}$$

$$g(k) = \int_{-\infty}^{\infty} d(k') f(k - k')\, dk'. \tag{4.7}$$

We also write

$$\mathbf{J}_{1B}(\mathbf{r}, z, t) = e^{-i\Omega t} \int_{-\infty}^{\infty} \mathbf{J}_{1B}(\mathbf{r}, k) e^{ik(z - vt)}\, dk, \tag{4.8}$$

$$\epsilon_{1B}(\mathbf{r}, z, t) = e^{-i\Omega t} \int_{-\infty}^{\infty} \epsilon_{1B}(\mathbf{r}, k) e^{ik(z - vt)}\, dk, \tag{4.9}$$

where, according to (3.11) and (3.12),

$$\mathbf{J}_{1B}(\mathbf{r}, k) = -eg(k)[\mathbf{v}(\hat{\mathbf{d}} \cdot \nabla_\perp) + i\Omega \hat{\mathbf{d}}] n(r), \tag{4.10}$$

$$\epsilon_{1B}(\mathbf{r}, k) = -eg(k)(\hat{\mathbf{d}} \cdot \nabla_\perp) n(r). \tag{4.11}$$

The solution of Maxwell's equations (4.1)–(4.4) can now be written

$$\mathbf{E}_1(\mathbf{r}, z, t) = e^{-i\Omega t} \int_{-\infty}^{\infty} \mathbf{E}_1(\mathbf{r}, k) e^{ik(z - vt)}\, dk, \tag{4.12}$$

$$\mathbf{B}_1(\mathbf{r}, z, t) = e^{-i\Omega t} \int_{-\infty}^{\infty} \mathbf{B}_1(\mathbf{r}, k) e^{ik(z - vt)}\, dk, \tag{4.13}$$

$$\begin{aligned}\mathbf{E}_1(\mathbf{r}, k) = \frac{4\pi e g(k)}{c^2} \Bigg\{ & \left[i(\Omega + kv)\mathbf{v} - \frac{kvc^2}{\Omega + kv + 4\pi i\sigma} \nabla \right] (\hat{\mathbf{d}} \cdot \nabla) \phi_{(1)}(r, k) \\ & - \Omega \left[(\Omega + kv)\hat{\mathbf{d}} + \frac{c^2}{\Omega + kv + 4\pi i \sigma} \times (\hat{\mathbf{d}} \cdot \nabla)\nabla \right] \phi_{(2)}(r, k) \Bigg\},\end{aligned} \tag{4.14}$$

$$\mathbf{B}_1(\mathbf{r}, k) = [4\pi e g(k)/c][-(\mathbf{v} \times \nabla)(\hat{\mathbf{d}} \cdot \nabla)\phi_{(1)}(r, k) - i\Omega(\hat{\mathbf{d}} \times \nabla)\phi_{(2)}(r, k)]. \tag{4.15}$$

The symbol ∇ now denotes the three-dimensional vector

$$\nabla = \nabla_\perp + ik\hat{\mathbf{v}}. \tag{4.16}$$

The two "potentials" $\phi_{(1)}$ and $\phi_{(2)}$ both satisfy

$$[\nabla_\perp^2 + q^2(k)]\phi_{(i)}(r, k) = n(r), \tag{4.17}$$

$$q^2(k) \equiv -k^2 + (\Omega + kv)(\Omega + kv + 4\pi i\sigma)/c^2. \tag{4.18}$$

It is necessary to distinguish $\phi_{(1)}$ and $\phi_{(2)}$, because in general they are subject to different boundary conditions outside the beam.[3] The most general solutions of (4.17) that are finite at $r=0$ are

$$\phi_{(i)}(r,k)=\int_0^\infty G_{(i)}(r,r')r'n(r')\,dr', \qquad (4.19)$$

$$G_{(i)}(r,k)=-\tfrac{1}{2}i\pi J_0(q(k)r_<)H_0(q(k)r_>) - \alpha_{(i)}(k)J_0(q(k)r)J_0(q(k)r'), \qquad (4.20)$$

where i is 1 or 2, and $q(k)$ is the root of (4.18) with

$$\operatorname{Im} q(k)>0. \qquad (4.21)$$

The boundary-condition parameters $\alpha_{(i)}(k)$ are to be determined by continuing the solutions (4.20) out to infinity or to the system walls, just as we did for the unperturbed beam in Sec. 2. However, the calculation of α is now much more complicated, and we have relegated it to an Appendix. We will just mention here that if the plasma conductivity σ is constant out to a distance R which is so large that

$$R\,|\operatorname{Im} q(k)|\ll 1,$$

then the plasma is effectively infinite and

$$\alpha_{(1)}(k)\cong\alpha_{(2)}(k)\cong 0. \qquad (4.22)$$

5. FORCES ON THE BEAM

Originally the beam was kept in equilibrium by the balance between $\mathbf{E}_0$, $\mathbf{B}_0$ and pressure forces. Hence when displaced it becomes subject to forces of two different kinds: the perturbed fields $\mathbf{E}_1$, $\mathbf{B}_1$ act on the unperturbed charge and current densities ϵ_0, $\mathbf{J}_0$; and the original fields $\mathbf{E}_0$, $\mathbf{B}_0$ act on the perturbed charge and current densities $\epsilon_{1\mathrm{B}}$, $\mathbf{J}_{1\mathrm{B}}$. So the force per unit beam length is a sum of two terms

$$\mathbf{F}(z,t)=\mathbf{F}_{01}(z,t)+\mathbf{F}_{10}(z,t), \qquad (5.1)$$

$$\mathbf{F}_{01}(z,t)=\int d^2r\,en_0(r,z,t) \times\{\mathbf{E}_1(\mathbf{r},z,t)+(\mathbf{v}/c)\times\mathbf{B}_1(\mathbf{r},z,t)\}, \qquad (5.2)$$

$$\mathbf{F}_{10}(z,t)=\int d^2r\{\epsilon_{1\mathrm{B}}(\mathbf{r},z,t)\mathbf{E}_0(\mathbf{r},z,t) + (1/c)\mathbf{J}_{1\mathrm{B}}(\mathbf{r},z,t)\times\mathbf{B}_0(\mathbf{r},z,t)\}. \qquad (5.3)$$

From (2.1) and (4.12)–(4.15) we find

$$\begin{aligned}\mathbf{F}_{01}(z,t)=\ &\frac{4\pi e^2}{c^2}e^{-i\Omega t}f(z-vt)\int_{-\infty}^{\infty}dk\,g(k)e^{ik[z-vt]}\\ &\times\int d^2\mathbf{r}\,n(r)\bigg(\bigg\{i\Omega\mathbf{v}+\bigg[v^2-\frac{kvc^2}{\Omega+kv+4\pi i\sigma}\bigg]\boldsymbol{\nabla}\bigg\}\\ &\times(\hat{\mathbf{d}}\cdot\boldsymbol{\nabla})\phi_{(1)}(r,k)+\bigg\{-\Omega^2\hat{\mathbf{d}}-\frac{c^2\Omega}{\Omega+kv+4\pi i\sigma}\\ &\times(\hat{\mathbf{d}}\cdot\boldsymbol{\nabla})\boldsymbol{\nabla}\bigg\}\phi_{(2)}(r,k)\bigg),\end{aligned} \qquad (5.4)$$

while (2.10), (2.11), and (4.8)–(4.11) give

$$\begin{aligned}&\mathbf{F}_{10}(z,t)\\ &=4\pi e^2e^{-i\Omega t}d(z-vt)f(z-vt)\int_{-\infty}^{\infty}dk\,f(k)e^{ik[z-vt]}\\ &\times\int d^2r\bigg((\hat{\mathbf{d}}\cdot\boldsymbol{\nabla})n(r)\bigg\{-\bigg(1+\frac{4\pi i\sigma}{kv}\bigg)^{-1}\\ &\times(\boldsymbol{\nabla}_\perp-i\hat{\mathbf{v}}q_0^2(k)/k)+\frac{v^2}{c}\boldsymbol{\nabla}_\perp\bigg\}\phi_0(r,k)\\ &-\frac{i\Omega\mathbf{v}}{c^2}n(r)(\hat{\mathbf{d}}\cdot\boldsymbol{\nabla})\phi_0(r,k)\bigg).\end{aligned} \qquad (5.5)$$

Observe that the fields $\mathbf{E}_0$, $\mathbf{B}_0$, $\mathbf{E}_1$, $\mathbf{B}_1$ are needed here only within the beam, where σ is constant and our formulas are reliable.

We can now use the cylindrical symmetry of the beam density $n(r)$ to simplify (5.4) by making the replacement

$$\boldsymbol{\nabla}_\perp(\hat{\mathbf{d}}\cdot\boldsymbol{\nabla}_\perp)\to\tfrac{1}{2}\hat{\mathbf{d}}\nabla_\perp^2, \qquad (5.6)$$

while all terms linear in $\boldsymbol{\nabla}_\perp$ integrate to zero. We have

$$\begin{aligned}\mathbf{F}_{01}(z,t)&=4\pi e^2\hat{\mathbf{d}}e^{-i\Omega t}f(z-vt)\\ &\times\int_{-\infty}^{\infty}dk\,g(k)e^{ik[z-vt]}\int d^2\mathbf{r}\,n(r)\\ &\times\bigg(\frac{1}{2}\bigg\{\frac{v^2}{c^2}-\frac{kv}{\Omega+kv+4\pi i\sigma}\bigg\}\nabla_\perp^2\phi_{(1)}(r,k)\\ &-\bigg\{\Omega^2+\frac{1}{2}\frac{\Omega c^2}{\Omega+kv+4\pi i\sigma}\nabla_\perp^2\bigg\}\phi_{(2)}(r,k)\bigg).\end{aligned} \qquad (5.7)$$

After integrating by parts we can simplify (5.5) in the same way, and obtain

$$\begin{aligned}\mathbf{F}_{10}(z,t)&=4\pi e^2\hat{\mathbf{d}}e^{-i\Omega t}f(z-vt)d(z-vt)\\ &\times\int_{-\infty}^{\infty}dk\,f(k)e^{ik[z-vt]}\,\frac{1}{2}\int d^2\mathbf{r}\,n(r)\\ &\times\bigg[\bigg(1+\frac{4\pi i\sigma}{kv}\bigg)^{-1}-\frac{v^2}{c^2}\bigg]\nabla_\perp^2\phi_0(r,k).\end{aligned} \qquad (5.8)$$

It will prove very convenient at this point to introduce three dimensionless functions of k that completely characterize the size and shape of both the beam and the plasma channel:

$$I_0(k)\equiv q_0(k)^2\int d^2\mathbf{r}\,n(r)\phi_0(r,k)\bigg/\int d^2\mathbf{r}\,n(r)^2$$

[3] That is, there are two modes, "TE_1" and "TM_1" coupled by the discontinuity conditions at $r=R$. It is therefore necessary to use two distinct potentials in (4.14) and (4.15), though the particular separation of $\phi_{(1)}$ and $\phi_{(2)}$ used here is not unique.

$$= q_0(k)^2 \frac{\int_0^\infty r\,dr \int_0^\infty r'\,dr' n(r) G_0(r, r'; k) n(r')}{\int_0^\infty n(r)^2 r\,dr}, \quad (5.9)$$

$$I_{(i)}(k) \equiv q(k)^2 \int d^2\mathbf{r}\, n(r)\phi_{(i)}(r, k) \Big/ \int d^2\mathbf{r}\, n(r)^2$$

$$= q(k)^2 \frac{\int_0^\infty r\,dr \int_0^\infty r'\,dr' n(r) G_{(i)}(r, r'; k) n(r')}{\int_0^\infty n(r)^2 r\,dr}$$

$$(i = 1, 2). \quad (5.10)$$

Using (2.15) and (4.20) allows us to rewrite these functions as

$$I_0(k) = F(q_0^2(k)) - \alpha_0(k) q_0(k)^2 \rho(q_0(k))^2, \quad (5.11)$$

$$I_{(i)}(k) = F(q^2(k)) - \alpha_{(i)}(k) q(k^2) \rho(q(k))^2, \quad (5.12)$$

where $\rho(q)$ is the Bessel-transformed beam shape

$$\rho(q) = \int_0^\infty n(r) J_0(qr) r\,dr \Big/ \left[\int_0^\infty n(r)^2 r\,dr\right]^{\frac{1}{2}} \quad (5.13)$$

and $F(q^2)$ is the "beam form-factor"

$$F(q^2) = -\frac{i\pi}{2} q^2$$

$$\times \frac{\int_0^\infty r\,dr \int_0^\infty r'\,dr' n(r) J_0(qr_<) H_0^{(1)}(qr_>) n(r')}{\int_0^\infty n(r)^2 r\,dr}. \quad (5.14)$$

The functions $\rho(q)$ and $F(q^2)$ are discussed in detail in Appendix B.

Using these definitions and the field equations (4.17) and (2.13), we can now rewrite (5.7) and (5.8):

$$\mathbf{F}_{01}(z, t) = -4\pi e^2 \hat{\mathbf{d}} \left[\int d^2\mathbf{r}\, n(r)^2\right] e^{-i\Omega t} f(z - vt)$$

$$\times \int_{-\infty}^{\infty} A(k) g(k) e^{ik[z - vt]}\, dk \quad (5.15)$$

$$\mathbf{F}_{10}(z, t) = -4\pi e^2 \hat{\mathbf{d}} \left[\int d^2\mathbf{r}\, n(r)^2\right] e^{-i\Omega t}$$

$$\times f(z - vt) d(z - vt) \int_{-\infty}^{\infty} A_0(k) f(k) e^{ik[z - vt]}\, dk, \quad (5.16)$$

where

$$A(k) = \frac{1}{2}\left\{\frac{v^2}{c^2} - \frac{kv}{\Omega + kv + 4\pi i\sigma}\right\}$$

$$\times [I_{(1)}(k) - 1] + \left(\frac{\Omega}{q(k)c}\right)^2 I_{(2)}(k)$$

$$- \frac{1}{2} \frac{\Omega}{\Omega + kv + 4\pi i\sigma} [I_{(2)}(k) - 1] \quad (5.17)$$

$$A_0(k) = \frac{1}{2}\left\{\left(1 + \frac{4\pi i\sigma}{kv}\right)^{-1} - \frac{v^2}{c^2}\right\}[I_0(k) - 1]. \quad (5.18)$$

Putting the two forces (5.15) and (5.16) together, and using (4.5) and (4.7), gives the total force per unit beam length as

$$\mathbf{F}(z, t)$$

$$= -4\pi e^2 \hat{\mathbf{d}} \left[\int d^2\mathbf{r}\, n(r)^2\right] e^{-i\Omega t} f(z - vt) \int_{-\infty}^{\infty} dk e^{ik[z - vt]}$$

$$\times \int_{-\infty}^{\infty} dk' f(k - k') d(k') [A(k) + A_0(k - k')].$$

$$(5.19)$$

6. THE EXACT DISPERSION RELATION

Newton's second law reads here

$$\mathbf{F}(z, t) = \left(\frac{\partial}{\partial t} + v\frac{\partial}{\partial z}\right)^2 \mathbf{d}(z, t) \int d^2\mathbf{r}\, m\gamma n_0(\mathbf{r}, z, t)$$

$$= -m\gamma\Omega^2 f(z - vt) d(z - vt) e^{-i\Omega t} \hat{\mathbf{d}} \int d^2\mathbf{r}\, n(r). \quad (6.1)$$

Comparing (5.19) with (6.1) yields the dispersion relation as an integral equation for the unknown function $d(k)$,

$$\Omega^2 d(k) = \omega_T^2 \int_{-\infty}^{\infty} dk' f(k - k') d(k')$$

$$\times [A(k) + A_0(k - k')]. \quad (6.2)$$

Here ω_T^2 is the "transverse beam plasma frequency"

$$\omega_T^2 \equiv 4\pi \langle n\rangle e^2 / m\gamma, \quad (6.3)$$

$$\langle n\rangle \equiv \int d^2\mathbf{r}\, n(r)^2 \Big/ \int d^2\mathbf{r}\, n(r).$$

If we assume that the beam modulation function $f(z - vt)$ is *periodic*, with period $2\pi/p$, we may express it as a Fourier series:

$$f(z - vt) = \sum_s f_s \exp[isp(z - vt)], \quad (6.4)$$

the sum running over a finite or infinite set of positive and negative integers. We will normalize $n(r)$ so that the beam modulation function $f(z - vt)$ has average value unity, and so

$$f_0 = 1. \quad (6.5)$$

Also $f(z - vt)$ must be real, so

$$f_s^* = f_{-s}. \quad (6.6)$$

The Fourier transform of (6.4) is

$$f(k) = \sum_s f_s \delta(k - sp) \quad (6.7)$$

and the dispersion relation (6.2) becomes

$$\Omega^2 d(k) = \omega_T^2 \sum_s f_s d(k - sp)[A(k) + A_0(sp)]. \quad (6.8)$$

Equation (6.8) provides an infinite set of coupled linear equations for the unknowns $d(k)$, $d(k \pm p)$, $d(k \pm 2p)$, etc. We can define the wavenumber κ for a particular normal mode by specifying that the only nonzero values of $d(k)$ are $d(\kappa)$, $d(\kappa \pm p)$, $d(\kappa \pm 2p)$, $\cdots$. This definition only involves the value of κ mod p, so we can always take

$$0 \leq \kappa < \mathrm{p}.$$

Replacing k in (6.8) by $\kappa - sp$ gives the explicit eigenvalue equation

$$\Omega^2 d(\kappa - sp) = \omega_T^2 \sum_{s'} f_{s'-s} d(\kappa - s'p) \times [A(\kappa - sp) + A_0([s' - s]p)]. \qquad (6.9)$$

The general exact dispersion relation for $\Omega(\kappa)$ is obtained by solving the secular equation

$$\det [\Omega^2 - \omega_T^2 M(\kappa)] = 0 \qquad (6.10)$$

$$M_{s's}(\kappa) \equiv f_{[s'-s]}[A(\kappa - sp) + A_0([s' - s]p)]. \qquad (6.11)$$

This would be a formidable task without further approximations. [It may be well to remind the reader at this point that $I_{(1)}(\kappa)$, $I_{(2)}(\kappa)$, $A(\kappa)$, and $M(\kappa)$ all have an implicit Ω dependence.]

7. THE HIGH- PLASMA-CONDUCTIVITY APPROXIMATION

We will now assume that the plasma conductivity σ is so much higher than either Ω, kv, or pv, that we may neglect all terms in (5.17) and (5.18) of order $1/\sigma$. They then become

$$A(k) = \tfrac{1}{2}(v^2/c^2)[I_{(1)}(k) - 1], \qquad (7.1)$$

$$A_0(k) = -\tfrac{1}{2}(v^2/c^2)[I_0(k) - 1]. \qquad (7.2)$$

In evaluating $I_{(1)}(k)$ and $I_0(k)$ we are to use (5.11) and (5.12), with $q(k)$ and $q_0(k)$ given by the limiting formulas for $\sigma \to \infty$

$$q(k)^2 = 4\pi i\sigma(\Omega + kv)/c^2, \qquad (7.3)$$

$$q_0(k)^2 = 4\pi i\sigma kv/c^2. \qquad (7.4)$$

However, we will *not* set q and q_0 infinite within the argument of F and ρ, because we do not necessarily want to assume that σ is so large that $4\pi\sigma\ |\Omega + kv|\ a^2c^2 \gg 1$ or that $4\pi\sigma pva^2/c^2 \gg 1$. Furthermore, in evaluating $\alpha(k)$ and $\alpha_0(k)$ we shall specifically assume that R is of order $1/\sigma^{\frac{1}{2}}$, in the sense that $4\pi\sigma\ |\Omega + kv|\ R^2/c^2$ and $4\pi\sigma pvR^2/c^2$ are of order unity. This last assumption is certainly not satisfied for all interesting values of $|\Omega + kv|$ and pv, but when $|qR| \gg 1$ or $|q_0R| \gg 1$ the parameters α and α_0 will be exponentially small, and any approximation which reproduces this exponential fall-off will be adequate.

It is shown in Appendix A that the limit $\sigma \to \infty$, σR^2 constant, gives the boundary value parameter $\alpha_{(1)}(k)$ as

$$\alpha_{(1)}(k) \to -\tfrac{1}{2}i\pi H_0^{(1)}(q(k)R)/J_0(q(k)R). \qquad (7.5)$$

The function $\beta_0(k)$ appearing in $\alpha_0(k)$ is given in this limit by (2.32) as

$$\beta_0(k) \to \frac{q_0(k)\gamma^2 v}{4\pi i\sigma kR \ln (CkR/2\gamma)} = \frac{\gamma^2\beta^2}{q_0(k)R \ln (CkR/2\gamma)},$$

$$\beta = v/c; \qquad \ln C = 0.5772 \cdots \qquad (7.6)$$

Using (7.6) in (2.33) gives

$$\alpha_0(k) = \left(-\frac{i\pi}{2}\right) \times \frac{q_0(k)R \ln (CkR/2\gamma) H_0^{(1)\prime}(q(k)R) + \gamma^2\beta^2 H_0^{(1)}(q_0(k)R)}{q_0(k)R \ln (CkR/2\gamma) J_0'(q(k)R) + \gamma^2\beta^2 J_0(q_0(k)R)}. \qquad (7.7)$$

We can therefore distinguish two simple cases

(a) *Extreme Relativistic* (ER): $\gamma \gg 1$, $\beta \cong 1$,

$$\alpha_0(k) \cong -\tfrac{1}{2}i\pi[H_0^{(1)}(q_0(k)R)/J_0(q_0(k)R)]. \qquad (7.8)$$

(b) *Nonrelativistic* (NR): $\gamma \cong 1$, $\beta \ll 1$,

$$\alpha_0(k) \cong -\tfrac{1}{2}i\pi[H_0^{(1)\prime}(q_0(k)R)/J_0'(q_0(k)R)]. \qquad (7.9)$$

The NR approximation cannot be used for $k \to 0$, in which case we have already seen that

$$\alpha_0(k) \to (1/\gamma^2\beta^2 + \tfrac{1}{2}) \ln k. \qquad (2.36)$$

However the ER approximation reproduces this behavior, except for the term $1/\gamma^2\beta^2$. For $|q_0(k)|\,R \gg 1$ both the ER and NR approximations give the correct factor

$$\exp [-2 \operatorname{Im} q_0(k)R], \qquad (7.10)$$

which in this limit makes $\alpha_0(k)$ negligible.

Using (7.5) and (7.3) in (5.12), we see that $I_{(1)}(k)$ is now a function only of the frequency $\Omega + kv$:

$$I_{(1)}(k) = \mathfrak{F}(\Omega + kv), \qquad (7.11)$$

where

$$\mathfrak{F}(\omega) \equiv F(q^2) - \tfrac{1}{2}i\pi[H_0^{(1)}(qR)/J_0(qR)]q^2\rho(q)^2, \qquad (7.12)$$

$$q^2(\omega) = 4\pi i\sigma\omega/c^2; \qquad \operatorname{Im} q(\omega) > 0. \qquad (7.13)$$

Also, (5.11), (7.4), and (7.8) show that in the ER limit $I_0(k)$ becomes equal to the *same* function of kv:

$$I_0(k) = \mathfrak{F}(kv) \qquad \text{(ER)}. \qquad (7.14)$$

In the NR limit, $I_0(k)$ is a slightly different function of kv:

$$I_0(k) = \mathfrak{F}_0(kv) \qquad \text{(NR)}, \qquad (7.15)$$

$$\mathfrak{F}_0(\omega) \equiv F(q^2) - \tfrac{1}{2}i\pi[H_0^{(1)\prime}(qR)/J_0'(qR)]q^2\rho(q)^2. \quad (7.16)$$

Of course $I_{(1)}(k)$ has had an explicit Ω dependence all along, which we have now made explicit. The form factor $F(q^2)$ and Bessel transform $\rho(q)$ are discussed in Appendix B.

We will now use these results in the exact dispersion relation derived in the last section:

$$\Omega^2 d(\kappa - sp) = \omega_T^2 \sum_{s'} f_{s'-s} d(\kappa - s'p) \times [A(\kappa - sp) + A_0([s' - s]p)]. \quad (6.9)$$

We have defined $f_0 = 1$, and will show in the next section that

$$A_0(0) = \tfrac{1}{2}\beta^2.$$

Hence (6.9) may be rewritten to isolate the term $s' = s$:

$$[\Omega^2 - \omega_T^2(A(\kappa - sp) + \tfrac{1}{2}\beta^2)]d(\kappa - sp) = \omega_T^2 \sum_{s'\neq s,0} f_{s'-s} d(\kappa - s'p) \times [A(\kappa - sp) + A_0([s' - s]p)]. \quad (7.17)$$

Using (7.1), (7.2), (7.11), and (7.14) in (7.17) gives the dispersion relation for the high σ extreme relativistic case in its final form:

$$[\Omega^2 - \omega_\beta^2\mathfrak{F}(\Omega + \kappa v - spv)]d(\kappa - sp) = \omega_\beta^2 \sum_{s'\neq s,0} f_{s'-s} d(\kappa - s'p) \times [\mathfrak{F}(\Omega + \kappa v - spv) - \mathfrak{F}([s' - s]pv)], \quad (7.18)$$

where ω_β is the betatron frequency

$$\omega_\beta^2 = \tfrac{1}{2}\beta^2\omega_T^2 = 2\pi\langle n\rangle e^2\beta^2/m\gamma. \quad (7.19)$$

In the nonrelativistic case, the last $\mathfrak{F}$ in (7.18) must be replaced by $\mathfrak{F}_0$.

8. THE UNMODULATED BEAM

Before trying to solve the general dispersion relation (6.8) or (7.18), we will first examine the important special case of an unmodulated beam. Here $f(z - vt)$ is constant, and consequently all f_s vanish, except for $f_0 = 1$. The exact dispersion relation (6.8) becomes

$$\Omega^2 = \omega_T^2[A(k) + A_0(0)]. \quad (8.1)$$

We will now show that

$$A_0(0) = v^2/2c^2. \quad (8.2)$$

In the limit $k \to 0$, Eq. (5.18) gives

$$A_0(k) \to -(v^2/2c^2)[I_0(k) - 1]. \quad (8.3)$$

Also, as $k \to 0$, Eqs. (2.12), (2.36), (B9), and (B4) give

$$q_0(k)^2 \to 4\pi i\sigma kv/c^2, \quad (8.4)$$

$$\alpha_0(k) \to (1/\gamma^2\beta^2 + \tfrac{1}{2}) \ln k, \quad (8.5)$$

$$F(q_0^2(k)) \to -\tfrac{1}{4}(4\pi i\sigma kva^2/c^2) \ln k, \quad (8.6)$$

$$\rho(q_0(k)) \to a/\sqrt{2}. \quad (8.7)$$

Therefore (5.11) shows that, as $k \to 0$,

$$I_0(k) \to (4\pi i\sigma kva^2/c^2)[-\tfrac{1}{2} - 1/\gamma^2\beta^2] \ln k; \quad (8.8)$$

so

$$I_0(0) = 0, \quad (8.9)$$

and (8.3) gives (8.2).

The exact dispersion relation for an unmodulated beam is therefore

$$\Omega^2 = \omega_T^2[A(k) + (v^2/2c^2)], \quad (8.10)$$

or, using (5.17),

$$\Omega^2 = \omega_T^2\left\{\frac{1}{2}\left(\frac{v}{c}\right)^2 I_{(1)}(k) + \left[\frac{\Omega}{q(k)c}\right]^2 I_{(2)}(k) - \frac{1}{2}\frac{kv}{\Omega + kv + 4\pi i\sigma}[I_{(1)}(k) - 1] - \frac{1}{2}\frac{\Omega}{\Omega + kv + 4\pi i\sigma}[I_{(2)}(k) - 1]\right\}. \quad (8.11)$$

If we now make the high-σ approximation described in the last section [either directly in (8.11), or using (8.10) and (7.1)] we find that the dispersion relation simplifies to

$$\Omega^2 = \omega_\beta^2\mathfrak{F}(\Omega + kv), \quad (8.12)$$

with $\mathfrak{F}(\omega)$ given by (7.12) and (7.13), and the betatron frequency ω_β given by (7.19).

We now distinguish three limiting cases of interest, characterized by the relative magnitudes of the transverse wavelength $|q(\Omega + kv)|^{-1}$, the plasma channel radius R, and the mean beam radius:

$$a \equiv \sqrt{2}\int_0^\infty n(r)r\,dr \bigg/ \left[\int_0^\infty n(r)^2 r\,dr\right]^{\frac{1}{2}}. \quad (8.13)$$

$$(1) \quad |q(\Omega + kv)|\,a \ll 1; \quad |q(\Omega + kv)|\,R \ll 1 \quad (8.14)$$

In this case the beam form factor $F(q)$ is given by (B9) and the Bessel transform $\rho(q)$ is given by (B4), so (7.12) gives

$$\mathfrak{F}(\omega) = \tfrac{1}{4}q(\omega)^2a^2 \ln (4r_0^2/C^2R^2), \quad (8.15)$$

where

$$C = 1.78 \cdots,$$

and r_0 is another effective radius given by (B10)–(B12) for any $n(r)$. For example:

$$n(r) = n \exp(-r^2/2b^2), \quad a = 2b, \quad r_0 = 0.90a;$$

$$n(r) = n \exp(-r/b), \quad a = 2.43b, \quad r_0 = 0.41a;$$

$$n(r) = \begin{cases} n & (r < a), \\ 0 & (r > a), \end{cases} \quad r_0 = 0.69a.$$

The dispersion relation (8.12) is now a quadratic equation

$$\Omega^2 = i\omega_0(\Omega + kv), \tag{8.16}$$

where

$$\omega_0 = (\omega_\beta^2 \pi\sigma a^2/c^2) \ln(C^2R^2/4r_0^2) > 0. \tag{8.17}$$

The roots are

$$\Omega_\pm = (-\tfrac{1}{2}i\omega_0)[1 \pm [1 + 4ikv/\omega_0]^{\frac{1}{2}}]. \tag{8.18}$$

The root $\Omega_+(k)$ corresponds to a decaying mode for all k, with limiting behavior

$$\begin{aligned} \Omega_+(k) &\to -i\omega_0 && (|kv| \ll \omega_0), \\ &\to (1 - i)(|kv|\ \tfrac{1}{2}\omega_0)^{\frac{1}{2}} && (|kv| \gg \omega_0). \end{aligned} \tag{8.19}$$

The other root $\Omega_-(k)$ corresponds to a growing mode, with limiting behavior

$$\begin{aligned} \Omega_-(k) &\to -kv + i(kv)^2\omega_0 && (|kv| \ll \omega_0), \\ &\to (-1 + i)(|kv\ \tfrac{1}{2}\omega_0)^{\frac{1}{2}} && (|kv| \gg \omega_0). \end{aligned} \tag{8.20}$$

The fastest growing mode evidently occurs for k as large as possible. If $|kv| \gg \omega_0$ then

$$q^2(\Omega + kv) \to 4\pi\sigma kv/c^2; \tag{8.20}$$

so (8.14) sets an upper limit on kv and hence on the growth rate.

$$(2) \quad |q(\Omega + kv)|\, a \ll 1, \quad |q(\Omega + kv)|\, R \gg 1 \tag{8.21}$$

The beam form factor $F(q^2)$ is again given by (B9), but now the second "boundary-value" term in (7.12) is exponentially small [because Im $q > 0$]; so (7.12) gives

$$\mathfrak{F}(\omega) = \tfrac{1}{4}q(\omega)^2a^2 \ln[-q(\omega)^2r_0^2]. \tag{8.22}$$

The dispersion relation (8.12) is now rather complicated

$$\Omega^2 = (i\omega_\beta^2\pi\sigma a^2/c^2)(\Omega + kv) \times \ln[-4\pi i\sigma(\Omega + kv)r_0^2/c^2]. \tag{8.23}$$

However, for very small $|qa|$, the logarithm becomes essentially a real negative constant, and we find ourselves back in Case 1.

$$(3) \quad |q(\Omega + kv)a| \gg 1, \quad |q(\Omega + kv)R| \gg 1 \tag{8.24}$$

Now the beam form factor $F(q^2)$ is given by (B13) as unity, while the second term in (7.12) is again exponentially small. Hence (7.12) is now just

$$\mathfrak{F}(\omega) = 1, \tag{8.25}$$

and the dispersion relation (8.12) is

$$\Omega = \pm\omega_\beta. \tag{8.26}$$

There is no growing mode, but only a stable oscillation at the betatron frequency.

It may be illuminating to compare these results with the groundbreaking work of Rosenbluth.[1] In our notation, his result for the case $|qa| \ll 1$ is

$$\Omega^2 = \omega_\beta^2 \mathcal{G}(4\pi i\sigma(\Omega + kv)/c^2) \tag{8.27}$$

where $\mathcal{G}(q^2)$ is defined by

$$\mathcal{G}(q^2) = q^2 \int d^2r n(r)\psi(r) \Big/ \int d^2r n(r)^2, \tag{8.28}$$

$$\nabla_\perp^2 \psi(r) = n(r). \tag{8.29}$$

At first sight $\mathcal{G}(q^2)$ looks like a reasonable approximate version of our formula (5.10) for $I_{(1)}$, since $\psi(r)$ is just what $\phi_{(1)}(r, k)$ would be if we dropped the $q^2(k)$ term in the differential equation (4.17) for $\phi_{(1)}$. In physical terms, Rosenbluth's result differs from ours by the neglect of the effect of plasma currents on the perturbed fields.

But, as Rosenbluth recognized, this leads to mathematical nonsense. For example, if we calculate (8.28) for a square beam of radius a we find that

$$\mathcal{G}(q^2) = -\frac{q^2a^2}{2}\left(\frac{1}{4} + \int_a^\infty \frac{dr}{r}\right). \tag{8.30}$$

Rosenbluth guessed that the integral should be cut off at $r = r_C \approx |q|^{-1}$, giving

$$\mathcal{G}(q^2) = -\tfrac{1}{2}(q^2a^2)[\tfrac{1}{4} + \ln(r_C/a)]. \tag{8.31}$$

In contrast, our answer for $I_{(1)}$ is always perfectly finite.[2] In fact, a comparison with (8.22) and (B30) shows that our answer agrees with Rosenbluth's for $|qa| \ll 1$ and $|qR| \gg 1$ if we take the cutoff r_C as

$$r_C = (2i/Cq)(C = 1.78 \cdots). \tag{8.32}$$

In this case, Rosenbluth's guess $r_C \approx |q|^{-1}$ is almost perfect in magnitude but wrong in phase. On the other hand, if $|qa| \ll 1$ and $|qR| \ll 1$, then our result (8.15) agrees with Rosenbluth's if we take the cutoff r_C as the plasma channel radius,

$$r_C = R. \tag{8.33}$$

This is entirely reasonable, since with $|qR| \ll 1$ the limit on the magnetic field energy (Rosenbluth's W) is set by the finite channel size, rather than by the induced plasma currents.

9. LONG WAVELENGTHS

Let us now return to the general dispersion relation in the approximate form (7.18):

$$[\Omega^2 - \omega_\beta^2 \mathfrak{F}(\Omega + \kappa v - spv)]d(\kappa - sp) = \omega_\beta^2 \sum_{s' \neq s} f_{s'-s}[\mathfrak{F}(\Omega + \kappa v - spv) - \mathfrak{F}([s' - s]pv)]d(\kappa - s'p) \quad \text{(ER)}. \tag{9.1}$$

For $s = 0$ this gives

$$[\Omega^2 - \omega_\beta^2 \mathfrak{F}(\Omega + \kappa v)]d(\kappa) = \omega_\beta^2 \sum_{s \neq 0} f_s[\mathfrak{F}(\Omega + \kappa v) - \mathfrak{F}(spv)]d(\kappa - sp). \tag{9.2}$$

For $s \neq 0$, (9.1) can be rewritten to isolate the term with $s' = 0$:

$$[\Omega^2 - \omega_\beta^2 \mathfrak{F}(\Omega + \kappa v - spv)]d(\kappa - sp) = \omega_\beta^2 f_s^*[\mathfrak{F}(\Omega + \kappa v - spv) - \mathfrak{F}(-spv)]d(\kappa), + \omega_\beta^2 \sum_{s' \neq 0, s} f_{s'-s}[\mathfrak{F}(\Omega + \kappa v - spv) - \mathfrak{F}([s' - s]pv)]d(\kappa - s'p). \tag{9.3}$$

Equation (9.3) may be regarded as an inhomogeneous linear equation for the quantities $d(\kappa - sp)/d(\kappa)$ with $s \neq 0$. Using its solution in (9.2) will give the desired dispersion relation.

We will now specialize, and look for a mode with Ω and κv sufficiently small so that

$$4\pi\sigma a^2 \, |\Omega + \kappa v|/c^2 \ll 1, \tag{9.4}$$

$$|\Omega + \kappa v| \ll pv, \tag{9.5}$$

$$|\Omega^2| \ll \omega_\beta^2. \tag{9.6}$$

These conditions are imposed partly because of our experience in Sec. 8 with the unmodulated beam, but mostly because they are needed to put the dispersion relation into manageable form. Of course we will eventually have to check back to make sure that (9.4)–(9.6) are consistent with the dispersion relation.

We showed in Sec. 8 that assumption (9.4) allows us to approximate $\mathfrak{F}(\Omega + \kappa v)$ as

$$\mathfrak{F}(\Omega + \kappa v) \cong -\tfrac{1}{4}q(\Omega + \kappa v)^2 a^2 L = -i\tau(\Omega + \kappa v), \tag{9.7}$$

where

$$\tau = \pi\sigma a^2 L/c^2, \tag{9.8}$$

and L is a logarithm

$$L = \ln\left(\frac{c^2R^2}{4r_0^2}\right), \quad \frac{4\pi\sigma R^2 \, |\Omega + \kappa v|}{c^2} \gg 1, \tag{9.9}$$

$$L = -\ln\left[-\frac{4\pi i\sigma r_0^2}{c^2}(\Omega + \kappa v)\right], \quad \frac{4\pi\sigma R^2 \, |\Omega + \kappa v|}{c^2} \ll 1 \tag{9.10}$$

[see (8.15) or (8.22), and (8.4)]. Also, assumption (9.5) allows us to approximate $\mathfrak{F}(\Omega + \kappa v - spv)$, for $s \neq 0$, as

$$\mathfrak{F}(\Omega + \kappa v - spv) \cong \mathfrak{F}(-spv), \tag{9.11}$$

and also

$$\mathfrak{F}(\Omega + \kappa v - spv) - \mathfrak{F}(-spv) \cong (\Omega + \kappa v)\mathfrak{F}'(-spv). \tag{9.12}$$

Finally we note that all $\mathfrak{F}(-spv)$ for $s \neq 0$ are roughtly of order unity [unless $4\pi\sigma pva^2/c^2 \ll 1$, which is not expected to be the case]. Hence (9.6) and (9.11) give

$$|\Omega^2| \ll \omega_\beta^2 \, |\mathfrak{F}(\Omega + \kappa v - spv)| \quad (\text{all } s \neq 0). \tag{9.13}$$

If we now use (9.13), (9.12), and (9.11) in (9.3), we find that for all $s \neq 0$

$$-\mathfrak{F}(-spv)d(\kappa - sp) = (\Omega + \kappa v)f_s^*\mathfrak{F}'(-spv)d(\kappa) + \sum_{s' \neq 0, s} f_{s'-s} \times [\mathfrak{F}(-spv) - \mathfrak{F}([s' - s]pv)]d(\kappa - s'p). \tag{9.14}$$

Thus $d(\kappa - sp)$ has a very simple dependence on $\Omega + \kappa v$:

$$d(\kappa - sp) = (4\pi i\sigma a^2/c^2)(\Omega + \kappa v)u_s d(\kappa), \tag{9.15}$$

where u_s is a dimensionless constant coefficient defined by

$$-\mathfrak{F}(-spv)u_s = \frac{c^2}{4\pi i\sigma a^2} f_s^*\mathfrak{F}'(-spv) + \sum_{s' \neq 0, s} f_{s'-s}[\mathfrak{F}(-spv) - \mathfrak{F}([s' - s]pv)]u_{s'}. \tag{9.16}$$

It follows from (9.4) and (9.5) that

$$|\mathfrak{F}(\Omega + \kappa v)| \ll |\mathfrak{F}_0(spv)| \qquad (\text{all } s \neq 0). \tag{9.17}$$

Using (9.17) and (9.7) in (9.2) gives

$$[\Omega^2 + i\omega_0(\Omega + \kappa v)]d(\kappa) = -\omega_\beta^2 \sum_{s \neq 0} f_s\mathfrak{F}(spv)d(\kappa - spv), \tag{9.18}$$

where ω_0 is the characteristic frequency introduced in Sec. 8 for the unmodulated beam

$$\omega_0 = \tau\omega_\beta^2 = \pi\sigma a^2 L\omega_\beta^2/c^2. \tag{9.19}$$

The dispersion relation now follows immediately from (9.18) and (9.15):

$$\Omega^2 = i\omega_1(\Omega + \kappa v) \quad \text{(ER)}, \tag{9.20}$$

where ω_1 is the frequency

$$\omega_1 = \omega_0 + \frac{4\pi\sigma a^2\omega_\beta^2}{c^2} \sum_{s\neq 0} f_s u_s \mathfrak{F}(spv). \tag{9.21}$$

So we see that the dispersion relation in the high σ, extreme relativistic, low-frequency case is precisely the same in form as for the unmodulated beam.[4] Furthermore, the conditions (9.4)–(9.6) are evidently consistent with the dispersion relation, since (9.20) allows us to take Ω^2 and $\Omega + \kappa v$ as small as we like.

Matters are rather more complicated in the nonrelativistic case. The functions $\mathfrak{F}(\pm\, spv)$ in (9.2) and (9.3) must be replaced by $\mathfrak{F}_0(\pm\, spv)$. Making assumptions (9.4)–(9.6) then yields, in place of (9.14), for $s \neq 0$,

$$\begin{aligned} -\mathfrak{F}(-spv)d(\kappa - sp) &= f_s^*[\mathfrak{F}(-spv) - \mathfrak{F}_0(-spv) + (\Omega + \kappa v)\mathfrak{F}'(-spv)]d(\kappa) \\ &+ \sum_{s'\neq 0,s} f_{s'-s}[\mathfrak{F}(-spv) - \mathfrak{F}_0([s'-s]pv)]d(\kappa - s'p). \end{aligned} \tag{9.22}$$

We see now that $d(\kappa - sp)$ takes the form

$$d(\kappa - sp) = [v_s + (4\pi i\sigma a^2/c^2)(\Omega + \kappa v)u_s]d(\kappa), \tag{9.23}$$

where v_s is a second dimensionless coefficient

$$\begin{aligned} -\mathfrak{F}(-spv)v_s &= f_s^*[\mathfrak{F}(-spv) - \mathfrak{F}_0(-spv)] \\ &+ \sum_{s'\neq 0,s} f_{s'-s}[\mathfrak{F}(-spv) - \mathfrak{F}_0([s'-s]pv)]v_{s'}. \end{aligned} \tag{9.24}$$

Also, $u_{s'}$ is given by (9.16), with the last $\mathfrak{F}$ replaced by $\mathfrak{F}_0$, Using (9.23) in (9.18) [with $\mathfrak{F}(spv)$ replaced by $\mathfrak{F}_0(spv)$] gives the dispersion relation

$$\Omega^2 = -i\omega_1(\Omega + \kappa v) - \omega_R^2 \quad \text{(NR)}, \tag{9.25}$$

where ω_1 is given by (9.21) (replacing $\mathfrak{F}$ by $\mathfrak{F}_0$) and ω_R is a second characteristic frequency, given by

$$\omega_R^2 = \omega_\beta^2 \sum_{s\neq 0} f_s v_s \mathfrak{F}_0(spv). \tag{9.26}$$

Whether or not (9.25) is consistent with the assumptions (9.4) or (9.6) depends on how small ω_R is, a point that cannot be decided without detailed computation.

It should be mentioned that the ER and NR cases yield the same dispersion relation if pv and/or R is large enough to give

$$4\pi\sigma pvR^2/c^2 \gg 1. \tag{9.27}$$

In this case the boundary-value parameters α_0 and $\alpha_{(1)}$ are negligible, since they contain the factor

$$\exp\,[-2(4\pi\sigma R^2pv/c^2)^{\frac{1}{2}}]. \tag{9.28}$$

We therefore have

$$\mathfrak{F}(\omega) = \mathfrak{F}_0(\omega) = F(q^2(\omega)); \tag{9.29}$$

so (9.20) is the correct dispersion relation, with (9.29) used for $\mathfrak{F}(\omega)$ in (9.21) and (9.16).

We do not go into details here concerning the solution of Eq. (9.16) for u_s. However, if the modulation coefficients f_s for $s \neq 0$ are sufficiently small, then (9.16) may evidently be solved by iteration

$$u_s = -(c^2/4\pi i\sigma a^2)f_s^*\mathfrak{F}'(-spv)/\mathfrak{F}(-spv) + O(f^3). \tag{9.30}$$

Using (9.30) in (9.21) then gives

$$\omega_1 \cong \omega_0 + i\omega_\beta^2 \sum |f_s|^2\, \mathfrak{F}'(-spv)/\mathfrak{F}(-spv). \tag{9.31}$$

Preliminary calculations at Stanford Research Institute indicate that (9.30) may be a quite reasonable approximation even for strong modulation, and that the second term in (9.31) is a small real correction to ω_0.

10. FAST CHOPPING

As an example, let us now consider the case of large modulation frequency pv, with

$$\epsilon \equiv c^2/4\pi\sigma pvb^2 \ll 1, \tag{10.1}$$

where b is the effective beam radius defined by

$$b^2 = \int_0^\infty n(r)^2 r\, dr \bigg/ \int_0^\infty \left[\frac{dn(r)}{dr}\right]^2 r\, dr. \tag{10.2}$$

This assumption implies the validity of (9.27), so the dispersion relation is given here by (9.20), with ω_1 given by (9.21), (9.16), and (9.29).

Assumption (10.1) also lets us evaluate $F(q_s^2)$ and $F'(q_s^2)$ by employing the asymptotic limit (B13),

$$F(q^2) \cong 1 + 1/q^2b^2 \qquad (|qb|^2 \gg 1). \tag{10.3}$$

Letting $\gamma_s = -(4\pi\sigma pvab/c^2)^2u_s$, (9.21) and (9.16) now become

$$\omega_1 = \omega_0 - \frac{\epsilon\omega_\beta^2}{(pv)^2} \sum_{s\neq 0} f_s\gamma_s, \tag{10.4}$$

$$\gamma_s = (f_s^*/s^2) - i\epsilon \sum_{s'\neq 0,s} f_{s'-s}\left(\frac{1}{s} - \frac{1}{s-s'}\right)\gamma_{s'}. \tag{10.5}$$

Clearly, (10.1) lets us solve (10.5) by iteration,

$$\begin{aligned} \gamma_s &= (f_s^*/s^2) \\ &+ i\epsilon \sum_{s'\neq 0,s} (f_{s'-s}f_s^*/ss'(s-s')) + \cdots . \end{aligned} \tag{10.6}$$

[4] A. Sessler has shown that the second term in ω_1 is real, so that modulating the beam can, at most, affect the scale of the growth rate.

Using only the first term in (10.4) gives

$$\omega_1 = \omega_0 - \frac{\epsilon\omega_\beta^2}{(pv)^2}\sum_{s\neq 0}\frac{|f_s|^2}{s^2} + O(\epsilon^2)$$

$$= \omega_0\left[1 - \frac{4}{L}\left(\frac{b}{a}\right)^2\epsilon^2\sum_{s\neq 0}\frac{|f_s|^2}{s^2}\right]. \qquad (10.8)$$

Hence the condition on the chopping rate for the dispersion relation to be unaffected by the modulation is

$$\epsilon \ll \left[\frac{4}{L}\left(\frac{b}{a}\right)^2\sum_{s\neq 0}\frac{|f_s|^2}{s^2}\right]^{-\frac{1}{2}}, \qquad (10.9)$$

or explicitly

$$pv \gg \frac{c^2}{(2\pi\sigma ba)L^{\frac{1}{2}}}\left[\sum_{s\neq 0}\frac{|f_s|^2}{s^2}\right]^{+\frac{1}{2}}. \qquad (10.10)$$

This condition can be rather stringent if the modulation is either very strong (i.e., large f_s) or very complicated (i.e., many f_s).

ACKNOWLEDGMENTS

It is a pleasure to thank K. M. Watson for his advice throughout this work. I am also grateful to H. Chang, N. Christofilos, H. Lewis, M. Rosenbluth, and A. Sessler, for their valuable comments.

APPENDIX A. BOUNDARY CONDITIONS ON THE PERTURBED FIELDS

We suppose that the plasma conductivity σ is constant out to some plasma radius R, and then drops sharply to zero. For $r < R$, the fields are given by (4.14)–(4.21). It is our task here to find the parameters $\alpha_{(1)}(k)$ and $\alpha_{(2)}(k)$ appearing in (4.20), by connecting this interior solution to an appropriate solution in the vacuum $r > R$, just as we did for the unperturbed beam in Sec. 2.

For $r < R$, but outside the beam, the fields are determined by (4.14) and (4.15), with $\phi_{(1)}$ and $\phi_{(2)}$ given by (4.19) as

$$\phi_{(i)}(r,k) = \left[-\frac{i\pi}{2}H_0^{(1)}(q(k)r) - \alpha_{(i)}(k)J_0(q(k)r)\right]\int_0^\infty J_0(q(k)r')n(r')r'\,dr'. \qquad (A1)$$

For $r > R$ the fields are again determined by (4.14) and (4.15), but with σ now set equal to zero, and with $\phi_{(1)}$ and $\phi_{(2)}$ proportional to the exponentially decaying solution of Bessel's equation:

$$\phi_{(i)}(r,k) = \xi_{(i)}(k)H_0^{(1)}(h(k)r), \qquad (A2)$$

$$h(k)^2 \equiv -k^2 + (\Omega + kv)^2/c^2, \quad \operatorname{Im} h(k) > 0. \qquad (A3)$$

The conditions for connecting (A1) with (A2) are again provided by the discontinuity relations (2.23)–(2.28). The relations $\Delta B_z = 0$, $\Delta B_r = 0$, $\Delta B_\theta = 0$, and $\Delta E_z = 0$ give, respectively,

$$\Delta\phi'_{(2)} = 0, \qquad (A4)$$

$$\Delta[\phi'_{(1)} - i(k\Omega/v)\phi_{(2)}R] = 0, \qquad (A5)$$

$$\Delta[v\phi'_{(1)}/R + ik\Omega\phi_{(2)} + vq^2(k)\phi_{(1)}] = 0, \qquad (A6)$$

and

$$\Delta\left[\frac{q^2(k)\phi'_{(1)} - (k\Omega/v)\phi'_{(2)}}{\Omega + kv + 4\pi i\sigma}\right] = 0. \qquad (A7)$$

The relation $\Delta E_\theta = 0$ gives a condition which can be deduced from (A5) and (A7), while $\Delta(\Omega + kv + 4\pi i\sigma)E_r = 0$ gives a condition which follows from (A4) and (A6). In all of these conditions, σ and q^2 are to be taken, respectively, as 0 and $+h^2$ for r just outside R. Condition (A6) can be rewritten with the aid of (A5) as a condition on $\phi_{(1)}$ alone:

$$\Delta[2\phi'_{(1)}/R + q^2(k)\phi_{(1)}] = 0. \qquad (A6')$$

In order to solve for $\alpha_{(1)}$ and $\alpha_{(2)}$ it will be convenient to introduce coefficients

$$a_1(k) = 2q(k)J_0'(q(k)R) + q(k)^2J_0(q(k)R)R, \qquad (A8)$$

$$a_2(k) = q(k)J_0'(q(k)R), \qquad (A9)$$

$$a_3(k) = (k\Omega/v)J_0(q(k)R)R, \qquad (A10)$$

$$a_4(k) = [q(k)^3/(\Omega + kv + 4\pi i\sigma)]J_0'(q(k)R), \qquad (A11)$$

$$a_5(k) = -\left(\frac{k\Omega}{v}\right)\frac{q(k)}{\Omega + kv + 4\pi i\sigma}J_0'(q(k)R). \qquad (A12)$$

We will also define another five coefficients b_1, b_2, b_3, b_4, b_5 by the same equations, making the replacement:

$$J_0 \to -\tfrac{1}{2}(i\pi)H_0^{(1)}. \qquad (A13)$$

And we will define a further five coefficients c_1, c_2, c_3, c_4, c_5 by the same equations (A8)–(A12), but now making the replacement

$$J_0 \to H_0^{(1)}, \quad q(k) \to h(k), \quad \sigma \to 0. \qquad (A14)$$

Then (A6′) gives the normalization of $\phi_{(1)}$ outside the plasma by

$$\xi_{(1)}(k) = \left(\frac{b_1(k) - a_1(k)\alpha_{(1)}(k)}{c_1(k)}\right)\times\int_0^\infty J_0(q(k)r')n(r')r'\,dr', \qquad (A15)$$

and (A4) gives the normalization of $\phi_{(2)}$ outside the plasma channel as

$$\xi_{(2)}(k) = \left(\frac{b_2(k) - a_2(k)\alpha_{(2)}(k)}{c_2(k)}\right) \times \int_0^\infty J_0(q(k)r')n(r')r'\,dr'. \quad \text{(A16)}$$

Using these definitions and (A15) and (A16) now allows us to write (A5) and (A7) as two linear inhomogeneous equations in $\alpha_{(1)}$ and $\alpha_{(2)}$:

$$b_2 - \alpha_{(1)}a_2 + b_3 - \alpha_{(2)}a_3 = \left[\frac{b_1 - a_1\alpha_{(1)}}{c_1}\right]c_2 + \left[\frac{b_2 - a_2\alpha_{(2)}}{c_2}\right]c_3$$

$$b_4 - \alpha_{(1)}a_4 + b_5 - \alpha_{(2)}a_5 = \left[\frac{b_1 - a_1\alpha_{(1)}}{c_1}\right]c_4 + \left[\frac{b_2 - a_2\alpha_{(2)}}{c_2}\right]c_5.$$

[We are omitting the argument k because of fatigue.] The solution for $\alpha_{(1)}$ and $\alpha_{(2)}$ is thus

$$\alpha_{(1)} = \frac{\left[b_2 - \frac{b_1c_2}{c_1} + b_3 - \frac{b_2c_3}{c_2}\right]\left[a_5 - \frac{a_2c_5}{c_2}\right] - \left[b_4 - \frac{b_1c_4}{c_1} + b_5 - \frac{b_2c_5}{c_2}\right]\left[a_3 - \frac{a_2c_3}{c_2}\right]}{\left[a_2 - \frac{a_1c_2}{c_1}\right]\left[a_5 - \frac{a_2c_5}{c_2}\right] - \left[a_4 - \frac{a_1c_4}{c_1}\right]\left[a_3 - \frac{a_2c_3}{c_2}\right]}, \quad \text{(A16)}$$

$$\alpha_{(2)} = \frac{\left[b_4 - \frac{b_1c_4}{c_1} + b_5 - \frac{b_2c_5}{c_2}\right]\left[a_2 - \frac{a_1c_2}{c_1}\right] - \left[b_2 - \frac{b_1c_2}{c_1} + b_3 - \frac{b_2c_3}{c_2}\right]\left[a_5 - \frac{a_2c_5}{c_2}\right]}{\left[a_2 - \frac{a_1c_2}{c_1}\right]\left[a_5 - \frac{a_2c_5}{c_2}\right] - \left[a_4 - \frac{a_1c_4}{c_1}\right]\left[a_3 - \frac{a_2c_3}{c_2}\right]}. \quad \text{(A17)}$$

We can see immediately that $\alpha_{(1)}$ and $\alpha_{(2)}$ will be exponentially small if

$$\operatorname{Im} q(k)R \gg 1, \quad \text{(A18)}$$

for then all the a's will be of order exp $[\operatorname{Im} q(k)R]$, while all the b's will be of order exp $[-\operatorname{Im} q(k)R]$, and hence $\alpha_{(1)}$ and $\alpha_{(2)}$ will be of order exp $(-2\operatorname{Im} q(k)R]$. This conclusion does not depend on the particular boundary conditions chosen at R, since a plasma radius R satisfying (A18) is effectively infinite from the point of view of the beam.

We will be primarily interested in having formulas for $\alpha_{(1)}$ and $\alpha_{(2)}$ in the "high-σ approximation" described in Sec. 7; we let $\sigma \to \infty$, so that $q(k)$ is of order $\sigma^{\frac{1}{2}}$:

$$q^2(k) \to 4\pi i\sigma(\Omega + kv)/c^2, \quad \text{(A19)}$$

but we keep $|q(k)R|$ of order unity. In this case a_1, a_2, a_4, b_1, b_2, and b_4 are of order q, while a_3, a_5, b_3, and b_5 are of order $1/q$. Furthermore, c_1, c_2, c_4, and c_5 are of order $1/R \sim q$, while c_3 is of order $R \sim 1/q$. Neglecting a_3, a_5, b_3, b_5, and c_3 in (A16) gives

$$\alpha_{(1)} \to \left[b_2 - \frac{b_1c_2}{c_1}\right]\Big/\left[a_2 - \frac{a_1c_2}{c_1}\right] \to \frac{[b_2 - b_1/2]}{[a_2 - a_1/2]} = -\frac{i\pi}{2}\frac{H_0^{(1)}(qR)}{J_0(qR)}. \quad \text{(A20)}$$

The same approximations made in (A17) give

$$\alpha_{(2)} \to \frac{\left[b_4 - \frac{b_1c_4}{c_1} - \frac{b_2c_5}{c_2}\right]\left[a_2 - \frac{a_1c_2}{c_1}\right] - \left[b_2 - \frac{b_1c_2}{c_1}\right]\left[-\frac{a_2c_5}{c_2}\right]}{\left[a_2 - \frac{a_1c_2}{c_1}\right]\left[-\frac{a_2c_5}{c_2}\right]}$$

$$\to -\alpha_{(1)} - \frac{i\pi}{2}\frac{[(\Omega + kv)H_0^{(1)\prime}(qR) - (h^2vqR/2k)H_0^{(1)}(qR)]}{\Omega J_0'(qR)}. \quad \text{(A21)}$$

Our result (A20) for $\alpha_{(1)}$ just expresses the fact that $\phi_{(1)}$, given by (A1), must vanish for $r = R$. This could have been seen more directly by letting $q \to \infty$ in Eq. (A6′), which gives

$$q^2\phi'_{(1)}(R - \epsilon) = h^2\phi'_{(1)}(R + \epsilon);$$

so $\phi'_{(1)}(R - \epsilon)$ is of order $1/q^2$. We will not need $\alpha_{(2)}$ in this article, but we have calculated it here because it might be needed in further work.

For $|qR| \gg 1$, (A20) and (A21) give $\alpha_{(1)}$ and $\alpha_{(2)}$ of order exp $(-2 \operatorname{Im} qR)$. This is in agreement with the estimate we made before using the high-σ approximation. Hence we may safely use (A20) and (A21) for any $|qR|$ not much less than one, even though they were only derived for $|qR|$ of order unity, since when $|qR| \gg 1$ both $\alpha_{(1)}$ and $\alpha_{(2)}$ are so small that they make no contribution to the fields or to the dispersion relation.

APPENDIX B. THE BEAM FORM FACTOR

The size and shape of the beam enter into the dispersion relation only in two functions introduced in Sec. 5: the Bessel transform of the beam shape

$$\rho(q) = \int_0^\infty n(r)J_0(qr)r\,dr \Big/ \left[\int_0^\infty n(r)^2 r\,dr\right]^{\frac{1}{2}} \tag{B1}$$

and the "beam form factor"

$$F(q^2) = \frac{-\frac{1}{2}(i\pi)q^2 \int_0^\infty r\,dr \int_0^\infty r'\,dr' n(r)H_0^{(1)}(qr_>)J_0(qr_<)n(r')}{\int_0^\infty n(r)^2 r\,dr}. \tag{B2}$$

[We are interested in these functions when q takes the complex value $q_0(k)$ or $q(k)$, with $\operatorname{Im} q > 0$, so this definition of $\rho(q)$ is really only adequate if $n(r)$ falls off faster than any exponential as $r \to \infty$. However, if $n(r)$ behaves like $\exp(-r/a)$ then $\rho(q)$ may usually be defined by analytic continuation from the values given by (B1) for $\operatorname{Im} q < a^{-1}$.]

The properties of $\rho(q)$ are easy to deduce. It is normalized so that

$$\int_0^\infty \rho(l)^2 l\,dl = 1. \tag{B3}$$

For $q \to 0$ it approaches a finite limit

$$\rho(0) = a/\sqrt{2}, \tag{B4}$$

where a is a mean beam radius defined by

$$a = \sqrt{2}\int_0^\infty n(r)\,r dr \Big/ \left[\int_0^\infty n(r)^2 r\,dr\right], \tag{B5}$$

the factor $\sqrt{2}$ being inserted to agree with the radius of a "square" beam. Finally, it is real, in the sense that

$$\rho(q)^* = \rho(q^*). \tag{B6}$$

The general properties of $F(q)$ are easiest to establish if we re-express it in terms of $\rho(l)$. We use the well-known formula

$$-\tfrac{1}{2}(i\pi)H_0^{(1)}(qr_>)J_0(qr_<) = \int_0^\infty \frac{J_0(lr)J_0(lr')}{q^2 - l^2}\,l\,dl, \tag{B7}$$

and obtain

$$F(q^2) = q^2\int_0^\infty \frac{\rho(l)^2}{q^2 - l^2}\,l\,dl. \tag{B8}$$

A number of its properties are now apparent:

(i) As $|q| \to 0$,

$$F(q^2) \to \tfrac{1}{4}q^2a^2 \ln(-q^2r_0^2) + O(q^4), \tag{B9}$$

where a is the mean beam radius (B5), and r_0 is another characteristic beam radius. We can get (B9) either from (B8) and (B4), or directly from (B2); the latter method also gives an explicit formula for r_0,

$$\ln r_0^2 = \int_0^\infty \ln\left(\frac{C^2r^2}{4}\right) d\mu(r), \tag{B10}$$

where

$$\mu(r) \equiv \left[\int_0^r n(r')r'\,dr' \Big/ \int_0^\infty n(r)r\,dr\right]^2 \tag{B11}$$

$$\ln C = 0.5772 \cdots. \tag{B12}$$

(ii) As $|q| \to \infty$,

$$F(q^2) \to 1 + q^{-2}b^{-2}, \tag{B13}$$

where b is yet another effective beam radius, defined by

$$b^{-2} = \int_0^\infty \rho(l)^2 l^3\,dl. \tag{B14}$$

Using (B1) gives

$$b^2 = \int_0^\infty n(r)^2 r\,dr \Big/ \int_0^\infty \left[\frac{dn(r)}{dr}\right]^2 r\,dr. \tag{B15}$$

This incidentally shows how important it is to avoid basing qualitative conclusions on unrealistic examples like the "square" beam; the discontinuity in $n(r)$ at the beam edge would give $b = 0$!

(iii) For q^2 real and negative,

$$F(q^2) > 0. \tag{B16}$$

(iv) For q^2 complex, the imaginary part of $F(q^2)$ cannot vanish, and in fact must have the same sign as $-\operatorname{Im} q^2$.

We now turn to specific examples. The most realistic one we will consider is the Gaussian beam shape

$$n(r) = n\exp(-r^2/2b^2). \tag{B17}$$

The radius b here is chosen to agree with that defined by (B15). The Bessel transform (B1) is

$$\rho(q) = \sqrt{2}\,b\exp[-\tfrac{1}{2}(b^2q^2)]. \tag{B18}$$

The form factor (B2) can be written as an integral

$$F(q^2) = q^2b^2 \int_0^\infty \frac{e^{-x}\,dx}{q^2b^2 - x}. \tag{B19}$$

The effective radii appearing in the low-q limit formula (B9) are

$$a = 2b, \tag{B20}$$

$$r_0 = C^{\frac{1}{2}}b = 1.80b. \tag{B21}$$

The Gaussian beam has the disadvantage of not allowing the expression of $F(q)$ in elementary functions. We can do better with an exponential beam shape

$$n(r) = n \exp(-r/b). \tag{B22}$$

Again, we are choosing b in (B22) to agree with the formula (B15). The Bessel transform (B1) is

$$\rho(q) = 4b[1 + b^2q^2]^{-\frac{3}{2}}. \tag{B23}$$

The beam form factor is

$$F(q^2) = 2q^2b^2\left[\frac{1}{2(1+q^2b^2)} + \frac{1}{(1+q^2b^2)^2} + \frac{\ln(-q^2b^2)}{(1+q^2b^2)^3}\right]. \tag{B24}$$

[Despite appearances $F(q^2)$ is analytic at $q^2b^2 = -1$, taking the finite value $F(-b^{-2}) = \frac{2}{3}$.] The effective beam radii in (B9) are

$$a = 8^{\frac{1}{2}}b, \tag{B25}$$

$$r_0 = b. \tag{B26}$$

It is easy to check the properties (i)–(iv) for both the Gaussian and exponential beams.

For the purpose of comparison with previous work,[2] we will also give some results for the "square" beam

$$n(r) = \begin{cases} n & r < a \\ 0 & r > a. \end{cases} \tag{B27}$$

The Bessel-transformed beam shape is

$$\rho(q) = \sqrt{2}\, J_1(aq)/q, \tag{B28}$$

and the beam form-factor is

$$F(q^2) = 1 - i\pi J_1(qa)H_1^{(1)}(qa). \tag{B29}$$

The behavior of $F(q)$ as $q \to 0$ is given by (B9), with

$$r_0 = (C/2e^{\frac{1}{4}})a = 0.694a. \tag{B30}$$

But as already remarked, (B15) gives $b = 0$. In fact, the sharp beam edge invalidates (B13), and instead we have here, for $|q| \to \infty$,

$$F(q^2) \to 1 - i/qa. \tag{B31}$$

There is no doubt that (B13) will be more reliable in practice than (B31).

One remaining touchy point is the asymptotic behavior of $\rho(q)$ as $|q| \to \infty$. It is obvious from (B1) that $\rho(q)$ will vanish as $q \to \infty$ along the real axis, but the case of greatest physical interest is for $q^2 \to \infty$ along the imaginary axis. In order to find the asymptotic behavior of $\rho(q)$ in this case we would in general need to use the method of steepest descent, which depends very much on the particular form of $n(r)$. The best we can do is to look at our most realistic example, the Gaussian beam. In this case the $\rho(q)$ given by (B18) just oscillates as $q^2 \to i\infty$. We only need to assume that $\rho(q)$ grows no faster than a polynomial in q. [The exponential beam is less realistic than the Gaussian, because its sharp cusp at $r = 0$ can give a spurious behavior as $|q| \to \infty$. The square beam is even worse.]

Magnetic Configuration of a Cylinder with Infinite Conductivity

GRAZIELLA BARDOTTI AND BRUNO BERTOTTI
Laboratorio Gas Ionizzati (EURATOM-C.N.E.N.), Frascati, Rome, Italy
AND
LAURA GIANOLIO
Applicazioni e Ricerche Scientifiche, Milan, Italy
(Received 31 March 1964)

The magnetic configuration of a cylinder with infinite conductivity is computed numerically for different values of the height $2h$. The calculation indicates that the current density $j(z)$ diverges at the edges as $I(h)/(h^2-z^2)^{\frac{1}{2}}$ and permits the evaluation of the function $I(h)$. The regular part of $j(z)$ seems to be approximated by an elliptical profile. The case of a cut cylinder is also discussed.

I. INTRODUCTION

In a number of experiments on magnetic compression of a plasma, a cylindrical coil is used, in which the current circulates in the (x, y) sections (Fig. 1).[1] Its conductivity can be considered infinite at the frequency used. The current density $j(z)$, therefore, is determined by the condition that the radial component B_ρ of the magnetic field vanish on the conductor. We would like to compute $j(z)$ for a given total current and show the relevant points in the magnetic configuration. In the last paragraph we discuss the case of two (or more) coaxial cylinders, whose separation is much smaller than their height and their radius.

When R/h reaches the extreme values $(0, \infty)$, it is possible to integrate analytically Laplace's equation for the magnetic potential φ with suitable boundary conditions. When $R/h \to \infty$, the problem corresponds to plane, incompressible flow around a strip (Fig. 2); its solution is[2]:

$$\varphi(s, z) = 2I_0 \operatorname{Re} \{\sin^{-1} [(z + is)/h]\}; \quad s \geq 0, \tag{1.1}$$

where I_0 is the total current flowing.

The current density, therefore, is singular at the end points $|z| = h$:

$$j(z) = \frac{1}{4\pi} \frac{d}{dz} [\varphi(0^+, z) - \varphi(0^-, z)]$$
$$= \frac{1}{2\pi} \frac{d}{dz} \varphi(0^+, z) = \frac{I_0}{\pi(h^2 - z^2)^{\frac{1}{2}}}. \tag{1.2}$$

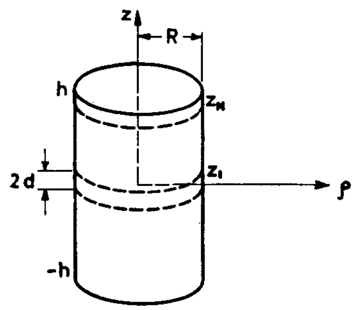

FIG. 1. The coil geometry.

When $R/h \to 0$, the problem, in the neighborhood of the edge, corresponds to that of a cylinder extending from $z = 0$ to $z = -\infty$ and can be solved by the method of Wiener–Hopf.[3] If the current density for $z \to -\infty$ is j_0 (which in our case is $I_0/2h$), for small negative values of z we find

$$j(z) \simeq (-z)^{-\frac{1}{2}} j_0/(\pi)^{\frac{1}{2}}. \tag{1.3}$$

II. THE INTEGRAL EQUATION FOR $j(z)$

The magnetic field $\mathbf{B}(\rho, z)$ can be represented as an integral of the elementary field $\mathbf{B}^0(\rho, z)$ created by a coil of radius R, in the plane $z = 0$, carrying a unitary current:

$$\mathbf{B}(\rho, z) = \int_{-h}^{h} dz' j(z') \mathbf{B}^0(\rho, z - z'). \tag{2.1}$$

$j(z)$ is then determined by the integral equation

$$B_\rho(R, z) = 0 = \int_{-h}^{h} dz' j(z') B_\rho^0(R, z - z'); \qquad |z| \leq h, \tag{2.2}$$

[1] See, e.g., W. E. Quinn, "Studies of Plasmas in Fast Magnetic Compression Experiments," in *Plasma Physics Thermonuclear Research*, edited by C. L. Langmuire, J. L. Tuck, and W. B. Thompson (Pergamon Press, Inc., New York, 1963). The same mathematical problem arises, of course, in hydrodynamics and corresponds to the flow around a cylindrical vortex.

[2] E. Durand, *Electrostatique et Magnetostatique* (Masson & Cie, Paris, 1953), p. 302.

[3] B. Noble, *Methods based on the Wiener–Hopf Technique* (Pergamon Press, Inc., New York, 1958), p. 110; H. Levine and J. Schwinger, Phys. Rev. **73**, 383 (1948).

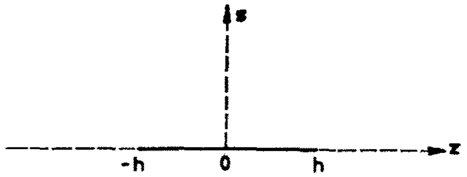

FIG. 2. The limiting case of a current strip.

with the normalization condition

$$\int_{-h}^{h} dz' j(z') = 1. \tag{2.3}$$

In the following, the radius R is taken equal to unity.

$\mathbf{B}^0(\rho, z)$ is known and can be expressed in terms of elliptic functions[4]

$$B_\rho^0(\rho, z) = B_z^0(0,0)\frac{z}{\pi\rho[(1+\rho)^2+z^2]^{\frac{1}{2}}} \times\left[-K(k)+\frac{1+\rho^2+z^2}{(1-\rho)^2+z^2}E(k)\right],$$

$$B_z^0(\rho, z) = B_z^0(0,0)\frac{1}{\pi[(1+\rho)^2+z^2]^{\frac{1}{2}}} \times\left[K(k)+\frac{1-\rho^2-z^2}{(1-\rho)^2+z^2}E(k)\right], \tag{2.4}$$

$$k^2 = \frac{4\rho}{(1+\rho)^2+z^2}.$$

The integral equation (2.2) is solved numerically.

Considering the singularities at the edges when $R/h \to \infty$ we can reasonably write

$$j(z) = I/(h^2-z^2)^{\frac{1}{2}} + j^*(z), \tag{2.5}$$

where I is a constant to be determined and $j^*(z)$ is the new unknown function; at the end, in fact, we find that $j^*(z)$ is numerically everywhere bounded.

To obtain an approximation to the integral equation (2.2), we demand that the magnetic flux be zero on a succession of elementary cylinders of height $2d$, centered at $z = z_s = (2s-1)d$ (Fig. 1). They are $N = h/2d$ in number; j^* is supposed to be constant on each interval. From (2.2), (2.3), and (2.5) we obtain, then, a system of $N+1$ equations in the $N+1$ unknowns $j^*(z_s) = j_s^*$ and I:

$$IC_\alpha + \sum_1^N {}_s\, j_s^*(C_{\alpha,s}+C_{\alpha,-s}) + \sum_1^N {}_s\,(\epsilon_{\alpha,s}+\epsilon_{\alpha,-s}) = 0,$$

$$4d\sum_1^N {}_s\, j_s^* + \pi I + 2\sum_1^N {}_s\,\epsilon_s = 1, \tag{2.6}$$

$$\alpha = 1, 2, \cdots, N.$$

[4] W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company, Inc., New York, 1950), p. 270.

The terms ϵ_s, $\epsilon_{\alpha,s}$ stem from the neglect of variations of $j^*(z)$ within each interval $2d$; they can be considered as errors of the numerical method and are neglected. The coefficients C_α, $C_{\alpha,s}$ are given by the formulas

$$C_\alpha = \int_{z_\alpha-d}^{z_\alpha+d} dz^* \int_{-h}^{h} dz \frac{\gamma B_\rho^0(1, z^*-z)}{(h^2-z^2)^{\frac{1}{2}}}$$

$$= \int_{z_\alpha-d}^{z_\alpha+d} dz^* \int_{-h}^{h} dz \frac{\left[\gamma B_\rho^0(1,z^*-z) - \dfrac{1}{z^*-z}\right]}{(h^2-z^2)^{\frac{1}{2}}}$$

$$\simeq 2d\int_{-h}^{h} dz \frac{\left[\gamma B_\rho^0(1,z_\alpha-z) - \dfrac{1}{z_\alpha-z}\right]}{(h^2-z^2)^{\frac{1}{2}}}, \tag{2.7}$$

$$\gamma = \frac{\pi}{B_z^0(0,0)},$$

$$C_{\alpha,s} = \int_{z_\alpha-d}^{z_\alpha+d} dz^* \int_{z_s-d}^{z_s+d} dz\gamma B_\rho^0(1,z^*-z)$$

$$= \int_{z_\alpha-d}^{z_\alpha+d} dz^* \int_{z_s-d}^{z_s+s} dz\left[\gamma B_\rho^0(1,z^*-z) - \frac{1}{z^*-z}\right] + \int_{z_\alpha-d}^{z_\alpha+d} dz^* \int_{z_s-d}^{z_s+d} dz\frac{1}{z^*-z}$$

$$\simeq 2d\int_{z_s-d}^{z_s+d} dz\left[\gamma B_\rho^0(1,z_\alpha-z) - \frac{1}{z_\alpha-z}\right] + \int_{z_s-d}^{z_s+d} dz \ln\left|\frac{z_\alpha+d-z}{z_\alpha-d-z}\right| \tag{2.8}$$

$$\simeq 4d^2\left[\gamma B_\rho^0(1,z_\alpha-z_s) - \frac{1}{z_\alpha-z_s}\right] + [(z_\alpha-z_s+2d)\ln|z_\alpha-z_s+2d| - 2(z_\alpha-z_s)\ln|z_\alpha-z_s| + (z_\alpha-z_s-2d)\ln|z_\alpha-z_s-2d|].$$

The singular part of B_ρ^0 has been subtracted out and integrated analytically; in the case of C_α it does not give any contribution.

The work is done in three steps: first, compute the integrals C_α, $C_{\alpha,s}$; next, solve the linear system (2.4); finally, show the relevant points in the magnetic configuration.

The integrals C_α cannot be computed numerically in a straightforward way, since the integrands are unbounded at $|z| = h$, and at $z = z_\alpha$ their second derivatives have a logarithmic singularity. The new variable of integration ϑ given by

$$z = h\sin(\sin^{-1} z_\alpha + \exp\vartheta) \quad \text{for} \quad z < z_\alpha,$$
$$z = h\sin(\sin^{-1} z_\alpha - \exp\vartheta) \quad \text{for} \quad z > z_\alpha, \tag{2.9}$$

removes both difficulties.

The method is intrinsically approximate to within terms of order d^2, which for $N = 80$ and $h = 5$, is 9.6×10^{-4}. The integrals occuring in the expression for C_α have been computed with a numerical percentage error not larger than 10^{-5}.

The system (2.6) was solved inverting the corresponding matrix with the "rank annihilation" method. The currents j_i^* turned out to be stable, while N varies up to 80, to within a percentage error of about 10^{-4} (as the constant I) at the center and about 10^{-2} at the edge.

III. RESULTS

The problem was solved on an IBM 7090 for some values of h; the function $j^*(z)$ was always found to have a maximum at the center and to vanish at the edges $|z| = h$; as expected, the contribution of $j^*(z)$ seems to be approximated with a good accuracy which improves with the decrease of h, by

$$j^*(z) = C(h^2 - z^2)^{\frac{1}{2}}, \tag{3.1}$$

where

$$C = [(1/\pi) - I](2/h^2).$$

The error is less than 1% for $h = 1$. Figure 3 shows the behavior of the current density $i(z) = j(z)2h$ when the mean current is kept equal to unity; the current profile approaches unity over most of the range as $h \to \infty$. In Fig. 4 are plotted, as function of h, the constant $I(h)$ and the coefficient of the singular part of the current $i(z)$, i.e., $I(2h)^{\frac{1}{2}}$; the numerical results check with theoretical values $\lim_{h\to 0} I(h) = 1/\pi$ and $\lim_{h\to\infty} (2h)^{\frac{1}{2}} I = 1/(\pi)^{\frac{1}{2}}$ that one can obtain from (1.3) and (1.4). Figure 5 shows that the longitudinal magnetic field is a monotone function of z along the axis and of ρ on the median plane; note that $(\partial B_z(\rho, z)/\partial\rho)_{\rho=1,|z|=h} = 0$, as it follows from $(\nabla \times \mathbf{B})_\varphi = 0$ and $B_\rho(1, z) = 0$ ($|z| = h$). Figure 6 shows how the inductance $\phi = 2\pi \int_0^1 \rho B_z(\rho, 0)\, d\rho$ varies as function of the height h, when the mean current is equal to unity.

IV. THE CASE OF A CUT CYLINDER[5]

Since the current flows in parallel planes, one would expect that the operation of cutting a cylinder at a given plane and leaving the same current in the two pieces would not change the magnetic configuration. We would like to know, however, what happens when in the two cylinders currents I_1 and I_2 are circulating, whose intensities are different than

[5] The content of this paragraph is due to Dr. J. E. Allen.

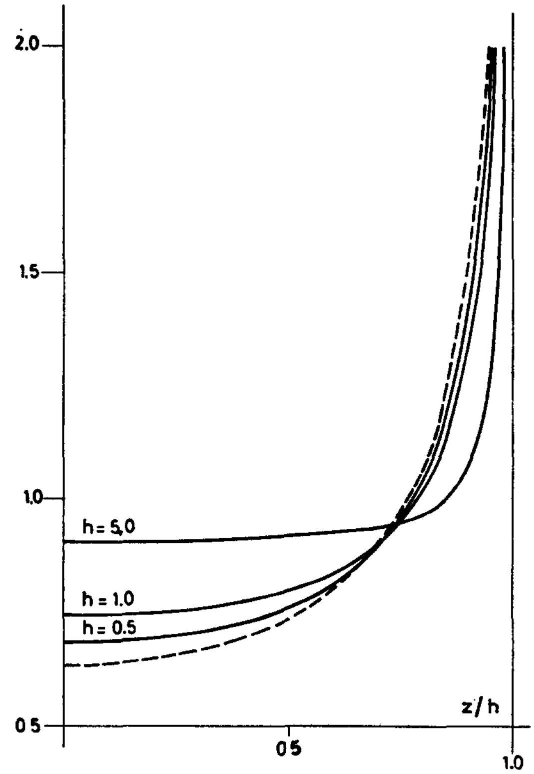

FIG. 3. The current density $i(z)$ as a function of z/h. The dashed curve represents the limiting case $h \to 0$.

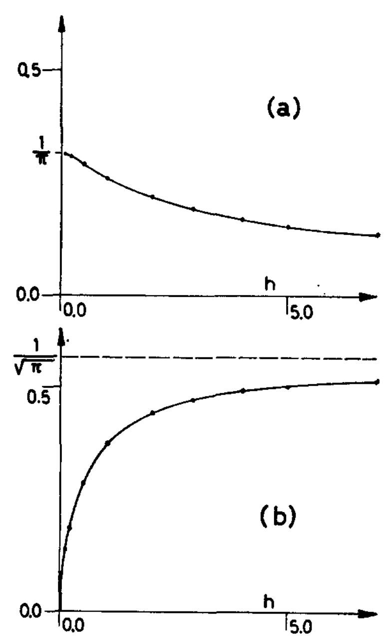

FIG. 4. (a) the behavior of $I(h)$; (b) the behavior of the coefficient of the singular part of $i(z)$, i.e., $(2h)^{\frac{1}{2}}I$, as function of h.

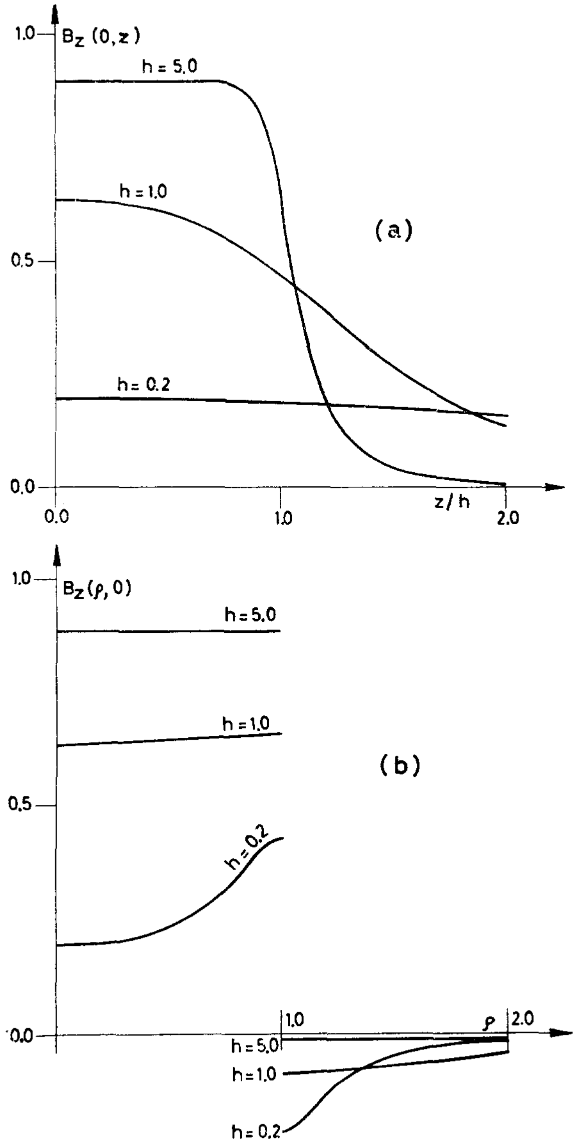

FIG. 5. The longitudinal component of the magnetic field along the axis (a), and on the median plane (b), normalized to magnetic field of a solenoid whose current density is equal to unity.

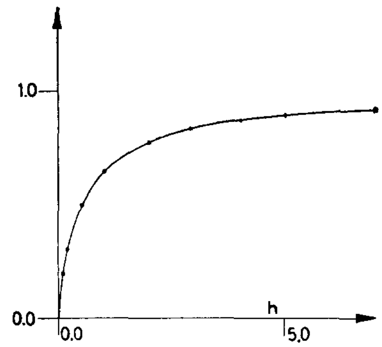

FIG. 6. The total flux through the cylinder.

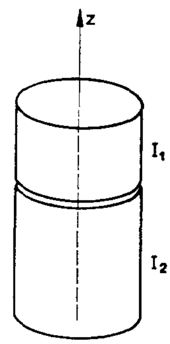

FIG. 7. Another configuration.

the fractions $\alpha(I_1 + I_2)$ and $(1 - \alpha)(I_1 + I_2)$ that one would have for a whole cylinder (Fig. 7). If the cut is small enough so as to prevent any magnetic leak, the solution should be the same as to the one computed before and one wonders where the extra currents are lost. Clearly there is no answer to the problem unless one admits the presence of two equal and opposite current lines at the facing edges of the cylinders, which neutralize each other as far as the magnetic field is concerned. If i_1 and $-i_1$ are their intensity, we must have

$$I_1 - i_1 = \alpha(I_1 + I_2); \quad I_1/(I_1 + I_2) > \alpha. \tag{4.1}$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 10 OCTOBER 1964

Complex Angular Momenta and Many-Particle States. I. Properties of Local Representations of the Rotation Group

M. ANDREWS AND J. GUNSON
Department of Mathematical Physics, University of Birmingham, Birmingham, England
(Received 20 March 1964)

The properties of the "local representations" of the rotation group corresponding to complex angular momentum are further developed. Completeness and bi-orthogonality relations are derived and a reduction of products is carried out, giving a generalization of the Clebsch-Gordan reduction. The connection with the representation theory of the group $SL(2,R)$ is considered and a generalization of Regge's use of the Sommerfeld-Watson transform is made to the case where three momentum transfer variables occur in the description of scattering amplitudes.

1. INTRODUCTION

RECENT investigations[1-3] into the connection between stable particles and Regge poles have suggested examination of the following problems: Given a certain set of particles all of which are Regge particles, i.e., they are represented by poles in scattering amplitudes, which lie on Regge trajectories, how can one define scattering amplitudes for these particles taking nonphysical values of the spin variable? Are these amplitudes uniquely defined, and how are the analytic and unitary properties related to those for physical scattering amplitudes?

In this series of papers, we develop first some of the basic techniques required in attacking these problems and use them to provide partial answers in certain important cases. The definition of an "off-spin-shell" amplitude is based here on the factorization of residues at Regge poles, in complete analogy with the corresponding definition of unstable particle amplitudes. (See Gunson's "Analyticity and Unitarity",[4] hereafter referred to as A). In one sense, these amplitudes may be regarded as "off-mass-shell" continuations, but differing from the more usual versions in that the spin is continued simultaneously with the mass, so as to follow a Regge trajectory. The other important difference is that of uniqueness. We expect the Regge particle amplitudes to be unique, being defined entirely in terms of on-mass-shell amplitudes without any extra external parameters (see Sec. 6.5 of A).

Before we can investigate these amplitudes in detail, we have to develop suitable mathematical tools for handling the transforms of many-particle amplitudes to and from the complex angular momentum planes. The two main obstacles met are (i) the extension in full of the representations of the rotation group to complex angular momenta and (ii) the problem of "complex singularities" in the momentum transfer variables. Our solutions to these problems are given in the first and second papers of this series. In the third paper, we treat unitarity and analyticity.

For a physical interpretation of states involving particles with nonphysical spins, we refer the reader to papers by Regge[5] and Ford and Wheeler.[6] In many senses, these states have the same status as unstable or virtual particle states. Both are somewhat ephemeral and wavefunctions representing them cannot be constructed without violating the usual boundary conditions. In the case of particles with nonphysical spins, the relevant condition is that of single (or double) valuedness over the unit sphere. However, it often occurs that the physical effect of the violation of the boundary conditions is small. The linking of mass and spin variables on a Regge trajectory demonstrates the complementary nature of these two forms of ephemeral state. In the boson case, the transformation properties under rotations of the states with well-defined complex spin j in the rest frame are those of the $Y_{j,m}(\theta, \varphi)$. These and related properties have been studied from a mathematical standpoint in terms of "local representations" of the rotation group in three dimensions in an appendix to A. In particular, definitions were given for functions $D_j^{m,m'}(R)$, of rotations R, which gave a natural extension of the well-known unitary representations to complex values of j—

[1] S. C. Frautschi, M. Gell-Mann, and F. Zachariasen, Phys. Rev. **126**, 2204 (1962).
[2] M. Gell-Mann, and M. L. Goldberger, Phys. Rev. Letters **9**, 275, (1962).
[3] M. Gell-Mann, *et al.*, Phys. Rev. **133**, B145, B161 (1964).
[4] J. Gunson "Unitarity and On-mass-shell Analyticity as a Basis for S-Matrix Theories," Birmingham Preprint (1962) (to be published).
[5] T. Regge, Nuovo Cimento **18**, 947, (1960).
[6] K. W. Ford, and J. A. Wheeler, Ann. Phys. (N. Y.) **7**, 272 (1959).

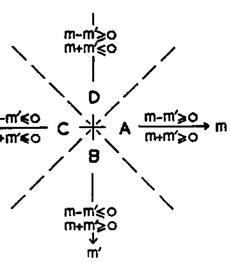

FIG. 1 Regions associated with the various expressions of $d^{m,m'}(z)$ in terms of hypergeometric functions.

where they are no longer unitary nor true representations since the matrix products do not converge over the whole group manifold; but when they do converge the group structure is preserved.[7] Also, functions of the second kind $E_j^{m,m'}(R)$ were defined which bear a similar relation to $D_j^{m,m'}(R)$ as do the Legendre functions $P_l(z)$ to those of the second kind $Q_l(z)$. In this first paper, we derive further relations for these representations and, in particular, develop further the observation made in A of the connection with the representation theory of the group $SL(2, R)$. The main results are:

(1) A generalization of the Regge–Sommerfeld–Watson transform [Eq. (14.6)] to the case where three momentum transfer variables necessarily occur in the description of scattering amplitudes. The restrictions on the nature and locations of singularities in these variables are too strong for immediate application but the result is generalized in Paper II. This transform is closely related to the Laplace transform on $SL(2, R)$.

(2) A generalization of the Clebsch–Gordan formula for the reduction of tensor products of representations in terms of irreducible representations (Sec. 13). We have found that the most convenient form is obtained in terms of products of functions of the second kind, expressed as a series of similar functions. This formula is closely connected with the usual reduction formula for the unitary representations of both $SO(3)$ [or $SU(2)$] and $SL(2, R)$. As the unitarity or nonunitarity of a representation is of no particular physical significance here, the series expansion possesses definite advantages over expressions involving contour integration over a continuous range of representations, as in the usual reduction formulas.[8] In particular, the domain of convergence is here much larger and the terms are uniquely determined.

Results equivalent to some of those presented in the early parts of this paper were given by Charap and Squires[9] and Kibble.[10]

As in A, rotations are parametrized by the Euler angles α, β, γ and the α and γ dependence of the (local) representation matrices $D_j^{m,m'}$ and those of the second kind $E_j^{m,m'}$ is factored out as follows:

$$\begin{aligned} D_j^{m,m'}(\alpha, \beta, \gamma) &= e^{im\alpha} d_j^{m,m'}(\cos\beta) e^{im'\gamma}, \\ E_j^{m,m'}(\alpha, \beta, \gamma) &= e^{im\alpha} e_j^{m,m'}(\cos\beta) e^{im'\gamma}. \end{aligned} \tag{1.1}$$

In all that follows, j and z are arbitrary complex numbers unless otherwise specified, while m and m' are either both integers or both half-odd integers. We often write M for max $(|m|, |m'|)$.

2. THE FUNCTIONS OF THE FIRST KIND

To give a definition of $d_j^{m,m'}(z)$ in terms of hypergeometric functions, it is appropriate to divide the space of m and m' into four regions A, B, C, D as indicated in Fig. 1.

In region A, the appropriate relation is

$$d_j^{m,m'}(z) = \left\{\frac{\Gamma(j+m+1)\Gamma(j-m'+1)}{\Gamma(j+m'+1)\Gamma(j-m+1)}\right\}^{\frac12}\left(\frac{1+z}{2}\right)^{\frac12(m+m')}\left(\frac{1-z}{2}\right)^{\frac12(m-m')} \times \frac{F(-j+m, j+m+1; 1+m-m' : \frac12 - \frac12 z)}{\Gamma(1+m-m')}. \tag{2.1}$$

Equivalent forms in the other regions are obtained by use of the symmetry relations given in (1).

For region B, use $d_j^{m,m'}(z) = (-1)^{m-m'} d_j^{m',m}(z)$.

For region C, use $d_j^{m,m'}(z) = (-1)^{m-m'} d_j^{-m,-m'}(z)$.

For region D, use $d_j^{m,m'}(z) = d_j^{-m',-m}(z)$.

In what follows, proofs are usually given for region A only. They can readily be extended to the other regions.

The function $d_j^{m,m'}(z)$ is branched and we must specify a principal sheet. The function is a product of the normalization term $\phi_j^{m,m'}$ (in region A, $\phi_j^{m,m'} =$

[7] On this point see, also, E. G. Beltrametti, and G. Luzzatto, Nuovo Cimento 29, 1003, (1963).

[8] L. Pukanszky, Trans. Am. Math. Soc. 100, 116, (1961).

[9] J. M. Charap, E. J. Squires, Ann. Phys. (N. Y.) 20, 145, (1962); 21, 8 (1963); 25, 143 (1963).

[10] T. W. B. Kibble, Phys. Rev. 131, 2282, (1963).

$\{\Gamma(j+m+1)\Gamma(j-m'+1)/\Gamma(j+m'+1)\Gamma(j-m+1)\}^{\frac12}$), which is branched as a function of j, and the remainder which is branched as a function of z but is an entire function of j.

In region A,

$$(\phi_j^{m,m'})^2 = (j+m)(j+m-1)\cdots(j+m'+1)$$
$$\times (j-m')(j-m'-1)\cdots(j-m+1).$$

so we can cut the j plane as shown in Fig. 2, and define the principal sheet to be positive for j large and positive. Since $(\phi_j^{m,m'})^2 = (\phi_{-j-1}^{m,m'})^2$ and $\phi_j^{m,m'} \sim j^{m-m'}$ for large $|j|$, it follows that $\phi_j^{m,m'} = (-1)^{m-m'}\phi_{-j-1}^{m,m'}$. Then since $F(-j+m, j+m+1; 1+m-m': \frac12 - \frac12 z)$ is unchanged under $j \to -j-1$, we have

$$d_j^{m,m'}(z) = (-1)^{m-m'}d_{-j-1}^{m,m'}(z) = d_{-j-1}^{-m,-m'}(z). \tag{2.2}$$

As a function of z, $(\frac12 + \frac12 z)^{\frac12}$ and $(\frac12 - \frac12 z)^{\frac12}$ have cuts which we extend outwards from -1 and $+1$; and we give these functions their positive values in the physical region $-1 < z < 1$. The hypergeometric function has a cut from -1 which we extend to $-\infty$, taking the principal sheet to be positive in the physical region.

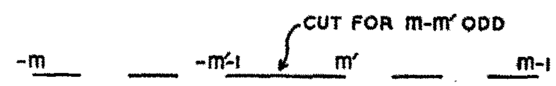

FIG. 2. The cut j-plane for region A.

3. THE FUNCTIONS OF THE SECOND KIND

From the $d_j^{m,m'}(z)$ which are analytic in $-1 < z < 1$, but which have a complicated branch at ∞, we can construct "functions of the second kind" $e_j^{m,m'}(z)$, which have simple behavior at $z = \infty$. In view of the linear relation between hypergeometric functions given, for example, in Sec. 2.9, Eq. (2.6) of HTF,[11] a suitable definition is

$$e_j^{m,m'}(z) = \frac{\pi}{2\sin\pi(j-m)}$$
$$\times \{e^{\mp i\pi(j-m)}d_j^{m,m'}(z) - d_j^{m,-m'}(-z)\}, \tag{3.1}$$

where we take $\mp$ for $\operatorname{Im} z \gtrless 0$. For then we have

$$e_j^{m,m'}(z) = \tfrac12\{\Gamma(j+m+1)\Gamma(j-m+1)\Gamma(j+m'+1)\Gamma(j-m'+1)\}^{\frac12}$$
$$\times (\tfrac12 + \tfrac12 z)^{\frac12(m+m')}(\tfrac12 - \tfrac12 z)^{-\frac12(m-m')}(\tfrac12 z - \tfrac12)^{-j-m'-1}\frac{F(j+m+1, j+m'+1; 2j+2: 2/1-z)}{\Gamma(2j+2)}, \tag{3.2}$$

where we put the same cuts in $(\frac12 + \frac12 z)^{\frac12}$ and $(\frac12 - \frac12 z)^{\frac12}$ as above, and cut the remainder from $+1$ to $-\infty$, giving it its positive value for $z > 1$.

From the hypergeometric identity

$$F(a, b; c : z) = (1-z)^{c-a-b}F(c-a, c-b; c : z)$$

we have

$$e_j^{m,m'}(z) = \tfrac12\{\Gamma(j+m+1)\Gamma(j-m+1)\Gamma(j+m'+1)\Gamma(j-m'+1)\}$$
$$\times (\tfrac12 + \tfrac12 z)^{-\frac12(m+m')}(\tfrac12 - \tfrac12 z)^{-\frac12(m-m')}(\tfrac12 z - \tfrac12)^{-j+m-1}\frac{F(j-m+1, j-m'+1; 2j+2: 2/1-z)}{\Gamma(2j+2)}. \tag{3.3}$$

From (3.2) and (3.3) one sees that

$$e_j^{m,m'}(z) = (-1)^{m-m'}e_j^{m',m}(z)$$
$$= (-1)^{m-m'}e_j^{-m,-m'}(z) = e_j^{-m',-m}(z). \tag{3.4}$$

From (2.2) and (3.1) we obtain the important relation

$$e_j^{m,m'}(z) - e_{-j-1}^{-m,-m'}(z) = \pi\cot\pi(j-m)d_j^{m,m'}(z). \tag{3.5}$$

Also, from (3.1) we have

$$e_j^{m,-m'}(-z) = -e^{\pm i\pi(j-m)}e_j^{m,m'}(z)$$
$$(\pm \quad \text{for} \quad \operatorname{Im} z \gtrless 0). \tag{3.6}$$

In the region $-1 < z < 1$, the discontinuity across the cut in z can be found directly from (3.1):

$$e_j^{m,m'}(z+i0) - e_j^{m,m'}(z-i0) = -i\pi d_j^{m,m'}(z). \tag{3.7}$$

4. RELATION TO JACOBI POLYNOMIALS AND LEGENDRE FUNCTIONS

For $j - m$ an integer, we can relate $d_j^{m,m'}(z)$ to the Jacobi polynomials $P_n^{\alpha,\beta}(z)$ as defined, for example, in Sec. 10.8 of HTF.

In region A, the relation is

$$d_j^{m,m'}(z)$$
$$= \{(j+m)!\,(j-m)!/(j+m')!\,(j-m')!\}^{\frac12}$$
$$\times (\tfrac12 + \tfrac12 z)^{\frac12(m+m')}(\tfrac12 - \tfrac12 z)^{\frac12(m-m')}P_{j-m}^{m-m',m+m'}(z),$$
$$j - m = 0, 1, 2, \cdots. \tag{4.1}$$

[11] Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdelyi, (McGraw-Hill Book Company, Inc., New York, 1953). (Referred to as HTF.)

TABLE I. Behavior for $j - m$ an integer.

	Region	$d_j^{m,m'}(z)$	$e_j^{m,m'}(z)$
I	$\lvert m\rvert, \lvert m'\rvert \le \lvert j+\frac{1}{2}\rvert - \frac{1}{2}$.	Finite: the series reduces to a polynominal.	Finite for $j \ge 0$. Pole of residue $d^{j\,m,m'}(z)$ for $j \le -1$.
II	$\lvert m\rvert, \lvert m'\rvert > \lvert j+\frac{1}{2}\rvert - \frac{1}{2}$ and $mm' > 0$.	Finite	Pole of residue $\frac{1}{2}d_j^{m,m'}(z)$.
III	$\lvert m\rvert, \lvert m'\rvert > \lvert j+\frac{1}{2}\rvert - \frac{1}{2}$ and $mm' < 0$.	Zero	Pole of residue $(-1)^{j-m+1}\frac{1}{2}d_j^{m,-m'}(-z)$.
IV	One of $\lvert m\rvert, \lvert m'\rvert \le \lvert j+\frac{1}{2}\rvert - \frac{1}{2}$ and the other $> \lvert j+\frac{1}{2}\rvert - \frac{1}{2}$	"Square-root zero"	"Square-root pole"

Similarly $e_j^{m,m'}(z)$ can be related to the Jacobi functions of the second kind $Q_n^{\alpha,\beta}(z)$. For region A,

$$e_j^{m,m'}(z) = (-1)^{m-m'}\{(j+m)!(j-m)!/(j+m')!(j-m')!\}^{\frac{1}{2}} \times (\tfrac{1}{2}+\tfrac{1}{2}z)^{\frac{1}{2}(m+m')}(\tfrac{1}{2}-\tfrac{1}{2}z)^{\frac{1}{2}(m-m')}Q_{j-m}^{m-m',m+m'}(z),$$
$$j - m = 0, 1, 2, \cdots . \quad (4.2)$$

When $m = m' = 0$ we obtain, for arbitrary complex j, the Legendre functions; thus

$$d_j^{0,0}(z) = P_j(z), \qquad e_j^{0,0}(z) = Q_j(z). \quad (4.3)$$

5. BEHAVIOR FOR INTEGER VALUES OF $j-m$

In considering the behavior of $d_j^{m,m'}(z)$ and $e_j^{m,m'}(z)$ for integer values of $j - m$, it is appropriate to divide the space of m and m' into regions I–IV as shown in Fig. 3. To cover the case of negative j as well as positive, we should replace j by $|j+\frac{1}{2}|-\frac{1}{2}$.

The hypergeometric function in (2.1) is analytic in j for all j and finite except for special values of z. The normalization factor can then be seen to give to $d_j^{m,m'}(z)$ the behavior indicated in Table I. It is clear from (3.2) that $e_j^{m,m'}(z)$ has no poles for $j \ge \mathrm{M}$, where $\mathrm{M} = \max(|m|, |m'|)$.

It follows from (3.1) that

$$d_j^{m,m'}(z) = (-1)^{j-m}d_j^{m,-m'}(-z),$$
$$j - \mathrm{M} = 0, 1, 2, \cdots . \quad (5.1)$$

From this, and (2.2), we can deduce that

$$d_j^{m,m'}(z) = -(-1)^{j-m}d_j^{m,-m'}(-z),$$
$$j + M = -1, -2, -3, \cdots . \quad (5.2)$$

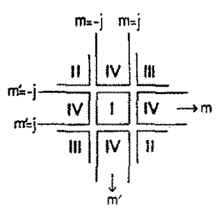

FIG. 3 Regions of different behavior for integer values of $j-m$.

Hence $e_j^{m,m'}(z)$ has a pole for $j - m$ an integer with $j \le -\mathrm{M} - 1$, the residue being $d_j^{m,m'}(z)$. The other regions can now be readily dealt with using (3.1) or (3.2). The results are given in Table I.

6. ASYMPTOTIC VALUES

The asymptotic values of $d_j^{m,m'}(z)$ and $e_j^{m,m'}(z)$ for fixed j, z, m, and large $|m'|$ were given in A.

For fixed j, m, m' the hypergeometric series of Eq. (3.2) immediately gives the behavior of $e_j^{m,m'}(z)$ for large $|z|$:

$$e_j^{m,m'}(z) \sim \tfrac{1}{2}\{\Gamma(j+m+1) \times \Gamma(j-m+1)\Gamma(j+m'+1)\Gamma(j-m'+1)\}^{\frac{1}{2}} \times e^{\pm i(\pi/2)(m-m')}(\tfrac{1}{2}z)^{-j-1}/\Gamma(2j+2)$$
$$(\pm \text{ for } \operatorname{Im} z \gtrless 0). \quad (6.1)$$

For $d_j^{m,m'}(z)$ we then use (3.5).

For fixed z, m, m' and large $|j|$ the relevant theory was given by Watson.[12] The result is quoted in Eq. (16) of Sec. 2.3 of HTF. From this we deduce

$$e_j^{m,m'}(z) \sim \left(\frac{\pi}{2}\right)^{\frac{1}{2}} \frac{e^{\pm i(\pi/2)(m-m')}}{j^{\frac{1}{2}}} \cdot \frac{[z-(z^2-1)^{\frac{1}{2}}]^{j+\frac{1}{2}}}{(z^2-1)^{\frac{1}{4}}},$$
$$-\pi + \epsilon < \arg j < \pi - \epsilon,$$
$$(\pm \text{ for } \operatorname{Im} z \gtrless 0). \quad (6.2)$$

The last term is cut along the $-ve$ z-axis only and is taken positive for $z > 0$.

Again, for $d_j^{m,m'}(z)$ we use (3.5).

7. OPERATORS WHICH CONSTRUCT THE REPRESENTATIONS FROM LEGENDRE FUNCTIONS

The functions $d_j^{m,m'}(z)$ and $e_j^{m,m'}(z)$ can be built up from $P_j(z)$ and $Q_j(z)$ by use of the operators I_z^μ and K_z^μ of Riemann–Liouville and Weyl fractional integration, respectively. These are defined

[12] G. N. Watson, Trans. Cambridge Phil. Soc. **22**, 277, (1918).

by the equations[13-14]

$$I_z^\mu[f(x)] = \frac{1}{\Gamma(\mu)} Pf \int_1^z f(y)(x-y)^{\mu-1}\,dx, \qquad \mu \neq 0, -1, -2, \cdots, \tag{7.1}$$

$$K_z^{-n}[f(x)] = I_z^{-n}[f(x)] = \left(\frac{d}{dx}\right)^n f(x), \qquad n = 0, 1, 2, \cdots, \tag{7.2}$$

$$K_z^\mu[f(x)] = \frac{1}{\Gamma(\mu)} Pf \int_z^\infty f(y)(y-x)^{\mu-1}\,dx, \qquad \mu \neq 0, -1, -2, \cdots. \tag{7.3}$$

We will show that, in region A,

$$\begin{aligned} d_j^{m,m'}(z) &= \{\Gamma(j+m+1) \\ &\times \Gamma(j-m'+1)/\Gamma(j+m'+1)\Gamma(j-m+1)\}^{\frac12} \\ &\times (\tfrac12+\tfrac12 z)^{\frac12(m+m')}(\tfrac12-\tfrac12 z)^{-\frac12(m-m')}(-\tfrac12)^{m-m'} \\ &\times I_z^{-m'}[(\tfrac12+\tfrac12 z)^{-m} I_z^m[P_j(z)]] \equiv I^{m,m'}[P_j(z)] \end{aligned} \tag{7.4}$$

and

$$\begin{aligned} e_j^{m,m'}(z) &= \{\Gamma(j-m+1) \\ &\times \Gamma(j+m'+1)/\Gamma(j+m+1)\Gamma(j-m'+1)\}^{\frac12} \\ &\times (\tfrac12+\tfrac12 z)^{-\frac12(m+m')}(\tfrac12-\tfrac12 z)^{\frac12(m-m')}(-2)^{m-m'} \\ &\times K_z^{m'}[(\tfrac12+\tfrac12 z)^{m} K_z^{-m}[Q_j(z)]] \equiv K^{m,m'}[Q_j(z)]. \end{aligned} \tag{7.5}$$

The steps involved in establishing (7.4) are as follows: Using the Riemann–Liouville transform given in Eq. (94) of Sec. 13.1 of IT[13] we find

$$I_z^m[P_j(z)] = (-2)^m(\tfrac12-\tfrac12 z)^m \times F(-j, j+1; 1+m : \tfrac12-\tfrac12 z)/\Gamma(1+m). \tag{7.6}$$

The simple hypergeometric identity given in HTF 2.9(2) leads to

$$(\tfrac12+\tfrac12 z)^{-m} I_z^m[P_j(z)] = (-2)^m(\tfrac12-\tfrac12 z)^m \times F(-j+m, j+m+1; 1+m : \tfrac12-\tfrac12 z)/\Gamma(1+m), \tag{7.7}$$

and then

$$\begin{aligned} I_z^{-m'}[(\tfrac12+\tfrac12 z)^{-m} I_z^m[P_j(z)]] &= (-2)^{m-m'}(\tfrac12-\tfrac12 z)^{m-m'} F(-j+m, j+m+1; \\ &\quad 1+m-m' : \tfrac12-\tfrac12 z)/\Gamma(1+m-m'), \end{aligned} \tag{7.8}$$

from which (7.4) follows.

To establish (7.5), we use Eq. (78) of Sec. 13.2 of IT. The equations corresponding to the above three are

$$\begin{aligned} K_z^{-m}[(Q_j z)] &= 2^{-(m+1)}\Gamma(j+1)\Gamma(j+m+1)(\tfrac12 z-\tfrac12)^{-j-m-1} \\ &\times F(j+m+1, j+1; 2j+2 : 2/1-z)/\Gamma(2j+2), \end{aligned} \tag{7.9}$$

$$\begin{aligned} (\tfrac12+\tfrac12 z)^m K_z^{-m}[Q_j(z)] &= 2^{-(m+1)}\Gamma(j+1)\Gamma(j+m+1)(\tfrac12 z-\tfrac12)^{-j-1} \\ &\times F(j-m+1, j+1; 2j+2 : 2/1-z)/\Gamma(2j+2), \end{aligned} \tag{7.10}$$

$$\begin{aligned} K_z^{m'}[(\tfrac12+\tfrac12 z)^m K_z^{-m}[Q_j(z)]] &= 2^{-(m-m'+1)} \\ &\times \Gamma(j-m'+1)\Gamma(j+m+1)(\tfrac12 z-\tfrac12)^{-j+m'-1} \\ &\times F(j-m+1, j-m'+1; 2j+2 : \\ &\quad 2/1-z)/\Gamma(2j+2). \end{aligned} \tag{7.11}$$

As usual, the symmetry relations lead to analogous expressions for regions B, C, and D.

Using these results, we can evaluate a certain integral. We will show that, in region A for example,

$$\int_1^\infty d_{j_1}^{m,m'}(z) e_{j_2}^{m,m'}(z)\,dz = \frac{\left\{\dfrac{\Gamma(j_1+m+1)\Gamma(j_1-m'+1)\Gamma(j_2-m+1)\Gamma(j_2+m'+1)}{\Gamma(j_1+m'+1)\Gamma(j_1-m+1)\Gamma(j_2+m+1)\Gamma(j_2-m'+1)}\right\}^{\frac12}}{(j_2-j_1)(j_2+j_1+1)}, \qquad \operatorname{Re} j_2 > \operatorname{Re} j_1 > 0. \tag{7.12}$$

Using (7.4) and (7.5) the integral becomes

$$\left\{\frac{\Gamma(j_1+m+1)\Gamma(j_1-m'+1)\Gamma(j_2-m+1)\Gamma(j_2+m'+1)}{\Gamma(j_1+m'+1)\Gamma(j_1-m+1)\Gamma(j_2+m+1)\Gamma(j_2-m'+1)}\right\}^{\frac12} \times \int_1^\infty K_z^{m'}[(\tfrac12+\tfrac12 z)^m K_z^{-m}[Q_{j_2}(z)]] I_z^{-m'}[(\tfrac12+\tfrac12 z)^{-m} I_z^m[P_{j_1}(z)]]\,dz.$$

[13] Bateman Manuscript Project, *Tables of Integral Transforms*, edited by A. Erdelyi, (McGraw-Hill Book Company, Inc., New York, 1954), Vol. II (referred to as IT).
[14] L. Schwartz, *Theorie des distributions* (Hermann & Cie, Paris, 1959).

Integrating by parts first for $K^{m'}$ and $I^{-m'}$, and then for K^{-m}, I^{m}, we obtain the required result on using

$$\int_1^\infty Q_{j_2}(z)P_{j_1}(z)\,dz = \{(j_2 - j_1)(j_2 + j_1 + 1)\}^{-1},$$

$$\operatorname{Re} j_2 > \operatorname{Re} j_1 > 0. \qquad (7.13)$$

which is Eq. (4) of HTF 3.12.

8. RECURRENCE RELATION

We first prove a relation between hypergeometric functions. For brevity, since we are here considering z, m, m' to be fixed, we write F_j for $F(j + m + 1, j + m' + 1; 2j + 2: 2/1 - z)$. We prove that

$$\begin{aligned}(j + m + 1)&(j + m' + 1)(j - m + 1)(j - m' + 1)\\ &\times \{j/(2j + 2)(2j + 3)\}(2/1 - z)F_{j+1}\\ &+ (j + 1)2j(2j + 1)\tfrac{1}{2}(1 - z)F_{j-1}\\ = (2j + 1)&\{2j(j + 1)\tfrac{1}{2}(1 - z) + mm' - j(j + 1)\}F_j.\end{aligned} \qquad (8.1)$$

This could presumably be derived by repeated application of the relations between contiguous hypergeometric functions, but we give a more direct proof. For $\frac{1}{2}|1 - z| > 1$ the hypergeometric series converge and equating coefficients of $(2/1 - z)^n$ it can be seen that the above identity is equivalent to the algebraic identity

$$\begin{aligned}(j + m + 1)&(j + m' + 1)jn/(2j + n + 2)\\ &+ (j - m)(j - m')(j + 1)(2j + n + 1)/(n + 1)\\ = (j - m &+ n + 1)(j - m' + n + 1)(2j + 2)\\ &\times 2j(j + 1)/(n + 1)(2j + n + 2)\\ &+ \{mm' - j(j + 1)\}(2j + 1).\end{aligned} \qquad (8.2)$$

This can readily be verified for $n + 1 = 0$, $2j + n + 2 = 0$, and $n = 0$ (three values of n clearly being sufficient).

From the relation (8.1) we obtain

$$\begin{aligned}\{(j + m + 1)&(j - m + 1)(j + m' + 1)\\ &\times (j - m' + 1)\}^{\frac{1}{2}}(j + 1)^{-1}e_{j+1}^{m,m'}(z)\\ &+ \{(j + m)(j - m)(j + m')(j - m')\}^{\frac{1}{2}}j^{-1}e_{j-1}^{m,m'}(z)\\ = (2j + 1)&\{z - mm'/j(j + 1)\}e_j^{m,m'}(z),\end{aligned} \qquad (8.3)$$

which is the required recurrence relation.

Changing j to $-j - 1$, m to $-m$, and m' to $-m'$, and using (3.5), one can see that $d_j^{m,m'}(z)$ satisfies the same recurrence relation.

9. COMPLETENESS RELATION

If we take first the recurrence relation (8.3) writing μ instead of j, secondly take the relation obtained from this by $\mu \to -\mu - 1$, $m \to -m$, $m' \to -m'$, $z \to t$, then multiply the first equation by $e_{-\mu-1}^{-m,-m'}(t)$, the second equation by $e_\mu^{m,m'}(z)$ and add, we obtain

$$(z - t)(2\mu + 1)e_\mu^{m,m'}(z)e_{-\mu-1}^{-m,-m'}(t) = R_{\mu+1}^{m,m'}(z, t) - R_\mu^{m,m'}(z, t), \qquad (9.1)$$

where

$$R_j^{m,m'}(z, t) = \{(j + m)(j - m)(j + m')(j - m')\}^{\frac{1}{2}}j^{-1} \times [e_j^{m,m'}(z)e_{-j}^{-m,-m'}(t) - e_{j-1}^{m,m'}(z)e_{-j-1}^{-m,-m'}(t)]. \qquad (9.2)$$

We can explicitly sum this for μ ranging in integer steps from say j to ν ($\nu - j$ an integer). Thus

$$(z - t)\sum_{\mu=j}^{\nu}(2\mu + 1)e_\mu^{m,m'}(z)e_{-\mu-1}^{-m,-m'}(t) = R_{\nu+1}^{m,m'}(z, t) - R_j^{m,m'}(z, t). \qquad (9.3)$$

In view of the asymptotic relation (6.2), we have for large $|\nu|$,

$$R_\nu^{m,m'}(z, t) \sim c(z, t)\left(\frac{z - (z^2 - 1)^{\frac{1}{2}}}{t - (t^2 - 1)^{\frac{1}{2}}}\right)^\nu. \qquad (9.4)$$

So if z lies outside the ellipse $E(t)$ which has foci at ±1 and which passes through t, then

$$\sum_{\mu=j}^{\infty}(2\mu + 1)e_\mu^{m,m'}(z)e_{-\mu-1}^{-m,-m'}(t) = -\frac{R_j^{m,m'}(z, t)}{z - t}, \qquad (9.5)$$

the convergence being uniform in z for z on any compact set lying wholly outside $E(t)$.

We will see that (9.5) can play the role of a completeness relation for a certain class of functions.

10. SOME PROPERTIES OF $R_j^{m,m'}(z,t)$

We first show that

$$\pi^{-1}\tan\pi(j - m)R_j^{m,m'}(z, z) = 1. \qquad (10.1)$$

Putting $z = t$ in (9.3), we see that

$$R_{j+1}^{m,m'}(z, z) = R_j^{m,m'}(z, z). \qquad (10.2)$$

When $j - m$ is $\frac{1}{2}$-odd-integer, it is clear from (3.5) that $e_j^{m,m'}(z) = e_{-j-1}^{-m,-m'}(z)$, and it follows that $R_j^{m,m'}(z, z)$ vanishes at these points. Therefore $\pi^{-1}\tan\pi(j - m)\,R_j^{m,m'}(z, z)$ is analytic for all finite j, except possibly for $-M < j < M$. But from the periodicity (10.2) it must then be analytic for all finite j. From the asymptotic form given in (6.2) we find that, for large $|j|$, $\pi^{-1}\tan\pi(j-m)R_j^{m,m'}\cdot(z, z) \sim 1$. A function analytic for all finite j and bounded for large $|j|$ must be constant and (10.1) follows.

Secondly we wish to evaluate $\pi^{-1}\tan\pi(j - m)\cdot$

$R_j^{m,m'}(z,t)$ for $j = M$, since this leads to a sum over physical values of j. Only the term in $e_{M-1}^{m;m'}(z)\cdot e_{-M-1}^{-m;-m'}(t)$ contributes, and the second factor of this has a pole of residue $-\pi\, d_M^{m,m'}(t)$. But in view of its relation (4.1) to $P_0^{|m-m'|,|m+m'|}(t)$ which is a constant, we can write

$$d_M^{m,m'}(t) = [(1+t)/(1+z)]^{\frac{1}{2}|m+m'|} \times [(1-t)/(1-z)]^{\frac{1}{2}|m-m'|} d_M^{m,m'}(z).$$

Then from the above property (10.1) we have

$$\pi^{-1}\tan\pi(j-m)\cdot R_j^{m,m'}(z,t) = [(1+t)/(1+z)]^{\frac{1}{2}|m+m'|} \times [(1-t)/(1-z)]^{\frac{1}{2}|m-m'|}, \quad \text{for } j = M. \tag{10.3}$$

The relation (9.5) then leads to

$$\sum_{j=M}^{\infty}(2j+1)e_j^{m,m'}(z)d_j^{m,m'}(t) = \left(\frac{1+t}{1+z}\right)^{\frac{1}{2}|m+m'|}\left(\frac{1-t}{1-z}\right)^{\frac{1}{2}|m-m'|}(z-t)^{-1}, \tag{10.4}$$

which agrees with Eq. B.33 of A.

11. EXPANSIONS IN SERIES OF FUNCTIONS OF THE SECOND KIND

Suppose $f_j(z)$ is such that $z^{j+1}f_j(z)$ is analytic in some neighborhood of $z = \infty$. Consider the integral

$$\frac{1}{2\pi i}\int_c f_j(t)\,\frac{\tan\pi(j-m)}{\pi}\,\frac{R_j^{m,m'}(z,t)}{z-t}\,dt,$$

where z and t lie outside the closed contour C and where ± 1 and all the singularities of $f_j(t)$ (except the branch at ∞) lie inside. It can be seen that $f_j(t)R_j^{m,m'}(z,t)$ is analytic for all t outside or on C, and tends to zero for large $|t|$. The value of the integral is therefore given by the residue at the pole for $t = z$. Thus, using (10.1), we see that value of the integral is just $f_j(z)$.

When the series given in (9.5) converges uniformly, we can insert it and integrate term by term. We thus have the following result.

Theorem. If $z^{j+1}f_j(z)$ is analytic for all z (including $z = \infty$) outside the ellipse E with foci at ± 1, then for z outside E,

$$f_j(z) = \sum_{\mu=j}^{\infty} a_\mu^{m,m'}e_\mu^{m,m'}(z),$$

where

$$a_\mu^{m,m'} = -(2\mu+1)\pi^{-1}\tan\pi(\mu-m)(2\pi i)^{-1} \times \int_c e_{-\mu-1}^{-m;-m'}(t)f_j(t)\,dt,$$

the contour C enclosing ± 1 and all the singularities of $z^{j+1}f_j(z)$.

The sum can be extended to negative $\mu - j$ since the coefficients a_μ vanish then, as can be seen by deforming the contour to infinity.

12. BI-ORTHOGONALITY RELATION

By the above theorem we can expand $e_{j+r}^{m;m'}(z)$, where r is any integer, as

$$e_{j+r}^{m;m'}(z) = \sum_{s=-\infty}^{\infty} a_{rs}e_{j+s}^{m,m'}(z) \qquad (s \text{ integer}).$$

But the functions $e_{j+r}^{m;m'}(z)$, $r = 0, \pm 1, \pm 2, \cdots$ are clearly linearly independant so we must have $a_{rs} = \delta_{rs}$. Hence

$$\frac{1}{2\pi i}\int_C e_{j+r}^{m,m'}(t)e_{-j-s-1}^{-m,-m'}(t)\,dt = -\delta_{rs}\frac{\pi\cot\pi(j-m)}{2(j+r)+1}, \tag{11.1}$$

where r and s are any integers, and the contour C encloses ± 1.

13. THE REDUCTION OF PRODUCTS OF FUNCTIONS OF THE SECOND KIND

The theorem of Sec. 11 immediately enables us to expand the product $e_{j_1}^{m_1,m_1'}(z)e_{j_2}^{m_2,m_2'}(z)$ as an infinite series over the functions $e_j^{m,m'}(z)$, where $m = m_1 + m_2$, $m' = m_1' + m_2'$ and $j - (j_1 + j_2) = 1, 2, 3, \cdots$, the expansion coefficients being given as a contour integral of a product of three functions of the second kind. The series converges over the whole cut z-plane. Burchnall and Chaundy[15] have conveniently given an explicit formula for the reduction of products for a class of hypergeometric functions which includes our functions of the second kind. This gives the expansion coefficients in terms of generalized hypergeometric functions (${}_3F_2$) of unit argument:

$$F(a,b;c:x)F(\alpha,\beta;\gamma:x) = \sum_{r=0}^{\infty}\frac{(a)_r(b)_r(\gamma)_r}{r!\,(c)_r(c+\gamma+r-1)_r} \times {}_3F_2\begin{bmatrix}\alpha, & 1-c-r, & -r\\ & \gamma, & 1-a-r\end{bmatrix} \times {}_3F_2\begin{bmatrix}\beta, & 1-c-r, & -r\\ & \gamma, & 1-b-r\end{bmatrix} \times x^rF(a+\alpha+r, b+\beta+r; c+\gamma+2r:x). \tag{13.1}$$

[15] J. L. Burchnall and T. W. Chaundy, Proc. London Math. Soc. (2) 50, 72, (1944).

From this it can be seen that the coefficients for the reduction of a product of two functions of the second kind can be written as a product of two terms, one involving the m's and the other the m''s. Thus

$$e_{j_1}^{m_1,m_1'}(z)e_{j_2}^{m_2,m_2'}(z) = \sum_{j=j_1+j_2+1}^{\infty} E(j_1, j_2; j : m_1, m_2)$$
$$\times E(j_1, j_2; j : m_1', m_2')e_j^{m,m'}(z). \tag{13.2}$$

Equation (13.1) then gives an algebraic expression for these E-coefficients. The occurrence of the negative integer $-r$ in the parameters of the ${}_3F_2$ means that it is no more than a finite sum over ratios of Γ-functions. We will see that the E-coefficients are closely related to a complex generalization of the Clebsch–Gordan coefficients. In fact,

$$E(j_1, j_2; j: m_1, m_2)$$
$$= \left\{\frac{\pi \tan \pi(j - m)}{\tan \pi(j_1 - m_1) \tan \pi(j_2 - m_2)}\right\}^{\frac{1}{2}}$$
$$\times C(-j_1 - 1, -j_2 - 1; -j - 1: -m_1, -m_2), \tag{13.3}$$

where

$$C(j_1, j_2; j: m_1, m_2)$$
$$= \left\{\frac{\Gamma(j_1 + m_1 + 1)\Gamma(j_2 - m_2 + 1)}{\Gamma(j_1 - m_1 + 1)\Gamma(j_2 + m_2 + 1)} \Gamma(j - m + 1)\Gamma(j + m + 1)\right\}^{\frac{1}{2}}$$
$$\times \left\{(2j + 1) \frac{\Gamma(j_1 - j_2 + j + 1)\Gamma(-j_1 + j_2 + j + 1)}{\Gamma(j_1 + j_2 - j + 1)\Gamma(j_1 + j_2 + j + 2)}\right\}^{\frac{1}{2}} \cdot \frac{1}{\Gamma(j - j_2 + m_1 + 1)\Gamma(j - j_1 - m_2 + 1)}$$
$$\times {}_3F_2\begin{bmatrix} -j_1 + m_1, & -j_2 - m_2, & -j_1 - j_2 + j \\ j - j_1 - m_2 + 1, & j - j_2 + m_1 + 1 & \end{bmatrix} \tag{13.4}$$
$$= \{\Gamma(j_1 + m_1 + 1)\Gamma(j_1 - m_1 + 1)\Gamma(j_2 + m_2 + 1)\Gamma(j_2 - m_2 + 1)\Gamma(j + m + 1)\Gamma(j - m + 1)\}^{\frac{1}{2}}$$
$$\times \left\{(2j + 1) \frac{\Gamma(j_1 + j_2 - j + 1)\Gamma(j_1 - j_2 + j + 1)\Gamma(-j_1 + j_2 + j + 1)}{\Gamma(j_1 + j_2 + j + 2)}\right\}^{\frac{1}{2}}$$
$$\times \sum_s (-1)^s \{s!\, \Gamma(j_1 + j_2 - j - s + 1)\Gamma(j_1 - m_1 - s + 1)\Gamma(j_2 + m_2 - s + 1)$$
$$\times \Gamma(j - j_2 + m_1 + s + 1)\Gamma(j - j_1 - m_2 + s + 1)\}^{-1}, \tag{13.5}$$

which reduces to Wigner's form for the Clebsch–Gordan coefficients for integer values of $j_1 - m_1$ and $j_2 - m_2$.

This relation can be derived from (13.1) by transforming the ${}_3F_2$ by means of the relation

$${}_3F_2\begin{bmatrix} \alpha, & 1 - c - r, & -r \\ \gamma, & 1 - a - r & \end{bmatrix} = \frac{(\gamma - \alpha)_r}{(\gamma)_r}$$
$$\times {}_3F_2\begin{bmatrix} \alpha, & c - a, & -r \\ 1 - \gamma + \alpha - r, & 1 - a - r & \end{bmatrix}. \tag{13.6}$$

This is given, for example, by Bailey.[16] It can be verified directly by using Gauss's formula [HTF, 2.8 (46)] to replace $(1 - c - r)_t/t!(1 - a - r)_t$ by $\sum_{s=0}^{t} (-1)^s(c - a)_s/s!(t - s)!\,(1 - a - r)_s$, then interchanging the order of summation (the sum being finite) and finally performing the t summation by Gauss's formula. Straightforward manipulation of the many Γ-functions involved in the resulting expression for the E-coefficients leads to the above relation to the C-coefficients.

The relation (13.3) shows that the functions

$$b_j^{m,m'}(z) = -\pi^{-1} \tan \pi(j - m)e_{-j-1}^{-m,-m'}(z) \tag{13.7}$$

reduce according to

$$b_{j_1}^{m_1,m_1'}(z)b_{j_2}^{m_2,m_2'}(z) = \sum_{j=-\infty}^{j_1+j_2} C(j_1, j_2; j: m_1, m_2)$$
$$\times C(j_1, j_2; j: m_1', m_2')b_j^{m,m'}(z). \tag{13.8}$$

This is a direct complex generalization of the Clebsch–Gordan reduction, since the $b_j^{m,m'}(z)$ reduce to the $d_j^{m,m'}(z)$ for positive integral values of $j - M$.

The reduction to the Clebsch–Gordan decomposition for the physical cases can also be seen from the integral expression

$$E(j_1, j_2; j: m_1, m_2)E(j_1, j_2; j: m_1', m_2')$$
$$= -(2j + 1) \frac{\tan \pi(j - m)}{\pi} \frac{1}{2\pi i}$$
$$\times \oint e_{j_1}^{m_1,m_1'}(z)e_{j_2}^{m_2,m_2'}(z)e_{-j-1}^{-m,-m'}(z)\, dz, \tag{13.9}$$

[16] W. N. Bailey, *Cambridge Tracts in Mathematics and Mathematical Physics* (Cambridge University Press, Cambridge, England, 1935), No. 32, p. 22.

which follows from the theorem of Sec. 11. For $j_1 + M_1$ and $j_2 + M_2$ negative integers, $e_{j_1}^{m_1,m_1'}(z)$ has a pole of residue $d_{j_1}^{m_1,m_1'} = d_{-j_1-1}^{-m_1,-m_1'}$. Furthermore the discontinuity of $e_{-j-1}^{-m,-m'}(z)$ across $(-1, 1)$ is $-i\pi\, d_{-j-1}^{-m,-m'}(z)$ [Eq. (3.7)].

So

$$\frac{\tan\pi(j_1 - m_1)\tan\pi(j_2 - m_2)}{\pi\tan\pi(j - m)}$$
$$\times E(j_1, j_2; j: m_1, m_2)E(j_1, j_2; j: m_1', m_2')$$
$$= -(2j + 1)(-1)^{m-m'}\tfrac{1}{2}$$
$$\times \int_{-1}^{1} d_{-j_1-1}^{-m_1,-m_1'}(z)\, d_{-j_2-1}^{-m_2,-m_2'}(z)\, d_{-j-1}^{m,m'}(z)\, dz$$
$$= C(-j_1 - 1, -j_2 - 1; -j - 1: -m_1, -m_2)$$
$$\times C(-j_1 - 1, -j_2 - 1; -j - 1: -m_1', -m_2'), \tag{13.10}$$

using the well known integral expression for the Clebsch–Gordan coefficients. (See, for example, Edmonds[17] 4.6.2).

All the above results can be trivially amended to apply to the functions $E_j^{m,m'}(\alpha, \beta, \gamma)$, $D_j^{m,m'}(\alpha, \beta, \gamma)$ instead of $e_j^{m,m'}(z)$, $d_j^{m,m'}(z)$.

14. RELATIONS WITH THE REPRESENTATION THEORY OF $SL(2,R)$

The unitary representations of the locally compact group $SL(2, R)$ have been much studied.[18–20,8] Those representations which occur in the expansion of an arbitrary square-integrable function defined on the group, the discrete series and the principal series, are closely related to the local representations of $SO(3)$ as noted in A. The connection is made more explicit by letting $SL(2, C)$ be parametrized in the following manner:

$$\begin{bmatrix} \cos\tfrac{1}{2}\beta\cdot e^{\frac{1}{2}i(\alpha+\gamma)} & \sin\tfrac{1}{2}\beta\cdot e^{\frac{1}{2}i(\alpha-\gamma)} \\ -\sin\tfrac{1}{2}\beta\cdot e^{-\frac{1}{2}i(\alpha-\gamma)} & \cos\tfrac{1}{2}\beta\cdot e^{-\frac{1}{2}i(\alpha+\gamma)} \end{bmatrix}, \tag{14.1}$$

α, z, γ arbitrarily complex, where $z = \cos\beta$. Then

(a) The subgroup $SU(2)$ is given by restriction to $-1 \le z \le 1$; $-2\pi \le \alpha + \gamma$, $\alpha - \gamma \le 2\pi$.
(b) The subgroup $SL(2, R)$ is given by restriction to $1 \le z < \infty$; $-\infty < i\alpha, i\gamma < \infty$.
(c) The subgroup G is given by restriction to $1 \le z < \infty$; $-2\pi \le \alpha + \gamma$, $\alpha - \gamma \le 2\pi$.

[17] A. R. Edmonds, *Angular Momentum in Quantum Mechanics*, (Princeton University Press, Princeton, New Jersey, 1957).
[18] V. Bargmann, Ann. Math. **48**, 568, (1947).
[19] R. A. Kunze, and E. M. Stein, Am. J. Math. **82**, 1, (1960).
[20] R. Takahashi, Japan J. Math. **31**, 55, (1961).

The group G is isomorphic to $SL(2, R)$[18] and is the most convenient choice to take here. If, in the group (local) representation formula (Eq. B.7 of A), we set Re $j = -\frac{1}{2}$, restrict $m - m'$, $m + m'$ to be integers and α, z, γ to lie in the manifold of G, then the convergence restrictions (Eq. B.14 of A) disappear. This gives directly the principal series of unitary representations of G. The discrete series appear in conjunction with the finite-dimensional (nonunitary) representations of G when we set $m - m'$, $m + m'$, $j - m$ integers. For this case the representations appear in the fully reduced form described in Sec. 5.

From the completeness relation of Sec. 9, we obtain an elementary proof of the completeness of the above unitary representations of G. Setting $1 < z < \infty$ in (10.4) and using (5.1) we can express the sum as a contour integral

$$\frac{\left(\dfrac{1+t}{1+z}\right)^{\frac{1}{2}|m+m'|}\left(\dfrac{1-t}{1-z}\right)^{\frac{1}{2}|m-m'|}}{z - t}$$
$$= \frac{1}{2i}\int_{M-\frac{1}{2}+i\infty}^{M-\frac{1}{2}-i\infty} dj(2j + 1)\,\frac{d_j^{m,-m'}(-t)e_j^{m,m'}(z)}{\sin\pi(j - m)}$$
$$[\arg(1 - t) < \pi], \tag{14.2}$$

of the Sommerfeld–Watson type.

Let us now take the difference of the limits on each side of the boundary of the region of convergence in t of this integral, given by $|\arg(1-t)| = \pi$. In the integrand the dependence on t is given by $d_j^{m,-m'}(-t)$ whose discontinuity in $1 < t < \infty$ is obtained immediately from (3.1) as

$$d_j^{m,-m'}(-t - i\epsilon) - d_j^{m,-m'}(-t + i\epsilon)$$
$$= 2i\sin\pi(j - m)d_j^{m,m'}(t), \quad 1 < t < \infty, \tag{14.3}$$

where we have ignored the cuts in $(1 + t)^{\frac{1}{2}}$ and $(1 - t)^{\frac{1}{2}}$ which clearly cancel in (14.2).

In view of the relation

$$\{z - (t + i\epsilon)\}^{-1} - \{z - (t - i\epsilon)\}^{-1} = 2i\pi\delta(z - t),$$

this leads to

$$\delta(z - t) = \frac{1}{2\pi i}\int_{M-\frac{1}{2}+i\infty}^{M-\frac{1}{2}-i\infty} dj(2j + 1)\, d_j^{m,m'}(t)e_j^{m,m'}(z),$$
$$z, t \in [1, \infty]. \tag{14.4}$$

Using Table I we have

$$\delta(z - t) = \frac{1}{2\pi i}\int_{-\frac{1}{2}+i\infty}^{-\frac{1}{2}-i\infty} dj(2j + 1)\, d_j^{m,m'}(t)e_j^{m,m'}(z)$$
$$+ \frac{1}{2}\sum_{\substack{j=0,1,2,\cdots \\ \text{or } 1/2,3/2,5/2,\cdots}}^{M-|m-m'|} (2j + 1)\, d_j^{m,m'}(t)\, d_j^{m,m'}(z)$$

$$= \frac{1}{4i} \int_{-\frac{1}{2}+i\infty}^{-\frac{1}{2}-i\infty} dj\,(2j+1)\cot\pi(j-m) d_j^{m,m'}(t) d_j^{m,m'}(z)$$

$$+ \frac{1}{2} \sum_{\substack{j=0,1,2,\cdots \\ \text{or } 1/2,3/2,5/2,\cdots}}^{M-|m-m'|} (2j+1) d_j^{m,m'}(t) d_j^{m,m'}(z). \quad (14.5)$$

This may be compared with the Plancherel formula for G.[19-21] We thus obtain the following fundamental theorem:

Theorem. Let $f(\alpha, z, \gamma)$ be a square-integrable function defined on G, and $f^{m,m'}(z)$ the Fourier component obtained by projection onto $e^{im\alpha}$ and $e^{im'\gamma}$. Then we have the expansion

$$f^{m,m'}(z) = \frac{1}{2\pi i} \int_{-\frac{1}{2}+i\infty}^{-\frac{1}{2}-i\infty} dj (2j+1) \tilde{f}^{m,m'}(j)\, d_j^{m,m'}(z)$$

$$+ \frac{1}{2} \sum_{\substack{j=0,1,2,\cdots \\ \text{or } 1/2,3/2,5/2,\cdots}}^{M-|m-m'|} (2j+1) \tilde{\chi}^{m,m'}(j)\, d_j^{m,m'}(z), \quad (14.6)$$

where

$$\tilde{f}^{m,m'}(j) = \int_1^\infty e_j^{m,m'}(z) f^{m,m'}(z)\, dz, \quad \operatorname{Re} j = -\tfrac{1}{2} \quad (14.7)$$

and

$$\tilde{\chi}^{m,m'}(j) = \int_1^\infty d_j^{m,m'}(z) f^{m,m'}(z)\, dz \quad (14.8)$$

are the projections onto the principal and discrete series, respectively.

In the second paper of this series the theorem is extended to cover a much larger class of generalized functions on G, so as to include cases of physical interest.

[21] L. Ehrenpreis, and F. Mautner, Ann. Math. **61**, 406, (1955); Trans. Am. Math. Soc. **84**, 1, (1957); **90**, 435, (1959).

Statistical Mechanics of Quenched Solid Solutions with Application to Magnetically Dilute Alloys*

TOHRU MORITA†
Department of Physics, The Catholic University of America, Washington, D. C.
(Received 28 May 1964)

The arrangement of atoms in solid solutions and alloys, prepared at high temperatures and cooled nonadiabatically, is not the one which is thermodynamically most stable. In establishing theories of phenomena related to the internal degrees of freedom of such a system, such as magnetism, one must be careful to account for this nonequilibrium distribution of atoms. In this paper, systems are treated with the aid of a fictitious equilibrium system. This fictitious system is constructed such that its thermal equilibrium properties are the same as the properties of the non-thermal-equilibrium system. Thus one can treat nonequilibrium systems by applying well known thermal equilibrium techniques to the fictitious system.

The method is illustrated via the example of a magnetically dilute alloy. Brout's result for a very dilute Ising system is obtained with the aid of the theory of classical fluids, without collecting diagrams. A method for applying the higher approximations developed for classical fluids to the present problem is suggested; calculations and discussions of which are retained for a forthcoming paper.

INTRODUCTION

THE motion of atoms in solid solutions, or alloys prepared at high temperatures and cooled nonadiabatically, will be quenched at a certain temperature. The atoms will not be able to attain their thermal-equilibrium distribution at lower temperatures, although the system will come to thermal equilibrium with respect to other degrees of freedom, such as magnetism or lattice vibrations.

In order to investigate such a system one has to know the distribution of atoms as well as the potential energy,[1] while for a classical system in its equilibrium state, the latter knowledge is sufficient. Let X and x be the sets of variables which describe the distribution of atoms and the internal degrees of freedom, respectively. Then the probability of a system in its equilibrium state having a configuration (X, x) is known to be proportional to $\exp -\beta\Phi(X, x)$, where $\Phi(X, x)$ is the potential energy associated with the configuration (X, x), and $\beta = 1/kT$; k is the Boltzmann constant and T is the temperature of the system. Thus the equilibrium properties of the system can be calculated.

The system of interest in the present calculation is not in thermal equilibrium with respect to the distribution of X; though it is in equilibrium with respect to x for a fixed value of X. Let $P(X)$ be the probability of the atoms in the system of interest having configuration X. Then the probability of the configuration (X, x) will be proportional to

$$P(X)\cdot[\exp -\beta\Phi(X, x)/\textstyle\sum_x \exp -\beta\Phi(X, x)].$$

The second factor is the conditional probability for variable x in thermal equilibrium for a fixed value of X. Brout[2] and Mazo[3] discussed the calculation of free energy of such a system. They pointed out that one must be very careful in averaging over X and x, because of the two factors in the probability.

The purpose of this paper is to present another approach to the investigation of such systems. Let us define $\Psi(X)$ by

$$\exp -\beta\Psi(X) = \text{const} \times P(X)/\textstyle\sum_x \exp -\beta\Phi(X, x),$$

and consider a fictitious system described by the potential energy $\Psi(X) + \Phi(X, x)$. Then the probability of the fictitious system having an equilibrium configuration (X, x) is the same as the probability of configuration (X, x) in the non-thermal-equilibrium system. Hence to investigate the non-thermal-equilibrium system one must determine $\Psi(X)$ and investigate the equilibrium properties of the fictitious system.[4] The general method of doing this will be sketched in Sec. I, and the application of the method to magnetically dilute alloys will be given in Sec. II.

I. GENERAL THEORY

Solid solutions and alloys in which the distribution of the constituent atoms is given, but in which the constituent atoms have internal degrees of freedom, are systems of the type to be considered.

Let X and x be the sets of coordinates which

* This work was supported by the U. S. Office of Aerospace Research under Contract No. AF-AFOSR-445-63.

† On leave of absence from Shizuoka University, Shizuoka, Japan.

[1] For clear understanding, a classical system is considered in this introduction.

[2] R. Brout, Phys. Rev. 115, 824 (1959).

[3] R. M. Mazo, J. Chem. Phys. 39, 1224 (1963).

[4] This avoids the need for special cautions in averaging over X and x.

assign the distribution of atoms and the internal degrees of freedom, respectively. For instance, if the system is composed of N_1 atoms of species 1, N_2 atoms of species 2, $\cdots$, N_σ atoms of species σ, X stands for the set of all $N(=\sum_{\nu=1}^{\sigma} N_\nu)$ atoms: $R_1, \cdots, R_{N_1}$ for atoms of species 1, $R_{N_1+1}, \cdots, R_{N_1+N_2}$ for atoms of species 2, $\cdots$, $R_{N-N_\sigma+1}, \cdots, R_N$ for atoms of species σ. In the following the probability of configuration X, $P(X)$, is assumed to be prescribed; and $\int dX$ stands for

$$\frac{1}{\prod_{\nu=1}^{\sigma} N_\nu!} \int dR_1 \int dR_2 \cdots \int dR_N$$

or

$$\frac{1}{\prod_{\nu=1}^{\sigma} N_\nu!} \sum_{R_1} \sum_{R_2} \cdots \sum_{R_N},$$

according as the coordinates are continuous or discrete variables.

First consider a classical system which is described by the potential energy $\Phi(X, x)$. If one knew the distribution of atoms, X, for the system, the expectation value of any physical quantity $A(X, x)$ would be calculated from

$$[1/Z(X)] \sum_x A(X, x) \exp -\beta\Phi(X, x),$$

where $\beta = 1/kT$ as usual and

$$Z(X) = \sum_x \exp -\beta\Phi(X, x). \quad (1.1)$$

However, only the probability distribution of X, $P(X)$, is given; so that the expectation value of $A(X, x)$ for our system is given by the average of the above expression:

$$\langle A \rangle = \int dX\, P(X) \frac{1}{Z(X)} \sum_x A(X, x) \times \exp -\beta\Phi(X, x). \quad (1.2)$$

Introduce $\Psi(X)$ by[5]

$$P(X)/Z(X) = (1/Z) \exp -\beta\Psi(X), \quad (1.3)$$

where Z is a constant to be determined by the normalization condition of $P(X)$, so that

$$Z = \int dX\, Z(X) \exp -\beta\Psi(X). \quad (1.4)$$

This determines $\Psi(X)$ uniquely except for an arbitrary additive constant. Substituting Eqs. (1.3) and (1.1) into Eqs. (1.2) and (1.4), respectively, one obtains the following expression for $\langle A \rangle$:

$$\langle A \rangle = \int dX \sum_x \rho(X, x) A(X, x), \quad (1.5)$$

where

$$\rho(X, x) = Z^{-1} \exp -\beta[\Psi(X) + \Phi(X, x)] \quad (1.6)$$

and

$$Z = \int dX \sum_x \exp -\beta[\Psi(X) + \Phi(X, x)]. \quad (1.7)$$

This is the average value of $A(X, x)$ for a system described by the potential energy $\Psi(X) + \Phi(X, x)$. The properties of this system can be discussed by applying the methods of classical statistical mechanics, once $\Psi(X)$ is known. In order to apply the techniques of classical statistical mechanics it is convenient to determine $\Psi(X)$ in the form

$$\Psi(X) = \sum_{i=1}^{N} \psi_{\nu_i}^{(1)}(R_i) + \sum_{N \geq i > j \geq 1} \psi_{\nu_i \nu_j}^{(2)}(R_i, R_j) + \cdots. \quad (1.8)$$

Now it is convenient to consider the variational problem where the integral

$$\int dX \sum_x \rho_t(X, x)[\beta\Phi(X, x) + \ln \rho_t(X, x)] \quad (1.9)$$

is minimized with respect to variations of $\rho_t(X, x)$, under the conditions that

$$1 = \int dX \sum_x \rho_t(X, x) \quad (1.10)$$

and

$$P(X) = \sum_x \rho_t(X, x). \quad (1.11)$$

The variation can be easily taken by introducing a Lagrange multiplier for each restriction. If the Lagrange multipliers for Eqs. (1.10) and (1.11) are denoted as $\ln Z^* - 1$ and $\beta\Psi^*(X)$, then as the result of the variation of

$$\int dX \sum_x \rho_t(X, x)\{\beta\Phi(X, x) + \ln \rho_t(X, x)\} + (\ln Z^* - 1)\left\{\int dX \sum_x \rho_t(X, x) - 1\right\} + \int dX \beta\Psi^*(X)\{\sum_x \rho_t(X, x) - P(X)\}, \quad (1.12)$$

one obtains

$$\rho(X, x) = (1/Z^*) \exp -\beta[\Phi(X, x) + \Psi^*(X)], \quad (1.13)$$

where $\rho(X, x)$ is the $\rho_t(X, x)$ which minimizes integral (1.12). The Lagrange multipliers, Z^* and

[5] One should not confuse Z with $Z(X)$; Z will be shown to be the partition function for the fictitious system to be introduced below.

$\beta\Psi^*(X)$, are determined by

$$Z^* = \int dX \textstyle\sum_x \exp -\beta[\Phi(X, x) + \Psi^*(X)] \quad (1.14)$$

and

$$P(X) = \frac{1}{Z^*} \textstyle\sum_x \exp -\beta[\Phi(X, x) + \Psi^*(X)]. \quad (1.15)$$

Comparing Eqs. (1.13), (1.14), and (1.15) with Eqs. (1.6), (1.7), and (1.3), it is obvious that $\beta\Psi(X)$ can be considered as the Lagrange multiplier which secures the distribution to be the prescribed one, i.e., $P(X)$.

When the one-, two-, $\cdots$, particle distribution functions, $\rho_\nu^{(1)}(R)$, $\rho_{\nu\nu'}^{(2)}(R, R')$, and so on, are given, instead of $P(X)$, restriction (1.11) in the above variational problem should be replaced by

$$\rho_\nu^{(1)}(R) = \int dX \textstyle\sum_x \rho(X, x) \sum_{i=1}^N \delta_{\nu,\nu_i}\, \delta_{R,R_i}, \quad (1.16)$$

$$\rho_{\nu\nu'}^{(2)}(R, R') = \int dX \textstyle\sum_x \rho(X,x) \displaystyle\sum_{i=1}^{N} \sum_{\substack{j=1 \\ (i\neq j)}}^{N} \delta_{\nu,\nu_i}\, \delta_{\nu',\nu_j}\, \delta_{R,R_i}\, \delta_{R',R_j},$$

and so forth. If Lagrange multipliers for these restrictions are denoted by

$$\psi_\nu^{(1)}(R), \qquad \psi_{\nu\nu'}^{(2)}(R, R'), \text{ etc.},$$

then $\rho(X, x)$, which is the $\rho_t(X, x)$ that minimizes integral (1.9), is given by Eq. (1.6), in which $\Psi(X)$ is given by Eq. (1.8). Hence $\psi_\nu^{(1)}(R)$, $\psi_{\nu\nu'}^{(2)}(R, R')$, etc., introduced by Eq. (1.8), can be taken to be Lagrange multipliers which secure the distributions $\rho_\nu^{(1)}(R)$, $\rho_{\nu\nu'}^{(2)}(R, R')$, etc., to be equal to the given ones.

As a consequence, one finds that to investigate the properties of a system in which the distributions are prescribed, one must calculate the properties and the distribution functions of the system which has potential energy given by

$$[\Psi(X) \text{ expressed by Eq. (1.8)}] + \Phi(X, x).$$

Then $\psi_\nu^{(1)}(R)$, $\psi_{\nu\nu'}^{(2)}(R, R')$, $\cdots$ must be determined such that the distribution functions obtained are equal to those which are prescribed.

In treating quantum systems the Hamiltonian, $H(X, x)$, which contains X as a parameter, will be given. For such a system one can argue in the same way as was done above by replacing $\Phi(X, x)$ by $H(X, x)$, and $\sum_x$ by tr_x. In this case one must investigate the properties and distribution functions of a system which has a Hamiltonian given by

$$[\Psi(X) \text{ expressed by Eq. (1.8)}] + H(X, x).$$

II. MAGNETICALLY DILUTE ALLOYS

As an example, an Ising system of N magnetic atoms randomly distributed in a paramagnetic lattice of L sites will be considered. X will be taken as the set of coordinates of $N(=\rho L)$ magnetic atoms, $R_1, \cdots, R_N$, each of which runs over the L lattice sites; and x as the set of their spin variables, $s_1, \cdots, s_N$, which take on the values $+1$ or -1. In this notation the N-particle distribution function, $P(X)$, is independent of $R_1, \cdots, R_N$; and hence the one-, two-, $\cdots$, particle distribution functions are given by

$$\rho(R) = \rho = N/L, \quad (2.1)$$
$$\rho^{(2)}(R, R') = \begin{cases} \rho^2 & \text{for } R \neq R', \\ 0 & \text{for } R = R', \end{cases}$$

and so on, and $\Phi(X, x)$ is

$$\Phi(X, x) = -\sum_{i=1}^N H(R_i)s_i - \sum_{N\geq i>j\geq 1} J(R_i, R_j)s_is_j. \quad (2.2)$$

Following Eq. (1.8), $\Psi(X)$ is introduced by as

$$\Psi(X) = \sum_{i=1}^N \psi^{(1)}(R_i) + \sum_{N\geq i>j\geq 1} \psi^{(2)}(R_i, R_j) + \cdots. \quad (2.3)$$

Now, following Sec. I, the thermodynamic properties and distribution functions of the fictitious system described by the potential energy

$$[\Psi(X) \text{ given by Eq. (2.3)}] + [\Phi(X, x) \text{ given by Eq. (2.2)}]$$

must be investigated. Then $\psi^{(1)}(R)$, $\psi^{(2)}(R, R'), \cdots$, must be determined such that the distribution functions obtained are equal to those given by Eq. (2.1).

The grand partition function, Ξ, for the fictitious system is given by[6]

$$\Xi = \sum_{N=0}^\infty \frac{z^N}{N!} \sum_{R_1=1}^L \cdots \sum_{R_N=1}^L \sum_{s_1=\pm1} \cdots \times \sum_{s_N=\pm1} \exp -\beta[\Psi(X) + \Phi(X, x)], \quad (2.4)$$

where z is the fugacity, which is to be determined such that the total number of particles calculated is equal to the given value, N. The functions $z_s^*(R)$, $z^*(R)$ and $b_{ss'}(R, R')$ are introduced by

$$N \ln z - \beta[\Psi(X) + \Phi(X, x)] = \sum_{i=1}^N \ln z_{s_i}^*(R_i) + \sum_{N\geq i>j\geq 1} \ln [b_{s_is_j}(R_i, R_j) + 1] + \cdots, \quad (2.5)$$

[6] For mathematical convenience, the grand canonical ensemble is used in this part.

$$\begin{aligned}\ln z_s^*(R) &= \ln z^*(R) + \beta H(R)s,\\ \ln z^*(R) &= \ln z - \beta\psi^{(1)}(R),\end{aligned} \tag{2.6}$$

$$\ln [b_{ss'}(R,R')+1] = -\beta\psi^{(2)}(R,R') + \beta J(R,R')ss',$$

$$b_{ss'}(R,R') = \exp -\beta\psi^{(2)}(R,R') \cosh \beta J(R,R') \times [1 + ss' \tanh \beta J(R,R')] - 1. \tag{2.7}$$

It is known that, using Eq. (2.4), the one- and two-particle distribution functions can be expressed as[7]

$$\rho_s(R) = \delta \ln \Xi / \delta \ln z_s^*(R) \tag{2.8}$$

and

$$\rho_{ss'}^{(2)}(R,R') = \delta \ln \Xi / \delta \ln [b_{ss'}(R,R')+1]. \tag{2.9}$$

It is also known that this set can be transformed to[7]

$$\ln \Xi = \sum_R \sum_s \rho_s(R)[\ln z_s^*(R) - \ln \rho_s(R) + 1]$$

+ Sum of all the more than singly connected diagrams composed of black circles $^\rho\bullet$ and bonds $\underline{\ \ b\ \ }$, (2.10)

$$0 = \delta \ln \Xi / \delta \rho_s(R), \tag{2.11}$$

and

$$\rho_{ss'}^{(2)}(R,R') = \delta \ln \Xi / \delta \ln [b_{ss'}(R,R')+1] \tag{2.12}$$

[cf. Eqs. (4.6)–(4.8) of CF]. The functions $\rho(R)$ and $\lambda(R)$ are introduced by

$$\rho(R) = \sum_s \rho_s(R) \quad \text{and} \quad \lambda(R) = \sum_s \rho_s(R)s \tag{2.13}$$

or

$$\rho_s(R) = \tfrac{1}{2}[\rho(R) + s\lambda(R)]. \tag{2.13'}$$

Using Eqs. (2.13) and (2.13′), Eq. (2.8) can be replaced by

$$\rho(R) = \delta \ln \Xi / \delta \ln z^*(R) \tag{2.8'}$$

and

$$\lambda(R) = \delta \ln \Xi / \delta\beta H(R),$$

and Eq. (2.11) can be replaced by

$$0 = \delta \ln \Xi / \delta\rho(R) \quad \text{and} \quad 0 = \delta \ln \Xi / \delta\lambda(R). \tag{2.11'}$$

Retaining only that diagram in Eq. (2.10) corresponding to a bond is analogous to retaining terms up to the second virial coefficient in the nonideal-gas expansion. In this approximation, substituting Eqs. (2.13′) and (2.7), one obtains

$$\begin{aligned}\ln \Xi &= \sum_R \sum_s \rho_s(R)[\ln z_s^*(R) - \ln \rho_s(R) + 1]\\ &\quad + \tfrac{1}{2} \sum_R \sum_{R'} \sum_s \sum_{s'} \rho_s(R)\rho_{s'}(R')b_{ss'}(R,R')\\ &= \sum_R \rho(R) \ln z^*(R) + \sum_R \beta H(R)\lambda(R)\\ &\quad - \tfrac{1}{2} \sum_R \sum_s [\rho(R) + s\lambda(R)]\\ &\quad \times \{\ln [\rho(R) + s\lambda(R)] - \ln 2 - 1\}\\ &\quad + \tfrac{1}{2} \sum_R \sum_{R'} \exp -\beta\psi^{(2)}(R,R') \cosh \beta J(R,R')\\ &\quad \times [\rho(R)\rho(R') + \lambda(R)\lambda(R') \tanh \beta J(R,R')]\\ &\quad - \tfrac{1}{2} \sum_R \sum_{R'} \rho(R)\rho(R').\end{aligned} \tag{2.14}$$

Substituting this expression into (2.11′) and (2.12), one gets

$$\begin{aligned}0 &= \ln z - \beta\psi^{(1)}(R) - \tfrac{1}{2} \ln [\rho(R)^2 - \lambda(R)^2]\\ &\quad + \ln 2 + 1 + \sum_{R'} \rho(R') \exp -\beta\psi^{(2)}(R,R')\\ &\quad \times \cosh \beta J(R,R') - \sum_{R'} \rho(R'),\end{aligned} \tag{2.15}$$

$$0 = \beta H(R) - \frac{1}{2} \ln \frac{\rho(R) + \lambda(R)}{\rho(R) - \lambda(R)} + \sum_{R'} \lambda(R') \times \exp -\beta\psi^{(2)}(R,R') \sinh \beta J(R,R'), \tag{2.16}$$

and

$$\begin{aligned}\rho_{ss'}^{(2)}(R,R') &= \rho_s(R)\rho_{s'}(R')[1 + b_{ss'}(R,R')]\\ &= \tfrac{1}{4}[\rho(R) + s\lambda(R)][\rho(R') + s'\lambda(R')]\\ &\quad \times \exp -\beta\psi^{(2)}(R,R') \cosh \beta J(R,R')\\ &\quad \times [1 + ss' \tanh \beta J(R,R')].\end{aligned} \tag{2.17}$$

The conditions that $H(R) = H$, $\lambda(R) = \lambda$ and

$$\rho(R) = \sum_s \rho_s(R) = \rho \tag{2.18}$$

and

$$\rho^{(2)}(R,R') = \sum_s \sum_{s'} \rho_{ss'}^{(2)}(R,R') = \begin{cases}\rho^2 & \text{for } R \neq R',\\ 0 & \text{for } R = R',\end{cases} \tag{2.19}$$

determine $\psi^{(2)}(R,R')$ as

$$\exp \beta\psi^{(2)}(R,R') = \cosh \beta J(R,R') \times [1 + (\lambda^2/\rho^2) \tanh \beta J(R,R')], \tag{2.20}$$

for $R \neq R'$ and $\psi^{(2)}(R,R) = \infty$. Substituting the thus determined $\psi^{(2)}(R,R')$ into Eq. (2.16), one obtains

$$0 = \beta H - \frac{1}{2} \ln \frac{1 + \lambda/\rho}{1 - \lambda/\rho} + \lambda \sum_{\substack{R' \\ (R' \neq R)}} \frac{\tanh \beta J(R,R')}{1 + (\lambda/\rho)^2 \tanh \beta J(R,R')}, \tag{2.21}$$

which is the expression determining the magnetization λ as a function of the external field H. This is

[7] T. Morita and K. Hiroike, Progr. Theoret. Phys. (Kyoto) **25**, 537 (1961), to be referred to as CF.

the same expression Brout[2] obtained by collecting many diagrams. The value of $\psi^{(1)}(R)$ can be determined by substituting the results of Eqs. (2.20) and (2.21) into Eq. (2.15).

The above approximation is considered applicable only to very dilute systems in which $J(R, R')$ is of short range. In order to consider systems of higher densities, or in which $J(R, R')$ is of longer range, one must use the higher approximations applicable to less dilute gases. This can readily be done if $\psi^{(3)}$, $\psi^{(4)}$, etc. are neglected. In this approximation the above system is analogous to a multicomponent system in which the equations determining the one- and two-particle distribution functions $\rho_s^{(1)}(R)$ and $\rho_{ss'}^{(2)}(R, R')$ are given in the form

$$\ln z - \beta\varphi_s(R) = F_s^{(1)}\{R;\, \rho_s^{(1)}(R),\, \rho_{ss'}^{(2)}(R, R')\}, \quad (2.22)$$

$$-\beta\varphi_{ss'}(R, R') = F_{ss'}^{(2)}\{R, R';\, \rho_s^{(1)}(R),\, \rho_{ss'}^{(2)}(R, R')\}, \quad (2.23)$$

where $\varphi_s(R)$ and $\varphi_{ss'}(R, R')$ are the potential energies of the multicomponent system, and $F_s^{(1)}(R)$ and $F_{ss'}^{(2)}(R, R')$ are some functionals of $\rho_s^{(1)}(R)$ and $\rho_{ss'}^{(2)}(R, R')$. Applying this to the present problem, in which $\varphi_s(R) = -H(R)s - \psi^{(1)}(R)$ and $\varphi_{ss'}^{(2)}(R, R') = \psi^{(2)}(R, R') - J(R, R')ss'$, one obtains

$$\ln z + \beta H(R)s - \beta\psi^{(1)}(R) = F_s^{(1)}\{R;\, \rho_s^{(1)}(R),\, \rho_{ss'}^{(2)}(R, R')\}, \quad (2.24)$$

$$\beta J(R, R')ss' - \beta\psi^{(2)}(R, R') = F_{ss'}^{(2)}\{R, R';\, \rho_s^{(1)}(R),\, \rho_{ss'}^{(2)}(R, R')\}. \quad (2.25)$$

In addition to these six equations, one has

$$\rho(R) = \textstyle\sum_s \rho_s^{(1)}(R), \quad (2.26)$$

$$\rho^{(2)}(R, R') = \textstyle\sum_s \sum_{s'} \rho_{ss'}^{(2)}(R, R'). \quad (2.27)$$

The eight functions $\rho_s^{(1)}(R)$, $\rho_{ss'}^{(2)}(R, R')$, $\psi^{(1)}(R)$, and $\psi^{(2)}(R, R')$ can be determined from these eight equations.[8] In practice, it is more convenient to use the set of equations which are obtained by applying $\sum_s$ and $\sum_s s$ to Eq. (2.24) and $\sum_s \sum_{s'}$, $\sum_s \sum_{s'} s$, $\sum_s \sum_{s'} s'$ and $\sum_s \sum_{s'} ss'$ to Eq. (2.25). The equations which are obtained by applying $\sum_s$ and $\sum_s \sum_{s'}$ to Eq. (2.24) and Eq. (2.25), respectively, are equations determining $\psi^{(1)}(R)$ and $\psi^{(2)}(R, R')$. The other six equations determine $\rho_s^{(1)}(R)$ and $\rho_{ss'}^{(2)}(R, R')$. The expressions for $F_s^{(1)}(R)$ and $F_{ss'}^{(2)}(R, R')$ are obtained from the right-hand sides of Eqs. (4.22) and (4.18) of CF by using the translation indicated in Sec. 6 of CF, which must be done to apply them to the case of a multicomponent system. The Bethe approximation is obtained by neglecting the contribution from the diagrams in Eqs. (4.22) and (4.18) of CF. The hypernetted chain approximation can be obtained by taking into account the contributions from the chain and ring diagrams in these equations. Any approximation in which equations determining $\rho_s^{(1)}(R)$ and $\rho_{ss'}^{(2)}(R, R')$ can be written in the forms of Eqs. (2.22) and (2.23) is applicable to the present problem. Calculations and discussions of such improved approximations will be the theme of a forthcoming paper.

ACKNOWLEDGMENTS

The author wishes to acknowledge his indebtedness to Professor T. Matsubara. Discussions with Professor Matsubara when the author was in the Research Institute for Fundamental Physics, Kyoto University, suggested the use of a fictitious system in investigating magnetically dilute alloys. The author also wishes to express his appreciation to Professor T. Tanaka for many helpful discussions and to Mr. R. Farrell for his help in preparing the manuscript.

[8] It should be noted that, up to this point, the considerations of this paragraph are not restricted to the Ising model of spin ½, but are also applicable to the Ising model of spin greater than ½. Moreover, this result is applicable for the usual Ising system by putting $\rho(R) = 1$ and $\rho^{(2)}(R, R') = 1 - \delta_{R,R'}$. This case will be discussed in a separate paper.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 10 OCTOBER 1964

Remarks on the Polynomial Boundedness in the Mandelstam Representation*

Y. S. JIN AND A. MARTIN†
Institute for Advanced Study, Princeton, New Jersey
(Received 19 May 1964)

The Mandelstam representation is a statement about the region of analyticity and asymptotic behavior (polynomial boundedness) of a scattering amplitude. In virtue of the unitarity condition, however, these two are not completely independent. Some physical consequences, e.g., uniqueness, polynomial boundedness of the total cross section, etc., which have been already derived from the Mandelstam representation, are shown to be preserved, even if the polynomial boundedness is replaced by a somewhat weaker assumption. By making use of unitarity, analyticity, and crossing symmetry, the following type of scattering amplitude $F = E + M$, where E is an entire function in both variables s and t, while M denotes a Mandelstam-type function with finite number of subtraction, is shown to be ruled out. Similarly, $F = EM$ is also ruled out, if one imposes the additional restriction that E should increase less fast than an exponential in one variable while the other is finite.

I. INTRODUCTION

IN the Mandelstam representation,[1] the following two distinct assumptions are made:

(A) analyticity of the scattering amplitude in two complex variables s and t, with singularities only at the poles given by

$$s = \text{const}, \qquad t = \text{const}, \quad \text{and} \quad u = \text{const},$$

and cuts given by

$$\begin{aligned} s &= 4\mu^2 + \lambda_1, \\ t &= 4\mu^2 + \lambda_2, \quad \lambda_1, \lambda_2, \lambda_3 \text{ real} > 0, \qquad (1) \\ u &= 4\mu^2 + \lambda_3, \end{aligned}$$

where $u = 4\mu^2 - s - t$. For simplicity, we consider the case of spinless equal-mass particles; however, the results of this paper can easily be extended to more general cases.

(B) boundedness of the scattering amplitude by a polynomial in s, t, and u for large values of two or three variables among s, t, and u. This latter assumption enables one to write explicitly the Mandelstam representation, as a sum of three double dispersion integrals plus a finite number of subtraction terms including one-dimensional integrals and polynomial. It is clear that these two assumptions are independent. However, the physical amplitude being submitted to the unitarity condition, they are no longer completely independent of each other, e.g., the polynomial boundedness in the physical region follows from a much weaker assumption than (B).

The question, which we wish to investigate in this paper, is to see to what extent one can release partly or completely the assumption (B). In fact, Mandelstam[2] himself has recently found a class of graphs whose sum does not satisfy (B). It has also been suggested by Freund and Oehme[3] to introduce an entire function (with faster growth than any polynomial in some directions) in the Mandelstam representation in order to explain high-energy scattering data.

Of course, we can no longer write down explicitly the Mandelstam representation, if the assumption (B) is dropped. However, by making use of the assumption (A) together with the unitarity, many of the physical consequences of the Mandelstam representation turn out to be still valid.[4] For instance, the scattering amplitude is still analytic in the cut t-plane, for fixed s. Then, for fixed s real $> 4\mu^2$, we can define an absorptive part of the scattering amplitude $A_s(s, t)$ and define the double spectral function in $s > 0$, $t > 0$ as the discontinuity of $A_s(s, t)$ across the positive t-cut.

The unitarity condition, for s below the first inelastic threshold,

$$A_s(s, \cos\theta_{12}) = \int d\Omega_3 F(s, \cos\theta_{13}) F^*(s, \cos\theta_{32}), \qquad (2)$$

can still be analytically continued outside the physical region, by inserting in the right-hand side a finite contour integral for F. Thus the Mandelstam equation[1] for the spectral function in the s elastic strip will be preserved. Furthermore, the double spectral

* This study was supported by the U. S. Air Force Office of Scientific Research, Grant No. AF-AFSOR-42-64.
† On leave of absence from CERN Theory Division, Geneva, Switzerland.
[1] S. Mandelstam, Phys. Rev. **112**, 1344 (1958).
[2] S. Mandelstam, Nuovo Cimento **30**, 1148 (1963).
[3] P. G. O. Freund and R. Oehme, Phys. Rev. Letters **10**, 450 (1963).
[4] See, for instance, A. J. Dragt and R. Karplus, J. Math. Phys. **5**, 120 (1964).

function defined to be the discontinuity of $A_s(s, t)$ for $t > 0$ will still coincide with that one defined as the discontinuity of $A_t(s, t)$ for $s > 0$. So we still get the familiar picture of the support of the spectral functions, with the same boundary curves as in the ordinary Mandelstam representation.

In view of the vast possibilities of replacing the condition (B) by a weaker one, we want to consider the following two simple ways of alteration of (B), which we nevertheless hope, would turn out to be useful in further discussions of more general situations:

(1) We release the polynomial boundedness assumption only when all three variables s, t, and u go to fininity. We maintain that for fixed t the scattering amplitude is still bounded by a polynomial in s, etc. Then, most of the consequences of the Mandelstam representation are saved. For instance, the one-dimensional dispersion relations hold, the Froissart bound[5] holds, and the theorem on the uniqueness of the scattering amplitude,[6] when the spectral function is given in a substrip of the elastic strip of one channel, is still valid. An open question is whether one can still write the partial-wave dispersion relations.

(2) We release (B) by introducing an entire function. It turns out that this kind of modification is very difficult to make in such a way that the unitarity is preserved. So far, we have only been able to treat the following two rather extreme cases:

(a) $$F = E + M,$$

(b) $$F = EM,$$

where F denotes the scattering amplitude and M denotes ordinary Mandelstam-like expression, while E is an entire function in both variables s and t. (a) will turn out to be inconsistent with the elastic unitarity, and should be rejected. (b) will also be rejected, if $|E(s, t)| < \exp \lambda |s|^{\alpha} (\alpha < 1)$ for $t < t_0 (t_0 > 0)$ and the same inequalities with circular permutation among s, t, and u hold. In the proof, the positiveness of the absorptive part of the forward scattering amplitude plays an essential role and only restricted analyticity in an ellipse in the t plane is used. Clearly, these two rather extreme cases do not cover all the possibilities of introducing entire functions; however, more conscientious exploitation of the given information might lead to a further limitation on the possibility of modifying the Mandelstam representation.

[5] M. Froissart, Phys. Rev. **123**, 1053 (1961).
[6] A. Martin, Phys. Rev. Letters **9**, 410 (1962).

II. POLYNOMIAL BOUNDEDNESS IN ONE VARIABLE

First of all, let us state that the Froissart bounds[5] for the scattering amplitude:

$$\begin{aligned} |F(s, \cos\theta = 1)| &< Cs \log^2 s, \\ |F(s, \cos\theta)| &< Cs^{\frac{3}{4}}(\log s)^{\frac{3}{4}}, \quad \text{for} \quad \epsilon < \theta < \pi - \epsilon \end{aligned} \tag{3}$$

are preserved, if, for $t = t_0 (0 < t_0 < 4\mu^2)$, the scattering amplitude is bounded by a polynomial in s. This is obvious from the derivation of the Froissart bounds presented by one of us.[7] However, the more refined nonforward bound[8]

$$|F(s, \cos\theta)| < (\log s)^{\frac{3}{2}} \quad \text{for} \quad \epsilon < \theta < \pi - \epsilon \tag{4}$$

no longer holds, since it is assumed in its derivation that the polynomial boundedness in s holds for $|\cos\theta| = |1 + t/2k^2| < C (C > 1)$. This assumption implies that the amplitude should be polynomial bounded, when both $|s|$ and $|t| \to \infty$ with $|s| \propto |\frac{1}{2}t|$.

Now we come to the uniqueness theorem.[6] Let us remind the readers that the scattering amplitude will be completely fixed if there are pure elastic regions in two channels, and if the double spectral functions $\rho(s, t)$ and $\sigma(s, u)$ are known in $4\mu^2 < s < 4\mu^2 + \epsilon$.

The first step of the proof of the uniqueness consists in showing that, apart from a finite polynomial in t, the scattering amplitude is completely fixed. This is established by making use of the fixed energy dispersion relation which still holds here, and the analyticity of the partial-wave amplitude in the angular momentum plane for $\mathrm{Re}\, l > L$. This follows also from the fixed energy dispersion relation and therefore is still preserved. In the course of the proof, an ambiguity is removed by using fixed t dispersion relation for the absorptive part in the t channel. This is again the case here.

The second step consists in showing that the arbitrary polynomial in t with coefficients depending on s, has a degree at most one in t, by making use of the Froissart bound in that channel. Hence, it causes no problem here. Thus the arbitrary polynomial is now reduced to the form $A + Bs + Ct + Dst$, by making use of the crossing symmetry. Now, from the unitarity in terms of partial-wave amplitudes, B, C, and D are shown to be zero. Besides, A being nonzero leads to an anomalous singularity of the scattering amplitude at $s = 0$. At this point,

[7] A. Martin, in Strong Interactions and High Energy Physics, *Lecture Notes at the Scottish Universities Summer School*, edited by R. G. Moorehouse (Oliver and Boyd, Ltd., Edinburgh, 1963); A. Martin, Phys. Rev. **129**, 1432 (1963).
[8] T. Kinoshita, J. J. Loeffel, and A. Martin, Phys. Rev. Letters **10**, 460 (1963).

it is now clear that in the proof the polynomial boundedness in both s and t has never been used. Thus the uniqueness theorem is also valid in the more general framework presented here.

An open question is whether the partial-wave amplitudes satisfy dispersion relations with a finite number of subtractions. The left-hand-cut discontinuity can still be calculated from the double spectral functions; however, its energy dependence is not at all transparent.

III. ADDITIVE ENTIRE FUNCTIONS

In this section, we shall show that a scattering amplitude $F(s, t, u)$ can not have the form

$$F(s,t,u) = E(s,t,u) + M(s,t,u), \tag{5}$$

where $E(s, t, u)$ is an entire function in both s and t, and $M(s, t, u)$ in an ordinary Mandelstam-like function. Of course, F *must satisfy the unitarity, but not* M. In fact M could be of the type considered in the preceding section, i.e., bounded by a polynomial only when one variable is held fixed, as it will be seen in the course of the following proof.

First of all, we notice that E is a *real* entire function, i.e., for s and t real, E is real, since F has to be real in the so-called "Euclidian" region below the thresholds of each channel and M is also supposed to be real there.

Assuming that there exists a pure elastic interval $s_0 < s < s_1$ in the s channel, the unitarity reads

$$\operatorname{Im} F(s, \cos\theta_{12}) = \int d\Omega_3 \{M(s, \cos\theta_{13}) + E(s, \cos\theta_{13})\}\{M^*(s,\cos\theta_{32}) + E(s,\cos\theta_{32})\} \tag{6}$$

in this interval. From the reality of $E(s, t, u)$, we have

$$\operatorname{Im} F(s, \cos\theta_{12}) = \operatorname{Im} M(s,\cos\theta_{12}). \tag{7}$$

It is clear from (6) that Im F can be analytically continued into the entire cut t plane as well as the absorptive part $A_s(s, t)$ in the ordinary Mandelstam representation. In addition to this, Im F is bounded by a polynomial in the cut t plane, as it follows from (7).

Let us now expand M and E in the partial waves:

$$\begin{aligned} M(s,\cos\theta) &= (s^{\frac{1}{2}}/k)\sum(2l+1)m_l(s)P_l(\cos\theta), \\ E(s,\cos\theta) &= (s^{\frac{1}{2}}/k)\sum(2l+1)e_l(s)P_l(\cos\theta). \end{aligned} \tag{8}$$

Then we have

$$\overline{\lim_{l\to\infty}}\,(m_l)^{1/l} = \alpha, \qquad \overline{\lim_{l\to\infty}}\,(e_l)^{1/l} = 0, \tag{9}$$

where α is related to the distance of the nearest singularity of M to the physical region in the $\cos\theta$ plane. The second equation of (9) follows from the fact that the partial wave expansion of E converges in an arbitrarily large ellipse, since E is an entire function. Substituting (8) into (6), we get

$$\operatorname{Im} F(s,\cos\theta_{12}) = \int d\Omega_3 M(s,\cos\theta_{13})M^*(s,\cos\theta_{32}) + \sum_l (2l+1)[m_l(s) + m_l^*(s) + e_l(s)]e_l(s)P_l(\cos\theta_{12}). \tag{10}$$

From (9) it is clear that the series in the right-hand side converges in the entire $\cos\theta_{12}$-plane. The integral is the standard one encountered in the ordinary Mandelstam representation. M being analytic and polynomial bounded in the cut plane, the integral can be shown to represent an analytic function bounded by a polynomial in the same cut plane, while by (7) the left-hand side is also polynomial-bounded. Consequently, the series in (10) which is an entire function in $\cos\theta_{12}$ must be also polynomial-bounded in all complex directions of $\cos\theta_{12}$. Hence, it is in fact a polynomial and the series must stop, i.e.,

$$(m_l + m_l^* + e_l)e_l = 0, \quad \text{for} \quad l > L. \tag{11}$$

We now proceed to show that from (11) follows necessarily $e_l = 0$. For this purpose, let us consider the following two situations separately:

(a) The amplitude has a pole at fixed t. Then, for a sufficiently large l, it can be shown that the pole contribution to 2 Re m_l will dominate over that from the cut, so that $\lim_{l\to\infty}|m_l|^{1/l} = \alpha$, $\alpha \neq 0$. However, since $\lim_{l\to\infty}(e_l)^{1/l} = 0$, we must conclude that $e_l = 0$ for sufficiently large l.

(b) There is no fixed pole. Then, from the positiveness of the absorptive part, it is obvious that there exists a finite interval starting from the boundary of the spectral function, where $\rho(s, t)$ is positive.[9] In the completely symmetric case like ours, we have

$$\operatorname{Im} m_l(s) = C(s)\int Q_l(1 + t/2k^2)\rho(s,t)\,dt, \tag{12}$$

and from the polynomial boundedness of $\rho(s, t)$ for fixed s, we can show that for large enough l the contribution from the region of small t values dominates in (12). From this, it is not difficult to get

$$\lim_{l\to\infty}(\operatorname{Im} m_l)^{1/l} = \alpha. \tag{13}$$

On the other hand, from unitarity,

$$(\operatorname{Re} m_l + e_l)^2 = \operatorname{Im} m_l - (\operatorname{Im} m_l)^2. \tag{14}$$

[9] G. Mahoux and A. Martin, Nuovo Cimento. **33**, 883 (1964).

Consequently, we get $\lim_{l\to\infty} (m_l + m_l^* + e_l)^{1/l} \neq 0$, and it follows $e_l = 0$ for sufficiently large l. Hence, $E(s, t)$ is actually a polynomial in t, for any $s_0 < s < s_1$. By analyticity, it is also true for any s. Then, with the assumption of the existence of a pure elastic t region, by crossing $E(s, t)$ is shown to be polynomial also in s. Therefore, any additive entire function should be rejected.

IV. MULTIPLICATIVE ENTIRE FUNCTIONS

In this section, we shall investigate the following form of the scattering amplitude

$$F(s, t, u) = E(s, t, u)M(s, t, u). \tag{15}$$

It appears that we are forced to restrict ourselves to the case where E is order less than one in s for fixed t, if we want to make use of the Phragmén–Lindelöf theorem. Here again, $E(s, t, u)$ is a *real* entire function of s and t. To be definite, we shall assume

$$|E(s, t, u)| < \exp \lambda |s|^\alpha \quad \text{with } \alpha < 1, \quad \text{for } |t| < t_0, \tag{16}$$

where $0 < t_0 < 4\mu^2$, while $|M(s, t, u)| < |s|^N$ and circular permutations of these conditions. We now start with the upper bound on the forward scattering amplitude, using essentially the same technique as the one used to get the Froissart bound in the case of $E \equiv 1$.[7]

(a) Upper bound on $F(s, 0, 4\mu^2 - s)$. First of all, we make the following partial wave expansions:

$$F(s, t, u) = \frac{s^{\frac{1}{2}}}{k} \sum_l (2l + 1)f_l(s)P_l\left(1 + \frac{t}{2k^2}\right), \tag{17}$$

$$\operatorname{Im} F(s, t, u) = \frac{s^{\frac{1}{2}}}{k} \sum_l (2l + 1) \operatorname{Im} f_l(s)P_l\left(1 + \frac{t}{2k^2}\right), \tag{18}$$

where k is the center-of-mass momentum. From (16), we get

$$(2l + 1) \operatorname{Im} f_l < \frac{|s|^N \exp \lambda |s|^\alpha}{P_l(1 + t/2k^2)}, \tag{19}$$

while the unitarity condition gives

$$|f_l|^2 < \operatorname{Im} f_l < 1. \tag{20}$$

Combining (19) and (20), we establish that only $Cs^{3/2}$ partial waves contribute appreciably to the forward scattering amplitude, and consequently we obtain

$$|F(s, 0, 4\mu^2 - s)| < C |s|^3, \tag{21}$$

and more generally

$$|F(s, t, 4\mu^2 - s - t)| < C |s|^3, \quad \text{for } t < 0.$$

By applying the Phragmén–Lindelöf theorem, this helds aslo for all compex s directions.

(b) Lower bound on $M(s, 0, 4\mu^2 - s)$. In order to get a lower bound on $M(s, 0, 4\mu^2 - s)$, we shall use the technique of the Herglotz function.[10] To begin with, we prove that $\operatorname{Im} F(s, 0, 4\mu^2 - s)$ for $s > 4\mu^2$ never vanishes. If it does so, by unitarity, the amplitude itself for any scattering angle at that energy must vanish. However, from the unitarity in the t-channel, we can show that there is a region where the double spectral functions are positive-definite for each energy.[9] Hence, we have

$$\operatorname{Im} F(s, 0, 4\mu^2 - s) > 0 \quad \text{for} \quad s > 4\mu^2.$$

In addition to this, from (15) and the reality of $E(s, t, u)$, we have

$$\operatorname{Im} F(s, t, u) = E(s, t, u) \operatorname{Im} M(s, t, u). \tag{22}$$

Therefore, $E(s, t, u)$ can never vanish for $t = 0$ and $4\mu^2 < s \leq \infty$, and

$$\operatorname{Im} M(s, 0, 4\mu^2 - s) > 0 \quad \text{for} \quad s > 4\mu^2.$$

By the same argument in the u channel, it turns out that $\operatorname{Im} M(s, 0, 4\mu^2 - s)$ keeps a constant sign for $s < 0$ too. As shown in a previous paper,[11] $M(s, 0, 4\mu^2 - s)$ being polynomial bounded and having a discontinuity with constant sign across each cut, this can now be written essentially as the product of a Herglotz function by a polynomial, namely,

$$M(s, 0, 4\mu^2 - s) = H(s)P_N(s), \tag{23a}$$

for the case where the discontinuities on both cuts have the same sign, and

$$M(s, 0, 4\mu^2 - s) = (1/s)H(s)P_N(s), \tag{23b}$$

otherwise, where $H(s)$ is a Herglotz function with $\operatorname{Im} H(s)/\operatorname{Im} s > 0$ and $P_N(s)$ denotes a polynomial in s. Consequently, we get the following lower bound:

$$|M(s, 0, 4\mu^2 - s)| > C \frac{|\operatorname{Im} s|}{|s|^3} \quad \text{for} \quad |s| > |s_0|, \tag{24}$$

where s_0 is a sufficiently large number, since any Herglotz function satisfies[10]

$$|H(s)| > C \frac{|\operatorname{Im} s|}{|s|^2}. \tag{25}$$

Now combining (21) and (24), we get

$$E(s, 0, 4\mu^2 - s) < C \frac{|s|^6}{|\operatorname{Im} s|},$$

[10] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments* (American Mathematical Society, New York, 1943), p. 23.

[11] Y. S. Jin and A. Martin, Phys. Rev. **135**, B 1369 (1964).

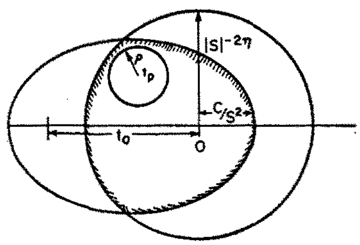

FIG. 1.

and it follows that

$$E(s,\,0,\,4\mu^2 - s) < C\,|s|^5, \tag{26}$$

in any complex direction in the s plane. Hence, $E(s, 0, 4\mu^2 - s)$ is a polynomial in s of degree at most 5.

(c) Power series expansion of $E(s, t)$. The result obtained above can now be extended to all the derivatives of $E(s, t)$ with respect to t at $t = 0$. From the definition (15), we have

$$\left(\frac{\partial F}{\partial t}\right)_s = M\left(\frac{\partial E}{\partial t}\right)_s + E\left(\frac{\partial M}{\partial t}\right)_s,$$

for $s > 4\mu^2$. Evidently, $(\partial M/\partial t)_s$ is polynomial-bounded. Moreover, $(\partial F/\partial t)_s$ can be shown to be also polynomial bounded, by means of summing the partial-wave expansion. For $s < 0$, we use

$$M\left(\frac{\partial E}{\partial t}\right)_s = M\left(\frac{\partial E}{\partial t}\right)_u - M\left(\frac{\partial E}{\partial u}\right)_t.$$

Then, in both cases, it is now clear that $M(\partial E/\partial t)_s$ is polynomial bounded. Hence, from the lower bound on M given by (24) in any complex direction, we conclude that $(\partial E/\partial t)_s$ is a polynomial in s. One can repeat the same argument to all the derivatives. Thus the power-series expansion of $E(s, t)$ around $t = 0$ turns out to be

$$E(s,\,t) = \sum_n t^n P_n(s), \tag{27}$$

where the P_n's are polynomials in s. By crossing, we get also

$$E(s,\,t) = \sum_m s^m Q_m(t), \tag{28}$$

where the Q_m's are polynomials in t. This result, unfortunately, is not yet sufficient to show that $E(s, t)$ is actually a polynomial in both variables.

(d) Bound on $E(s, t)$ for complex s. First of all, we shall show that there exists an ellipse in the t plane with foci $(-t_0, 0)$ and semimajor axis $\frac{1}{2}t_0 + C/|s|^2$, in which $F(s, t)$ is bounded by $|s|^3$. From the assumption (16), $|F(s, t)| < \exp \lambda\,|s|^\alpha$ inside an ellipse with foci $(-t_0, 0)$ and semimajor axis $\frac{1}{2}t_0$. Mapping the segment $(-t_0, 0)$, where (21) holds, on a circle, we map simultaneously the ellipse on a concentric circle, and we are in a position to apply Hadamard's three-circle theorem.[12] Thus we can find an intermediate circle on which the upper bound of F is $|s|^3$. By mapping back this circle to the t plane, we get the desired ellipse.

We now proceed to prove the following theorem.

Theorem. In the intersection of the ellipse with foci $(-t_0, 0)$ and semimajor axis $\frac{1}{2}t_0 + C/2|s|^2$ with a circle $|t| < |s|^{-\epsilon}$, we have

$$|E(s,\,t)| < |s|^Q$$

for $|s| > s_0$ and $|\mathrm{Im}\,s| > |s|^{-N}$.

Proof: From (24), it is clear that

$$|M(s,\,0)| > |s|^{-P} \text{ for } |s| > s_0 \text{ and } |\mathrm{Im}\,s| > |s|^{-P+2},$$

and $|M(s, t)| < s^N$ for $|t| < t_0$ by assumption. Consider the intersection of the ellipse in which $|F| < |s|^3$ holds and the circle $|t| < |s|^{-2\eta}$, where η is arbitrary but sufficiently small positive number. Then, around any point t_p whose minimum distance to the boundary of the intersection is larger than $|s|^{-3}$, we can draw a circle with radius $\rho > |s|^{-3}$, lying entirely inside the intersection (see, Fig. 1). Now, according to Theorem A2 in the Appendix, there exists a circle $|t - t_p| = r < \rho$, on which $|M(s, t)| > |s|^{-Q'}$. Hence we have

$$|E(s,\,t)| = |F(s,\,t)/M(s,\,t)| < |s|^{3+Q'}.$$

This implies, by the maximum modulus principle, that at the center of the circle $E(s, t) < |s|^{3+Q'}$. It is evident for $|s| > s_0$ and $|\mathrm{Im}\ s| > |s|^{-P+2}$ that the region in the t plane, where $|E(s, t)| < |s|^{3+Q'}$ holds, contains the intersection of the circle $|t| < |s|^{-3\eta}$ and the ellipse with foci $(-t_0, 0)$ and semimajor axis $1 + C/2s^2$. Thus, in particular, we have

$$E(s,\,-1/|s|^\epsilon) < |s|^Q$$

for $|s| > s_0$ and $|\mathrm{Im}\ s| > |s|^{-P'}$.

If it were not for the restriction in the s plane, i.e., $|\mathrm{Im}\,s| > |s|^{-P'}$, this would have been sufficient to our purpose. In the next subsection, we shall remove this restriction.

(e) Bound on $E(s, t)$ for real s. We now consider the function $E(s, -1/s^\epsilon)$ in the right half-plane $\mathrm{Re}\ s > 0$. This function is certainly an analytic function of s for $|s| > s_0$ and $\mathrm{Re}\ s > 0$. Let us con-

[12] E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, Oxford, England, 1939), p. 172.

sider this function in the region $|s| > s_0$ and $|\arg s| < |s|^{-P}$. On the boundary of the region $\arg s = \pm|s|^{-P}$, this function becomes

$$E(|s|\, e^{\pm i|s|^{-P}}, \ -|s|^{-\epsilon}\, e^{\mp i\epsilon|s|^{-P}}).$$

By choosing P larger than 2, we can manage to get t inside the region of polynomial boundedness. We are now in a very comfortable position to be able to apply the Phragmen–Lindelöf theorem to this region, since both boundary curves in the s plane join together at infinity and E is bounded by $\exp \lambda\ |s|^{\alpha}$ ($\alpha < 1$). Thus it follows immediately

$$E(s, \ -1/|s|^{\epsilon}) < |s|^{Q} \quad \text{also for arg} \quad s = 0.$$

Exactly the same argument can be extended to $0 < |\arg s| < |s|^{-P}$. The left half-plane can also be treated in the same manner. Our final conclusion is

$$|E(s, \ -1/|s|^{\epsilon})| < |s|^{Q} \quad \text{for all} \quad |s| > s_0. \tag{29}$$

With this, we are now finally in a position to prove that $E(s, t)$ is in fact a polynomial in both variables. Let us consider the expansion

$$E(s, \ -|s|^{-\epsilon}) = \sum_n Q_n(|s|^{-\epsilon})s^n, \tag{30}$$

where the Q_n's are polynomials. Applying the Cauchy theorem, we get

$$\frac{1}{2\pi}\int_0^{2\pi} e^{-in\theta}E(se^{i\theta}, \ -|s|^{-\epsilon})\, d\theta = |s|^n Q_n(|s|^{-\epsilon}) < |s|^{Q}. \tag{31}$$

We now expand $Q_n(t)$ in power series, i.e.,

$$Q_n(t) = \sum_m a_{mn}t^m. \tag{32}$$

$Q_n(t)$ being a polynomial in t, the only way to satisfy (31) is to have

$$a_{mn} = 0 \quad \text{for} \quad n > m\epsilon + Q. \tag{33}$$

However, by simply exchanging s and t, we also get

$$a_{mn} = 0 \quad \text{for} \quad m > n\epsilon + Q. \tag{34}$$

We can see that (33) and (34) lead to a contradiction, since we may choose $\epsilon < 1$, i.e., suppose $a_{mn} \neq 0$ for sufficiently large m or n, then

$$m\epsilon + Q > n > m/\epsilon - Q/\epsilon,$$

and hence

$$m(1/\epsilon - \epsilon) < Q(1 + 1/\epsilon), \ \ n(1/\epsilon - \epsilon) < Q(1 + 1/\epsilon).$$

Therefore, $E(s, t)$ is a polynomial in both s and t.

V. CONCLUDING REMARKS

It is very hard to conclude, from the last two sections, whether an entire function may be introduced to the Mandelstam representation in a reasonable way, because those examples which were excluded do not cover the most general case. However, it is rather remarkable that to exclude both $F = E + M$ and $F = EM$ very little has really been used of the fact that M satisfies the Mandelstam representation. It is sufficient, for instance, that M be polynomial bounded in one variable, when the other is fixed. It is also clear that the presence of complex singularities, for $|s|$ and $|t|$ both larger than $16\mu^2$, will not alter the proofs that E is a polynomial in s and t, since in that case the one-dimensional dispersion relation in s for $t < t_0(0 < t < 4\mu^2)$ and the elastic unitarity in both channels are unaffected.

The main impression one gets from the present work is that one should be very cautious before proposing an alteration of polynomial boundedness in the Mandelstam representation. For instance, a multiplicative entire function must be of the order at least one, thereby excluding any possibilities of replacing the polynomial boundedness by a *slightly weaker* one, e.g., $s^{\log s}$, etc.

Another striking fact is the immense usefulness of the positiveness of the absorptive part of the forward scattering amplitude, which is implied by the unitarity. This point has already been emphasized elsewhere in different connections[11,13]; however, it is our feeling that the exploitation of this information has not yet been exhausted in previous works.

ACKNOWLEDGMENTS

The authors wish to thank Professor R. Oehme for his comments on the introduction of entire functions. We gratefully acknowledge the hospitality extended to us by Professor J. R. Oppenheimer at the Institute for Advanced Study.

APPENDIX

In this Appendix, we shall present the following two theorems, which are necessary for the proof of the theorem in the Sec. IV(d) in the text.

Theorem A1. $f(z)$ is analytic and bounded by $|s|^N$ in the unit circle $|z| < 1$, and $|f(0)| > |s|^{-P}$. Then, in the circle $|z| < \frac{1}{2}|s|^{-\eta}$ with arbitrarily small $\eta > 0$, we have

$$|f(z)| > |s|^{-Q} \prod_i |z - z_i|,$$

where the product $\prod_i$ extends over all zeros inside $|z_i| < |s|^{-\eta}$, which are shown to be *finite* bounded by a number *independent of* s.

[13] Y. S. Jin and A. Martin, Phys. Rev. **135**, B 1375 (1964).

Proof: First, from the Jensen theorem,[14] we have

$$\prod_i |z_i| > |s|^{-(N+P)}, \tag{A1}$$

where $z_1, z_2, \cdots, z_i, \cdots$ are all the zeros inside $|z| < 1$. It follows also that the number of zeros in the circle of radius r is given by

$$n(r) < \log\left(\max\left|\frac{f(e^{i\theta})}{f(0)}\right|\right)(-\log r)^{-1}. \tag{A2}$$

As we have $|f(z)| < |s|^N$ and $|f(0)| > |s|^{-P}$, it is clear that

$$n(|s|^{-\eta}) < \frac{(N+P)\log|s|}{\eta\log|s|} = K$$

$$\text{(const } \textit{independent} \text{ of } s). \tag{A3}$$

For the circle with $r = \frac{1}{2}$, we get

$$n(\tfrac{1}{2}) < \frac{(N+P)\log|s|}{\log 2}. \tag{A4}$$

Following Carathéodory,[15] we now decompose $f(z)$ as

$$f(z) = \phi(z)\prod_{i=1}^{p}\frac{z - z_i}{\frac{1}{2} - 2zz_i^*}, \tag{A5}$$

where the product extends over the zeros inside $|z| < \frac{1}{2}$. Then, it is evident that $|f(z)| = |\phi(z)|$ on the circle $|z| = \frac{1}{2}$ and $|\phi(z) < |s|^N$, while at the origin we have

$$\phi(0) = \frac{|f(0)|}{2^p\,|z_1 z_2 \cdots z_p|} > |f(0)| > |s|^{-P}. \tag{A6}$$

Hence, from Carathéodory's inequality,[16] it follows

$$|\phi(z)| > |s|^{-P}, \quad \text{for} \quad |z| < \tfrac{1}{4}, \tag{A7}$$

since $\phi(z)$ has no zeros in $|z| < \frac{1}{2}$. It is also clear that

$$\prod_{i=1}^{p}|\tfrac{1}{2} - 2zz_i^*| < (\tfrac{1}{2} + |s|^{-\eta})^{(N+P)\log|s|/\log 2}$$

$$= |s|^{-(N+P)}(1 + 2|s|^{-\eta})^{(N+P)\log|s|/\log 2} \tag{A8}$$

for $|z| < |s|^{-\eta}$ and $|s| > s_0$, so it is bounded by a polynomial in s. On the other hand, from (A1) and (A4), we obtain

$$\prod_{i=p_0}^{p}|z - z_i| > \prod_{p_0}^{p}(|z_i| - |z|)$$

$$= \prod_{p_0}^{p}\left(1 - \frac{|z|}{|z_i|}\right)\prod_{p_0}^{p}|z_i|$$

$$> |s|^{-N-P}(1 - \tfrac{1}{2})^{(N+P)\log|s|/\log 2}, \tag{A9}$$

where the product $\prod_{p_0}^{p}$ runs over all zeros z_i in the ring $|s|^{-\eta} < |z_i| < \frac{1}{2}$. Substituting (A7)–(A9) into (A5), we finally arrive at

$$|f(z)| > |s|^{-Q}\prod_i|z - z_i| \quad \text{for} \quad |z| < \tfrac{1}{2}|s|^{-\eta}, \tag{A10}$$

where the product is extended over the finite number of zeros, bounded by a fixed number for $|s| > s_0$, inside the circle $|z| < |s|^{-\eta}$. By making use of this theorem we shall now prove the following theorem.

Theorem A2. Given any point P inside the circle $|z| < |s|^{-2\eta}$ for $|s| > s_0$ and given ρ such that $|s|^{-3} < \rho < |s|^{-2\eta}$, there exists a circle of radius $\rho(1-\epsilon) < r < \rho$ around P, on which $|f(z)| > |s|^{-Q'}$ holds, where Q' depends only on η.

Proof: Take such a point P and draw the circles $|z - z_P| = \rho$ and $|z - z_P| = \rho(1-\epsilon)$. Since the number of zeros inside $|z| < |s|^{-\eta}$ is bounded by K defined by (A3) which is independent of s, we shall draw $(K+1)$ equidistant concentric circles around P between $|z - z_P| = \rho$ and $|z - z_P| = (1-\epsilon)\rho$. Then, in at least one of the rings so obtained, there are clearly no zeros. Let us now evaluate $f(z)$ on the median circle of this ring. For the minimum distance of z on this circle to a zero is $\epsilon\rho/2(K+1)$, by applying Theorem A1, we get

$$|f(z)| > |s|^{-Q}\prod_i|z - z_i|$$

$$> |s|^{-Q}\left(\frac{\epsilon\,|s|^{-3}}{2(K+1)}\right)^K \equiv Cs^{-Q}. \tag{A11}$$

since we have already taken the precaution to impose $\rho > |s|^{-3}$.

[14] E. C. Titchmarsh, Ref. 12, p. 125.
[15] C. Carathéodory, *Funktionentheorie* (Verlag Birkhäuser, Basel, 1950), Vol. II, p. 11.
[16] R. P. Boas, *Entire Functions* (Academic Press Inc., New York, 1954), p. 3.

Multiple Scattering of Waves. II. "Hole Corrections" in the Scalar Case*

J. G. FIKIORIS AND P. C. WATERMAN
Research and Advanced Development Division, Avco Corporation, Wilmington, Massachusetts
(Received 25 February 1964; final manuscript received 24 April 1964)

Scalar multiple scattering effects due to a random distribution of spheres are considered in detail. Transformation from a volume to a surface integral allows one to take account of the "hole corrections" involved in the equation of multiple scattering, and yields a secular equation for the propagation constant K of the composite medium. In the low-frequency limit a result is given which appears to be exact over the entire range $0 \leq \delta \leq 1$, where δ is the fractional volume occupied by scatterers. Also in this limit, the boundary conditions appropriate to the boundary of the composite medium are established from examination of the total transmitted and reflected fields.

I. INTRODUCTION

THE problem of multiple scattering of waves by a random spatial array of scatterers has been the subject of numerous investigations[1–5]; an excellent review of the literature has been given by Twersky.[3e] In 1945 Foldy[1] introduced the concept of "configurational average" and established the basis for nearly all subsequent formulations. His successful treatment of isotropic point scatterers was extended by Lax to anisotropic point scatterers[2] and later to finite scattering regions randomly distributed throughout a region of space.[3] Throughout this work, one attempts to describe multiple-scattering phenomena by the average values of certain quantities of interest, and this leads naturally to the concept of a composite or scattering medium characterized by certain bulk parameters, although such a concept has so far not been rigorously demonstrated in general. An heuristic integral equation for the exciting field was first given by Lax,[2b] employing a truncation approximation somewhat akin to the approximation employed by Foldy for the simpler case of isotropic scatterers.[1] Waterman and Truell[5] attempted to put the derivation on a somewhat more quantitative basis, and gave also a prescription for obtaining the total field by quadrature, once the exciting field has been found, including the presence of nonzero volume scattering regions. They then applied Lax's integral equation to the case of normal incidence on a half-space containing randomly distributed spheres. Kasterin's representation was employed for spherical waves in terms of the zeroth Hankel function,[6] and the "random analog" [(3.4) and (3.13) of I] of Kasterin's formalism for the periodic lattice of spheres was obtained.

The integral equation was evaluated employing a "disc-shaped exclusion region" (with thickness going to zero), leading to a result [(3.25) of I] for the complex propagation constant which had also been obtained earlier by Urick and Ament.[4] Twersky derived a corresponding result for arbitrary angle of incidence and arbitrary scatterers[3a] and discussed its various limitations; we mention other limitations subsequently.

The present paper starts with Lax's equation, employing spherical wavefunctions and correlations in sphere positions described by a distance of closest approach b of sphere centers, i.e., essentially with (3.4), (3.16), and (3.17) of I, where $P(\mathbf{r}_1)$ now means exclusion of a sphere of radius b when integrating. Because of the ensuing rather complicated transition region behavior at the boundary of the volume accessible to spheres, we obtain only an approximate solution of the integral equation, which can be regarded as the leading term in an iteration approach. This solution has the form of a system of algebraic equations from which the bulk propagation constant and expansion coefficients of the exciting field may be determined.

The acoustic limiting case of small spheres is then considered, in the process taking $kb \ll 1$. This limit corresponds to employing point scatterers with a spherical excluded volume, and we note that the argument following (3.17) of I, in favor of a disc rather than sphere exclusion, appears to be incorrect. There is thus no *a priori* reason for choos-

* Work supported by the Ballistic Systems Division of the United States Air Force.

[1] L. L. Foldy, Phys. Rev. **67**, 107 (1945).
[2] M. Lax, (a) Rev. Mod. Phys. **23**, 287 (1951); (b) Phys. Rev. **85**, 621 (1952).
[3] V. Twersky, J. Math. Phys. **3**, (a) 700; (b) 716; (c) 724 (1962); (d) J. Opt. Soc. Am. **52**, 145 (1962); (e) J. Res. NBS **64D**, 715 (1960); (f) J. Acoust. Soc. Am. **36**, 1314 (1964).
[4] R. J. Urick and W. S. Ament, J. Acoust. Soc. Am. **21**, 115 (1949).
[5] P. C. Waterman and R. Truell, J. Math. Phys. **2**, 512 (1961). Hereafter referred to as I.
[6] N. Kasterin, Amst. Koning. Akd. Wetens. Versl. **6**, 460 (1898).

ing one over the other, and one must ultimately resolve the question by appeal to experiment. The total field is also examined in this limiting example, where complexities of boundary region behavior can be neglected, and a complete medium description is found to be possible in terms of an effective density and compressibility.

The results obtained below for the propagation constant in this limit are in agreement with the earlier work of Rayleigh[7] and Kasterin,[6] both of whom considered periodic arrays of spheres. The formalism developed by Twersky employing a "schizoid" single scatterer is also applicable to the present problem and yields the same results in the low-frequency limit, provided that Twersky's generally unknown "available volume" V_a is identified with the volume left unoccupied by scatterers.

The vector extension of the present work has been performed, in application to the electromagnetic case. The procedure and results of this extension will be described in a subsequent paper.

II. SOLUTION OF THE INTEGRAL EQUATION

Waterman and Truell[5] examined the problem of scalar multiple scattering in terms of Lax's[2b] heuristic integral equation for the configurational average of the exciting field $\langle\psi^E(\mathbf{r}\mid\mathbf{r}_1)\rangle$ at field point $\mathbf{r}$, acting on a scatterer at $\mathbf{r}_1$;

$$\langle\psi^E(\mathbf{r}\mid\mathbf{r}_1)\rangle = \psi^{\text{inc}}(\mathbf{r}) + \int_\tau d\tau' n(\mathbf{r}'\mid\mathbf{r}_1)T(\mathbf{r}')\langle\psi^E(\mathbf{r}\mid\mathbf{r}')\rangle. \tag{2.1}$$

In this equation, $n(\mathbf{r}\mid\mathbf{r}_1)$ is the conditional number density of scatterers at $\mathbf{r}$ if a scatterer is known to be at $\mathbf{r}_1$; $T(\mathbf{r}_1)\langle\psi^E(\mathbf{r}\mid\mathbf{r}_1)\rangle$ is the field scattered by a single scatterer at $\mathbf{r}_1$ when excited by $\langle\psi^E(\mathbf{r}\mid\mathbf{r}_1)\rangle$, and can be obtained from the presumably known scattering behavior of a single scatterer. The integral is taken over the whole volume τ accessible to scatterers.

The specific problem under consideration consists of a uniform and random array of spheres (radius a, constant density n_0) with centers in the semi-infinite region $z \geq 0$. A plane wave $\psi^{\text{inc}}(\mathbf{r}) = e^{ikz}$ is impinging normally from the left. The conditional density is chosen to be

$$n(\mathbf{r}'\mid\mathbf{r}_1) = \begin{cases} n_0 \text{ for } |\mathbf{r}'-\mathbf{r}_1| > b,\ z' > 0 \text{ and } z_1 > 0 \\ 0 \text{ otherwise,} \end{cases} \tag{2.2}$$

where $b = 2a$ if interpenetration is excluded; more generally $b(\geq 2a)$ represents the distance of closest approach between centers of adjacent spheres.

For any given "frozen" configuration of scatterers, $\psi^E(\mathbf{r}\mid\mathbf{r}_1)$ is a regular solution of the unperturbed Helmholtz equation in k and $\mathbf{r}$ provided $\mathbf{r}$ lies in the spherical region $|\mathbf{r}-\mathbf{r}_1| \leq b-a$. Assuming that the Helmholtz operator commutes with the averaging process, $\langle\psi^E(\mathbf{r}\mid\mathbf{r}_1)\rangle$ will have the above properties. Invoking also the planar symmetry of the problem, both exciting and scattered fields can be written in the general form

$$\langle\psi^E(\mathbf{r}\mid\mathbf{r}_1)\rangle = \sum_{n=0}^{\infty} i^n(2n+1)A_n(z_1)j_n(k|\mathbf{r}-\mathbf{r}_1|) \times P_n(\cos\theta_{rr_1}); \qquad |\mathbf{r}-\mathbf{r}_1| \leq b-a \tag{2.3}$$

$$T(\mathbf{r}')\langle\psi^E(\mathbf{r}\mid\mathbf{r}')\rangle = \sum_{j=0}^{\infty} i^j(2j+1)A_j(z')B_jh_j(k|\mathbf{r}-\mathbf{r}'|) \times P_j(\cos\theta_{rr'}); \qquad |\mathbf{r}-\mathbf{r}'| \geq a, \tag{2.4}$$

where the B_n are the presumably known scattering coefficients of a single scatterer.[5] In particular,

$$f(\theta) = \frac{1}{ik}\sum_{n=0}^{\infty}(2n+1)B_nP_n(\cos\theta) \tag{2.5}$$

is the far field amplitude of a single sphere. In Eq. (2.3) the $A_n(z_1)$ give the intrinsic dependence of $\langle\psi^E\rangle$ on scatterer position.

Equations (2.2), (2.3), (2.4), and the well-known expansion

$$\psi^{\text{inc}}(\mathbf{r}) = e^{ikz_1}e^{ik(z-z_1)} = e^{ikz_1}\sum_{n=0}^{\infty} i^n(2n+1)j_n(k|\mathbf{r}-\mathbf{r}_1|)P_n(\cos\theta_{rr_1}) \tag{2.6}$$

may now be inserted in Eq. (2.1). The spherical wave $h_j(k|\mathbf{r}-\mathbf{r}'|)P_j(\cos\theta_{rr'})$ inside the integral is then re-expanded in terms of spherical waves about the fixed origin $\mathbf{r}_1$. This may be accomplished using translational addition theorems for spherical wavefunctions available in the literature,[8] applied specifically to the translation $\mathbf{r}-\mathbf{r}' = (\mathbf{r}-\mathbf{r}_1) - (\mathbf{r}'-\mathbf{r}_1)$. In Eq. (2.1) $\mathbf{r}$ is restricted to lie within the sphere $|\mathbf{r}-\mathbf{r}_1| \leq b-a$, and for all points $\mathbf{r}'$ in the volume of integration the condition $|\mathbf{r}-\mathbf{r}_1| < |\mathbf{r}'-\mathbf{r}_1|$ is satisfied, allowing a single expansion to be used. Owing to the axial symmetry of the volume of integration only terms independent of the azimuthal angle $\phi_{r'r_1}$ contribute. The following terms remain inside the volume integral in place

[7] Lord Rayleigh, Phil. Mag. **34**, 481 (1892).

[8] O. R. Cruzan, Quart. J. Appl. Math. **20**, 33 (1962). See also S. Stein, *ibid.* **19**, 15 (1961); B. Friedman and J. Russek, *ibid.* **12**, 13 (1954).

of the above original wave:

$$i^{-j} \sum_{n=0}^{\infty} i^n (2n+1) j_n(k|\mathbf{r}-\mathbf{r}_1|) P_n(\cos\theta_{rr_1})$$
$$\times \sum_p (-i)^p a(0, j|0, n|p) h_p(k|\mathbf{r}'-\mathbf{r}_1|) P_p(\cos\theta_{r'r_1}),$$

where $p = j+n, j+n-2, \cdots, |j-n|$, and, in agreement with Cruzan's notation,[8]

$$a(0, j|0, n|p) = (2p+1)\frac{(j+n-p)!(j+p-n)!(n+p-j)!}{(j+n+p+1)!} \times \left[\frac{[\frac{1}{2}(j+n+p)]!}{[\frac{1}{2}(j+n-p)]![\frac{1}{2}(j+p-n)]![\frac{1}{2}(n+p-j)]!}\right]^2. \quad (2.7)$$

Note that these coefficients appear also in the expansion

$$P_j(x)P_n(x) = \sum_p a(0, j|0, n|p) P_p(x) \quad (2.8)$$

for the product of two Legendre polynomials.[8]

The wavefunctions $j_n(k|\mathbf{r}-\mathbf{r}_1|)P_n(\cos\theta_{rr_1})$ may next be factored out of the three terms of Eq. (2.1) and each coefficient in this sum set equal to zero, because of the orthogonality of these functions on a spherical surface with center at $\mathbf{r}_1$. As a result the infinite set of coupled integral equations

$$A_n(z_1) = e^{ikz_1} + n_0 \sum_{j=0}^{\infty} (2j+1) B_j \sum_p (-i)^p a(0, j|0, n|p) \times \int_{\tau-\tau_e} d\tau' \, A_j(z') h_p(k|\mathbf{r}'-\mathbf{r}_1|) P_p(\cos\theta_{r'r_1});$$
$$n = 0, 1, 2, \cdots \quad (2.9)$$

is obtained for the determination of $A_n(z)$. The volume τ_e, excluded from the integration, is shown in Fig. 1. For $z_1 \geq b$ it is a complete sphere of radius b, center at $\mathbf{r}_1$; for $z_1 < b$, however, a truncated sphere is excluded whose size depends on z_1. This complication strongly indicates the necessity of a separate treatment of the region $0 \leq z_1 \leq b$ for finite scatterers. The above equation is equivalent to Eq. (3.12) in I where Kasterin's[6] representation for spherical waves in terms of h_0 was used to carry out the re-expansion of h_jP_j. The alternate version given here requires fewer steps and leads to more compact results.

In order to solve the above set of equations three heuristic assumptions are made: (1) all $A_n(z)$ are expressible in terms of the same function $g(z)$

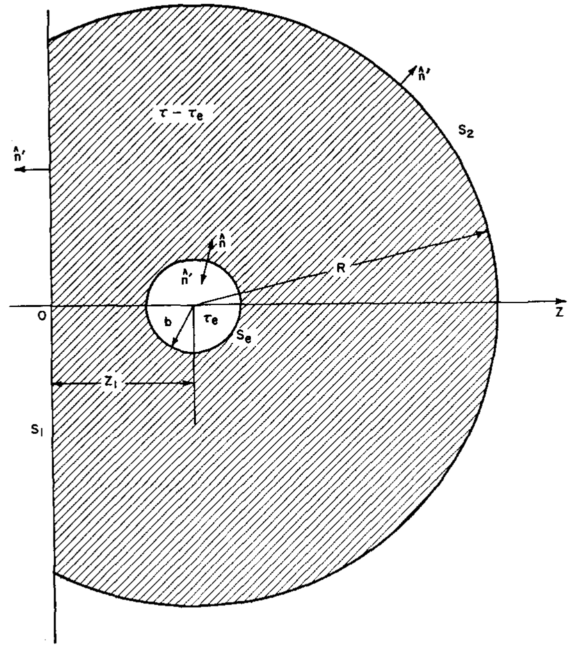

FIG. 1. Geometry of the half-space. The volume of integration for Eq. (2.9) consists (in the limit $R \to \infty$) of the half-space $z > 0$ less a sphere of radius b centered at the point $\mathbf{r}_1$.

which satisfies the wave equation

$$(\nabla^2 + K^2) g(z) = 0, \quad (2.10)$$

where K is a constant to be determined. This assumption makes $\langle \psi^E(\mathbf{r} \mid \mathbf{r}_1) \rangle$ equal to the product of two functions satisfying wave equations in K, z_1 and k, $\mathbf{r}$. (2) This representation is assumed to be valid for all positive values of z_1, even in the region $0 \leq z_1 \leq b$, which was seen to call for special treatment and where a more complicated dependence on z_1 should be expected. (3) A third assumption is made in the course of evaluating the integral equations, in that the truncated shape of the excluded sphere for $0 \leq z_1 \leq b$ is neglected. These assumptions are plausible and will be justified explicitly in the limit kb, $|Kb| \ll 1$. For larger scatterers, they may be considered as first approximations in an iteration scheme which will be outlined later in an attempt to obtain higher order corrections.

On the basis of these assumptions an explicit solution can be obtained. From the planar symmetry of the semi-infinite region and Eq. (2.10) the relations

$$A_n(z_1) = A_n^0 e^{iKz_1}; \quad n = 0, 1, 2, \cdots \quad (2.11)$$

are obtained, where A_n^0 are constants. Substitution into Eq. (2.9) and integration in accordance with the last two assumptions yields an infinite set of linear algebraic equations for the determination of K and the coefficients A_n^0. The steps are shown in the Appendix, and lead to the results

$$A_n^0 = n_0 \sum_{j=0}^{\infty} (2j+1)B_j A_j^0 \sum_p (-i)^p a(0, j\,|0, n|\, p) \times d_p(k, K \mid b); \qquad n = 0, 1, 2, \cdots \tag{2.12}$$

$$K = k + \frac{2\pi n_0}{k} F(0) = k + \frac{2\pi n_0}{ik^2} \sum_{j=0} (2j+1)B_j A_j^0: \quad \text{extinction theorem,} \tag{2.13}$$

where $F(0)$ is the forward-scattered amplitude with multiple scattering effects included, and

$$d_p(k, K \mid b) = -\frac{4\pi b^2}{K^2 - k^2} i^p[kh_p'(kb)j_p(Kb) - Kh_p(kb)j_p'(Kb)]. \tag{2.14}$$

The set of equations (2.12) consists of an infinite number of homogeneous linear equations for the A_n^0. For a nontrivial solution to exist the determinant must vanish, providing the secular equation from which K is determined. The A_n^0 are then determined using the additional (inhomogeneous) Eq. (2.13).

Because of the three assumptions used in arriving at these results, they can only be considered as the zeroth-order approximation in the following iteration scheme: The zeroth-order result (2.11) is inserted in the right-hand side of Eq. (2.9) and new $A_n(z_1)$ are evaluated by taking into account the truncated shape of the excluded sphere in the region $0 \leq z_1 \leq b$. That would be the procedure if the third assumption was not made. It may also be noticed that these first-order results add a correction term to each $A_n(z_1)$ only in the interface region $0 \leq z_1 \leq b$, leaving them unchanged for $z_1 \geq b$. This step can be carried out explicitly and gives rather compact results for low orders (n, $j = 0, 1$). For higher n, j, explicit expressions become very complicated. In principle the first-order values of $A_j(z_1)$ can be reinserted in the right-hand side of Eq. (2.9) and new $A_n(z_1)$ evaluated, taking again into account the truncated shape of the excluded region. It may be noticed that this and similar subsequent steps affect the values of $A_n(z_1)$ for $z_1 > b$. Investigation of the convergence of such an iteration scheme in the general case appears formidable and has not been attempted.

III. LOW-FREQUENCY APPROXIMATION

For large scatterers explicit results, even to zeroth order, cannot be obtained. However, in the limit kb, $|Kb| \ll 1$ it is known that only the monopole and dipole terms,

$$B_0 \cong \frac{i(ka)^3}{3}\left(\frac{M}{M'} - 1\right); \quad B_1 \cong \frac{i(ka)^3}{3} \frac{\rho' - \rho}{\rho + 2\rho'} \tag{3.1}$$

are significant, the rest of the B_n depending on $(ka)^n$, $n \geq 5$. In Eq. (3.1) the explicit formulas of an acoustic problem were used: ρ', M', $k' = \omega(\rho'/M')^{\frac{1}{2}}$ are, respectively, the density, reciprocal compressibility, and propagation constant of the scatterer material (oil or nonresonant air bubbles in water, for example); ρ, M, $k = \omega(\rho/M)^{\frac{1}{2}}$ are those of the supporting medium. In the same limit

$$d_p(k, K \mid b) \cong \frac{4\pi i^p}{ik(K^2 - k^2)}\left(\frac{K}{k}\right)^p \tag{3.2}$$

as can be seen from Eq. (2.14) using

$$j_p(x) \cong 2^p p! x^p/(2p+1)!,$$
$$h_p(x) \cong -[(2p)!/2^p p!](i/x^{p+1}).$$

Substituting into Eqs. (2.12) and (2.13) and neglecting all B_j for $j = 2, 3, \cdots$, yields

$$A_0^0 = \frac{4\pi n_0}{ik(K^2 - k^2)}[B_0 A_0^0 + 3B_1 A_1^0(K/k)], \tag{3.3a}$$

$$A_1^0 = \frac{4\pi n_0}{ik(K^2 - k^2)}[B_0 A_0^0(K/k) + B_1 A_1^0(1 + 2K^2/k^2)], \tag{3.3b}$$

$$ik^2(K - k)/2\pi n_0 = B_0 A_0^0 + 3B_1 A_1^0. \tag{3.4}$$

From the first two of these equations K/k is obtained. Then A_0^0 and A_1^0 can be determined using Eq. (3.4), while A_n^0, for $n = 2, 3, \cdots$, may be obtained from Eq. (2.12), in which $j = 0, 1$ are the only terms retained. The results for K, A_0^0 and A_1^0 are

$$\left(\frac{K}{k}\right)^2 = \frac{(1 + (4\pi n_0/ik^3)B_0)(1 + (4\pi n_0/ik^3)B_1)}{1 - (8\pi n_0/ik^3)B_1} \tag{3.5}$$

$$A_0^0 = \frac{2(K/k)}{K/k + 1 + (4\pi n_0/ik^3)B_0}, \quad A_1^0 = \frac{A_0^0(K/k)}{1 + (4\pi n_0/ik^3)B_1}. \tag{3.6}$$

In the acoustic problem, using Eqs. (3.1) and in terms of the fractional volume $\delta = \frac{4}{3}\pi n_0 a^3$ occupied by scatterers,

$$(K/k)^2 = (1 - \delta + \delta M/M') \times \left[\frac{1 - \delta + (2 + \delta)\rho'/\rho}{1 + 2\delta + 2(1 - \delta)\rho'/\rho}\right]. \tag{3.7}$$

As mentioned earlier, basically the same result was obtained by Rayleigh and Kasterin,[6] who considered periodic rather than random arrays. This is a positive definite form of the physically allowable values of M, M', ρ, ρ', δ. For $\delta = \binom{0}{1}$ the result $K = \binom{k}{k'}$ is obtained, correct at both limits. For small δ, and up to the first power in δ, Eq. (3.7) yields

$$\left(\frac{K}{k}\right)^2 = \left(1 - \delta + \frac{\delta M}{M'}\right)\left(1 + 3\delta\,\frac{\rho' - \rho}{\rho + 2\rho'}\right), \quad (3.8)$$

i.e., the result obtained if formula (3.25) in Ref. 5 is used. This formula had been derived previously[4] and later extended to arbitrary angle of incidence and arbitrary scatterers by Twersky,[3a] who also considered its various limitations. Certain failings of Eq. (3.8) are also apparent: Whenever $\rho \neq \rho'$, this formula fails at the upper limit $\delta = 1$. Moreover with no intrinsic losses, i.e., with real parameters for both media, and for certain values $\rho' < \rho$ it yields the embarrassing result $(K/k)^2 < 0$; in other words, it is not a positive definite form of the allowed values of δ and ρ'/ρ as would be expected on physical grounds. The validity of the previous result, (3.25) of I, is thus restricted to sparse concentrations.

Application of Twersky's schizoid formalism[3b-d,f] for the special case ψ = pressure, $A = A' = 1$, $B = \rho/\rho_{\rm eff}$, $B' = \rho/\rho'$ yields the same result (3.7) provided that Twersky's choice, $V_a = V_0(1 - \delta)$, is made for the unknown "available volume" V_a, where V_0 is the total volume available to the distribution and δ is the fractional volume, as used in this paper. That is, V_a is identified with the volume left unoccupied by scatterers in the distribution, as suggested by Twersky.

The next step of the iteration procedure described previously yields for the lowest-order coefficients $(n, j = 0, 1)$

$$A_0^1(z_1) = A_0^0 e^{iKz_1} + B_0 A_0^0 f_0(b - z_1) + 3B_1 A_1^0 f_1(b - z_1), \quad (3.9a)$$

$$A_1^1(z_1) = A_1^0 e^{iKz_1} + B_0 A_0^0 f_1(b - z_1) - B_1 A_1^0 f_2(b - z_1), \quad (3.9b)$$

where, for kb, $|Kb| \ll 1$,

$$f_0(x) = \frac{\pi n_0}{ik^3}(kx)^2[1 + O(kx)], \quad (3.10a)$$

$$f_1(x) = \frac{\pi n_0}{k^3}\frac{(kx)^2}{kb}[1 + O(kx)], \quad (3.10b)$$

$$f_2(x) = \frac{2\pi n_0}{ik^3}(kx)^2\left[\frac{3i(1 + K/k)}{kKb^2} + \frac{kx}{k^3b^3}\right][1 + O(kx)]. \quad (3.10c)$$

The above equations define f_0, f_1, f_2 for $0 \leq x \leq b$; outside this interval they are identically 0. They can be used up to $x = b$, the remainders being of higher order in kb than the terms given explicitly. It is obvious that the first correction to $A_0(z_1)$ is small of order kb, while for $A_1(z_1)$ it is of order unity but confined to the layer of thickness kb near the interface. Thus, in the present limit, a volume integral over the scattered field, given in Eq. (2.4), will not be affected if either the zeroth or first order values are used for $A_j(z')$. The assumptions made in Sec. II are thus justified explicitly in the limit $kb \ll 1$.

IV. BOUNDARY CONDITIONS AND THE TOTAL FIELD

In acoustics, pressure is given by the product of ψ with density, which of course may differ for scatterer and supporting medium. The total average pressure $\langle p(\mathbf{r})\rangle$ is given rigorously in terms of $\langle \psi^E(\mathbf{r} \mid \mathbf{r}_1)\rangle$ by the formula[5]

$$\langle p(\mathbf{r})\rangle = \rho\psi^{\rm inc} + \rho \int_{\mathbf{r}\,\text{"outside"}\,\mathbf{r}'} d\tau' n(\mathbf{r}')T(\mathbf{r}')\langle\psi^E(\mathbf{r}\mid\mathbf{r}')\rangle + \int_{\mathbf{r}\,\text{"inside"}\,\mathbf{r}'} d\tau' n(\mathbf{r}')[\rho' T^I(\mathbf{r}') - \rho]\langle\psi^E(\mathbf{r}\mid\mathbf{r}')\rangle, \quad (4.1)$$

where the various terms are properly normalized in the sense described in I. The notation r "outside" r′ means that integration is to be carried over all points, such that r is outside the scatterer having center r′, while r "inside" r′ is the complementary statement. According to this interpretation, for points r external to the scattering medium and sufficiently away from the interface only the first two terms remain in Eq. (4.1), and the integral is carried over the whole volume accessible to scatterers. For points r inside this region, all three terms are retained in Eq. (4.1). For spheres, in particular, a spherical region with center at r and radius a (equal to the scatterer radius) is excluded from the first integral, while the second is taken over this region alone. Obviously, the slab region $-a \leq z \leq a$ near and on both sides of the interface requires special treatment. $T^I(\mathbf{r}')\langle\psi^E(\mathbf{r} \mid \mathbf{r}')\rangle$, the field inside the scatterer at r′ when excited by $\langle\psi^E(\mathbf{r} \mid \mathbf{r}')\rangle$, is defined, in analogy with Eq. (2.4), by

$$T^I(\mathbf{r}')\langle\psi^E(\mathbf{r}\mid\mathbf{r}')\rangle = \sum_{n=0}^{\infty} i^n(2n + 1)A_n(z') \times B_n' j_n(k'|\mathbf{r} - \mathbf{r}'|)P_n(\cos\theta_{rr'}), \quad (4.2)$$

where B_n' are the partial wave coefficients of the internal field of a single scattering sphere. For

$ka \ll 1$ in particular

$$B'_n \cong \frac{2n+1}{(n+1)(\rho'/\rho)+n}(k/k')^2; \quad n = 0, 1, 2, \cdots . \tag{4.3}$$

Evaluation of the total field inside and outside the scattering region will be based on the assumptions made in Sec. II for $\langle\psi^E\rangle$ and, in addition, transition region complications of Eq. (4.1) will be ignored. These assumptions were justified in the limit $kb \ll 1$; additional justification will be obtained shortly.

For $z < -a$, Eq. (4.1) yields for the reflected field

$$\langle p(\mathbf{r})\rangle^{\text{refl}} = \rho \int_\tau d\tau' n(\mathbf{r}')T(\mathbf{r}')\langle\psi^E(\mathbf{r} \mid \mathbf{r}')\rangle. \tag{4.4}$$

As shown in the Appendix, substitution of Eqs. (2.4), (2.11), and integration yields

$$\langle p(\mathbf{r})\rangle^{\text{refl}} = \rho R e^{-ikz}, \tag{4.5}$$

where the reflection coefficient R is given by

$$R = -\frac{2\pi n_0}{ik^2(K+k)}\sum_{n=0}^{\infty}(-1)^n(2n+1)A_n^0 B_n$$

$$= -\frac{2\pi n_0}{k(K+k)}F(\pi), \tag{4.6}$$

$F(\pi)$ being the backward multiple scattering amplitude [see Eq. (2.13)].

For $z > a$, all terms are retained in Eq. (4.1). Use of Eqs. (2.3), (2.4), (2.11), and (4.2) results in the equation

$$\langle p(\mathbf{r})\rangle^{\text{tr}} = \rho e^{ikz} + n_0 e^{iKz}\sum_{n=0}^{\infty}(-i)^n(2n+1)A_n^0$$

$$\times\left[\rho B_n \int_{\tau-\tau_a} h_n(k\,|\mathbf{r}'-\mathbf{r}|) + \rho' B'_n \int_{\tau_a} j_n(k'\,|\mathbf{r}'-\mathbf{r}|)\right.$$

$$\left. - \rho\int_{\tau_a} j_n(k\,|\mathbf{r}'-\mathbf{r}|)\right]e^{iK(z'-z)}P_n(\cos\theta_{r'r})\,d\tau', \tag{4.7}$$

where τ_a is the volume of a sphere of radius a with center at $\mathbf{r}$. Evaluation of this expression is carried out in the Appendix yielding

$$\langle p(\mathbf{r})\rangle^{\text{tr}} = \rho_{\text{eff}} T e^{iKz}, \tag{4.8}$$

where the (suggestively written) constant $\rho_{\text{eff}}T$ is given by

$$\rho_{\text{eff}}T = 4\pi a^2 n_0 \sum_{n=0}^{\infty}(2n+1)A_n^0 J_n$$

$$+ n_0\rho\sum_{n=0}^{\infty}(-i)^n(2n+1)A_n^0 B_n d_n(k, K \mid a), \tag{4.9}$$

$$J_n \equiv \rho' B'_n \frac{k' j'_n(k'a)j_n(Ka) - K j_n(k'a)j'_n(Ka)}{K^2 - k'^2}$$

$$- \rho\frac{k j'_n(ka)j_n(Ka) - K j_n(ka)j'_n(Ka)}{K^2 - k^2}. \tag{4.10}$$

Under the assumptions made, the above results for the total reflected and transmitted fields can be extended in the transition regions $-a \le z \le 0$ and $0 \le z \le a$, respectively. The boundary conditions at $z = 0$, continuity of pressure and normal velocity, would then require

$$\rho(1+R) = \rho_{\text{eff}}T, \tag{4.11a}$$

$$k(1-R) = KT. \tag{4.11b}$$

The first equation provides a check on the assumptions that have been made. If it is found to hold effective parameters ρ_{eff}, M_{eff} can be determined and a full "medium" description attained. Elimination of T between the above expressions yields ρ_{eff}, while $M/M_{\text{eff}} = (K/k)^2(\rho_{\text{eff}}/\rho)^{-1}$.

In the limit $ka \ll 1$,

$$R \cong -\frac{2\pi n_0}{ik^3(1+K/k)}(B_0A_0^0 - 3B_1A_1^0) = A_0^0 - 1 \tag{4.12}$$

and Eq. (4.11a) reduces to $\rho_{\text{eff}}\,T \cong \rho A_0$. To the same order $d_n(k, K \mid a) \cong d_n(k, K \mid b)$ and Eq. (2.12) shows that the second term in Eq. (4.9) is simply equal to ρA_0. Thus,

$$\rho_{\text{eff}}T \cong \frac{3\delta}{a}\sum_{n=0}^{\infty}(2n+1)A_n^0 J_n + \rho A_0^0. \tag{4.13}$$

In the same limit Eqs. (4.3), (4.10) yield

$$J_n \cong \left[\frac{2^n n!}{(2n+1)!}\right]^2\frac{(Kk)^n}{2n+3}$$

$$\times\,\rho(\rho'-\rho)\frac{n}{(n+1)\rho'+n\rho}a^{2n+1}. \tag{4.14}$$

Thus $J_0 \cong 0$, while $J_n \sim (Kka^2)^n a$. So $\rho_{\text{eff}}\,T \cong \rho A_0^0$ and Eq. (4.11a) is satisfied. In the low-frequency limit the assumptions are completely justified and effective parameters can be determined:

$$\frac{\rho_{\text{eff}}}{\rho} = \frac{1-\delta+(2+\delta)\rho'/\rho}{1+2\delta+2(1-\delta)\rho'/\rho},$$
$$M/M_{\text{eff}} = 1-\delta+\delta M/M'. \tag{4.15}$$

These results are correct at both limits, i.e., for $\delta = \binom{0}{1}$: $\rho_{\text{eff}} = \binom{\rho}{\rho'}$, $M_{\text{eff}} = \binom{M}{M'}$.

It is tempting to check Eq. (4.11a) to the next higher order, i.e., to order $(kb)^2$, by retaining

B_0, B_1, B_2, but still using $A_n(z) = A_n^0 e^{iKz}$. The attempt fails, indicating the necessity of using corrected values for $A_n(z)$, $\langle\psi^E\rangle$, and $\langle\psi\rangle$, if a consistent treatment to higher order is desired. The expressions become very complicated, however, and have not been carried out to the end.

APPENDIX

Substituting Eq. (2.11) into (2.9), and making the change of variable $\mathbf{r}' - \mathbf{r}_1 = \mathbf{r}$ inside the integral yields

$$A_n^0 e^{iKz_1} = e^{ikz_1} + n_0 e^{iKz_1} \sum_{j=0}^{\infty} (2j+1) B_j A_j^0$$
$$\times \sum_p (-i)^p a(0, j\,|0, n|\, p) \int_{\tau-\tau_e} d\tau e^{iKz} w_p(kr), \quad \text{(A1)}$$

where

$$w_p(kr) = h_p(kr) P_p(\cos\theta) = (-i)^p P_{pz} h_0(kr). \quad \text{(A2)}$$

In the second equality, Kasterin's representation is used to express the spherical wave (see also Ref. 5) with P_{pz} standing for $P_p(1/ik\ \partial/\partial z)$. Since $(\nabla^2 + k^2) w_p(kr) = 0$ and $(\nabla^2 + K^2) e^{iKz} = 0$ it is possible to write

$$e^{iKz} w_p(kr) = \frac{1}{K^2 - k^2} (e^{iKz} \nabla^2 w_p - w_p \nabla^2 e^{iKz}) \quad \text{(A3)}$$

and transform the volume integral in Eq. (4.1) into surface integrals, using Green's theorem:

$$\int_{\tau-\tau_e} d\tau e^{iKz} w_p(kr) = \frac{1}{K^2 - k^2} \int_{S+S_e} ds$$
$$\times \left[e^{iKz} \frac{\partial w_p}{\partial n'} - w_p \frac{\partial e^{iKz}}{\partial n'} \right]. \quad \text{(A4)}$$

Here $S = \lim_{R\to\infty} [S_1(z_1) + S_2]$, S_e is a complete spherical surface of radius b (see assumptions in Sec. II), while n' is the normal outward unit vector as shown in Fig. 1. The above surface integral can be split into three integrals over S_1, S_2, and S_e. The first is best calculated in cylindrical coordinates $r = (\rho^2 + z^2)^{\frac{1}{2}}$, yielding

$$\frac{1}{K^2 - k^2} \int_{S_1} = \frac{2\pi}{K^2 - k^2}$$
$$\times \left[e^{iKz} \int_{\rho=0}^{R} \left(iKw_p - \frac{\partial w_p}{\partial z} \right) \rho\, d\rho \right]_{z=-z_1}$$
$$= -\frac{2\pi e^{-iKz_1}}{k^2(K^2 - k^2)} (-i)^p \left(iK + \frac{\partial}{\partial z_1} \right) P_{p,-z_1}$$
$$\times \left(e^{ik(R^2+z_1^2)^{\frac{1}{2}}} - e^{ikz_1}\right). \quad \text{(A5)}$$

Letting $R \to \infty$ it is noticed that the oscillating term $e^{ik(R^2+z_1^2)^{\frac{1}{2}}}$ disappears; also $P_{p,-z_1} e^{ikz_1} = P_p(-1) e^{ikz_1} = (-1)^p e^{ikz_1}$. Therefore,

$$\frac{1}{K^2 - k^2} \int_{S_1} = \frac{2\pi i}{k^2(K - k)} i^p e^{i(k-K)z_1}. \quad \text{(A6)}$$

The contribution of this integral to the right-hand side of Eq. (A1) can be writen

$$\frac{2\pi n_0 i e^{ikz_1}}{k^2(K - k)} \sum_{j=0}^{\infty} (2j+1) B_j A_j^0, \quad \text{(A7)}$$

because, as seen from Eq. (2.8) for $x = 1$, $\sum_p a(0, j\mid 0, n\mid p) = 1$.

The remaining integrals

$$d_p(k, K\mid b) = -1/(K^2 - k^2)$$
$$\times \int_{S_e} ds \left[e^{iKz} \frac{\partial w_p}{\partial n} - w_p \frac{\partial e^{iKz}}{\partial n} \right], \quad \text{(A8)}$$

$$g_p = \frac{1}{K^2 - k^2} \lim_{R\to\infty} \int_{S_2} ds \left[e^{iKz} \frac{\partial w_p}{\partial R} - w_p \frac{\partial e^{iKz}}{\partial R} \right] \quad \text{(A9)}$$

are independent of z_1, and are just constants depending on K, k. On the spherical surface $r = b$, with $\cos\theta = x$,

$$d_p(k, K\mid b) = -\frac{2\pi b^2}{K^2 - k^2} \int_{-1}^{1} e^{iKbx} P_p(x) [kh_p'(kb)$$
$$- h_p(kb) iKx]\, dx, \quad \text{(A10)}$$

while on S_2, for large R and with $h_p(kR) \approx i^{-p}(ikR)^{-1} e^{ikR}$, the corresponding result is

$$g_p = \frac{2\pi i^{-p}}{K^2 - k^2} \lim_{R\to\infty} \left[R e^{ikR} \int_0^1 dx e^{iKRx} P_p(x) \left(1 - \frac{K}{k} x \right) \right]. \quad \text{(A11)}$$

In the latter equation $P_p(x)(1 - (K/k)x) = q(x)$ is a polynomial of degree $p + 1$. Repeated integrations by parts yield

$$g_p = \frac{2\pi i^{-p}}{K^2 - k^2} \lim_{R\to\infty} R e^{ikR}$$
$$\times \left[\frac{e^{iKR} q(1) - q(0)}{iKR} + \frac{e^{iKR} q'(1) - q'(0)}{(KR)^2} + \cdots \right]$$
$$= \frac{2\pi i^{-p}}{K^2 - k^2} \lim_{R\to\infty} \frac{e^{ikR}}{iK} [e^{iKR} q(1) - q(0)] = 0, \quad \text{(A12)}$$

the last result following in the usual manner by introducing a small positive imaginary part in k.

The integrals in Eq. (A10) can be evaluated explicitly with the use of the identities[9]

[9] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), p. 1575.

$$\int_{-1}^{1} e^{ixz}P_\nu(x)\,dx = 2i^\nu j_\nu(z),$$
$$\int_{-1}^{1} ixe^{ixz}P_\nu(x)\,dx = 2i^\nu j'_\nu(z). \tag{A13}$$

The result is given in Eq. (2.14). Substituting all the above results into Eq. (A1) and equating separately the factors multiplying e^{iKz_1} and e^{ikz_1} yields Eqs. (2.12) and (2.13).

For the reflected field, use of Eqs. (2.4), (2.11) in (4.4) yields

$$\langle p(\mathbf{r})\rangle^{\text{refl}} = \rho e^{iKz} n_0 \sum_{n=0}^{\infty} (-i)^n(2n+1)A_n^0 B_n$$
$$\times \int_\tau d\tau' e^{iK(z'-z)} h_n(k\,|\mathbf{r}'-\mathbf{r}|)P_n(\cos\theta_{r'r}). \tag{A14}$$

The integration can be carried out as before, using Green's theorem, to get

$$\int_\tau = \int_{S_1} = \frac{e^{-iKz}}{K^2-k^2}\frac{2\pi(-i)^n}{k^2}$$
$$\times \left[\left(-\frac{\partial}{\partial z'} + iK\right)P_{nz'}e^{ik|z'-z|}\right]_{z'=0}. \tag{A15}$$

With $z' > z$ the result is

$$(2\pi i(-i)^n/k^2)(e^{-i(K+k)z}/(K+k));$$

substitution into Eq. (A14) leads at once to Eqs. (4.5), (4.6).

For the transmitted field the three volume integrals in Eq. (4.7) are transformed into surface integrals. The first splits into two integrals over S_1 and S_a, the last two go over to integrals over S_a. From S_1 the result (A15) is obtained, which, with $z' < z$, reduces to $(e^{i(k-K)z}/(K-k))(2\pi i/k^2)i^n$. The corresponding term in Eq. (4.7) becomes

$$-\rho e^{ikz}\frac{2\pi n_0}{ik^2(K-k)}\sum_{n=0}^{\infty}(2n+1)A_n^0B_n$$

and with use of the extinction theorem, Eq. (2.13), is seen to exactly cancel the incident field in Eq. (4.7). There remain three surface integrals over S_a, on which $|\mathbf{r}'-\mathbf{r}| = a$. Therefore,

$$\langle p(\mathbf{r})\rangle^{\text{tr}} = 2\pi z^2 n_0 e^{iKz}\sum_{n=0}^{\infty}(-i)^n(2n+1)A_n^0$$
$$\times \int_{-1}^{1} dx e^{iKax}P_n(x)\left[-\rho B_n\frac{kh_n'(ka) - iKxh_n(ka)}{K^2-k^2}\right.$$
$$\left. + \rho' B_n'\frac{k'j_n'(k'a) - iKxj_n(k'a)}{K^2-k'^2} - \rho\frac{kj_n'(ka) - iKxj_n(ka)}{K^2-k^2}\right].$$

Integration according to the formulas (A13) results in Eqs. (4.8), (4.9), and (4.10).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 10 OCTOBER 1964

An Integral Equation for the Associated Legendre Function of the First Kind

ATTIA A. ASHOUR*
High Altitude Observatory, Boulder, Colorado
(Received 26 May 1964)

The solution of the Fredholm homogeneous equation

$$\psi(\xi) = \lambda \int_0^1 M(\xi, \xi')\psi(\xi')\, d\xi',$$

where

$$M(\xi, \xi') = \int_0^\pi \frac{\cos m\phi \, d\phi}{(x^2 - 2xx' \cos\phi + x'^2)^{\frac{1}{2}}}$$

and $x = (1 - \xi^2)^{\frac{1}{2}}$ is found to be the associated Legendre function $P_n^m(\xi)$, $n + m$ even, and the characteristic numbers of this kernel are obtained. The solution of the corresponding equation of the second kind is also found. The kernel of the homogeneous equation whose solution is $P_n^m(\xi)$, $n + m$ odd, is obtained.

1. INTRODUCTION

MANY of the integral equations arising in electrostatics (Collins[1]), electromagnetic induction (Ashour[2,3]) and in diffraction theory (Noble[4]), have as kernel either the function

$$K(x, x') = \int_0^\pi \frac{\cos m\phi \, d\phi}{(x^2 - 2xx' \cos\phi + x'^2)^{\frac{1}{2}}} \qquad (1)$$
$$(m \text{ integer})$$

or a function which does not differ from $K(x, x')$ except by a simple factor (depending on x and x'). Hence, it is of interest to find an exact solution for an integral equation whose kernel is simply related to that given by Eq. (1). It can easily be seen that $K(x, x')$ has an infinity of order $\log |x - x'|$ when $|x - x'| \to 0$. This kernel may also be expressed (Eason, Noble, and Sneddon[5]) as

$$K(x, x') = \pi \int_0^\infty J_m(x\, u) J_m(u\, x')\, du, \qquad (2)$$

where $J_m(x)$ is the Bessel function of order m in x. In this paper, the solution of the integral equation

$$\psi(\xi) = f(\xi) + \lambda \int_0^1 M(\xi, \xi')\psi(\xi')\, d\xi' \quad (0 \le \xi \le 1), \quad (3)$$

when $f(\xi)$ is even, and

$$\xi = (1 - x^2)^{\frac{1}{2}}, \qquad M(\xi, \xi') = K(x, x'), \qquad (4)$$

is obtained. In particular, the solution of Eq. (3) when $f(\xi) = 0$ is found to be simply the associated Legendre function of the first kind $P_n^m(\xi)$ with $n + m$ even, and the characteristic numbers λ_n of the kernel are obtained. The kernel of the homogeneous equation whose solution is $P_n^m(\xi)$, $n + m$ odd, is also obtained and the solution of the corresponding nonhomogeneous equation with $f(\xi)$ an odd function of ξ is found.

2. AN EXPANSION FOR $M(\xi, \xi')$

We first obtain a Fourier series expansion for $(x^2 - 2xx' \cos\phi + x'^2)^{-\frac{1}{2}}$ symmetrical in x and x'. Hobson[6] gave a Fourier series expression which is not symmetrical in x and x'. Sack[7] obtained an expansion as a series in the Legendre functions $P_l(\cos\phi)$, the coefficients being symmetrical in x and x'. Neither of these expressions is suitable for the present purpose.

Let R denote the distance between two points (ξ, ζ, ϕ) and (ξ', ζ', ϕ'), where ξ, ζ, ϕ are oblate spheroidal coordinates given in terms of cylindrical polar coordinates z, ρ, ϕ as

$$\left.\begin{aligned} z &= a\xi\zeta \\ \rho &= a\{(1 - \xi^2)(1 + \zeta^2)\}^{\frac{1}{2}} \end{aligned}\right\} \begin{aligned} 0 &\le \zeta \le \infty, \\ -1 &\le \xi \le 1. \end{aligned} \qquad (5)$$

Hence,

$$a/R = [2 + \zeta^2 + \zeta'^2 - \xi^2 - \xi'^2 - 2\xi\xi'\zeta\zeta' - 2\{(1 - \xi^2)(1 - \xi'^2)(1 + \zeta^2)(1 + \zeta'^2)\}^{\frac{1}{2}} \times \cos(\phi - \phi')]^{-\frac{1}{2}}. \qquad (6)$$

* On leave from the Applied Mathematics Department, Faculty of Science, Cairo University, Cairo, Egypt.
[1] W. D. Collins, Proc. Cambridge Phil. Soc. 55, 377 (1955).
[2] A. A. Ashour, Quart. J. Mech. Appl. Math. 3, 119 (1950).
[3] A. A. Ashour, Quart. J. Mech. Appl. Math. (to be published).
[4] B. Noble, in *Electromagnetic Waves*, edited by Rudolf E. Langer, University of Wisconsin Press, Madison, Wisconsin (1962), p. 323.
[5] G. Eason, B. Noble, and I. N. Sneddon, Phil. Trans. Roy. Soc. (London) A247, 529 (1955).
[6] E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, (Cambridge University Press, New York 1951), p. 443.
[7] R. A. Sack, J. Math. Phys. 5, 245 (1964).

The expression for a/R as a series of oblate spheroidal harmonics has been given by Hobson (Ref. 6, p. 430). With the present notation (which is the same as that used by Smythe[8]), this expression is

$$\frac{a}{R} = \sum_{n=0}^{\infty}\sum_{m=0}^{n} N_{mn}Q_n^m(i\zeta)P_n^m(\xi)\cos m(\phi-\varphi') \qquad \zeta \geq \zeta' \geq 0, \tag{7}$$

where

$$N_{mn} = i(2-\delta_{m0})(-1)^m(2n+1) \times \{(n-m)!/(n+m)!\}^2 P_n^m(i\zeta')P_n^m(\xi'), \tag{8}$$

δ_{m0} being the Kronecker symbol and $P_n^m(u)$, $Q_n^m(u)$ are the associated Legendre functions of the first and second kinds in u. If now in Eqs. (6), (7), and (8) we take $\zeta = \zeta' = \phi' = 0$ and note that

$$\begin{aligned} P_n^m(i\cdot 0) &= 0 && n+m \text{ odd} \\ &= (-1)^{n/2}\frac{1\cdot 3\cdot 5\cdots(n+m-1)}{2\cdot 4\cdot 6\cdots(n-m)} && n+m \text{ even} \end{aligned} \tag{9}$$

$$Q_n^m(i\cdot 0) = -\tfrac{1}{2}i\pi P_n^m(i\cdot 0) \qquad n+m \text{ even}$$

it is found from (6), (7), (8), and (9) that

$$\{x^2 - 2xx'\cos\phi + x'^2\}^{-\frac{1}{2}} = \frac{\pi}{2}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(2-\delta_{m0}) \times (4n+2m+1)\left\{\frac{1\cdot 3\cdot 5\cdots 2n-1}{2\cdot 4\cdot 6\cdots 2n+2m}\right\}^2 P_{2n+m}^m(\xi) \times P_{2n+m}^m(\xi')\cos m\phi. \tag{10}$$

This is the required expression. From Eqs. (1) and (10) we now readily obtain

$$M(\xi,\xi') = \frac{\pi^2}{2}\sum_{n=0}^{\infty}(4n+2m+1) \times \left\{\frac{1\cdot 3\cdot 5\cdots 2n-1}{2\cdot 4\cdot 6\cdots 2n+2m}\right\}^2 P_{2n+m}^m(\xi)P_{2n+m}^m(\xi'). \tag{11}$$

If in (10), ϕ is replaced by ω where

$$\cos\omega = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi-\phi'), \tag{12}$$

we obtain at once an expression (which appears to be new) for the reciprocal of the distance between the two points whose spherical polar coordinates are (ax, θ, ϕ), (ax', θ', ϕ') $(x, x' \leq 1)$.

3. SOLUTION OF THE FREDHOLM INTEGRAL EQUATION

We first consider the integral equation (3) with $f(\xi) = 0$. From Eq. (11), and the orthogonality[9] of the Legendre functions, we immediately see that the solution in this case is

$$\psi(\xi) = P_{2n+m}^m(\xi) \tag{13}$$

provided that λ is one of the numbers λ_n (n a positive integer or zero) given by

$$\lambda_n = \frac{2}{\pi^2}\frac{2\cdot 4\cdot 6\cdots 2n}{1\cdot 3\cdot 5\cdots 2n-1}\cdot\frac{2\cdot 4\cdot 6\cdots 2n+2m}{1\cdot 3\cdot 5\cdots 2n+2m-1}. \tag{14}$$

To obtain the solution of the general equation (3), we assume that $f(\xi)$ may be expanded as a series of Legendre functions of even parity all of the same order m [note that $f(\xi)$ is even], i.e.,

$$f(\xi) = \sum_{n=0}^{\infty} A_n P_{2n+m}^m(\xi), \tag{15}$$

where

$$A_n = \frac{(4n+2m+1)(2n)!}{(2n+2m)!}\int_0^1 f(\xi)P_{2n+m}^m(\xi)\,d\xi. \tag{16}$$

Then, using Eqs. (13), (14), and (15) we obtain the solution of Eq. (3) as

$$\psi(\xi) = \sum_{n=0}^{\infty}\frac{A_n P_{2n+m}^m(\xi)}{(1-\lambda/\lambda_n)}. \tag{17}$$

In the same way, it can be proved that the equation

$$\int_0^1 \psi(\xi')M(\xi,\xi')\,d\xi' = f(\xi), \tag{18}$$

where $f(\xi)$ is given by Eq. (15), has the solution

$$\psi(\xi) = \sum_0^{\infty}\lambda_n A_n P_{2n+m}^m(\xi). \tag{19}$$

4. AN INTEGRAL EQUATION FOR $P_n^m(\xi)$, $n+m$ ODD

If $\{\partial^2 R^{-1}/\partial\zeta\ \partial\zeta'\}_{\zeta=\zeta'=\phi'=0}$ is found using the two expressions (6) and (7) for R^{-1}, we obtain after inserting the numerical values for $P_n^{m\prime}(i\cdot 0)$ and $Q_n^{m\prime}(i\cdot 0)$:

$$\xi\xi'(x^2 - 2xx'\cos\phi + x'^2)^{-\frac{3}{2}} = -\frac{\pi}{2}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(2-\delta_{m0}) \times (4n+2m+3)\left\{\frac{1\cdot 3\cdot 5\cdots 2n+1}{2\cdot 4\cdot 6\cdots 2n+2m}\right\}^2 \times P_{2n+m+1}^m(\xi)P_{2n+m+1}^m(\xi')\cos m\phi. \tag{20}$$

[8] W. R. Smythe, *Static and Dynamic Electricity*, McGraw-Hill Book Company, Inc., New York (1950), pp. 146, 148, 152, 158, 166.

[9] $$\begin{aligned}\int_0^1 P_{2n+m}^m(\xi)\,P_{2n'+m}^m(\xi)\,d\xi &= 0 && (n \neq n') \\ &= \frac{(2n+2m)!}{(2n)!\,(4n+2m+1)} && (n = n').\end{aligned}$$

If in Eq. (20) ϕ is replaced by ω [given by Eq. (12)], we obtain a Fourier series expansion symmetrical in x, x' for the inverse third power of the distance between two points in spherical polar coordinates, which again appears to be new.[10]

We now define a new kernel $G(\xi, \xi')$ as

$$G(\xi, \xi') = \xi\xi' \int_0^\pi \frac{\cos m\phi \, d\phi}{(x^2 - 2xx' \cos\phi + x'^2)^{\frac{3}{2}}}. \tag{21}$$

From Eqs. (20) and (21) we obtain

$$G(\xi, \xi') = -\frac{\pi^2}{2} \sum_{n=0}^{\infty} (4n + 2m + 3) \left\{\frac{1\cdot3\cdot5 \cdots 2n+1}{2\cdot4\cdot6 \cdots 2n+2m}\right\}^2 P_{2n+m+1}^m(\xi) P_{2n+m+1}^m(\xi'). \tag{22}$$

Hence, $P_{2n+m+1}^m(\xi)$ satisfies the integral equation

$$\psi(\xi) = \mu \int_0^1 \psi(\xi') G(\xi, \xi') \, d\xi', \tag{23}$$

provided that μ equals one of the characteristic numbers μ_n given by

$$\mu_n = -\frac{2}{\pi^2} \frac{2\cdot4\cdot6 \cdots 2n}{1\cdot3\cdot5 \cdots 2n+1} \cdot \frac{2\cdot4\cdot6 \cdots 2n+2m}{1\cdot3\cdot5 \cdots 2n+2m+1}. \tag{24}$$

The solutions of the equations

$$\psi(\xi) = g(\xi) + \mu \int_0^1 \psi(\xi') G(\xi, \xi') \, d\xi' \tag{25}$$

and

$$\int_0^1 \psi(\xi') G(\xi, \xi') \, d\xi' = g(\xi), \tag{26}$$

where $g(\xi)$ is odd and can be expanded in the form

$$g(\xi) = \sum_{n=0}^{\infty} B_n P_{2n+m+1}^m(\xi), \tag{27}$$

can be found in the same way as before. They are

$$\sum_0^\infty \frac{B_n}{(1 - \mu/\mu_n)} P_{2n+m+1}^m(\xi) \quad \text{and} \quad \sum_{n=0}^{\infty} \mu_n B_n P_{2n+m+1}^m(\xi),$$

respectively.

Note added in proof: Another integral equation satisfied by $P_n^m(\xi)$ (n, m integers) is

$$\psi(\xi) = \nu \int_{-1}^{+1} H(\xi, \xi') \psi(\xi') \, d\xi' \tag{28}$$

where

$$H(\xi, \xi') = \int_0^\pi \frac{\cos m\phi \, d\phi}{(1 - \xi\xi' - xx' \cos\phi)^{\frac{1}{2}}} = \pi\sqrt{2} \sum_{n=m}^{\infty} \frac{(n-m)!}{(n+m)!} P_n^m(\xi) P_n^m(\xi'). \tag{29}$$

This can be proved by taking $r = r' = 1$, $\phi' = 0$, $\cos\theta = \xi$, $\cos\theta' = \xi'$ in the expression for the inverse distance R^{-1} between two points r, θ, ϕ and r', θ', ϕ'.[11] The characteristic numbers ν_n are then readily found to be

$$\nu_n = (2n+1)/2\sqrt{2}\,\pi, \; n = 0, 1, \cdots. \tag{30}$$

The solutions of equations of the form (3) or (18), with kernel $H(\xi, \xi')$ and integration interval $-1 \to 1$, can be obtained in exactly the same manner as before.

[10] The Fourier series expansion of the general power of the distance between two points, symmetrical in x and x' will be discussed in another communication [J. Math. Phys. (to be published)].

[11] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), p. 1274, No. 10.3.3.37.

Application of Operational Methods to the Analysis of the Motion of Rigid Bodies

SAUL SILVEN
Lockheed Missiles and Space Company, Sunnyvale, California
(Received 5 May 1964; final manuscript received 9 May 1964)

Various types of gyroscopic precessional motion are discussed by the application of operational methods to the vector Euler equation describing the motion of a rigid body about a fixed point. Discussed herein are free precession and forced precessions due to external torques both fixed in the body and fixed in space.

INTRODUCTION

THE well-known Euler equations, describing the motion of a rigid body, may be written as

$$\mathsf{A}\cdot\dot{\boldsymbol{\omega}}' + \boldsymbol{\omega}' \times \mathsf{A}\cdot\boldsymbol{\omega}' = \mathbf{L}, \tag{1}$$

where A is the inertia tensor, $\mathbf{L}$ is the external torque, and $\boldsymbol{\omega}'$ is the angular velocity. The frame of reference in which the quantities appearing in Eq. (1) are measured is attached to the body and the origin is either at a fixed point of the body, if one exists, or at the center of mass.[1]

If the body has a constant angular velocity $\boldsymbol{\Omega}$ and is in dynamic equilibrium ($\mathbf{L} = \mathbf{0}$), then

$$\boldsymbol{\Omega} \times \mathsf{A}\cdot\boldsymbol{\Omega} = \mathbf{0}. \tag{2}$$

Equation (2) shows that the rotation must be about a principal axis, since the angular momentum, $\mathsf{A}\cdot\boldsymbol{\Omega}$, must be parallel to the angular velocity.[2] This axis will be referred to as the axis of spin of the rigid body.

If the body were disturbed from its equilibrium so that the angular velocity became

$$\boldsymbol{\omega}'(t) = \boldsymbol{\Omega} + \boldsymbol{\omega}(t), \tag{3}$$

the equation of motion, if the disturbance $\boldsymbol{\omega}(t)$ remained small with respect to $\boldsymbol{\Omega}$, would be

$$\mathsf{A}\cdot\dot{\boldsymbol{\omega}}(t) + \Omega[\mathbf{i} \times \mathsf{A} - (\mathbf{i}\cdot\mathsf{A}) \times]\boldsymbol{\omega}(t) = \mathbf{L}(t), \tag{4}$$

where $\mathbf{i}$ is the unit vector along the axis of spin and $\Omega\mathbf{i} = \boldsymbol{\Omega}$. The reader's attention is called to the fact that the second-order term $\boldsymbol{\omega} \times \mathsf{A}\cdot\boldsymbol{\omega}$ is assumed to be negligible. The linear transform method discussed in this paper cannot be used in the analysis of nonlinear problems.

The linear transform method discussed in this paper is an extension of the Laplace transform to an equation in which the dependent variables are vector functions of time. This method was used by the author in the analysis of uniform plasmas.[3] A vector function of time $\mathbf{f}(t)$ transforms to a new vector function $\mathbf{F}(\mathsf{S})$, via

$$\mathbf{F}(\mathsf{S}) = \int_0^\infty \exp(-\mathsf{S}t)\mathbf{f}(t)\,dt, \tag{5}$$

where S represents the tensor time-derivative operator.

DESCRIPTION OF THE METHOD

Equation (4) can be written as

$$\dot{\boldsymbol{\omega}} - \Omega\mathsf{K}\cdot\boldsymbol{\omega} = \mathsf{A}^{-1}\cdot\mathbf{L}, \tag{6}$$

where $\mathsf{K} = (\mathsf{A})^{-1}[(\mathbf{i}\cdot\mathsf{A}) \times - \mathbf{i} \times \mathsf{A}]$. If the time-derivative operation is represented by the tensor S, Eq. (6) becomes

$$(\mathsf{S} - \Omega\mathsf{K})\boldsymbol{\omega}(\mathsf{S}) = \mathsf{A}^{-1}\cdot\mathbf{L}(\mathsf{S}) + \boldsymbol{\omega}_0, \tag{7}$$

where $\boldsymbol{\omega}$ and $\mathbf{L}$ are now functions of S, and $\boldsymbol{\omega}_0$ represents the initial perturbation. The perturbation thus has a pole[4] at $\mathsf{S} = \Omega\mathsf{K}$; i.e., has a functional behavior of the form $\exp(\mathsf{K}\Omega t)$, where

$$\exp(\mathsf{K}\Omega t) = \mathsf{I} + \textstyle\sum_n (n!)^{-1}(\mathsf{K}\Omega t)^n, \quad n = 1, 2, 3, \cdots, \tag{8}$$

and I is the idemtensor.

From the above definition of K it can be shown that[5]

$$(\mathsf{K})^2 = -k^2(\mathsf{I} - \mathbf{ii}) \tag{9}$$

where

$$(A_{22})^{\frac{1}{2}}k = (A_1^2 - A_1A_{21} + A_{22})^{\frac{1}{2}}. \tag{10}$$

In the above expression $A_1 = \mathbf{i}\cdot\mathsf{A}\cdot\mathbf{i}$ and A_{21} and A_{22} are the first and third scalar invariants of the

---

[1] R. Becker, *Introduction to Theoretical Mechanics* (McGraw-Hill Book Company, Inc., New York, 1954), p. 290.
[2] J. Slater and N. Frank, *Mechanics* (McGraw-Hill Book Company, Inc., New York, 1947), p. 100.
[3] S. Silven, J. Math. Phys. 5, 557 (1964).
[4] By the use of this term, it is intended to show the analogy with the analysis of scalar differential equations in the complex frequency domain.
[5] See Appendix.

tensor $\mathsf{A}_2 = \mathsf{A} - A_1\mathbf{ii}$. The quantity A_1 is, of course, the moment of inertia about the axis of spin. Making use of Eq. (9) and noting that $(\mathsf{K})^3 = -k^2\mathsf{K}$, the expansion of Eq. (8) leads to

$$\exp(\mathsf{K}\Omega t) = \mathbf{ii} + (\mathsf{I} - \mathbf{ii}) \cos k\Omega t + (\mathsf{K}/k) \sin k\Omega t, \tag{11}$$

provided, of course, that k is real. The effect of an imaginary k is discussed later in this paper.

TORQUE-FREE PRECESSION

If the external torque $\mathbf{L}(\mathsf{S})$ were zero, Eq. (6) would become

$$\boldsymbol{\omega}(\mathsf{S}) = (\mathsf{S} - \Omega\mathsf{K})^{-1}\boldsymbol{\omega}_0 \tag{12}$$

which has the inverse

$$\boldsymbol{\omega}(t) = (\mathbf{i}\cdot\boldsymbol{\omega}_0)\mathbf{i} + [\boldsymbol{\omega}_0 - (\mathbf{i}\cdot\boldsymbol{\omega}_0)\mathbf{i}] \cos k\Omega t + [(\mathsf{K}\cdot\boldsymbol{\omega}_0)/k] \sin k\Omega t. \tag{13}$$

The first term on the right indicates that perturbations parallel to the axis of spin remain constant, merely causing a change in the magnitude of the component of angular velocity along the axis of spin. The second and third terms indicate an elliptically polarized precession; i.e., the axis of spin generates, in space, an elliptic cone about an axis of precession, which coincides with the position of the axis of spin prior to the application of the source of the disturbance.

If the rigid body were symmetric about the axis of spin, the inertia tensor could be written as

$$\mathsf{A} = A_1\mathbf{ii} + A_2(\mathsf{I} - \mathbf{ii}), \tag{14}$$

from which it follows that

$$\mathsf{K} = k\mathbf{i}\times, \tag{15}$$

where

$$A_2k = A_1 - A_2. \tag{16}$$

Then Eq. (12) becomes

$$\boldsymbol{\omega}(t) = (\mathbf{i}\cdot\boldsymbol{\omega}_0)\mathbf{i} + [\boldsymbol{\omega}_0 - (\mathbf{i}\cdot\boldsymbol{\omega}_0)\mathbf{i}] \cos k\Omega t + (\mathbf{i}\times\boldsymbol{\omega}_0) \sin k\Omega t. \tag{17}$$

In this case the precession is circularly polarized.

THE EFFECT OF A TORQUE FIXED IN THE BODY

If a torque $\mathbf{L}$, constant in the frame of reference attached to the body, is applied to a body previously rotating with angular velocity $\boldsymbol{\Omega}$ about a principal axis, the transform of the equation of motion becomes

$$\boldsymbol{\omega}(\mathsf{S}) = (\mathsf{S} - \Omega\mathsf{K})^{-1}(\mathsf{A})^{-1}(\mathsf{S})^{-1}\mathbf{L}, \tag{18}$$

$$= (\mathsf{S} - \Omega\mathsf{K})^{-1}(\mathsf{A})^{-1}(\mathsf{S})^{-1}\mathbf{L}_2 + (\mathsf{S})^{-1}(\mathsf{A})^{-1}(\mathsf{S})^{-1}\mathbf{L}_1, \tag{18a}$$

where $\mathbf{L}_1$ is the component of $\mathbf{L}$ along the axis of spin and $\mathbf{L}_2 = \mathbf{L} - \mathbf{L}_1$.[6] The presence of $\mathbf{L}_1$ causes the angular velocity about the axis of spin to increase without limit. If $\mathbf{L}_1 = \mathbf{0}$, a "partial fraction" expansion of Eq. (18) yields

$$\boldsymbol{\omega}(\mathsf{S}) = [(\mathsf{S} - \Omega\mathsf{K})^{-1}\mathsf{B} - \mathsf{B}(\mathsf{S})^{-1}]\mathbf{L}, \tag{19}$$

where B is the tensor defined by $\Omega k^2\mathsf{B} = -\mathsf{K}(\mathsf{A})^{-1}$. The inverse of Eq. (19) is

$$\boldsymbol{\omega}(t) = -\mathsf{B}\cdot\mathbf{L}(1 - \cos k\Omega t) + (\mathsf{K}/k)\mathsf{B}\cdot\mathbf{L} \sin k\Omega t. \tag{20}$$

The body thus executes an elliptic precession. The axis of precession passes through the origin of the system and is parallel to the original direction, in space, of the vector $-\mathsf{B}\cdot\mathbf{L}$. Again, if the body were symmetric, the precession would be circular.

THE EFFECT OF A TORQUE FIXED IN THE PRECESSION FRAME

If an applied torque has a constant value of $\mathbf{L}_0$ in the frame of reference attached to the axis of spin and having an angular velocity equal to the disturbance $\boldsymbol{\omega}(t)$, it can be represented by

$$\mathbf{L}(t) = (\mathbf{ii})\mathbf{L}_0 + (\mathsf{I} - \mathbf{ii})\mathbf{L}_0 \cos \Omega t - \mathbf{i}\times\mathbf{L}_0 \sin \Omega t,$$

$$= \exp(-\Omega t\mathbf{i}\times)\mathbf{L}_0, \tag{21}$$

in the frame attached to the body. Under the action of such a torque, the transform of the equation of motion becomes

$$\boldsymbol{\omega}(\mathsf{S}) = (\mathsf{S} - \Omega\mathsf{K})^{-1}(\mathsf{A})^{-1}(\mathsf{S} + \Omega\mathbf{i}\times)^{-1}\mathbf{L}_0. \tag{22}$$

If $\mathbf{L}_0$ has no component along the axis of spin, the expansion of Eq. (22) is

$$\boldsymbol{\omega}(\mathsf{S}) = [(\mathsf{S} - \Omega\mathsf{K})^{-1}\mathsf{C} - \mathsf{C}(\mathsf{S} + \Omega\mathbf{i}\times)^{-1}]\mathbf{L}_0, \tag{23}$$

where it is found that $\Omega(1 - k^2)\mathsf{C} = \mathsf{K}(\mathsf{A})^{-1} - (\mathsf{A})^{-1}\mathbf{i}\times$. Again, the only effect of a spin-axis component would be an increase without limit of the spin-axis component of $\boldsymbol{\omega}$. Taking the inverse of Eq. (23), we get

$$\boldsymbol{\omega}(t) = \mathsf{C}\cdot\mathbf{L}_0 \cos k\Omega t + (\mathsf{K}/k)\mathsf{C}\cdot\mathbf{L}_0 \sin k\Omega t - \mathsf{C}\cdot\mathbf{L}_0 \cos \Omega t + \mathsf{C}\cdot\mathbf{i}\times\mathbf{L}_0 \sin \Omega t. \tag{24}$$

---

[6] Since K is degenerate, it yields a zero vector when operating on a vector parallel to the axis of spin. It is thus easily verified that the terms on the right in Eq. (18a) are "partial fractions" of Eq. (18).

The appearance of two elliptic precessions of different frequencies indicates that the axis of spin generates a cone whose right section is a Lissajous figure centered on the axis of precession.

THE EFFECT OF AN IMAGINARY k

It is noted from Eq. (10) that it is possible for k to be imaginary. If the quantity k_1 is defined by

$$(A_{22})^{\frac{1}{2}}k_1 = (-A_1^2 + A_1A_{21} - A_{22})^{\frac{1}{2}}, \tag{25}$$

then

$$\mathsf{K}^2 = k_1^2(\mathsf{I} - \mathbf{ii}). \tag{26}$$

Making use of Eq. (26), and noting that $(\mathsf{K})^3 = k_1^2\mathsf{K}$, the expansion of Eq. (8) leads to

$$\exp(\mathsf{K}\Omega t) = \mathbf{ii} + (\mathsf{I} - \mathbf{ii})\cosh k_1\Omega t + (\mathsf{K}/k_1)\sinh k_1\Omega t. \tag{27}$$

It is obvious that Eq. (27) indicates motion which increases exponentially with time. Physically, this means that the original dynamic equilibrium was unstable.

APPENDIX

If the tensors A and $\mathbf{i}\times$ are written in matrix form with respect to a rectangular coordinate system, i.e.,

$$\mathsf{A} = \begin{Vmatrix} A_1 & 0 & 0 \\ 0 & B & -D \\ 0 & -D & C \end{Vmatrix} \qquad \mathbf{i}\times = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{Vmatrix},$$

the following relationships are readily derived:

$$\mathsf{A}\mathbf{i}\times + \mathbf{i}\times\mathsf{A} = A_{21}\mathbf{i}\times, \tag{28}$$

$$\mathsf{A}\mathbf{i}\times\mathsf{A} = A_{22}\mathbf{i}\times. \tag{29}$$

From the definition of K, and recalling that $\mathbf{i}\cdot\mathsf{A} = A_1\mathbf{i}$, we have

$$\mathsf{K} = A_1(\mathsf{A})^{-1}\mathbf{i}\times - (\mathsf{A})^{-1}\mathbf{i}\times\mathsf{A}. \tag{30}$$

Applying Eq. (29) to the first term on the right and Eq. (30) to the second yields

$$\begin{aligned} A_{22}\mathsf{K} &= A_1\mathbf{i}\times\mathsf{A} - A_{22}A_{21}(\mathsf{A})^{-1}\mathbf{i}\times + A_{22}\mathbf{i}\times, \\ &= (A_1 - A_{21})\mathbf{i}\times\mathsf{A} + A_{22}\mathbf{i}\times. \end{aligned} \tag{31}$$

Upon squaring both sides of Eq. (31), it is seen that

$$\begin{aligned} A_{22}^2(\mathsf{K})^2 &= (A_1 - A_{21})^2\mathbf{i}\times\mathsf{A}\mathbf{i}\times\mathsf{A} \\ &\quad + A_{22}(A_1 - A_{21})\mathbf{i}\times(\mathsf{A}\mathbf{i}\times + \mathbf{i}\times\mathsf{A}) + A_{22}^2\mathbf{i}\times\mathbf{i}\times \\ &= A_{22}(A_1 - A_{21})^2\mathbf{i}\times\mathbf{i}\times \\ &\quad + A_{21}A_{22}(A_1 - A_{21})\mathbf{i}\times\mathbf{i}\times + A_{22}^2\mathbf{i}\times\mathbf{i}\times, \end{aligned} \tag{32}$$

or

$$A_{22}(\mathsf{K})^2 = -(A_1^2 - A_1A_{21} + A_{22})(\mathsf{I} - \mathbf{ii}), \tag{33}$$

from which Eqs. (9) and (10) follow.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 10 OCTOBER 1964

An Iterative Solution of the N/D Equation*

MYRON BANDER
Stanford Linear Accelerator Center, Stanford University, Stanford, California
(Received 19 May 1964)

An iterative scheme is presented for solving the N/D equations in the case where the left-hand cut consists entirely as a sum of poles. At no step is recourse to a matrix inversion of an algebraic system required. A scheme for approximating arbitrary cuts by sequences of poles is also presented.

I. INTRODUCTION

In many applications of the partial wave dispersion relations one is faced with nonlinear integral equations which may be linearized by the N/D method.[1] A left-hand cut discontinuity for the amplitude is assumed to be of some given form, and likewise some assumptions are made about the unitarity on the right-hand cut; usually elastic unitarity is assumed or the inelastic contributions are approximated by several two-body channels.[2] The resulting integral equations are either of the Fredholm type or have a kernel which is singular due only to infinite integration ranges. This difficulty is usually overcome by the introduction of a cutoff and the resulting equations lend themselves to standard numerical solutions, which in practice usually require the use of a high-speed computer.

Often a more drastic assumption is made about the left-hand cut, and it is replaced by a series of poles.[3] The integral equations are then reducible to a linear algebraic system. If the number of poles is large, we are then faced with the inversion of matrices of large order, which, even if the problems are solved with the aid of computers, it is the matrix inversion which consumes most of the time.

In this article an iterative method is presented by which we may go from an n-pole to an $(n + 1)$-pole problem directly without ever introducing the necessity of matrix inversion. This method is also applicable to the following case. Suppose that we have the solution for a certain left-hand cut, then we may immediately obtain the solution for a new cut which is equal to the old one plus a finite number of pole terms. This method is possible due to the fact that we have a freedom of choosing arbitrarily a subtraction point where the D function is normalized to unity.

Section II is devoted to the derivation of the iteration scheme starting from a general left-hand cut, and adding an arbitrary number of poles. In Sec. III this scheme is specialized to the left-hand cut consisting entirely of poles. In Sec. IV a discussion is given on how to obtain a sequence of poles approximation to any cut. All the results are given for a single-channel case, although they may easily be generalized to a many-channel problem.

II. ITERATION ON THE RESOLVENT KERNELS

In the N/D method we write the amplitude as a ratio of two functions N and D which have discontinuities, respectively, on the left or, respectively, right-hand cuts only. We choose to write an integral equation for N and express D in terms of N.

$$N(x; x_p) = B(x) + \frac{1}{\pi}\int \frac{\rho(z)N(z; x_p)}{(z-x)(z-x_p)} \times [B(z)(z-x_p) - B(x)(x-x_p)]\,dz, \quad \text{(II 1)}$$

$$D(x; x_p) = 1 - \frac{x-x_p}{\pi}\int \frac{\rho(z)N(z; x_p)}{(z-x)(z-x_p)}\,dz, \quad \text{(II 2)}$$

where $B(z)$ is a known function with only left-hand discontinuities and $\rho(z)$ is a phase space factor, which is a known kinematical function. The integrals in the above equation run over only the positive real axis from the start of the elastic cut. x_p is an arbitrary point at which we may normalize $D(x; x_p)$ to unity. The ratio $N(x; x_p)/D(x; x_p)$ is independent of x_p,[4] and in terms of the function $\rho(x)$ equals $\exp[i\delta(x)]\sin\delta(x)/\rho(x)$.

Instead of Eq. (II 1) let us consider an equation with the inhomogeneous term $B(x)$ replaced by an arbitrary function, $f(x)$.

$$N_f(x; x_p) = f(x) + \frac{1}{\pi}\int \frac{\rho(z)N_f(z; x_p)}{(z-x)(z-x_p)} \times [B(z)(z-x_p) - B(x)(x-x_p)]\,dz. \quad \text{(II 3)}$$

* Supported by the U. S. Atomic Energy Commission.
[1] G. F. Chew and S. Mandelstam, Phys. Rev. 119, 467 (1960).
[2] J. Bjorken, Phys. Rev. Letters 4, 473 (1960).
[3] See for example: H. P. Noyes and D. Y. Wong, Phys. Rev. Letters 3, 191 (1959); W. Frazer and J. Fulco, *ibid.* 2, 365 (1959).
[4] This result has been obtained by A. W. Martin (unpublished).

We assume that the above integral equation has a solution; namely, there exists a resolvent kernel, $G(x, y; x_p)$ such that

$$N_f(x; x_p) = \int G(x, y; x_p) f(y)\, dy. \quad \text{(II 4)}$$

The first step in this iteration procedure is to find the transformation between a resolvent kernel $G(x, y; x_p)$ for a subtraction point at x_p and the resolvent kernel $G(x, y; x_q)$ for another subtraction point x_q. Adding and subtracting to Eq. (II 3)

$$-\frac{1}{\pi}\int \frac{\rho(z) N_f(z; x_p)}{(z-x)(z-x_q)} B(x)(x - x_q)\, dz, \quad \text{(II 5)}$$

we obtain:

$$N_f(x; x_p) = f(x) + \frac{1}{\pi}\int \frac{\rho(z) N_f(z, x_p)}{(z-x)(z-x_q)} \times [B(z)(z-x_q) - B(x)(x-x_q)]\, dz + \frac{B(x)}{\pi}(x_p - x_q)\int \frac{\rho(z)N_f(z; x_p)}{(z-x_q)(z-x_p)}\, dz. \quad \text{(II 6)}$$

As the last term of Eq. (II 6) is $CB(x)$, where C is independent of x, we may write

$$N_f(x; x_p) = \int G(x, y; x_q)[f(y) + CB(y)]\, dy, \quad \text{(II 7)}$$

where

$$C = \frac{x_p - x_q}{\pi}\int \frac{\rho(z) N_f(z; x_p)}{(z-x_q)(z-x_p)}\, dz. \quad \text{(II 8)}$$

Substituting Eq. (II 7) into Eq. (II 8) we obtain C:

$$C = \frac{\int H(y; x_p, x_q) f(y)\, dy}{1 - \int H(y; x_p, x_q) B(y)\, dy}, \quad \text{(II 9)}$$

$$H(y; x_p, x_q) = \frac{x_p - x_q}{\pi} \times \int \frac{\rho(z)}{(z-x_p)(z-x_q)} G(z, y; x_q)\, dz. \quad \text{(II 10)}$$

The sought for relation between $G(x, y; x_p)$ and $G(x, y; x_q)$ is

$$G(x, y; x_p) = G(x, y; x_q) + \frac{\left[\int G(x, y'; x_q) B(y')\, dy'\right] H(y; x_p, x_q)}{1 - \int H(y'; x_p, x_q) B(y')\, dy'}. \quad \text{(II 11)}$$

Now the iteration scheme may be outlined. Suppose we add to $B(x)$ an extra pole,

$$\tilde{B}(x) = B(x) + \gamma_1/(x - x_1). \quad \text{(II 12)}$$

Equation (II 3) for an arbitrary subtraction point x_p becomes

$$\tilde{N}_f(x; x_p) = f(x) + \frac{1}{\pi}\int \frac{\rho(z)}{(z-x)(z-x_p)} \times \left[B(z)(z-x_p) - B(x)(x-x_p) + \gamma_1\left(\frac{z-x_p}{z-x_1} - \frac{x-x_p}{x-x_1}\right)\right]\tilde{N}_f(z; x_p)\, dz. \quad \text{(II 13)}$$

If we choose $x_p = x_1$ the kernel of Eq. (II 13) reduces to the kernel of Eq. (II 3). Thus

$$\tilde{G}(x, y; x_1) = G(x, y; x_1). \quad \text{(II 14)}$$

The iteration scheme we propose is as follows: Given $G(x, y; x_p)$ we obtain via Eq. (II 11) $G(x, y; x_1)$ which equals $\tilde{G}(x, y; x_1)$ which we may again via Eq. (II 11) transform to another subtraction point x_2 and add a pole at x_2. If we denote by $G^{(n)}(x, y; x_p)$ the resolvent kernel with n extra poles located at x_i, with residues γ_i, $i = 1, \cdots, n$, the scheme may be outlined as

$$G^{(n)}(x, y; x_n) \to G^{(n)}(x, y; x_{n+1}) = G^{(n+1)}(x, y; x_{n+1}) \to G^{(n+1)}(x, y; x_{n+2}) = G^{(n+2)}(x, y; x_{n+2}) \to \cdots, \quad \text{(II 15)}$$

where the arrows indicate application of Eq. (II 11) with appropriate $B(x)$, namely in going from $G^{(l)}(x, y; x_l)$ to $G^l(x, y; x_{l+1})$ the $B(x)$ that enters into Eq. (II 11) is

$$B(x) + \sum_{i=1}^{l} \frac{\gamma_i}{x - x_i}. \quad \text{(II 16)}$$

Thus, each step of the iteration procedure is reduced to quadratures and at no point do we encounter a problem of matrix inversion.

Before proceeding further we show that the amplitude $N(x; x_p)/D(x, x_p)$ is independent of x_p. By definition

$$N(x; x_p) = \int G(x, y; x_p) B(y)\, dy, \quad \text{(II 17)}$$

$$D(x_p; x_q) = 1 - \int H(y; x_p, x_q) B(y)\, dy. \quad \text{(II 18)}$$

From Eq. (II 11) we obtain

$$N(x; x_p) = N(x; x_q)/D(x_p; x_q). \quad \text{(II 19)}$$

As $D(x, x_q)/D(x_p, x_q)$ is one for $x = x_p$ and its imaginary part equals $-\rho(x)N(x; x_p)$, it satisfies

Eq. (II 2) with the subtraction point at x_p, and thus $N(x; x_p)/D(x; x_p) = N(x; x_q)/D(x; x_q)$.[4]

III. ITERATION OF THE n-POLE PROBLEM

If $B^n(x)$ consists entirely as a sum of pole terms

$$B^n(x) = \sum_{i=1}^{n} \frac{\gamma_i}{x - x_i}, \qquad \text{(III 1)}$$

the iteration scheme and especially Eq. (II 11) take on a much simpler form. In this case Eq. (II 3) may be reduced to

$$\begin{aligned} N_f(x; x_p) &= f(x) - \frac{1}{\pi} \sum_{i=1}^{n} \frac{\gamma_i(x_i - x_p)}{x - x_i} \\ &\quad \times \int \frac{\rho(z) N_f(z; x_p)}{(z - x_p)(z - x_i)} dz, \\ &= f(x) - \sum_{i=1}^{n} \frac{n_i(x_p)}{x - x_i}, \end{aligned} \qquad \text{(III 2)}$$

where the $n_i(x_p)$ are the solutions of a linear algebraic system;

$$\begin{aligned} n_i(x_p) &= \frac{1}{\pi} \gamma_i (x - x_p) \int \frac{\rho(z) f(z)}{(z - x_p)(z - x_i)} dz \\ &\quad - \frac{1}{\pi} \gamma_i (x_i - x_p) \int \frac{\rho(z)}{(z - x_p)(z - x_i)} \sum_{j=1}^{n} \frac{n_j(x_p)}{z - x_j} dz. \end{aligned} \qquad \text{(III 3)}$$

There exists a resolvent matrix $G_{ij}^{(n)}(x_p)$ such that

$$\begin{aligned} n_i(x_p) &= \sum_{j=1}^{n} G_{ij}(x_p) \gamma_j \frac{(x_j - x_p)}{\pi} \\ &\quad \times \int \frac{\rho(z) f(z)}{(z - x_p)(z - x_j)} dz. \end{aligned} \qquad \text{(III 4)}$$

Going through a procedure analogous to that of Sec. II we derive a transformation between $G_{ij}(x_p)$ and $G_{ij}(x_q)$

$$G_{ij}(x_p) = G_{ij}(x_q) + \frac{G_{im}(x_q)\gamma_m K(x_p, x_q)_l G_{lj}(x_q)}{1 - K(x_p, x_q)_l G_{lk}(x_q)\gamma_k}, \qquad \text{(III 5)}$$

where

$$K(x_p, x_q)_l = \frac{1}{\pi} (x_p - x_q) \int \frac{\rho(z)\, dz}{(z - x_p)(z - x_q)(z - x_l)}, \qquad \text{(III 6)}$$

and the summation convention has been adopted. The iterative scheme analogous to Eq. (II 15) is:

$$\begin{aligned} &G^{(k)}(x_p) \to G^{(k)}(x_{k+1}), \\ &G^{(k+1)}(x_{k+1}) = G^{(k)}(x_{k+1}) \oplus 1, \\ &G^{(k+1)}(x_{k+1}) \to G^{(k+1)}(x_{k+2}), \\ &\quad \vdots \end{aligned} \qquad \text{(III 7)}$$

where the arrow indicates an application of Eq. (III 5), and the notation $A \oplus 1$ means that if A is a $k \times k$ matrix, $A \oplus 1$ is a $(k + 1) \times (k + 1)$ matrix with

$$\begin{aligned} (A \oplus 1)_{ij} &= A_{ij}, \quad i, j \leq k, \\ &= 0, \quad i = k + 1, \quad j \leq k \\ &\qquad \text{or} \quad j = k + 1, \quad i \leq k, \\ &= 1; \quad i = j = k + 1. \end{aligned} \qquad \text{(III 8)}$$

IV. APPROXIMATION OF CUTS BY SEQUENCES OF POLES[5]

The functions $B(x)$ appearing in the previous equations are analytic functions cut along curves in the complex plane which are disjointed from the right-hand unitarity cut. In the equal mass case it is generally a single cut running along the negative real axis. More generally it may have additional cuts along finite curves.[6] In most applications $B(x)$ has at least one cut extending to infinity. To discuss any approximation technique it is convenient to make a change of variables such as

$$x = a/(u + b), \qquad \text{(IV 1)}$$

which makes all integration ranges finite. Such a change leaves the integral equations (II 1)–(II 3) of the same form. Now the problem is to approximate the transformed kernel $B(u)$ by a sequence of pole. The general expression for $B(u)$ will be of the form,

$$B(u) = (u + b) \sum_i \int_{c_i} \frac{\omega_i(z)}{z - u} dz. \qquad \text{(IV 2)}$$

Baker, Gammel, and Wills[7] have suggested a scheme using the Padé approximants. Their scheme consists of expanding $B(u)$ as a power series in u^{-n}, and writing each partial sum as a Padé approximant. These authors have shown that the sequence of approximants converges to the desired function under very general conditions. The cuts of Eq.

[5] All the results of this section are based on the analytic theory of continued fractions and on the theory of orthogonal polynomials. General references for this section are: H. S. Wall, *Analytic Theory of Continued Fractions* (D. Van Nostrand, Inc., Princeton, New Jersey, 1948); J. A. Shohat and J. D. Tamarkin, *The Problem of Moments* (American Mathematical Society, New York, New York, 1943), Chaps. I and II; O. Perron, *Die Lehre von den Kettenbrüchen* (B. G. Teubner, Leipzig and Berlin, 1913); G. Szegö, *Orthogonal Polynomials* (American Mathematical Society, New York, New York, 1959), specifically Chaps. II, III, and XVI.

[6] In problems involving fermions it is convenient to work in the energy instead of energy-squared plane, where the unitarity cut runs along the real axis excluding an interval about zero, and the function $B(x)$ has a cut along the imaginary axis.

[7] G. A. Baker, J. L. Gammel, and J. G. Wills, J. Math. Anal. Appls. **2**, 21 (1961).

(IV 2) are approximated by a sequence of poles whose positions approach the cuts themselves.

We shall present a method, which, although much more tedious, has the advantage that it makes few assumptions on the function $B(u)$ and likewise shows that the poles of the approximating sequences are not on the unitarity cut. In actual practice the scheme of Ref. 7 is strongly recommended. Although both methods are derived from the analytic structure of continuous fractions, the exact relation between them is not investigated.

The assumptions we shall make on the function $B(u)$ is that the contours c_i are rectifiable; that there exists a convex region containing each c_i,[8] such that it does not intersect the unitarity cut, and that the functions $\omega_i(z)$ are of bounded variation. Under these assumptions, let us rewrite Eq. (IV 2) as

$$B(u) = \sum_i \eta_i \int_{c_i} \frac{v_i(z)\,|dz|}{u-z}, \qquad \text{(IV 3)}$$

where $|dz|$ is the arc length along the curves c_i, and the factors η_i are chosen so that $v_i(z)$ is a nonnegative real function.

Now we are in a position to approximate each term in Eq. (IV 3) by a sequence of poles. As $v_i(z)$ is a nonnegative function, we may define a set of orthogonal polynomials with $v_i(z)$ as a weight function, i.e.,

$$\int_{c_i} p_n^{(i)*}(u) p_m^{(i)}(u) v_i(u)\,|du| = \delta_{mn}. \qquad \text{(IV 4)}$$

The coefficients of the $p_n^{(i)}(u)$ are real, and the zeros lie in the least convex region containing the curve c_i, and approach c_i for n sufficiently large. The zeros of $p_n^{(i)}(u)$ do not lie along the end points of c_i for any finite n.

Let

$$\int_{c_i} v_i(u) u^{*n} u^m\,|du| = c_{mn} \qquad \text{(IV 5)}$$

and

$$Q_n^{(i)}(x) = \int_{c_i} \frac{p_n^{(i)}(u) - p_n^{(i)}(x)}{u-x} v_i(u)\,|du|.$$

The $Q_n^{(i)}(x)$ are polynomials of degree $n-1$. By a theorem due to Markoff we have our result:

$$\int \frac{v_i(u)\,|du|}{u-x} = -\lim_{n\to\infty} c_{00}^2 (c_{00}c_{11} - c_{10}c_{01})^{-\frac{1}{2}} \frac{Q_n^{(i)}(x)}{P_n^{(i)}(x)},$$

which by partial fractions may be expressed as a sum of poles.

[8] The end of the transformed unitarity cut may coincide with the end of one of the cuts. This will cause no difficulty.

Duality Invariance and Riemannian Geometry

R. PENNEY
Scientific Laboratory, Ford Motor Company, Dearborn, Michigan
(Received 21 April 1964)

It is shown that the postulate of indistinguishability of the Maxwell field tensor from its dual leads to the concept of the electromagnetic field tensor as a spinor component in dual space. The demand for algebraic consistency dictates a unique connection with the gravitational field. The Maxwell field must be viewed as a set of potentials, and the necessity for a duality gauge condition excludes the existence of magnetic monopoles.

I. INTRODUCTION

It is the purpose of the present paper to exploit the underlying symmetry properties of the Maxwell–Einstein field to show that one is led, solely by algebraic considerations, to field equations relating the electromagnetic field to geometry.

We shall explicitly assume the following simple postulates: *Postulate* I. Space–time is Riemannian due to the presence of electromagnetic fields. *Postulate* II. There exists in space–time a second rank antisymmetric tensor, $F_{\mu\nu}$, called the Maxwell field tensor, which contains all of the information concerning the electromagnetic field; that is, we know of no restrictions on $F_{\mu\nu}$. *Postulate* III. The field equations relating the geometry to the electromagnetic field must be *algebraic*. *Postulate* IV. The laws of physics must be either first or second order, and, at most, linear in the second-order derivatives of physical quantities.

In Sec. II, we will apply these postulates to the Maxwell field and thereby show that $F_{\mu\nu}$ must be considered as a "potential", which must be subjected to gauge conditions. In Sec. III we will examine the fundamental tensors of Riemannian geometry and exhaust the possibilities for construction of field equations. In Sec. IV, we shall derive the field equations resulting from our postulates, and in Sec. V examine some of the consequences of them.

At this point, it seems prudent to explain our choice of postulates. The first postulate is in accord with Einstein's original ideas, and, classically at least, is an expression of the hope that one may understand physics on the basis of the Maxwell–Einstein field, without extraneous elements. The second postulate is taken as fairly self-evident, since we are only assuming existence. Essentially, it is a renunciation of any *a priori* knowledge. The third postulate is not as arbitrary as it may seem. Essentially, it is an expression of doubt. For example, as is well known, Noether's theorem shows that the conservation of energy and momentum is a consequence of the assumption of translational invariance of physical laws, which assumption is not necessarily valid globally. Thus, if we are to maintain maximum generality in our laws, we must invoke only symmetry conditions which are independent of the structure of the universe; we must not invoke any laws which involve derivatives of the $F_{\mu\nu}$. The last postulate is demanded by the necessity of physical interpretation, and seems to be a requisite for all physical theories.

We shall adopt the Einstein summation convention throughout, and make no distinction between Latin or Greek indices.

II. DUALITY SPACE

Postulate I automatically implies the existence not only of the normal Einstein pseudometric $g_{\mu\nu}$, but also the existence of the Riemannian–Christoffel tensor which obeys the relations:

$$R_{\mu\nu\sigma\rho} = -R_{\nu\mu\sigma\rho} = -R_{\mu\nu\rho\sigma} = R_{\sigma\rho\mu\nu}, \tag{1}$$

$$R_{\mu\nu\sigma\rho} + R_{\mu\rho\nu\sigma} + R_{\mu\sigma\rho\nu} = 0, \tag{2}$$

and, thereby, has *20* independent components. This tensor, as is well known, contains the maximal information about the space, while some information is lost in contracting it. Therefore, if we are to fully understand the geometry-field connection, we must obtain algebraic connections using the full Riemann–Christoffel tensor, not its contractions.

Postulate IV implies that we must construct our field equations using only $R_{\mu\nu\sigma\rho}$ and $g_{\lambda\beta}$, since, as is well known, these tensors and their combinations exhaust the possibilities of fundamental tensors.[1] That is, any other tensors in Riemannian geometry which are at most linear in the second

[1] A. S. Eddington, *Theory of Relativity* (Cambridge University Press, London, England, 1924), 2nd ed., p. 79.

derivatives of the $g_{\lambda\beta}$ may be constructed from the $R_{\mu\nu\sigma\rho}$.

We see, therefore, that, in general, our field equations must be a relation between fourth-rank tensors of the Maxwell field and the Einstein field.

According to Postulate II, the maximum knowledge concerning the electromagnetic field is given by $F_{\mu\nu}$, or, equivalently, by $*F_{\mu\nu}$ where the dual tensor is defined by

$$*F_{\mu\nu} \equiv \tfrac{1}{2}\eta_{\mu\nu\lambda\rho}F^{\lambda\rho}, \tag{3}$$

where we have used the "permutation tensor,"[2]

$$\eta_{\mu\nu\lambda\rho} \equiv -(-g)^{\frac{1}{2}}\epsilon_{\mu\nu\lambda\rho}, \tag{4}$$

$$\eta^{\mu\nu\lambda\rho} \equiv +(-g)^{-\frac{1}{2}}\epsilon_{\mu\nu\lambda\rho}. \tag{5}$$

$\epsilon_{\mu\nu\lambda\rho}$ is the usual Levi–Civita symbol, and g is the determinant of the $g_{\mu\nu}$.

A priori, we have no way of distinguishing $F_{\mu\nu}$ from $*F_{\mu\nu}$, so that the electromagnetic field must be invariant under the transformation:

$$\begin{aligned} F'_{\mu\nu} &= aF_{\mu\nu} + b\,*F_{\mu\nu}, \\ *F'_{\mu\nu} &= cF_{\mu\nu} + d\,*F_{\mu\nu}, \end{aligned} \tag{6}$$

where the a, b, c, d are Lorentz-invariant parameters. The coefficients must be *real*, since the field tensor is, and further, since

$$**F_{\mu\nu} = -F_{\mu\nu}, \tag{7}$$

we see that we must have

$$a = d \qquad c = -b. \tag{8}$$

We have, therefore, that the Maxwell field must be invariant under the transformation

$$\begin{bmatrix} F'_{\mu\nu} \\ *F'_{\mu\nu} \end{bmatrix} = M\begin{bmatrix} F_{\mu\nu} \\ *F_{\mu\nu} \end{bmatrix}, \tag{9}$$

where M is a *unitary* matrix. We see by inspection that the set of matrices forms a group. The invariance of the Maxwell field under M transformations we will call "duality invariance," following accepted convention.[3]

To proceed, we note the well-known fact that $F_{\mu\nu}$ has two Lorentz invariants associated with it, viz.,

$$\begin{aligned} I_1 &\equiv F_{\mu\nu}\ \ F^{\mu\nu}, \\ I_2 &\equiv F_{\mu\nu}\,*F^{\mu\nu}. \end{aligned} \tag{10}$$

Indeed, the field can be characterized by these invariants. Yet, these invariants are *not* duality invariant. However, one can easily deduce that

$$I^2 \equiv I_1^2 + I_2^2 \tag{11}$$

is both Lorentz invariant *and* duality invariant provided that

$$\det M = 1 \tag{12}$$

(det M cannot be -1 because of the reality of its components).

Since it is irrelevant whether we characterize the field by I_1, I_2 or by I, θ, where

$$\begin{aligned} I_1 &= I\cos 2\theta, \\ I_2 &= I\sin 2\theta, \end{aligned} \tag{13}$$

we lose no generality in taking

$$\begin{aligned} a &= \pm\cos\theta, \\ b &= \pm\sin\theta, \end{aligned} \tag{14}$$

so that

$$M = \pm\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}. \tag{15}$$

We now see that M is a unitary unimodular matrix. We further see that we may give a very picturesque meaning to the duality transformations. If we consider a two-dimensional space in which I_1, I_2 are Cartesian coordinates, or, alternatively, I and θ are polar coordinates, then the entity

$$\begin{bmatrix} F_{\mu\nu} \\ *F_{\mu\nu} \end{bmatrix}$$

is a "spinor" in this space. Of course, the entity

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix}$$

is a "vector" in the space.

In agreement with the concept of the field tensor as a spinor component, we note that the "value" of this spinor is zero,

$$(F_{\mu\nu}\ *F_{\mu\nu})\begin{pmatrix} F^{\mu\nu} \\ *F^{\mu\nu} \end{pmatrix} \equiv 0, \tag{16}$$

as follows from the definition of the dual. We have used the fact that the contravariant, or mixed, field tensor is also a spinor.

We remark further that the duality transformation applied to $F_{\mu\nu}$ yields a new spinor which is always orthogonal to the old, i.e.,

$$(F'_{\mu\nu}\ *F'_{\mu\nu})\begin{pmatrix} F^{\mu\nu} \\ *F^{\mu\nu} \end{pmatrix} \equiv 0. \tag{17}$$

[2] J. L. Synge, *Relativity, The General Theory* (North-Holland Publishing Company, Amsterdam, 1960), p. 18.
[3] J. A. Wheeler, *Geometrodynamics* (Academic Press Inc., New York, 1961), p. 239.

Using the properties thus far deduced, we note that we may unambiguously derive a second-rank tensor which is both Lorentz covariant, and duality invariant, viz.,

$$t_\mu^\alpha \equiv (F_{\mu\nu}\ {}^*F_{\mu\nu})\begin{pmatrix} F_{\alpha\nu} \\ {}^*F^{\alpha\nu}\end{pmatrix} = F_{\mu\nu}\ F^{\alpha\nu} + {}^*F_{\mu\nu}\ {}^*F^{\alpha\nu} \quad (18)$$

and this is the only such nonzero tensor available. It is perhaps not too surprising that t_μ^α is, apart from a factor of 2, the same as the usual energy–momentum tensor of the Maxwell field. The duality invariance of t_μ^α, which is a scalar in dual space, is obvious from the fact that it is a spinor dot-product, but one may also check its invariance directly.

It is important to note that the trace of t_μ^α vanishes identically, as can be seen from its definition.

A further remarkable property of t_μ^α is that

$$t_\mu^\alpha t_\alpha^\lambda = \tfrac{1}{16} I^2 \delta_\mu^\lambda, \quad (19)$$

a well-known fact which may be verified directly by using the two identities,

$$F_{\mu\beta}\ {}^*F^{\lambda\beta} \equiv \tfrac{1}{4} I_2 \delta_\mu^\lambda, \quad (20)$$

$$F^{i\alpha}\ F_{k\alpha} - {}^*F^{i\alpha}\ {}^*F_{k\alpha} \equiv \tfrac{1}{2} I_1 \delta_k^i, \quad (21)$$

as shown in the Appendix. We now see that it was *necessary* to take

$$\det M = 1,$$

since the duality invariance of t_μ^α implied the duality invariance of I^2.

The next lowest rank duality invariant tensor is given by

$$T_{\mu\nu\alpha\rho} = (F_{\mu\nu}\ {}^*F_{\mu\nu})\begin{pmatrix} F_{\alpha\rho} \\ {}^*F_{\alpha\rho}\end{pmatrix} = F_{\mu\nu}\ F_{\alpha\rho} + {}^*F_{\mu\nu}\ {}^*F_{\alpha\rho} \quad (22)$$

and this is the only fourth-rank tensor available, aside from duals of the same tensor. We have, then, that the only fourth-rank tensors available are

$$T_{\mu\nu\alpha\rho}, \quad {}^*T_{\mu\nu\alpha\rho}, \quad T^*_{\mu\nu\alpha\rho}, \quad {}^*T^*_{\mu\nu\alpha\rho}, \quad (23)$$

where we define the left-dual and the right-dual by

$${}^*T_{\mu\nu\alpha\rho} \equiv \tfrac{1}{2}\eta_{\mu\nu ab}T^{ab}{}_{\alpha\rho}, \quad (24)$$

$$T^*_{\mu\nu\alpha\rho} \equiv \tfrac{1}{2}\eta_{\alpha\rho cd}T_{\mu\nu}{}^{cd}. \quad (25)$$

We see by direct calculation that

$$\begin{aligned} {}^*T_{\mu\nu\alpha\rho} &= -T^*_{\mu\nu\alpha\rho} = -{}^*T_{\alpha\rho\mu\nu}, \\ T^*_{\mu\nu\alpha\rho} &= -{}^*T_{\alpha\rho\mu\nu}. \end{aligned} \quad (26)$$

Thus, we have only two *distinct* fourth-rank Maxwell tensors, and they have the symmetry properties:

$$T_{\mu\nu\alpha\rho} = -T_{\nu\mu\alpha\rho} = -T_{\mu\nu\rho\alpha} = T_{\alpha\rho\mu\nu}, \quad (28)$$

$$T_{\mu\nu\alpha\rho} + T_{\mu\alpha\rho\nu} + T_{\mu\rho\nu\alpha} = 0, \quad (29)$$

$${}^*T_{\mu\nu\alpha\rho} = -{}^*T_{\nu\mu\alpha\rho} = -{}^*T_{\mu\nu\rho\alpha} = -{}^*T_{\alpha\rho\mu\nu}. \quad (30)$$

It is seen that $T_{\mu\nu\alpha\rho}$ has exactly the same symmetry properties as $R_{\mu\nu\alpha\rho}$, but ${}^*T_{\mu\nu\alpha\rho}$ does not. The cyclic identity for $T_{\mu\nu\alpha\rho}$ is not obvious, but is shown in the Appendix. We note here some additional properties of $T_{\mu\nu\alpha\rho}$ which are proved in the Appendix, and will be important later:

$$g^{\mu\rho}T_{\mu\nu\alpha\rho} = -t_{\nu\alpha}, \quad (31)$$

$$g^{\mu\rho}g^{\nu\alpha}T_{\mu\nu\alpha\rho} \equiv 0, \quad (32)$$

$$T^{i\alpha\beta\gamma}T_{k\alpha\beta\gamma} = \tfrac{1}{4}(T^{abcd}T_{abcd})\delta_k^i, \quad (33)$$

$$2T_{\mu\nu\alpha\rho} \equiv t_{\mu\alpha}g_{\nu\rho} + t_{\nu\rho}g_{\mu\alpha} - t_{\mu\rho}g_{\nu\alpha} - t_{\nu\alpha}g_{\mu\rho}. \quad (34)$$

III. FUNDAMENTAL RIEMANNIAN TENSORS

Postulate IV allows us to exhaust the possibilities of geometric tensors available for the construction of field equations. At our disposal, we have only $g_{\mu\nu}$ and $R_{\mu\nu\alpha\rho}$, plus tensors derivable from these by operations involving contractions or dualism. Any other operations must necessarily increase the order of the derivatives of the $g_{\mu\nu}$, or produce tensors which are not linear in the second derivatives of the $g_{\mu\nu}$.

The *common* property of the two Maxwell tensors is a double antisymmetry in the indices. Certainly any Riemannian tensor which we may use in field equations must also possess this property.

We may now exhaust all the possibilities for fourth rank fundamental tensors by applying the dual operation and the contraction operations to the tensors $g_{\mu\nu}$ and $R_{\mu\nu\alpha\rho}$. Without contraction, we may form the tensors

$$R_{\mu\nu\alpha\rho}, \quad {}^*R_{\mu\nu\alpha\rho}, \quad R^*_{\mu\nu\alpha\rho}, \quad {}^*R^*_{\mu\nu\alpha\rho}. \quad (35)$$

With a single contraction, we have the tensors

$$V_{ijkm}, \quad {}^*V_{ijkm}, \quad V^*_{ijkm}, \quad {}^*V^*_{ijkm}, \quad (36)$$

$$W_{ijkm}, \quad {}^*W_{ijkm}, \quad W^*_{ijkm}, \quad {}^*W^*_{ijkm}, \quad (37)$$

and with a double contraction we have

$$G_{ijkm} \quad {}^*G_{ijkm} \quad G^*_{ijkm} \quad {}^*G^*_{ijkm}, \quad (38)$$

where we have defined

$$V_{ijkm} \equiv R_{ik}g_{jm} + R_{jm}g_{ik} - R_{im}g_{jk} - R_{jk}g_{im}, \quad (39)$$

$$W_{ijkm} \equiv R(g_{ik}g_{jm} - g_{jk}g_{im}), \quad (40)$$

$$G_{ijkm} \equiv g_{ik}g_{im} - g_{ik}g_{im}. \quad (41)$$

Apparently, we have 16 possible tensors for use in field equations. However, there are only 8 *distinct* tensors because of the identities

$$^*R_{ijkm} - R^*_{ijkm} = {}^*V_{ijkm} - \tfrac{1}{2}{}^*W_{ijkm}, \quad (42)$$

$$R^*_{ijkm} - {}^*R_{ijkm} = V^*_{ijkm} - \tfrac{1}{2}W^*_{ijkm}, \quad (43)$$

$$R_{ijkm} + {}^*R^*_{ijkm} = V_{ijkm} - \tfrac{1}{2}W_{ijkm}, \quad (44)$$

$$^*W_{ijkm} = W^*_{ijkm}, \quad (45)$$

$$^*W^*_{ijkm} = -W_{ijkm}, \quad (46)$$

$$^*G_{ijkm} = \eta_{ijkm}, \quad (47)$$

$$G^*_{ijkm} = \eta_{ijkm}, \quad (48)$$

$$^*G^*_{ijkm} = -G_{ijkm}, \quad (49)$$

which are easily verified by direct calculation.

Solely as a matter of convenience, we will rename some of the 8 distinct tensors by defining

$$A_{ijkm} \equiv R_{ijkm} + {}^*R^*_{ijkm}, \quad (50)$$

$$S_{ijkm} \equiv R_{ijkm} - {}^*R^*_{ijkm}, \quad (51)$$

$$P_{ijkm} \equiv {}^*R_{ijkm} + R^*_{ijkm}, \quad (52)$$

$$Q_{ijkm} \equiv {}^*R_{ijkm} - R^*_{ijkm}. \quad (53)$$

We note then the following properties of the tensors, ignoring indices for now,

$$\begin{array}{ll} ^*A = A, & A^* = -Q, \\ ^*Q = -A, & Q^* = A, \\ ^*S = P, & S^* = P, \\ ^*P = -S, & P^* = -S, \\ ^*G = \eta, & G^* = \eta, \\ ^*\eta = -G, & \eta^* = -G. \end{array} \quad (54)$$

Finally, we see that Q is antisymmetric for the double interchange of indices, e.g.,

$$Q_{ijkm} = -Q_{kmij}, \quad (55)$$

while all others are symmetric under the double interchange.

IV. FIELD EQUATIONS

On the basis of our postulates, the most general possible field equations we may have are

$$aA_{ijkm} + sS_{ijkm} + pP_{ijkm} + qQ_{ijkm} + g(R+\lambda)G_{ijkm} + \eta(R+\beta)\eta_{ijkm} = tT_{ijkm} + {}^*t^*T_{ijkm}. \quad (56)$$

But, if we consider the double interchange of indices we see that the field equations must split,

$$aA_{ijkm} + sS_{ijkm} + pP_{ijkm} + g(R+\lambda)G_{ijkm} + \eta(R+\beta)\eta_{ijkm} = tT_{ijkm}, \quad (57)$$

$$qQ_{ijkm} = {}^*t^*T_{ijkm}, \quad (58)$$

where the coefficients are (universal) constants. Further, if we take the left and right dual we see that we must have a further splitting,

$$aA_{ijkm} = tT_{ijkm}, \quad (59)$$

$$sS_{ijkm} + pP_{ijkm} + g(R+\lambda)G_{ijkm} + \eta(R+\beta)\eta_{ijkm} = 0, \quad (60)$$

$$aQ_{ijkm} = {}^*t^*T_{ijkm}, \quad (58)$$

but we see that the relation between Q_{ijkm} and $^*T_{ijkm}$ is already implicit in the field equation for T_{ijkm}.

We now readily see that our postulates demand the single set of field equations:

$$A_{ijkm} = kT_{ijkm}. \quad (61)$$

The other implied set of field equations must be trivially satisfied by vanishing coefficients, else we would have 19 field equations for the $g_{\mu\nu}$ and $F_{\mu\nu}$, and this cannot be. For a generally covariant theory, we must have no more than $(n-4)$ field equations for n field functions. Furthermore, the other set of field equations would imply a geometry unaffected by the sources, $F_{\mu\nu}$.

We now must ask whether our field equations are consistent, at least mathematically. First of all, we note that

$$^*A^*_{ijkm} \equiv A_{ijkm}, \quad (62)$$

$$^*T^*_{ijkm} \equiv T_{ijkm}, \quad (63)$$

$$g^{im}g^{ik}A_{ijkm} \equiv 0, \quad (64)$$

$$g^{im}g^{ik}T_{ijkm} \equiv 0, \quad (65)$$

$$A^{i\alpha\beta\gamma}A_{k\alpha\beta\gamma} \equiv \tfrac{1}{4}(A^{abcd}A_{abcd})\delta^i_k, \quad (66)$$

$$T^{i\alpha\beta\gamma}T_{k\alpha\beta\gamma} \equiv \tfrac{1}{4}(T^{abcd}T_{abcd})\delta^i_k, \quad (67)$$

where the relations follow solely from the definitions of the tensors, and are proved in the Appendix. Secondly, the two tensors agree in all symmetry properties.

It follows from the identities and the symmetry properties that our field equations constitute *nine* independent relations, which may be more clearly appreciated by realizing that, as Einstein[4] first

[4] A. Einstein, Math. Ann. 97, 99 (1926).

showed,

$$A_{ijkm} \equiv \mathring{R}_{ik}g_{jm} + \mathring{R}_{jm}g_{ik} - \mathring{R}_{im}g_{jk} - \mathring{R}_{jk}g_{im}, \quad (68)$$

where

$$\mathring{R}_{\mu\nu} \equiv R_{\mu\nu} - \tfrac{1}{4}Rg_{\mu\nu}. \quad (69)$$

Using this, we see that the field equations are completely reducible to

$$\mathring{R}_{\mu\nu} = -kt_{\mu\nu}, \quad (70)$$

which is obviously a set of *nine* independent equations, since the trace relation is missing.

We have been impelled by purely algebraic and group-theoretic arguments to a single set of field equations relating the geometry to the electromagnetic fields.

How must we interpret the field equations? We see from the postulate of duality invariance that the $F_{\mu\nu}$ are not uniquely determined by the geometry. That is, as far as $t_{\mu\nu}$ is concerned, the $F_{\mu\nu}$ *are potentials*, a fact which is easily understood from the realization that duality invariance is simply a form of gauge invariance for the potentials. On the basis of our postulates, therefore, we have no recourse but to accept the role of the $F_{\mu\nu}$ as completely analogous to that of the $g_{\mu\nu}$, viz., as potentials.

With the realization that the $F_{\mu\nu}$ are potentials, we see that we cannot in general have a determinate problem unless we pick a gauge. We must have some way of specifying the specific gauge for the $F_{\mu\nu}$ just as it is necessary in Lorentz electrodynamics to impose the Lorentz condition. By choosing a particular gauge, we must "destroy" the duality invariance.

How can we pick a meaningful gauge condition? We assume that our gauge condition must be globally valid and generally covariant, and therefore must be independent of the geometry. Thus, our gauge condition must involve only ordinary derivatives, *not* covariant differentiation. We are thereby forced to adopt the gauge condition:

$$\frac{\partial F_{\mu\nu}}{\partial x^\rho} + \frac{\partial F_{\rho\mu}}{\partial x^\nu} + \frac{\partial F_{\nu\rho}}{\partial x^\mu} = 0, \quad (71)$$

which is the only possible condition which does not involve geometrical quantities, yet is tensorial in character!

We see that the only possible way of picking a (global) gauge for the $F_{\mu\nu}$ automatically excludes the existence of magnetic monopoles, a very interesting fact, indeed.

We now note that the adoption of our gauge yields a mathematically determinate problem. We have the field equations:

$$\mathring{R}_{\mu\nu} = -kt_{\mu\nu}, \quad (72)$$

$$*F^{\mu\nu}{}_{|\mu} \equiv 0, \quad (73)$$

where we have written the gauge condition in an equivalent way, using the dual. We readily count $9 + 3 = 12$ independent equations since the identity:

$$*F^{\mu\nu}{}_{|\mu\nu} \equiv 0 \quad (74)$$

reduces the gauge condition to 3 relations. We also count *16* field functions,

$$g_{\mu\nu} \qquad F_{\mu\nu}$$

so, with the choice of 4 coordinate conditions, we have a determinate system.

V. CONSEQUENCES OF THE FIELD EQUATIONS

Let us examine the field equations derived in Part IV. We have

$$\mathring{R}_{\mu\nu} = -kt_{\mu\nu},$$

$$*F^{\mu\nu}{}_{|\nu} = 0,$$

and the application of the Bianchi identities leads to the relation

$$\frac{1}{4k}\frac{\partial R}{\partial x^\mu} = F_{\mu\nu}F^{\mu\lambda}{}_{|\lambda} \equiv F_{\mu\nu}J^\nu \quad (75)$$

and we see, therefore, that in a region of space-time which has a constant Riemannian curvature, the "current" vanishes, and $t_{\mu\nu}$ is conserved,

$$F^{\nu\lambda}{}_{|\lambda} = 0, \quad (76)$$

$$t^{\mu\nu}{}_{|\nu} = 0. \quad (77)$$

For such a region, the field equations may be written as

$$G_{\mu\nu} + \tfrac{1}{4}R_0 g_{\mu\nu} = -kt_{\mu\nu}, \quad (78)$$

where $G_{\mu\nu}$ is the usual Einstein tensor, and R_0, which plays the role of cosmological constant, is the constant value of R, the Ricci scalar. In general, the energy and momentum of the electromagnetic field is conserved only for regions of uniform curvature. More generally, we may consider an assemblage of incoherent electromagnetic fields such that the average "Lorentz force" is zero

$$\langle F_{\mu\nu}J^\nu\rangle = 0, \quad (79)$$

and it follows that the energy momentum tensor of such an assemblage will be conserved. For such gross matter, the field equations will be completely equivalent to the usual Einstein equations.

Thus, we see that the field equations derived by our postulates are quite equivalent to the macroscopic field equations of Einstein when we consider uncharged bodies. The real difference arises in treating those regions of space–time which one normally refers to as "charged." For such regions, the field equations imply that disruptive Coulomb forces will be balanced by a gravitational pressure, in accord with Einstein's original conceptions[5] of an "electron."

It is not difficult to show that a spherically symmetric static solution of our field equations, with an arbitrary charge distribution, is stable and, furthermore, transforms properly, i.e., its energy and momentum form a 4-vector. Thus a gravitational pressure, as prescribed by the field equations, is a suitable Poincaré "glue."

VI. CONCLUSIONS

We have shown that our postulate concerning the Maxwell field, which is essentially an expression of ignorance, leads one to a unique connection with the gravitational field, without invoking the conservation of energy. Since we cannot be sure of the isotropy of space–time, it would seem preferable to establish this connection by the algebraic manner used, rather than to demand, as is usually done, the condition

$$T^{\mu\nu}{}_{|\nu} = 0. \tag{80}$$

One obvious advantage of the approach described in the present work is that the role of the $F_{\mu\nu}$ as potentials is clearly shown. The necessity for a gauge condition then excludes magnetic monopoles. Since we do not assume Maxwell's equations as usually done, we feel absolutely no compulsion toward "completing the symmetry" of the relations:

$$F^{\mu\nu}{}_{|\nu} = J^{\mu}, \tag{81}$$

$$*F^{\mu\nu}{}_{|\nu} = 0. \tag{82}$$

In fact, our approach shows the fallaciousness of such a view. The real symmetry of the Maxwell field lies in its duality invariance, which is of algebraic character, not of differential character.

For a space of constant Ricci curvature, the field equations we have found are precisely those of "geometrodynamics,"[3] with the welcome exception that one less condition is necessary. The vanishing of the trace of $\mathring{R}_{\mu}^{\nu}$ is already assured.

APPENDIX

We present proofs of various identities used in the main text. Some of the identities are well known in the literature, but there do not appear to exist explicit proofs of them, so that it seems desirable to collect the formal proofs here.

Theorem I. The usual Maxwell energy tensor obeys the relation

$$t_{\mu}^{\alpha}t_{\alpha}^{\lambda} = \tfrac{1}{16}I^2\delta_{\mu}^{\lambda}.$$

Proof: We use the two identities quoted in the text, viz., Eqs. (20) and (21). We then have directly,

$$\begin{aligned} t^{i\beta}t_{k\beta} &= (F^{i\lambda}F_{\lambda}^{\beta} + *F^{i\lambda}*F_{\lambda}^{\beta})(F_{k\rho}F_{\beta}^{\rho} + *F_{k\rho}*F_{\beta}^{\rho}), \\ &= F^{i\lambda}F_{\lambda}^{\beta}F_{k\rho}F_{\beta}^{\rho} + F^{i\lambda}F_{\lambda}^{\beta}*F_{k\rho}*F_{\beta}^{\rho} \\ &\quad + *F^{i\lambda}*F_{\lambda}^{\beta}F_{k\rho}F_{\beta}^{\rho} + *F^{i\lambda}*F_{\lambda}^{\beta}*F_{k\rho}*F_{\beta}^{\rho}, \\ &= F^{i\lambda}F_{\beta\lambda}(*F_{k\rho}*F^{\beta\rho} + \tfrac{1}{2}I_1\delta_k^{\beta}) + \tfrac{1}{4}I_2\delta_{\lambda}^{\rho}F^{i\lambda}*F_{k\rho} \\ &\quad + \tfrac{1}{4}I_2\delta_{\lambda}^{\rho}*F^{i\lambda}F_{k\rho} + *F^{i\lambda}*F_{\beta\lambda}(F_{k\rho}F^{\beta\rho} - \tfrac{1}{2}I_1\delta_k^{\beta}), \end{aligned}$$

collecting terms, we have

$$t^{i\beta}t_{k\beta} = \tfrac{1}{16}(I_1^2 + I_2^2)\delta_k^i,$$

and we have a direct algebraic proof which does not require any usage of the special properties of the $F_{\mu\nu}$, nor any use of Lorentz transformations.

Theorem II. The fourth-rank Maxwell tensor

$$T_{\mu\nu\alpha\rho} \equiv F_{\mu\nu}F_{\alpha\rho} + *F_{\mu\nu}*F_{\alpha\rho}$$

obeys the cyclic identity:

$$T_{\mu\nu\alpha\rho} + T_{\mu\alpha\rho\nu} + T_{\mu\rho\nu\alpha} = 0.$$

Proof: By direct calculation we have that

$$\begin{aligned} T_{1234} + T_{1423} + T_{1342} &= (F_{12}F_{34} + F_{14}F_{23} + F_{13}F_{42}) \\ &\quad + (*F_{12}*F_{34} + *F_{14}*F_{23} + *F_{13}*F_{42}). \end{aligned}$$

Also by direct calculation we see that

$$(\det F_{\mu\nu})^{\frac{1}{2}} = F_{12}F_{34} + F_{14}F_{23} + F_{13}F_{42},$$

$$(\det F^{\mu\nu})^{\frac{1}{2}} = F^{12}F^{34} + F^{14}F^{23} + F^{13}F^{42},$$

so that we have

$$T_{1234} + T_{1423} + T_{1342} = (\det F_{\mu\nu})^{\frac{1}{2}} - g(\det F^{\mu\nu})^{\frac{1}{2}} \equiv 0.$$

Theorem III. The fourth-rank Maxwell tensor obeys the relations

$$T^{i\alpha\beta\gamma}T_{k\alpha\beta\gamma} = \tfrac{1}{4}(T^{abcd}T_{abcd})\delta_k^i,$$

$$2T_{\mu\nu\alpha\rho} = t_{\mu\alpha}g_{\nu\rho} + t_{\nu\rho}g_{\mu\alpha} - t_{\mu\rho}g_{\nu\alpha} - t_{\nu\alpha}g_{\mu\rho}.$$

[5] A. Einstein, Sitzber. Preuss. Akad. Wiss., 349 (1919).

Proof: The first identity may be verified by direct calculation, viz.,

$$\begin{aligned} T^{i\alpha\beta\gamma}T_{k\alpha\beta\gamma} &= [F^{i\alpha}F^{\beta\gamma} + {}^*F^{i\alpha}\,{}^*F^{\beta\gamma}][F_{k\alpha}F_{\beta\gamma} + {}^*F_{k\alpha}\,{}^*F_{\beta\gamma}],\\ &= (F^{\beta\gamma}F_{\beta\gamma})[F^{i\alpha}F_{k\alpha} - {}^*F^{i\alpha}\,{}^*F_{k\alpha}]\\ &\quad + (F^{\beta\gamma}\,{}^*F_{\beta\gamma})[F^{i\alpha}\,{}^*F_{k\alpha} + {}^*F^{i\alpha}F_{k\alpha}],\\ &= \tfrac{1}{2}I_1^2\delta_k^i + \tfrac{1}{2}I_2^2\delta_k^i, \end{aligned}$$

while the second identity follows directly from Lanczos'[6] proof of the same relation for the Riemann tensor.

[6] C. Lanczos, Ann. Math. **39**, 842 (1938).

Theorem IV. The tensor $A_{i\alpha\beta\gamma}$ satisfies the relation

$$A^{i\alpha\beta\gamma}A_{k\alpha b\gamma} = \tfrac{1}{4}(A^{abcd}A_{abcd})\delta_k^i.$$

Proof: This relation follows directly from a theorem due to Lanczos,[6] which we may state as:

$$R^{i\alpha\beta\gamma}\,{}^*R^*_{k\alpha\beta\gamma} = \tfrac{1}{4}(R_{\alpha\beta\gamma\delta}\,{}^*R^{*\alpha\beta\gamma\delta})\delta_k^i$$

and from the relation:

$$R^{i\alpha\beta\gamma}R_{k\alpha\beta\gamma} + {}^*R^{*i\alpha\beta\gamma}\,{}^*R^*_{k\alpha\beta\gamma} = \tfrac{1}{2}(R^{abcd}R_{abcd})\delta_k^i,$$

which is easily verified by direct calculation.

Representations of the Inhomogeneous Lorentz Group in Terms of an Angular Momentum Basis: Derivation for the Cases of Nonzero Mass and Zero Mass, Discrete Spin

J. S. LOMONT
Polytechnic Institute of Brooklyn, Brooklyn, New York
H. E. MOSES
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts*
(Received 21 May 1964)

In a previous paper the authors showed how the infinitesimal generators of the proper, orthochronous, inhomogeneous Lorentz group acted in a basis in which the square of the angular momentum, the z component of the angular momentum, the helicity and the energy were diagonal for the irreducible representations which correspond to the cases of nonzero and zero mass, discrete spin. In that paper no derivation of the results were given. It was possible, however, to verify them directly. In the present paper we carry out the derivation.

I. INTRODUCTION

We showed[1] how the infinitesimal generators of the proper, orthochronous, inhomogeneous Lorentz group act in a basis in which the energy, square of the angular momentum, z component of the angular momentum and helicity are diagonal for the irreducible representations corresponding to the cases of nonzero and zero mass, discrete-spin case. The derivations were not given. However, it was possible to verify the form directly by showing that the infinitesimal generators satisfied the correct commutation rules and had the proper invariants or Casimir operators, and that the set of infinitesimal generators were irreducible.

In the present paper we propose to show how we obtained the basis. The basis is obtained from the commutation rules and a knowledge of the form of the representations given in a basis in which the components of the momentum are diagonal (see, for example Ref. 2).

It should be mentioned that Pauli[3] has a very terse derivation for the case of zero mass and discrete spin. We were not aware of this reference when we began our work. Hence, our notation and procedure differ considerably from his, though the spirit of the derivation is the same. Our derivation for the massless case appears as a simplified version of the far more complicated derivation for the nonzero mass case.

We shall use the commutation rules for the infinitesimal generators as given in Eqs. (1.4)–(1.10) in Part I. From these commutation rules one can show that the operators H, $\mathbf{J}^2$, J_3, and $w = [H^2 - \mu^2]^{-\frac{1}{2}}w_0$ commute with each other, where $\mathbf{J}^2 = \sum_i J_i^2$, $w_0 = \mathbf{P}\cdot\mathbf{J} = \sum_i P_iJ_i$ and is thus the helicity operator, and μ is the mass of the particle.

We shall assume that H, $\mathbf{J}^2$, J_3, and w form a complete set of commuting variables. We shall find that we can obtain all the irreducible representations which we wish in terms of a basis in which this set is diagonal.

Let us denote the eigenvalue of H by E and the eigenvalue of w by α. Then from the results on the irreducible representations of the inhomogeneous Lorentz group range of E is given by $\mu < E < \infty$ for the positive energy representations and $\infty - < E < - \mu$ for the negative energy representations. In the case of nonzero mass, α can take on the values $-s, -s+1, \cdots, s-1, s$ where s is the spin of the particle.

In the case of zero mass, discrete spin α can take on only one value; either $\alpha = s$ or $\alpha = -s$, where s is the spin. That is, w is a scalar in the mass zero case. However, for the sake of uniformity of notation, it will be convenient to consider w to be an operator with only one eigenvalue.

We denote the eigenvalue of J_3 by m and of $\mathbf{J}^2$ by r.

Hence, a set of eigenkets $|E, r, m, \alpha\rangle$ exists such that

$$\begin{aligned} H\,|E,r,m,\alpha\rangle &= E\,|E,r,m,\alpha\rangle, \\ \mathbf{J}^2\,|E,r,m,\alpha\rangle &= r\,|E,r,m,\alpha\rangle, \\ J_3\,|E,r,m,\alpha\rangle &= m\,|E,r,m,\alpha\rangle, \\ w\,|E,r,m,\alpha\rangle &= \alpha\,|E,r,m,\alpha\rangle. \end{aligned} \tag{1.1}$$

* Operated with support from the U. S. Advanced Research Projects Agency.
[1] J. S. Lomont and H. E. Moses, J. Math. Phys. 5 (294), (1964). Henceforth referred as Part I.
[2] J. S. Lomont and H. E. Moses, J. Math. Phys. 3, (405), (1962).
[3] W. Pauli, CERN Rept 56-31, Geneva (1956) (unpublished).

From the last of Eqs. (1.1) the helicity operator w_0 is given by

$$w_0 \,|E, r, m, \alpha\rangle = p\alpha \,|E, r, m, \alpha\rangle, \tag{1.2}$$

where

$$p = (E^2 - \mu^2)^{\frac{1}{2}}. \tag{1.2a}$$

The helicity operator is also diagonal in this representation.

II. REPRESENTATIONS OF THE ANGULAR MOMENTUM OPERATORS

In the present section we shall show that the eigenvalue r of $\mathbf{J}^2$ has the form $r = j\,(j+1)$, where j is either always a nonnegative integer or a positive half-odd integer. Furthermore, we show that for a fixed value of j, m takes on the $2j+1$ values $-j, -j+1, \cdots, j-1, j$. Finally, we shall show how J_1 and J_2 act in terms of a basis which we shall construct.

We shall adapt Dirac's familiar ladder technique[4] to the inhomogeneous Lorentz group in order to obtain the above results.

The following familiar identity will be useful. Let A and B be two operators. Then for any scalar r

$$e^{-rA}Be^{rA} = \sum_{n=0}^{\infty} \frac{\{B, A\}^{(n)}}{n!} r^n, \tag{2.1}$$

where $\{B, A\}^{(n)}$ is defined by induction by means of commutators

$$\begin{aligned} \{B, A\}^{(n)} &= [\{B, A\}^{(n-1)}, A], \\ \{B, A\}^{(0)} &= B. \end{aligned} \tag{2.1a}$$

Let D be any of the infinitesimal generators. Then from the commutation rules (1.4)–(1.10) of Part I and Eq. (2.1) above

$$\exp(-2\pi i J_3) D \exp(2\pi i J_3) = D, \tag{2.2}$$

for all D. Thus, $\exp(2\pi i J_3)$ is a scalar from Schur's lemma, since it commutes with each member of an irreducible set of operators. We write

$$\exp(2\pi i J_3) = \omega I, \tag{2.3}$$

where ω is a complex number of unit modulus and I is the identity.

Again by the use of the commutation relations and Eq. (2.1)

$$\exp(\pi i J_2) J_3 \exp(-\pi i J_2) = -J_3, \tag{2.4}$$

from which

$$\exp(\pi i J_2) \exp(2\pi i J_3) \exp(-\pi i J_2) = \exp(-2\pi i J_3) = \omega^* I. \tag{2.5}$$

Thus from (2.3) and (2.5)

$$\omega = \omega^*,$$

from which it follows that either $\omega = 1$ or $\omega = -1$. In the first case, all the eigenvalues m of J_3 are integers, while in the second the eigenvalues m are half-odd integers.

It also follows from (2.4) that if m is an eigenvalue of J_3, so is $-m$.

Let us introduce the operator N defined by

$$N = J_2 - iJ_1. \tag{2.6}$$

Then

$$\hat{N} = J_2 + iJ_1, \tag{2.6a}$$

where the carat denotes the Hermitian adjoint.

From the commutation relations and the definition it follows that

$$\begin{aligned} [J_3, N] &= N, \\ [J_3, \hat{N}] &= -\hat{N}, \\ \mathbf{J}^2 &= \hat{N}N + J_3^2 + J_3. \end{aligned} \tag{2.7}$$

Also the following commutation rules will be useful:

$$\begin{aligned} [H, N] &= 0, \\ [\mathbf{J}^2, N] &= 0, \\ [w, N] &= 0, \\ [J_3, N^m] &= mN^m. \end{aligned} \tag{2.7a}$$

Consider an eigenket $|E, r, m, \alpha\rangle$. From the positive definiteness of the operator $\mathbf{J}^2$, it follows that $m^2 \leq r$ where r is nonnegative. For a fixed value of r, let us denote the maximum value of m by j. Of course, j is a nonnegative integer or positive half-odd integer. The minimum value of m is $-j$.

From the commutation relations (2.7) and (2.7a) we see that for $m < j$, $N\,|E, r, m, \alpha\rangle$ is an eigenstate of H with the eigenvalue E, of $\mathbf{J}^2$ with the eigenvalue r, of w with the eigenvalue α, and of J_3 with the eigenvalue $m+1$.

However,

$$N\,|E, r, j, \alpha\rangle = 0, \tag{2.8}$$

for if it were not, then it would be an eigenstate of J_3 with the eigenvalue $j+1$ which is impossible, since j is the maximum eigenvalue of m for fixed r.

We can now find r in terms of j. For from the

[4] P. A. M. Dirac, *Principles of Quantum Mechanics* (Oxford University Press, Oxford, England, 1947), 3rd ed., p. 144 ff.

last of Eqs. (2.7) and from Eq. (2.8);

$$\mathbf{J}^2 |E,r,j,\alpha\rangle = r|E,r,j,\alpha\rangle = \hat{N}N|E,r,j,\alpha\rangle + J_3^2|E,r,j,\alpha\rangle + J_3|E,r,j,\alpha\rangle = j(j+1)|E,r,j,\alpha\rangle. \tag{2.9}$$

Hence,

$$r = j(j+1). \tag{2.10}$$

Henceforth, we shall denote the kets which form the basis by $|E, j, m, \alpha\rangle$ instead of $|E, r, m, \alpha\rangle$. In terms of the new notation then,

$$\begin{aligned} H|E,j,m,\alpha\rangle &= E|E,j,m,\alpha\rangle, \\ \mathbf{J}^2|E,j,m,\alpha\rangle &= j(j+1)|E,j,m,\alpha\rangle, \\ J_3|E,j,m,\alpha\rangle &= m|E,j,m,\alpha\rangle, \\ w|E,j,m,\alpha\rangle &= \alpha|E,j,m,\alpha\rangle. \end{aligned} \tag{2.11}$$

Having the basis formed by the kets $|E, j, m, \alpha\rangle$, we can construct new kets $|E, j, m, \alpha)$ which satisfy (2.11) by

$$|E,j,m,\alpha) = N^{m+j}|E,j,-j,\alpha\rangle. \tag{2.12}$$

Since, for a fixed set of values of E, j, α, we have for each possible value of m an eigenstate $|E, j, m, \alpha)$, this new set of kets spans the same space as the original one.

In terms of this new basis

$$N|E,j,m,\alpha) = |E,j,m+1,\alpha) \text{ for } m < j \tag{2.13}$$
$$N|E,j,j,\alpha) = 0.$$

To find how $\hat{N}$ acts in the new basis we use the last of equations (2.7) as follows: For $m < -j$

$$\begin{aligned} \hat{N}|E,j,m,\alpha) &= \hat{N}N|E,j,m-1,\alpha) \\ &= (\mathbf{J}^2 - J_3^2 - J_3)|E,j,m-1,\alpha) \\ &= (j+m)(j-m+1)|E,j,m-1,\alpha). \end{aligned} \tag{2.14}$$

Hence in the new basis we know how N and $\hat{N}$ act.

Since the new basis spans the Hilbert space, there must be a real positive weight function $W(E, j, m, \alpha)$ such that the resolution of the identity operator I is given by

$$I = \sum_{j,m,\alpha} \int dE\, |E,j,m,\alpha)W(E,j,m,\alpha)(E,j,m,\alpha|, \tag{2.15}$$

where the integration and summation go over the entire range of quantum numbers. The function W will be determined by the requirements that N and $\hat{N}$ as given by (2.13) and (2.14) are Hermitian adjoints.

From (2.13) and (2.15), we have

$$N = \sum_{j,m,\alpha} \int dE\, |E,j,m+1,\alpha)W(E,j,m,\alpha) \times (E,j,m,\alpha|. \tag{2.16}$$

Hence, from (2.16)

$$\begin{aligned} \hat{N} &= \sum_{j,m,\alpha} \int dE\, |E,j,m,\alpha)W(E,j,m,\alpha) \times (E,j,m+1,\alpha|, \\ &= \sum_{j,m,\alpha} \int dE\, |E,j,m-1,\alpha)W(E,j,m-1,\alpha) \times (E,j,m,\alpha|. \end{aligned} \tag{2.17}$$

But from (2.14) and (2.15)

$$\hat{N} = \sum_{j,m,\alpha} \int dE (j+m)(j-m+1)|E,j,m-1,\alpha) \times W(E,j,m,\alpha)(E,j,m,\alpha|. \tag{2.18}$$

Since the bras and kets are linearly independent, we have an equation for W, namely,

$$W(E,j,m-1,\alpha) = (j+m)(j-m+1)W(E,j,m,\alpha). \tag{2.19}$$

Let us define $W(E, j, \alpha)$ by

$$W(E,j,\alpha) = W(E,j,-j,\alpha).$$

Then the solution of (2.19) is, as can be proved by induction,

$$W(E,j,m,\alpha) = \frac{(j-m)!}{(j+m)!}\frac{1}{(2j)!}W(E,j,\alpha). \tag{2.20}$$

Henceforth, for simplicity of notation, we shall replace the round kets by the angular kets which are not to be confused with the original angular kets. To summarize: We have shown the existence of a set of kets such that

$$\begin{aligned} H|E,j,m,\alpha\rangle &= E|E,j,m,\alpha\rangle, \\ \mathbf{J}^2|E,j,m,\alpha\rangle &= j(j+1)|E,j,m,\alpha\rangle, \\ J_3|E,j,m,\alpha\rangle &= m|E,j,m,\alpha\rangle, \\ w|E,j,m,\alpha\rangle &= \alpha|E,j,m,\alpha\rangle, \\ (J_2 - iJ_1)|E,j,m,\alpha\rangle &= |E,j,m+1,\alpha\rangle, \\ (J_2 + iJ_1)|E,j,m,\alpha\rangle &= (j+m)(j-m+1)|E,j,m-1,\alpha\rangle, \\ I &= \sum_{j,m,\alpha}\int dE\, |E,j,m,\alpha\rangle \frac{(j-m)!}{(j+m)!}\frac{1}{(2j)!} \times W(E,j,\alpha)\langle E,j,m,\alpha|, \end{aligned} \tag{2.22}$$

for values of j and ranges of m discussed in the body of the section.

III. REPRESENTATION OF THE LINEAR MOMENTUM OPERATORS

In the present section we shall show how the operators P_i act and prove that for a fixed value of α in the set of kets $|E, j, m, \alpha\rangle$, the values of j which occur are $j = |\alpha|, |\alpha| + 1, |\alpha| + 2, \cdots$.

Our starting point is the observation that since P_3 commutes with H, J_3, and w, the vector $P_3 |E, j, m, \alpha\rangle$ is a linear combination of $|E, j, m, \alpha\rangle$ in terms of the quantum number j only. That is

$$P_3 |E, j, m, \alpha\rangle = \sum_{j'} A_{j,j'}(E, m, \alpha) |E, j', m, \alpha\rangle, \tag{3.1}$$

where $A_{j,j'}(E, m, \alpha)$ is a matrix element in the j variables only which depends on E, m, α as shown. Since the quantum number j is nonnegative and is denumerable, there will be a minimum value of j which we shall call j_0, i.e. $j \geq j_0 \geq 0$. The matrix element $A_{j,j'}$ is defined only for those values of j, j' which are equal to or greater than j_0. At times it will be convenient to regard $A_{j,j'}$ as being equal to zero if either j or j' are less than j_0.

The next point in our development is to use the commutation rules

$$\begin{aligned}[][\mathbf{J}^2, P_3] &= 2i(\mathbf{P} \times \mathbf{J})_3 + 2P_3, \\ [\mathbf{J}^2, (\mathbf{P} \times \mathbf{J})_3] &= -2iP_3\mathbf{J}^2 + 2iJ_3(H^2 - \mu^2)^{\frac{1}{2}}w,\end{aligned} \tag{3.2}$$

where

$$(\mathbf{P} \times \mathbf{J})_3 = P_1J_2 - P_2J_1. \tag{3.2a}$$

These commutation rules follow from the commutation rules for the infinitesimal generators given in Part I. [The first of Eqs. (3.2) appears misprinted on page 13 of Ref. 3.]

Let us define the vectors $|E, j, m, \alpha)_1$ and $|E, j, m, \alpha)_2$ by

$$\begin{aligned}|E, j, m, \alpha)_1 &= P_3 |E, j, m, \alpha\rangle, \\ |E, j, m, \alpha)_2 &= (\mathbf{P} \times \mathbf{J})_3 |E, j, m, \alpha\rangle.\end{aligned} \tag{3.3}$$

Then from (3.2) we obtain two simultaneous equations for $|E, j, m, \alpha)_{1,2}$. These equations are obtained by applying both sides of both Eqs. (3.2) to the vector $|E, j, m, \alpha\rangle$. The simultaneous equations are:

$$\begin{aligned}&[\mathbf{J}^2 - (j^2 + j + 2)] |E, j, m, \alpha)_1 \\ &\qquad - 2i |E, j, m, \alpha)_2 = 0, \\ &2ij(j + 1) |E, j, m, \alpha)_1 + [\mathbf{J}^2 - j(j + 1)] \\ &\qquad \times |E, j, m, \alpha)_2 = 2impa |E, j, m, \alpha\rangle.\end{aligned} \tag{3.4}$$

In Eq. (3.4), $p = (E^2 - \mu^2)^{\frac{1}{2}}$.

Multiply the top equation of (3.4) by $[\mathbf{J}^2 - j(j+1)]$ and the bottom by $2i$ and add.

One obtains

$$\begin{aligned}&\{[\mathbf{J}^2 - (j^2 + j + 2)][\mathbf{J}^2 - j(j + 1)] - 4j(j + 1)\} \\ &\qquad \times |E, j, m, \alpha)_1 = -4mp\alpha |E, j, m, \alpha\rangle.\end{aligned} \tag{3.5}$$

However,

$$\begin{aligned}&\{[\mathbf{J}^2 - (j^2 + j + 2)][\mathbf{J}^2 - j(j + 1)] - 4j(j + 1)\} \\ &\qquad = [\mathbf{J}^2 - j(j - 1)][\mathbf{J}^2 - (j + 1)(j + 2)].\end{aligned}$$

Hence, Eq. (3.5) is

$$\begin{aligned}&[\mathbf{J}^2 - j(j - 1)][\mathbf{J}^2 - (j + 1)(j + 2)] |E, j, m, \alpha)_1 \\ &\qquad = -4mp\alpha |E, j, m, \alpha\rangle,\end{aligned} \tag{3.6}$$

or

$$\begin{aligned}&[\mathbf{J}^2 - j(j - 1)][\mathbf{J}^2 - (j + 1)(j + 2)]P_3 |E, j, m, \alpha\rangle \\ &\qquad = -4mp\alpha |E, j, m, \alpha\rangle.\end{aligned} \tag{3.6a}$$

Hence, from the first of Eqs. (3.4),

$$\begin{aligned}&(\mathbf{P} \times \mathbf{J})_3 |E, j, m, \alpha\rangle \\ &\qquad = -\tfrac{1}{2}i [\mathbf{J}^2 - j(j + 1) - 2]P_3 |E, j, m, \alpha\rangle.\end{aligned} \tag{3.7}$$

On substituting (3.1) into (3.6a) we obtain an equation for $A_{j,j'}$:

$$\begin{aligned}&\sum_{j'} \{[j'(j' + 1) - j(j - 1)] \\ &\qquad \times [j'(j' + 1) - (j + 1)(j + 2)] \\ &\qquad \times A_{j,j'}(E, m, \alpha) |E, j'm, \alpha\rangle\} \\ &\qquad = -4mp\alpha |E, j, m, \alpha\rangle.\end{aligned} \tag{3.8}$$

Since the kets $|E, j, m, \alpha\rangle$ are linearly independent, we see that only $A_{j,j}$, $A_{j,j+1}$, $A_{j,j-1}$ do not vanish identically. In fact, for $j \neq 0$

$$A_{j,j}(E, m, \alpha) = mp\alpha/j(j + 1). \tag{3.9}$$

To find $A_{0,0}$ we note that for $j = 0$ in a ket, m must also be zero. Also $J_i |E, 0, 0, \alpha\rangle = 0$ for all i. Then

$$w_0 |E, 0, 0, \alpha\rangle = \mathbf{P} \cdot \mathbf{J} |E, 0, 0, \alpha\rangle = 0.$$

Hence,

$$w |E, 0, 0, \alpha\rangle = 0. \tag{3.10}$$

From (3.10) we see that the only value for α when $j = 0$ in a ket is $\alpha = 0$. Hence, the only kets which span the subspace for $j = 0$ are $|E, 0, 0, 0\rangle$.

Also $(\mathbf{P}\times\mathbf{J})_3\,|E,0,0,0\rangle = 0$. On writing

$$P_3\,|E,0,0,0\rangle = A_{0,0}(E,0,0)\,|E,0,0,0\rangle + A_{0,1}(E,0,0)\,|E,1,0,0\rangle, \qquad (3.11)$$

and on substituting into (3.7) we obtain

$$A_{0,0}(E,0,0) = 0. \qquad (3.12)$$

Equation (3.12) which gives $A_{0,0}$ may be regarded as a special case of (3.9) which gives $A_{j,j}$, if we always evaluate the numerator of the fraction in (3.9) first before evaluating the denominator. We shall adopt this convention and consider (3.9) as giving $A_{j,j}(E,m,\alpha)$ for all j.

The Hermitian character of P_3 provides a relation between $A_{j,j+1}(E,m,\alpha)$ and $A_{j,j-1}(E,m,\alpha)$. (We remember that the element $A_{j,j-1}$ is defined to be zero if $j = j_0$.) From the completeness relation which is the last of Eqs. (2.22), we obtain the orthonormality relations for the kets $|E,j,m,\alpha\rangle$, namely,

$$\langle E',j',m',\alpha' \mid E,j,m,\alpha\rangle = \delta_{j,j'}\delta_{m,m'}\delta_{\alpha,\alpha'}\frac{(j+m)!}{(j-m)!}\frac{(2j)!}{W(E,j,\alpha)}. \qquad (3.13)$$

The requirements that P_3 be Hermitian can be expressed as the requirement that the matrix element of P_3 satisfies

$$\langle E,j,m,\alpha|\,P_3\,|E',j',m',\alpha'\rangle = \langle E',j',m',\alpha'|\,P_3\,|E,j,m,\alpha\rangle^*, \qquad (3.14)$$

where the asterisk means complex conjugate.

On substituting (3.1) into (3.14) we obtain

$$A_{j+1,j}(E,m,\alpha) = A_{j,j+1}^*(E,m,\alpha)\frac{j+m+1}{j-m+1} \times 2(j+1)(2j+1)\frac{W(E,j,\alpha)}{W(E,j+1,\alpha)}. \qquad (3.15)$$

It will be convenient to define $B_{j+1}(E,m,\alpha)$ and $A_{j-1}(E,m,\alpha)$ by

$$\begin{aligned} B_{j+1}(E,m,\alpha) &= A_{j,j+1}(E,m,\alpha),\\ A_{j-1}(E,m,\alpha) &= A_{j,j-1}(E,m,\alpha). \end{aligned} \qquad (3.16)$$

Then we have

$$P_3\,|E,j,m,\alpha\rangle = \frac{mp\alpha}{j(j+1)}\,|E,j,m,\alpha\rangle + B_{j+1}(E,m,\alpha)\,|E,j+1,m,\alpha\rangle + A_{j-1}(E,m,\alpha)\,|E,j-1,m,\alpha\rangle, \qquad (3.17)$$

where A_j and B_j are related by (3.15):

$$A_j(E,m,\alpha) = B_{j+1}^*(E,m,\alpha)\frac{j+m+1}{j-m-1} \times 2(j+1)(2j+1)\frac{W(E,j,\alpha)}{W(E,j+1,\alpha)}. \qquad (3.18)$$

Our principal objectives will now be to find B_{j+1} and, hence, A_{j-1}. On combining (3.7) and (3.17) we have

$$(\mathbf{P}\times\mathbf{J})_3\,|E,j,m,\alpha\rangle = i\Big[\frac{mp\alpha}{j(j+1)}\,|E,j,m,\alpha\rangle - jB_{j+1}(E,m,\alpha)\,|E,j+1,m,\alpha\rangle + (j+1)A_{j-1}(E,m,\alpha)\,|E,j-1,m,\alpha\rangle\Big]. \qquad (3.19)$$

Now from the commutation rules of Part I

$$\begin{aligned} (P_1+iP_2) &= -i[(J_2-iJ_1),P_3],\\ (P_1-iP_2) &= -i[(J_2+iJ_1),P_3]. \end{aligned} \qquad (3.20)$$

On applying these relations to the kets $|E,j,m,\alpha\rangle$ on using the fifth and sixth equations (2.22), and on using (3.17), we obtain

$$\begin{aligned}(P_1+iP_2)\,|E,j,m,\alpha\rangle = -i\Big\{&-\frac{p\alpha}{j(j+1)}\,|E,j,m+1,\alpha\rangle\\ &+[B_{j+1}(E,m,\alpha)-B_{j+1}(E,m+1,\alpha)]\\ &\times|E,j+1,m+1,\alpha\rangle\\ &+[A_{j-1}(E,m,\alpha)-A_{j-1}(E,m+1,\alpha)]\\ &\times|E,j-1,m+1,\alpha\rangle\Big\},\end{aligned} \qquad (3.21)$$

$$\begin{aligned}(P_1-iP_2)\,|E,j,m,\alpha\rangle = -i\Big\{&\frac{(j+m)(j-m+1)p\alpha}{j(j+1)}\,|E,j,m-1,\alpha\rangle\\ &+[(j+m+1)(j-m+2)B_{j+1}(E,m,\alpha)\\ &-(j+m)(j-m+1)B_{j+1}(E,m-1,\alpha)]\\ &\times|E,j+1,m-1,\alpha\rangle\\ &+[(j+m-1)(j-m)A_{j-1}(E,m,\alpha)\\ &-(j+m)(j-m+1)A_{j-1}(E,m-1,\alpha)]\\ &\times|E,j-1,m-1,\alpha\rangle\Big\}.\end{aligned} \qquad (3.22)$$

In (3.21) and (3.22) there are terms with $j(j+1)$ in the denominator. These terms vanish when $j = 0$. Since when $j = 0$ also $\alpha = 0$ as shown earlier, these terms would also vanish if the numerator were evaluated first. Hence, our convention for $j = 0$ is still valid.

Now

$$(\mathbf{P} \times \mathbf{J})_3 = \tfrac{1}{2}[(P_1 - iP_2)(J_2 - iJ_1) - (P_1 + iP_2)(J_2 + iJ_1)]. \qquad (3.23)$$

Thus, (3.21) and (3.22) together with (2.22) yield the following results:

$$\begin{aligned}(\mathbf{P} \times \mathbf{J})_3 \, |E, j, m, \alpha\rangle = -\frac{i}{2}\Big\{ &-\frac{2mp\alpha}{j(j+1)} |E, j, m, \alpha\rangle \\ &+ [(j + m + 2)(j - m + 1)B_{j+1}(E, m + 1, \alpha) \\ &+ (j + m)(j - m + 1)B_{j+1}(E, m - 1, \alpha) \\ &- 2(j^2 + j - m^2)B_{j+1}(E, m, \alpha)] \, |E, j + 1, m, \alpha\rangle \\ &+ [(j + m)(j - m - 1)A_{j-1}(E, m + 1, \alpha) \\ &+ (j + m)(j - m + 1)A_{j-1}(E, m - 1, \alpha) \\ &- 2(j^2 + j - m^2)A_{j-1}(E, m, \alpha)] \, |E, j - 1, m, \alpha\rangle\Big\}.\end{aligned} \qquad (3.24)$$

Let us compare (3.24) with (3.19). Since the kets are linearly independent, we can equate coefficients of $|E, j + 1, m, \alpha\rangle$. We are thus lead to a recursion relation for $B_{j+1}(E, m, \alpha)$ in the variable m, namely,

$$\begin{aligned}(j - m + 1)(j + m + 2)B_{j+1}(E, m + 1, \alpha) &= 2(j^2 + 2j - m^2)B_{j+1}(E, m, \alpha) \\ &\quad - (j + m)(j - m + 1)B_{j+1}(E, m - 1, \alpha).\end{aligned} \qquad (3.25)$$

Let $R_{j+1}(E, \alpha)$ be defined by

$$R_{j+1}(E, \alpha) = B_{j+1}(E, -j, \alpha). \qquad (3.26)$$

Then one can show that the general solution of (3.22) is

$$B_{j+1}(E, m, \alpha) = \frac{j - m + 1}{2j + 1} R_{j+1}(E, \alpha). \qquad (3.27)$$

Equation coefficients of $|E, j - 1, m, \alpha\rangle$ leads to a recursion formula for $A_{j-1}(E, m, \alpha)$ in the variable m.

$$\begin{aligned}(j + m)(j - m - 1)A_{j-1}(E, m + 1, \alpha) &= 2(j^2 - m^2 - 1)A_{j-1}(E, m, \alpha) \\ &\quad - (j + m)(j - m + 1)A_{j-1}(E, m - 1, \alpha).\end{aligned} \qquad (3.28)$$

The general solution of (3.28) is

$$A_{j-1}(E, m, \alpha) = (j + m)S_{j-1}(E, \alpha). \qquad (3.29)$$

The relation (3.18) becomes

$$S_j^*(E, \alpha) = 2(j + 1) \frac{W(E, j, \alpha)}{W(E, j + 1, \alpha)} R_{j+1}(E, \alpha). \qquad (3.30)$$

Also (3.17), (3.21), and (3.22) becomes

$$\begin{aligned}P_3 \, |E, j, m, \alpha\rangle = &\frac{mp\alpha}{j(j + 1)} |E, j, m, \alpha\rangle \\ &+ \frac{j - m + 1}{2j + 1} R_{j+1}(E, \alpha) \, |E, j + 1, m, \alpha\rangle \\ &+ (j + m)S_{j-1}(E, \alpha) \, |E, j - 1, m, \alpha\rangle,\end{aligned} \qquad (3.31)$$

$$\begin{aligned}(P_1 + iP_2) \, |E, j, m, \alpha\rangle = -i\Big\{ &-\frac{p\alpha}{j(j + 1)} |E, j, m + 1, \alpha\rangle \\ &+ \frac{1}{2j + 1} R_{j+1}(E, \alpha) \, |E, j + 1, m + 1, \alpha\rangle \\ &- S_{j-1}(E, \alpha) \, |E, j - 1, m + 1, \alpha\rangle\Big\},\end{aligned} \qquad (3.32)$$

$$\begin{aligned}(P_1 - iP_2) \, |E, j, m, \alpha\rangle = -i\Big\{ &\frac{(j + m)(j - m + 1)p\alpha}{j(j + 1)} |E, j, m - 1, \alpha\rangle \\ &+ \frac{(j-m+2)(j-m+1)}{2j+1} R_{j+1}(E, \alpha) \, |E, j+1, m-1, \alpha\rangle \\ &- (j+m)(j+m-1)S_{j-1}(E, \alpha) \, |E, j-1, m-1, \alpha\rangle\Big\}.\end{aligned} \qquad (3.33)$$

We shall now use the fact that

$$H^2 - \mathbf{P}^2 = \mu^2. \qquad (3.34)$$

We write

$$\mathbf{P}^2 = (P_1 + iP_2)(P_1 - iP_2) + P_3^2, \qquad (3.35)$$

and apply both sides of Eq. (3.34) to the ket $|E, j, -j, \alpha\rangle$.

After using Eqs. (3.31)–(3.33) we obtain an equation of the form

$$C(E, j, \alpha) \, |E, j, -j, \alpha\rangle = 0, \qquad (3.36)$$

where

$$C(E, j, \alpha) = \frac{p^2\alpha^2}{(j + 1)^2} + (2j + 3)R_{j+1}(E, \alpha)S_j(E, \alpha) - p^2. \qquad (3.37)$$

Hence,

$$C(E, j, \alpha) = 0. \qquad (3.38)$$

Let $j = j_0 - 1$ in (3.37) and (3.38) where j_0 is the minimum value of j. Then, since $S_{j_0-1} = 0$, we obtain the extremely important result

$$j_0 = |\alpha|. \qquad (3.39)$$

Thus for $j \geq |\alpha|$

$$R_{j+1}(E, \alpha)S_j(E, \alpha) = \frac{p^2}{(2j+3)(j+1)^2}[(j+1)^2 - \alpha^2]. \quad (3.40)$$

From (3.30) we have

$$|R_{j+1}(E, \alpha)|^2 = \frac{p^2}{2(2j+3)(j+1)^3} \times [(j+1)^2 - \alpha^2]\frac{W(E, j+1, \alpha)}{W(E, j, \alpha)}. \quad (3.41)$$

At this point it would appear that we would have to use more of the commutation rules to obtain the phase of R_{j+1}. However, we shall show that the unknown phase can be obtained by changing the basis.

Let us define $T_j(\alpha)$ by

$$T_j(\alpha) = \left[\frac{j^2 - \alpha^2}{2j^3(2j+1)}\right]^{\frac{1}{2}}. \quad (3.42)$$

Then

$$R_j(\alpha) = pT_j(\alpha)\left[\frac{W(E, j, \alpha)}{W(E, j-1, \alpha)}\right]^{\frac{1}{2}} e^{i\omega(j)}, \quad (3.43)$$

and

$$S_j(E, \alpha) = 2p(j+1)T_{j+1}(\alpha) \times \left[\frac{W(E, j, \alpha)}{W(E, j+1, \alpha)}\right]^{\frac{1}{2}} e^{-i\omega(j+1)}, \quad (3.44)$$

where $\omega(j)$ is the phase angle in the polar description of $R_j(E, \alpha)$.

Let us introduce a new basis:

$$|E, j, m, \alpha\rangle = \left[\frac{W(E, j, \alpha)}{W(E, \alpha)}\right]^{\frac{1}{2}} e^{i\beta(j)} |E, j, m, \alpha). \quad (3.45)$$

In (3.45), $W(E, \alpha) = W(E, j_0, \alpha)$ and $\beta(j) = \sum_{j_0+1}^{j} \omega(j')$ for $j > j_0$, $\beta(j_0) = 0$. All of the Eqs. (2.2) except the last (the completeness relation) hold in terms of the round kets. Equations (3.31)–(3.33) hold with R_j replaced by $pT_j(\alpha)$ and S_j replaced by $2p(j+1)T_{j+1}(\alpha)$. The completeness relation which is the last of Eqs. (2.2) takes on a simpler form, which will be given below.

We shall replace the new round kets by angular kets and summarize our results thus far. We have shown that a basis exists such that

$$H\,|E, j, m, \alpha\rangle = E\,|E, j, m, \alpha\rangle,$$
$$\mathbf{J}^2\,|E, j, m, \alpha\rangle = j(j+1)\,|E, j, m, \alpha\rangle,$$
$$J_3\,|E, j, m, \alpha\rangle = m\,|E, j, m, \alpha\rangle,$$
$$w\,|E, j, m, \alpha\rangle = \alpha\,|E, j, m, \alpha\rangle,$$
$$(J_2 - iJ_1)\,|E, j, m, \alpha\rangle = |E, j, m+1, \alpha\rangle,$$
$$(J_2 + iJ_1)\,|E, j, m, \alpha\rangle = (j+m)(j-m+1)\,|E, j, m-1, \alpha\rangle,$$
$$P_3\,|E, j, m, \alpha\rangle = p\left\{\frac{m\alpha}{j(j+1)}\,|E, j, m, \alpha\rangle + \frac{j-m+1}{2j+1}T_{j+1}(\alpha)\,|E, j+1, m, \alpha\rangle + 2j(j+m)T_j(\alpha)\,|E, j-1, m, \alpha\rangle\right\},$$
$$(P_1 + iP_2)\,|E, j, m, \alpha\rangle = -ip\left\{-\frac{\alpha}{j(j+1)}\,|E, j, m+1, \alpha\rangle + \frac{1}{2j+1}T_{j+1}(\alpha)\,|E, j+1, m+1, \alpha\rangle - 2jT_j(\alpha)\,|E, j-1, m+1, \alpha\rangle\right\},$$
$$(P_1 - iP_2)\,|E, j, m, \alpha\rangle = -ip\left\{\frac{(j+m)(j-m+1)\alpha}{j(j+1)}\,|E, j, m-1, \alpha\rangle + \frac{(j-m+2)(j-m+1)}{2j+1}T_{j+1}(\alpha)\,|E, j+1, m-1, \alpha\rangle - 2j(j+m)(j+m-1)T_j(\alpha)\,|E, j-1, m-1, \alpha\rangle\right\}, \quad (3.46)$$

where $T_j(\alpha)$ is given by (3.42). Hence, in our basis we know how all the operators, except $\mathcal{J}_i$, operate explicitly.

The completeness relation in the new basis is

$$I = \sum_{j,m,\alpha}\int dE\,|E, j, m, \alpha\rangle \frac{(j-m)!}{(j+m)!}\frac{1}{(2j)}W(E, \alpha) \times \langle E, j, m, \alpha|. \quad (3.47)$$

IV. THE DERIVATION OF THE FORMS OF THE OPERATORS $\mathcal{J}_i$

Part I. m Dependence

The remaining part of the paper contains the derivation of the form of the operators $\mathcal{J}_i$. The derivation of these operators is extremely lengthy, especially for the case of nonzero mass. In this section we show how these operators affect the m quantum number and the E quantum number in terms of the angular-momentum basis.

First of all we shall prove that

$$(E - H)\frac{\partial}{\partial E}|E, j, m, \alpha\rangle = -|E, j, m, \alpha\rangle. \tag{4.1}$$

For

$$(E - H)\frac{\partial}{\partial E}|E, j, m, \alpha\rangle$$
$$= \lim_{\Delta\to 0}(E-H)(|E+\Delta, j, m, \alpha\rangle - |E, j, m, \alpha\rangle)(\Delta)^{-1},$$
$$= \lim_{\Delta\to 0}\frac{-\Delta}{\Delta}|E+\Delta, j, m, \alpha\rangle = -|E, j, m, \alpha\rangle,$$

as required.

Next, it will be useful to show how the operator $\mathbf{J}\cdot\boldsymbol{\mathfrak{g}} = \sum_i J_i\mathfrak{g}_i$ acts on the ket $|E, j, m, \alpha\rangle$. We have the following commutation rules:

$$[\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}, H] = iw_0, \tag{4.2}$$

$$[\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}, \mathbf{J}^2] = 0, \tag{4.3}$$

$$[\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}, J_i] = 0. \tag{4.4}$$

On applying both sides of (4.2) to the ket $|E, j, m, \alpha\rangle$, we have

$$(E - H)\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}|E, j, m, \alpha\rangle = ip\alpha|E, j, m, \alpha\rangle. \tag{4.5}$$

As an ansatz, we write

$$\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}|E, j, m, \alpha\rangle$$
$$= |E, j, m, \alpha\rangle' - ip\alpha\frac{\partial}{\partial E}|E, j, m, \alpha\rangle, \tag{4.6}$$

where $|E, j, m, \alpha\rangle'$ is to be determined and substitute into (4.5). On using (4.1) we find

$$H|E, j, m, \alpha\rangle' = E|E, j, m, \alpha\rangle'. \tag{4.7}$$

Equations (4.3), (4.4), and (4.6) lead to

$$\mathbf{J}^2|E, j, m, \alpha\rangle' = j(j+1)|E, j, m, \alpha\rangle', \tag{4.8}$$

$$J_3|E, j, m, \alpha\rangle' = m|E, j, m, \alpha\rangle'. \tag{4.9}$$

Hence $|E, j, m, \alpha\rangle'$ is a simultaneous eigenstate of H, $\mathbf{J}^2$, and J_3 with eigenvalues E, j, and m, respectively. Hence, we may write

$$|E, j, m, \alpha\rangle' = \sum_{\alpha'} K_{\alpha,\alpha'}(E, j, m)|E, j, m, \alpha'\rangle, \tag{4.10}$$

where the arguments of the matrix element $K_{\alpha,\alpha'}$ indicate the possible dependence on these eigenvalues.

The commutation relation (4.4) leads to

$$[\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}, J_2 - iJ_1] = 0. \tag{4.4a}$$

On applying both sides of (4.4a) to the ket $|K, j, m, \alpha\rangle$ and on, subsequently, using (4.6), (4.10), and (3.46) we obtain

$$\sum_{\alpha'}[K_{\alpha,\alpha'}(E, j, m+1) - K_{\alpha,\alpha'}(E, j, m)]$$
$$\times|E, j, m+1, \alpha'\rangle = 0. \tag{4.11}$$

Hence, $K_{\alpha,\alpha'}(E, j, m+1) = K_{\alpha,\alpha'}(E, j, m)$. From this statement it follows that $K_{\alpha,\alpha'}(E, j, m)$ is independent of m. Thus we may define a matrix element $K_{\alpha,\alpha'}(E, j)$ by

$$K_{\alpha,\alpha'}(E, j) \equiv K_{\alpha,\alpha'}(E, j, m), \tag{4.12}$$

and thus eliminate dependence upon the quantum variable m in the notation.

The requirement that $\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}$ be Hermitian will give a condition on the weight function $W(E, \alpha)$. We require

$$\langle E, j, m, \alpha|\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}|E', j', m', \alpha'\rangle$$
$$= \langle E', j', m', \alpha'|\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}|E, j, m, \alpha\rangle^*. \tag{4.13}$$

From (3.47)

$$\langle E, j, m, \alpha \mid E', j', m', \alpha'\rangle$$
$$= \frac{(j+m)!}{(j-m)!}\frac{(2j)!}{W(E, \alpha)}\delta(E - E')\delta_{j,j'}\delta_{m,m'}\delta_{\alpha,\alpha'}, \tag{4.14}$$

$$\langle E, j, m, \alpha|\frac{\partial}{\partial E'}|E', j', m', \alpha'\rangle$$
$$= -\frac{(j+m)!}{(j-m)!}\frac{(2j)!}{W(E, \alpha)}\delta'(E - E')\delta_{j,j'}\delta_{m,m'}\delta_{\alpha,\alpha'}, \tag{4.15}$$

where $\delta'(E)$ is the derivative of the δ function with respect to its argument.

Then on using (4.6), (4.10), (4.12), (4.14), and (4.15) we have

$$\langle E, j, m, \alpha|\mathbf{J}\cdot\boldsymbol{\mathfrak{g}}|E', j', m', \alpha'\rangle$$
$$= \frac{(j+m)!}{(j-m)!}\frac{(2j)!}{W(E, \alpha)}K_{\alpha',\alpha}(E, j)\delta(E - E')\delta_{j,j'}\delta_{m,m'}$$
$$+ ip'\alpha\frac{(j+m)!}{(j-m)!}\frac{(2j)!}{W(E, \alpha)}\delta'(E - E')\delta_{j,j'}\delta_{m,m'}\delta_{\alpha,\alpha'}, \tag{4.16}$$

where $p' = [(E')^2 - \mu^2]^{\frac{1}{2}}$.

We also note the familiar identity

$$\delta'(E - E')f(E') = f'(E)\delta(E - E') + f(E)\delta'(E - E'). \tag{4.17}$$

Then (4.13), (4.16), and (4.17) lead to

$$\frac{K_{\alpha',\alpha}(E,j)}{W(E,\alpha)}+i\frac{E\alpha}{pW(E,\alpha)}\delta_{\alpha,\alpha'}=\frac{K^*_{\alpha,\alpha'}(E,j)}{W(E,\alpha')}+ip\alpha\frac{d}{dE}\left[\frac{1}{W(E,\alpha)}\right]\delta_{\alpha,\alpha'}. \tag{4.18}$$

Equation (4.18) suggests that we choose

$$W(E,\alpha)=1/p. \tag{4.19}$$

Equation (4.18) then leads to the particularly simple relation

$$K_{\alpha,\alpha'}(E,j)=K^*_{\alpha',\alpha}. \tag{4.20}$$

That is, $K_{\alpha,\alpha'}$ is a Hermitian matrix in the quantum variables α.

Thus, to summarize,

$$\mathbf{J}\cdot\boldsymbol{\mathcal{J}}\,|E,j,m,\alpha\rangle=\sum_{\alpha'}K_{\alpha,\alpha'}(E,j)\,|E,j,m,\alpha'\rangle-ip\alpha\frac{\partial}{\partial E}|E,j,m,\alpha\rangle, \tag{4.21}$$

where $K_{\alpha,\alpha'}$ satisfies (4.20).

For a while, to simplify notation, it will be useful to suppress the appearance of the quantum number α. In terms of the modified notation Eq. (4.21) takes the form

$$\mathbf{J}\cdot\boldsymbol{\mathcal{J}}\,|E,j,m\rangle=K(E,j)\,|E,j,m\rangle-ip\alpha(\partial/\partial E)\,|E,j,m\rangle, \tag{4.21a}$$

where $K(E,j)$ is a Hermitian operator in α. To express the Hermitian character of K we write (4.21) as

$$K^\dagger(E,j)=K(E,j), \tag{4.20a}$$

where the dagger means Hermitian adjoint in the quantum variable α.

Consider the commutation rule

$$[\mathcal{J}_3,H]=iP_3, \tag{4.22}$$

and apply both sides to the ket $|E,j,m\rangle$. From (3.46) we obtain

$$(E-H)\mathcal{J}_3\,|E,j,m\rangle=ip\left[\frac{m\alpha}{j(j+1)}|E,j,m\rangle+\frac{(j-m+1)}{2j+1}T_{j+1}(\alpha)\,|E,j+1,m\rangle+2j(j+m)T_j(\alpha)\,|E,j-1,m\rangle\right]. \tag{4.23}$$

In (4.23) we have not suppressed the variable α in $T_j(\alpha)$ in order to emphasize that T_j is simply a function of α and not an operator in this variable. Alternatively one could introduce an operator T_j which is diagonal in the representation and has the eigenvalues $T_j(\alpha)$. It seems simpler, however, to use the first convention.

Let D be an operator in α. Then $DT_j(\alpha)\,|E,j,m\rangle$ has the meaning $\sum_{\alpha'}D_{\alpha,\alpha'}T_j(\alpha')\,|E,j,m,\alpha'\rangle$. By contrast $T_j(\alpha)\,D\,|E,j,m\rangle$ means

$$T_j(\alpha)\sum_{\alpha'}D_{\alpha,\alpha'}\,|E,j,m,\alpha'\rangle.$$

Hence, the order of T_j and D is important.

Let us make the ansatz

$$\mathcal{J}_3\,|E,j,m\rangle=|E,j,m\rangle-ip\left[\frac{m\alpha}{j(j+1)}\frac{\partial}{\partial E}|E,j,m\rangle+\frac{(j-m+1)}{2j+1}T_{j+1}(\alpha)\frac{\partial}{\partial E}|E,j+1,m\rangle+2j(j+m)T_j(\alpha)\frac{\partial}{\partial E}|E,j-1,m\rangle\right]. \tag{4.24}$$

The ket $|E,j,m\rangle$ also depends on the variable α which we have suppressed. $|E,j,m\rangle$ is to be determined. On substituting (4.24) into (4.23) we obtain

$$H\,|E,j,m\rangle=E\,|E,j,m\rangle. \tag{4.25}$$

Also from $[J_3,\mathcal{J}_3]=0$ we see that

$$J_3\,|E,j,m\rangle=m\,|E,j,m\rangle. \tag{4.26}$$

Hence, we see that $|E,j,m\rangle$ can be written as

$$|E,j,m\rangle=\sum_{j'}D_{j,j'}(E,m)\,|E,j',m\rangle, \tag{4.27}$$

where the quantities $D_{j,j'}$ are operators in the variable α. These operators will, in general, be functions also of E and m as indicated by the notation.

It will now be our objective to show that $D_{j,j'}=0$ unless $j'=j,\ j+1$, or $j-1$. We shall also obtain the m-dependence of $D_{j,j'}$.

From the commutation relations we have

$$[\mathbf{J}^2,\mathcal{J}_3]=2i(\boldsymbol{\mathcal{J}}\times\mathbf{J})_3+2\mathcal{J}_3, \tag{4.28}$$

$$[\mathbf{J}^2,(\boldsymbol{\mathcal{J}}\times\mathbf{J})_3]=-2i\mathcal{J}_3\mathbf{J}^2+2i(\mathbf{J}\cdot\boldsymbol{\mathcal{J}})J_3, \tag{4.29}$$

where

$$(\boldsymbol{\mathcal{J}}\times\mathbf{J})_3=\mathcal{J}_1J_2-\mathcal{J}_2J_1. \tag{4.30}$$

Let us define

$$\begin{aligned}|E,j,m\rangle_1&=\mathcal{J}_3\,|E,j,m\rangle,\\|E,j,m\rangle_2&=(\boldsymbol{\mathcal{J}}\times\mathbf{J})_3\,|E,j,m\rangle.\end{aligned} \tag{4.31}$$

Then on applying the operator relations (4.28) and

(4.29) to the ket $|E, j, m\rangle$ we obtain, on using (3.46),

$$[\mathbf{J}^2 - (j^2 + j + 2)]\,|E, j, m\rangle_1 - 2i\,|E, j, m\rangle_2 = 0,$$
$$2ij(j + 1)\,|E, j, m\rangle_1 + [\mathbf{J}^2 - j(j + 1)]\,|E, j, m\rangle_2 = 2im(\mathbf{J}\cdot\boldsymbol{\mathcal{J}})\,|E, j, m\rangle. \tag{4.32}$$

We can regard (4.32) as a pair of simultaneous equations for $|E, j, m\rangle_1$ and $|E, j, m\rangle_2$. On eliminating $|E, j, m\rangle_2$ in the usual way and on using (4.24), (4.27), and (4.21a) we obtain the following equation for $D_{j,j'}$:

$$\sum_{j'} D_{j,j'}(E, m)[j'(j' + 1) - j(j - 1)] \times [j'(j' + 1) - (j + 1)(j + 2)]\,|E, j', m\rangle = -4mK(E, j)\,|E, j, m\rangle. \tag{4.33}$$

It is seen that

$$D_{j,j'}(E, m) = 0 \quad \text{if} \quad j' \neq j, j + 1, j - 1, \tag{4.34}$$

$$D_{j,j}(E, m) = \frac{m}{j(j + 1)}\,K(E, j), \qquad (j \neq 0). \tag{4.35}$$

Without going into details, it can also be shown

$$D_{0,0}(E, 0) = 0. \tag{4.35a}$$

Hence, (4.35) includes (4.35a) if one remembers to evaluate the numerator of the fraction first.

Let

$$C_{j+1}(E, m) = D_{j,j+1}(E, m), \qquad D_{j-1}(E, m) = D_{j,j-1}(E, m). \tag{4.36}$$

Then

$$\mathcal{J}_3\,|E, j, m\rangle = \frac{m}{j(j + 1)}\,K(E, j)\,|E, j, m\rangle + C_{j+1}(E, m)\,|E, j + 1, m\rangle + D_{j-1}(E, m)\,|E, j - 1, m\rangle - ip\left[\frac{m\alpha}{j(j + 1)}\frac{\partial}{\partial E}\,|E, j, m\rangle + \frac{j - m + 1}{2j + 1}\,T_{j+1}(\alpha)\frac{\partial}{\partial E}\,|E, j + 1, m\rangle + 2j(j + m)T_j(\alpha)\frac{\partial}{\partial E}\,|E, j - 1, m\rangle\right]. \tag{4.37}$$

On substituting (4.37) into the first of equations (4.32) we also find

$$(\boldsymbol{\mathcal{J}} \times \mathbf{J})_3\,|E, j, m\rangle = i\left[\frac{m}{j(j + 1)}\,K(E, j)\,|E, j, m\rangle - jC_{j+1}(E, m)\,|E, j + 1, m\rangle + (j + 1)D_{j-1}(E, m)\,|E, j - 1, m\rangle\right] + p\left[\frac{m\alpha}{j(j + 1)}\frac{\partial}{\partial E}\,|E, j, m\rangle - j\frac{j - m + 1}{2j + 1}\,T_{j+1}(\alpha)\frac{\partial}{\partial E}\,|E, j + 1, m\rangle + 2j(j + 1)(j + m)T_j(\alpha)\frac{\partial}{\partial E}\,|E, j - 1, m\rangle\right]. \tag{4.38}$$

Now from the commutation relations,

$$(\mathcal{J}_1 + i\mathcal{J}_2) = -i[J_2 - iJ_1, \mathcal{J}_3], \qquad (\mathcal{J}_1 - i\mathcal{J}_2) = -i[J_2 + iJ_1, \mathcal{J}_3]. \tag{4.39}$$

Then on applying both sides of both equations of (4.39) to the ket $|E, j, m\rangle$ and on using (4.37) and (3.46), we have

$$(\mathcal{J}_1 + i\mathcal{J}_2)\,|E, j, m\rangle = i\left\{\frac{K(E, j)}{j(j + 1)}\,|E, j, m + 1\rangle + [C_{j+1}(E, m + 1) - C_{j+1}(E, m)]\,|E, j + 1, m + 1\rangle + [D_{j-1}(E, m+1) - D_{j-1}(E, m)]\,|E, j-1, m+1\rangle\right\} + p\left\{\frac{\alpha}{j(j + 1)}\frac{\partial}{\partial E}\,|E, j, m + 1\rangle - \frac{1}{2j + 1}\,T_{j+1}(\alpha)\frac{\partial}{\partial E}\,|E, j + 1, m\rangle + 2jT_j(\alpha)\frac{\partial}{\partial E}\,|E, j - 1, m + 1\rangle\right\}. \tag{4.40}$$

$$(\mathcal{J}_1 - i\mathcal{J}_2)\,|E, j, m\rangle = -i\left\{\frac{(j + m)(j - m + 1)}{j(j + 1)}\,K(E, j)\,|E, j, m - 1\rangle + [(j + m + 1)(j - m + 2)C_{j+1}(E, m) - (j + m) \times (j - m + 1)C_{j+1}(E, m - 1)]\,|E, j + 1, m - 1\rangle + [(j + m - 1)(j - m)D_{j-1}(E, m) - (j + m) \times (j - m + 1)D_{j-1}(E, m - 1)]\,|E, j - 1, m - 1\rangle\right\} - p\left\{\frac{(j + m)(j - m + 1)}{j(j + 1)}\,\alpha\frac{\partial}{\partial E}\,|E, j, m - 1\rangle + \frac{(j-m+2)(j-m+1)}{2j+1}\,T_{j+1}(\alpha)\frac{\partial}{\partial E}\,|E, j+1, m-1\rangle - 2j(j + m)(j + m-1)T_j(\alpha)\frac{\partial}{\partial E}\,|E, j-1, m-1\rangle\right\}. \tag{4.41}$$

Now

$$(\mathcal{J}\times\mathbf{J})_3=\tfrac{1}{2}[(\mathcal{J}_1-i\mathcal{J}_2)(J_2-iJ_1)+(\mathcal{J}_1+i\mathcal{J}_2)(J_2+iJ_1)]. \tag{4.42}$$

Then from (4.42), (4.40), (4.41), and (3.46)

$$\begin{aligned}(\mathcal{J}\times\mathbf{J})_3\,|E,j,m\rangle &= i\Big\{\frac{m}{j(j+1)}K(E,j)\,|E,j,m\rangle\\
&-\tfrac{1}{2}[(j+m+2)(j-m+1)C_{j+1}(E,m+1)\\
&+(j+m)(j-m+1)C_{j+1}(E,m-1)\\
&-2(j^2-m^2+j)C_{j+1}(E,m)]\,|E,j+1,m\rangle\\
&-\tfrac{1}{2}[(j+m)(j-m-1)D_{j-1}(E,m+1)\\
&+(j+m)(j-m+1)D_{j-1}(E,m-1)\\
&-2(j^2-m^2+j)D_{j-1}(E,m)]\,|E,j-1,m\rangle\Big\}\\
&+p\Big\{\frac{m\alpha}{j(j+1)}\frac{\partial}{\partial E}\,|E,j,m\rangle\\
&-j\frac{j-m+1}{2j+1}T_{j+1}(\alpha)\frac{\partial}{\partial E}\,|E,j+1,m\rangle\\
&+2j(j+1)T_j(\alpha)\frac{\partial}{\partial E}\,|E,j-1,m\rangle\Big\}.\end{aligned} \tag{4.43}$$

On comparing the expression (4.43) with (4.38) we obtain equations for C_j and D_j which are identical to those for B_j and A_j, respectively [Eqs. (3.25) and (3.28)]. The solutions of these equations have the form

$$\begin{aligned}C_{j+1}(E,m)&=\frac{(j-m+1)}{2j+1}Q_{j+1}(E),\\
D_{j-1}(E,m)&=(j+m)V_{j-1}(E).\end{aligned} \tag{4.44}$$

The quantities Q_j and V_j are matrices in the variables α and may be functions of the quantum number E.

The requirement that $\mathcal{J}_3$ be Hermitian leads to the relation

$$V_{j-1}(E)=2jQ_j^\dagger(E). \tag{4.45}$$

It will be useful now to summarize:

$$\begin{aligned}\mathcal{J}_3\,|E,j,m\rangle &= \frac{m}{j(j+1)}\left[K(E,j)-ip\alpha\frac{\partial}{\partial E}\right]|E,j,m\rangle\\
&+\frac{j-m+1}{2j+1}\left[Q_{j+1}(E)-ipT_{j+1}(\alpha)\frac{\partial}{\partial E}\right]|E,j+1,m\rangle\\
&+2j(j+m)\left[Q_j^\dagger(E)-ipT_j(\alpha)\frac{\partial}{\partial E}\right]|E,j-1,m\rangle;\end{aligned} \tag{4.46}$$

$$\begin{aligned}(\mathcal{J}_1+i\mathcal{J}_2)\,|E,j,m\rangle &= \frac{i}{j(j+1)}\left[K(E,j)-ip\alpha\frac{\partial}{\partial E}\right]|E,j,m+1\rangle\\
&-\frac{i}{2j+1}\left[Q_{j+1}(E)-ipT_{j+1}(\alpha)\frac{\partial}{\partial E}\right]|E,j+1,m+1\rangle\\
&+2ij\left[Q_j^\dagger(E)-ipT_j(\alpha)\frac{\partial}{\partial E}\right]|E,j-1,m+1\rangle;\end{aligned} \tag{4.47}$$

$$\begin{aligned}(\mathcal{J}_1-i\mathcal{J}_2)\,|E,j,m\rangle &= \frac{-i(j+m)(j-m+1)}{j(j+1)}\\
&\times\left[K(E,j)-ip\alpha\frac{\partial}{\partial E}\right]|E,j,m-1\rangle\\
&-\frac{i(j-m+2)(j-m+1)}{2j+1}\\
&\times\left[Q_{j+1}(E)-ipT_{j+1}(\alpha)\frac{\partial}{\partial E}\right]|E,j+1,m-1\rangle\\
&+2ij(j+m)(j+m-1)\\
&\times\left[Q_j^\dagger(E)-ipT_j(\alpha)\frac{\partial}{\partial E}\right]|E,j-1,m-1\rangle.\end{aligned} \tag{4.48}$$

We are now left to determine the matrix $Q_j(E)$. This determination is a formidable task in the case of nonzero mass and is connected closely with the problem of obtaining a representation of the spin matrices in the angular-momentum basis. In the case of zero mass the solution is comparatively simple.

V. THE SPIN OPERATORS AND THE w_i OPERATORS

The method which we shall use to obtain the operator Q_j and the operator K is to use known facts about the irreducible representations of the inhomogeneous Lorentz group which can be obtained from the knowledge of how these representations work in a basis in which the linear momenta are diagonal.

For this purpose it is useful to consider the forms of the representation as given in Ref. 2, for example. We introduce the operators $\mathbf{w}=(w_1,w_2,w_3)$ defined in Eqs. (1.8) of Ref. 2 which we repeat below

$$\mathbf{w}=-[H\mathbf{J}+(\mathbf{P}\times\mathcal{J})]. \tag{5.1}$$

For particles of zero mass and discrete spin it can be shown from (3.1) of Ref. 2 that

$$\mathbf{w}=-\alpha\lambda\mathbf{P}, \tag{5.2}$$

where λ is the sign of the energy. The quantity

α, we recollect, is a scalar: $\alpha = +s$ or $-s$ where s is the spin of the particle.

The case of nonvanishing mass is more complicated. For this case one can introduce spin operators. These are the operators S_i $(i = 1, 2, 3)$ of Eq. (4.2) of Ref. 2.

They satisfy the following commutation rules:

$$
\begin{aligned}
[S_1, S_2] &= iS_3,\\
[S_2, S_3] &= iS_1,\\
[S_3, S_1] &= iS_2,\\
[S_i, H] &= 0,\\
[S_i, P_j] &= 0,\\
[J_i, S_i] &= 0,\\
[J_1, S_2] &= iS_3 = [S_1, J_2],\\
[J_2, S_3] &= iS_1 = [S_2, J_2],\\
[J_3, S_1] &= iS_2 = [S_3, J_1].
\end{aligned}
\tag{5.3}
$$

In addition the spin operators must satisfy

$$S_1^2 + S_2^2 + S_3^2 = s(s+1)I, \tag{5.4}$$

where I is the identity operator and s, as usual, is the spin of the particle. The operators w_0, $\mathbf{w}$ can be expressed in terms of the spin operators

$$w_0 = (\mathbf{P}\cdot\mathbf{J}) = (\mathbf{P}\cdot\mathbf{S}), \tag{5.5}$$

$$\mathbf{w} = -\frac{Pw_0}{H+\mu} - \mu\mathbf{S}. \tag{5.6}$$

The bulk of the remainder of the paper will be devoted to showing that Eqs. (5.3) through (5.6) determine how the spin operator S_i act in the angular-momentum basis. The principle by which the operators Q_i and K will be found is that w_i will be expressed both in the form (5.1) and in the form (5.2) or (5.6). A comparison of results determines the unknown operators.

It will thus be necessary to evaluate $w_3\,|E, j, m\rangle$ where w_3 is given by (5.1). For simplicity, we shall take the quantum number $m = -j$ since the operators for which we are looking are independent of m.

We note

$$
\begin{aligned}
w_3 = &-\tfrac{1}{2}(\mathcal{J}_1 - i\mathcal{J}_2)(P_1 + iP_2)\\
&+ \tfrac{1}{2}(\mathcal{J}_1 + i\mathcal{J}_2)(P_1 - iP_2) - HJ_3.
\end{aligned}
\tag{5.7}
$$

On using (3.46), (4.47), and (4.48) we obtain, after some tedious calculation,

$$
\begin{aligned}
&w_3\,|E, j, -j\rangle\\
&= \left\{jE - ip\left[\frac{-2j(j+1)(2j+3)}{2j+1}T_{j+1}(\alpha)Q_{j+1}^\dagger(E)\right.\right.\\
&\quad \left.\left. + 2j^2T_j(\alpha)Q_j(E) + \frac{\alpha}{j(j+1)^2}K(E, j)\right]\right\}|E, j, -j\rangle\\
&\quad + ip\left\{\frac{1}{(j+1)}T_{j+1}(\alpha)K(E, j+1)\right.\\
&\quad \left. - \frac{\alpha}{(j+1)}Q_{j+1}(E)\right\}|E, j+1, -j\rangle.
\end{aligned}
\tag{5.8}
$$

VI. THE DERIVATION OF THE FORMS OF THE OPERATORS $\mathcal{J}_i$

Part II. Completion of the Zero-Mass Case

The finding of Q_i and K for the massless case is now straightforward. Since α takes on only one value $Q_i(E)$ is simply a number. $K(E, j)$ likewise is a *real* number, since as an operator K is Hermitian. Also $Q_j^\dagger = Q_j^*$ where the asterisk means complex conjugate.

From (5.2) $w_3 = -\alpha\lambda P_3$. Hence,

$$
\begin{aligned}
w_3\,|E, j, -j\rangle &= -\alpha\lambda P_3\,|E, j, -j\rangle,\\
&= \frac{\lambda p\alpha^2}{(j+1)}|E, j, -j\rangle - \alpha\lambda pT_{j+1}(\alpha)\,|E, j+1, -j\rangle,
\end{aligned}
\tag{6.1}
$$

where in the present massless case

$$p = \lambda E.$$

On comparing (6.1) with (5.8) we can equate coefficients of like kets and thus obtain

$$
i\left[\frac{1}{(j+1)}T_{j+1}(\alpha)K(E, j+1) - \frac{\alpha}{(j+1)}Q_{j+1}(E)\right] = -\alpha\lambda T_{j+1}(\alpha), \tag{6.2}
$$

$$
\begin{aligned}
&ij\left[\frac{2(j+1)(2j+3)}{2j+1}T_{j+1}(\alpha)Q_{j+1}^*(E) - 2jT_j(\alpha)Q_j(E) - \frac{\alpha}{j^2(j+1)^2}K(E, j)\right]\\
&= -\frac{\lambda}{(j+1)}[j(j+1) - \alpha^2].
\end{aligned}
\tag{6.3}
$$

From (6.2),

$$Q_{j+1}(E) = -i\lambda(j+1)T_{j+1}(\alpha) + \frac{K(E, j+1)}{\alpha}T_{j+1}(\alpha), \tag{6.4}$$

if $\alpha \neq 0$. On substituting into (6.3) one obtains

$$K(E, j+1) = K(E, j), \tag{6.5}$$

or $K(E, j)$ is independent of j. We set

$$K(E) = K(E, j) \tag{6.6}$$

to indicate this independence in the notation.

Thus,

$$Q_{j+1}(E) = -i\lambda(j+1)T_{j+1}(\alpha) + \frac{K(E)}{\alpha} T_{j+1}(\alpha) \tag{6.7}$$

for $\alpha \neq 0$.

We now show that $K(E)$ can be set equal to zero by changing the phase of the basis suitably. We recollect that $K(E)$ arises from (4.21) which we rewrite as follows:

$$\mathbf{J}\cdot\boldsymbol{\mathcal{J}}\,|E, j, m\rangle = K(E)\,|E, j, m\rangle - ip\alpha \frac{\partial}{\partial E}\,|E, j, m\rangle. \tag{6.8}$$

Let us introduce a new set of kets $|E, j, m)$ defined by

$$|E, j, m) = \exp\,[i\beta(E)]\,|E, j, m\rangle, \tag{6.9}$$

where

$$\beta'(E) = K(E)/p\alpha. \tag{6.10}$$

In terms of the new basis

$$\mathbf{J}\cdot\boldsymbol{\mathcal{J}}\,|E, j, m) = -ip\alpha \frac{\partial}{\partial E}\,|E, j, m). \tag{6.11}$$

Thus in the new basis the quantity corresponding to $K(E)$ equals zero. It is clear that all the operators act in the new basis precisely as in the old one except $K(E) = 0$.

Hence, we have in the new basis

$$\begin{aligned} K(E) &= 0, \\ Q_{j+1}(E) &= -i\lambda(j+1)T_{j+1}(\alpha). \end{aligned} \tag{6.12}$$

Equations (6.12) are valid for $\alpha \neq 0$. If $\alpha = 0$, it follows from (6.2) that $K(E, j) = 0$ for $j > 0$. From (4.21) and the fact that $J_i\,|E, 0, 0\rangle = 0$, it follows that $K(E, 0) = 0$ also. Hence,

$$K(E, j) = 0, \quad \text{for} \quad \alpha = 0 \quad \text{and all } j. \tag{6.13}$$

On substituting (6.12) into (6.3) we obtain an equation for $Q_j(E)$.

We make the ansatz

$$Q_j(E) = -i\lambda j T_j(0) + F_j, \tag{6.14}$$

and substitute into (6.3). We are led to an equation for F_j which has the solution

$$\begin{aligned} F_j &= [3/j(2j+1)]F \quad \text{for} \quad j \text{ odd}, \\ &= [3/j(2j+1)]F^* \quad \text{for} \quad j \text{ even}, \end{aligned} \tag{6.15}$$

where F is a number.

The requirement that $\mathcal{J}_3$ be Hermitian leads to the result that F is real.

On introducing a new basis

$$|E, j, m) = \exp\,[i\beta(E)]\,|E, j, m\rangle, \tag{6.16}$$

where

$$\beta'(E) = 6(F/p), \tag{6.16a}$$

and the operators act in the new basis in precisely the same way as in the old, where Q_j is given by the second of Eqs. (6.12).

To summarize: For the massless case and arbitrary discrete spin we have obtained a basis such that the infinitesimal generators of the Lorentz group operate as in (3.26) and (4.46)–(4.48). The numbers $K(E, j) = 0$. The numbers Q_j are given by the second of Eqs. (6.12). Finally the completeness relation is given by (3.47) and (4.19).

Thus we have essentially completed the derivation for the massless case. The case of finite mass will be undertaken in the remainder of the paper. As will be seen, the derivation parallels that for zero mass but is considerably longer.

VII. DERIVATION OF THE FORM OF THE SPIN OPERATORS FOR THE CASE OF NONZERO MASS

In this section we show how the spin operators S_i operate on the basis vectors $|E, j, m, \alpha\rangle$. It will be shown that the commutation rules (5.3), together with (5.4) and (5.5) and the requirement that the spin operators be Hermitian are sufficient to give S_i uniquely.

It will be convenient to introduce the operators P^2 and T_i ($i = 1, 2, 3$) which are defined by

$$P^2 = H^2 - \mu^2, \tag{7.1}$$

$$T_i = S_i - w_0(P^2)^{-1}P_i. \tag{7.2}$$

In terms of the operators T_i, the commutation relations (5.3) become

$$\begin{aligned} [T_1, T_2] &= iw_0(P^2)^{-1}P_3, \\ [T_2, T_3] &= iw_0(P^2)^{-1}P_1, \\ [T_3, T_1] &= iw_0(P^2)^{-1}P_2, \\ [T_i, H] &= 0, \\ [T_i, P_j] &= 0, \\ [J_i, T_i] &= 0, \\ [J_1, T_2] &= iT_3 = [T_1, J_2], \\ [J_2, T_3] &= iT_1 = [T_2, J_3], \\ J_3, T_1] &= iT_2 = [T_3, J_1]. \end{aligned} \tag{7.3}$$

From (5.5)

$$\mathbf{P}\cdot\mathbf{T} = 0. \tag{7.4}$$

Also (5.4) and (5.6) become

$$\mathbf{T}^2 = s(s+1)I - (w_0)^2(P^2)^{-1}, \tag{7.5}$$

$$w_i = -Hw_0(P^2)^{-1}P_i - \mu T_i. \tag{7.6}$$

The operators T_i are Hermitian. We shall find how the operators T_i act in the basis. From (7.2) it will then be possible to determine how the spin operators act.

From the fact that $[T_3, J_3] = 0$ and $[T_3, H] = 0$ we note that $T_3 \,|E, j, m, \alpha\rangle$ can be written as the following superposition of ket vectors

$$T_3\,|E, j, m, \alpha\rangle = \sum_{j',\alpha'} A_{j,\alpha:j',\alpha'}(E, m)\,|E, j', m, \alpha'\rangle. \tag{7.7}$$

We shall now show that the coefficients $A_{j,\alpha:j',\alpha'}$ vanish unless $\alpha' = \alpha+1, \alpha-1$ and $j' = j, j-1, j+1$. On applying the commutation relations

$$[w_0, T_3] = -i(\mathbf{P}\times\mathbf{T})_3, \qquad [w_0, (\mathbf{P}\times\mathbf{T})_3] = iP^2T_3, \tag{7.8}$$

to the ket $|E, j, m, \alpha\rangle$ we obtain the equations

$$\begin{aligned}(w_0 - \alpha p)T_3\,|E, j, m, \alpha\rangle &+ i(\mathbf{P}\times\mathbf{T})_3\,|E, j, m, \alpha\rangle = 0;\\ -ip^2T_3\,|E, j, m, \alpha\rangle + (w_0 - \alpha p) &\times (\mathbf{P}\times\mathbf{T})_3\,|E, j, m, \alpha\rangle = 0.\end{aligned} \tag{7.9}$$

On eliminating $(\mathbf{P}\times\mathbf{T})_3\,|E, j, m, \alpha\rangle$ from the Eqs. (7.9) we obtain

$$[w_0 - (\alpha+1)p][w_0 - (\alpha-1)p]T_3\,|E, j, m, \alpha\rangle = 0. \tag{7.10}$$

On substituting (7.7) into (7.10) one verifies the result that

$$A_{j,\alpha:j',\alpha'} = 0 \quad \text{if} \quad \alpha' \neq \alpha+1, \alpha-1.$$

From the commutation rules

$$[\mathbf{J}^2, \mathbf{T}_3] = 2i(\mathbf{T}\times\mathbf{J})_3 + 2T_3, \qquad [\mathbf{J}^2, (\mathbf{T}\times\mathbf{J})_3] = 2i(\mathbf{T}\cdot\mathbf{J})J_3 - T_3\mathbf{J}^2, \tag{7.11}$$

we obtain, on applying these equations to the ket $|E, j, m, \alpha\rangle$,

$$\begin{aligned}&[\mathbf{J}^2 - j(j+1) - 2]T_3\,|E, j, m, \alpha\rangle - 2i(\mathbf{T}\times\mathbf{J})_3\,|E, j, m, \alpha\rangle = 0,\\ &2ij(j+1)T_3\,|E, j, m, \alpha\rangle - [\mathbf{J}^2 - j(j+1)](\mathbf{T}\times\mathbf{J})_3\,|E, j, m, \alpha\rangle = 2im(\mathbf{T}\cdot\mathbf{J})\,|E, j, m, \alpha\rangle.\end{aligned}$$

On eliminating $(\mathbf{T}\times\mathbf{J})_3\,|E, j, m, \alpha\rangle$ from (7.12) we obtain the following:

$$[\mathbf{J}^2 - j(j-1)][\mathbf{J}^2 - (j+1)(j+2)]T_3\,|E, j, m, \alpha\rangle = -4m\mathbf{T}\cdot\mathbf{J}\,|E, j, m, \alpha\rangle. \tag{7.13}$$

Since $\mathbf{T}\cdot\mathbf{J}$ commutes with J_3 and with $\mathbf{J}^2$, $T\cdot J\,|E, j, m, \alpha\rangle$ will be a superposition of kets of the form

$$\mathbf{T}\cdot\mathbf{J}\,|E, j, m, \alpha\rangle = \sum_{\alpha'} B_{\alpha,\alpha'}(E, j, m)\,|E, j, m, \alpha'\rangle. \tag{7.14}$$

On substituting (7.14) and (7.7) into (71.3) we see that $A_{j,\alpha:j',\alpha'} = 0$, if $j' \neq j, j+1$, or $j-1$.

Hence we may write

$$\begin{aligned}T_3\,|E, j, m, \alpha\rangle = {}& A_j(m, \alpha)\,|E, j, m, \alpha-1\rangle\\ &+ B_{j+1}(m, \alpha)\,|E, j+1, m, \alpha-1\rangle\\ &+ C_{j-1}(m, \alpha)\,|E, j-1, m, \alpha-1\rangle\\ &+ D_j(m, \alpha)\,|E, j, m, \alpha+1\rangle\\ &+ E_{j+1}(m, \alpha)\,|E, j+1, m, \alpha+1\rangle\\ &+ F_{j-1}(m, \alpha)\,|E, j-1, m, \alpha+1\rangle.\end{aligned} \tag{7.15}$$

The coefficients A_j, B_{j+1}, etc., may also depend also on E, but we have suppressed this dependence in the notation, since this dependence will play no role for a time. From now on we shall be interested in obtaining the coefficients in (7.15) explicitly. It should be mentioned that our derivation, at times, will not be valid for special values of α, j, or m, for example for $\alpha = s$ or $j = 0$. However, the final formulas will be correct for all values of α, m, and j. One can derive the formulas for the special values of α, m, or j by using special tricks. We omit these tricks because inclusion of the special cases would add greatly to the length of this exposition. A simpler treatment is to derive the formulas for the general values of α, m, and j and verify by direct computation that even for the singular values of α, m, and j the formulas are valid. We take this latter point of view.

The condition that T_3 be Hermitian provides a relation between some of the coefficients which appear in (7.15). We require

$$\langle E, j, m, \alpha|\,T_3\,|E', j', m', \alpha'\rangle = \langle E', j', m', \alpha'|\,T_3\,|E, j, m, \alpha\rangle^*. \tag{7.16}$$

On substituting (7.15) into (7.16) and on using the orthogonality properties

$$\langle E, j, m, \alpha \mid E', j', m', \alpha' \rangle = p\frac{(j+m)!}{(j-m)!}(2j)!\delta(E-E')\delta_{j,j'}\delta_{m,m'}\delta_{\alpha,\alpha'}, \quad (7.17)$$

which can be derived from the completeness relation (3.47) and (4.19), we obtain

$$A_j(m,\alpha) = D_j^*(m, \alpha - 1),$$
$$B_{j+1}(m,\alpha) = \frac{j-m+1}{j+m+1}\frac{1}{2(j+1)(2j+1)}F_j^*(m,\alpha-1),$$
$$C_{j-1}(m,\alpha) = 2j(2j-1)\frac{j+m}{j-m}E_j^*(m,\alpha-1). \quad (7.18)$$

From the commutation rules we have

$$[P_3, T_3]\,|E, j, m, \alpha\rangle = 0, \quad (7.19)$$
$$[P_2 - iP_1, T_3]\,|E, j, m, \alpha\rangle = 0, \quad (7.20)$$
$$[P_2 + iP_1, T_3]\,|E, j, m, \alpha\rangle = 0. \quad (7.20a)$$

The left-hand sides of (7.19) and (7.20) will be a linear combination of basis vectors. Since these vectors are linearly independent, the coefficients of each will be zero. In (7.19) we set the coefficient of $|E, j+2, m, \alpha+1\rangle$ equal to zero and likewise in (7.20) we set the coefficient of $|E, j+2, m+1, \alpha+1\rangle$ to zero. We thus obtain the following equations:

$$\frac{E_{j+1}(m,\alpha)}{E_j(m,\alpha)} = \frac{j-m+1}{j-m}\frac{2j-1}{2j+1}\frac{T_{j+1}(\alpha+1)}{T_j(\alpha)}, \quad (7.21)$$

$$\frac{E_{j+1}(m+1,\alpha)}{E_j(m,\alpha)} = \frac{2j-1}{2j+1}\frac{T_{j+1}(\alpha+1)}{T_j(\alpha)}. \quad (7.22)$$

From (7.21) and (7.22)

$$\frac{E_j(m+1,\alpha)}{E_j(m,\alpha)} = \frac{j-m-1}{j-m}. \quad (7.23)$$

On setting $E_j(-j,\alpha) = E_j(\alpha)$ we see that the solution of the recursion relation (7.23) is

$$E_j(m,\alpha) = \frac{j-m}{2j-1}E_j(\alpha). \quad (7.24)$$

On substituting (7.24) into (7.22) we obtain an equation for $E_j(\alpha)$:

$$E_{j+1}(\alpha)/E_j(\alpha) = T_{j+1}(\alpha+1)/T_j(\alpha). \quad (7.25)$$

The solution of this recursion relation is

$$E_{j+1}(\alpha) = \nu(j,\alpha)E(\alpha), \quad (7.26)$$

where

$$E(\alpha) = E_{|\alpha|+1}(\alpha),$$

$$\nu(j,\alpha) = \left[\frac{2|\alpha|+3}{2j+3}\left(\frac{|\alpha|+1}{j+1}\right)^3 \times \frac{(j+\alpha+2)(j+\alpha+1)}{(|\alpha|+\alpha+2)(|\alpha|+\alpha+1)}\right]^{\frac{1}{2}}. \quad (7.26a)$$

Thus from (7.24) and (7.26)

$$E_{j+1}(m,\alpha) = [(j-m+1)/(2j+1)]\nu(j,\alpha)E(\alpha). \quad (7.27)$$

Also from the second of Eqs. (7.18)

$$C_{j-1}(m,\alpha) = 2j(j+m)\nu(j-1,\alpha-1)E^*(\alpha-1). \quad (7.28)$$

In (7.19) and (7.20) set the coefficients of $|E, j+2, m, \alpha-1\rangle$ and $|E, j+2, m+1, \alpha-1\rangle$, respectively, equal to zero:

$$\frac{B_{j+1}(m,\alpha)}{B_j(m,\alpha)} = \frac{j-m+1}{j-m}\frac{2j-1}{2j+1}\frac{T_{j+1}(\alpha-1)}{T_j(\alpha)}, \quad (7.29)$$

$$\frac{B_{j+1}(m+1,\alpha)}{B_j(m,\alpha)} = \frac{2j-1}{2j+1}\frac{T_{j+1}(\alpha-1)}{T_j(\alpha)}. \quad (7.30)$$

By the techniques used to find $E_j(m,\alpha)$ and $C_j(m,\alpha)$ we find that

$$B_{j+1}(m,\alpha) = \frac{j-m+1}{2j+1}\nu(j,-\alpha)B(\alpha), \quad (7.31)$$

$$F_{j-1}(m,\alpha) = 2j(j+m)\nu(j-1,-\alpha-1)B^*(\alpha+1), \quad (7.32)$$

where $B(\alpha) = B_{|\alpha|+1}(-j,\alpha)$.

In (7.19) and (7.20a) let us set the coefficients of $|E, j+1, m, \alpha+1\rangle$ and $|E, j+1, m-1, \alpha+1\rangle$, respectively, equal to zero. We obtain the equations

$$T_{j+1}(\alpha)D_{j+1}(m,\alpha) - T_{j+1}(\alpha+1)D_j(m,\alpha) = \frac{(j-2\alpha)(2j+1)}{j(j+1)(j+2)(j-m+1)}mE_{j+1}(m,\alpha). \quad (7.33)$$

$$T_{j+1}(\alpha+1)D_j(m,\alpha) - T_{j+1}(\alpha)D_{j+1}(m-1,\alpha) = -\frac{2j+1}{(j+1)(j-m+1)(j-m+2)} \times\left[\frac{(j+m+1)(j-m+2)(\alpha+1)}{j+2}E_{j+1}(m,\alpha) - \frac{(j+m)(j-m+1)\alpha}{j}E_{j+1}(m-1,\alpha)\right]. \quad (7.34)$$

On adding (7.33) and (7.34) and on using (7.26) and $E_j(m,\alpha)$, we obtain the following recursion relation for $D_j(m,\alpha)$:

$$T_{j+1}(\alpha)[D_{j+1}(m,\alpha) - D_{j+1}(m-1,\alpha)] = \frac{\nu(j,\alpha)(\alpha-j-1)}{(j+1)(j+2)}E(\alpha). \quad (7.35)$$

Let us define $D_{j+1}(\alpha)$ by $D_{j+1}(\alpha) = D_{j+1}(-j-1, \alpha)$. The solution of (7.35) is

$$D_{j+1}(m, \alpha) = D_{j+1}(\alpha) - \frac{(j+m+1)(j-\alpha+1)}{(j+1)(j+2)} \frac{\nu(j, \alpha)}{T_{j+1}(\alpha)} E(\alpha). \quad (7.36)$$

Let us substitute (7.36) into (7.33). We then obtain a recursion relation for $D_j(\alpha)$, namely,

$$D_{j+1}(\alpha) = \frac{T_{j+1}(\alpha+1)}{T_{j+1}(\alpha)} D_j(\alpha) + \frac{(\alpha+1)\nu(j, \alpha)}{(j+1)(j+2)T_{j+1}(\alpha)} E(\alpha). \quad (7.37)$$

If we define $D(\alpha) = D_{|\alpha|+1}(\alpha)$, then the solution of (7.37) is

$$D_j(\alpha) = \frac{T_{|\alpha|+1}(\alpha)}{T_j(\alpha)} \nu(j-1, \alpha) D(\alpha) + \frac{(j-|\alpha|-1)(\alpha+1)}{(|\alpha|+2)(j+1)} \frac{\nu(j-1, \alpha)}{T_j(\alpha)} E(\alpha). \quad (7.38)$$

Hence, from (7.36) and (7.38) we see that

$$D_j(m, \alpha) = -\frac{m(j-\alpha)}{j(j+1)} \frac{\nu(j-1, \alpha)}{T_j(\alpha)} E(\alpha) + R(j, \alpha), \quad (7.39)$$

where $R(j, \alpha)$ is a function of j and α and is independent of m.

Now in the expression

$$[P_3, T_3] \, |E, j, m, \alpha\rangle = 0,$$

let us set the coefficient of $|E, j, m, \alpha+1\rangle$ equal to zero.

Then we obtain

$$\begin{aligned} &\frac{m}{j(j+1)} D_j(m, \alpha) \\ &\quad + 2(j+1)(j+m+1)T_{j+1}(\alpha+1)E_{j+1}(m, \alpha) \\ &\quad + \frac{j-m}{2j+1} T_j(\alpha+1)F_{j-1}(m, \alpha) \\ &= \frac{j-m+1}{2j+1} T_{j+1}(\alpha)F_j(m, \alpha) \\ &\quad + 2j(j+m)T_j(\alpha)E_j(m, \alpha). \end{aligned} \quad (7.40)$$

Let us use Eqs. (7.27) and (7.32) in (7.40) for $E_j(m, \alpha)$ and $F_j(m, \alpha)$, respectively. After some rearrangement, (7.40) becomes

$$\begin{aligned} &\frac{m}{j(j+1)} D_j(m, \alpha) \\ &+ 2\left\{\frac{j+1}{2j+1} [(j+1)^2 - m^2]T_{j+1}(\alpha+1)\nu(j, \alpha) - \frac{j}{2j-1} (j^2 - m^2)T_j(\alpha)\nu(j, \alpha-1)\right\}E(\alpha) \\ &= 2\left\{\frac{j+1}{2j+1} [(j+1)^2 - m^2]T_{j+1}(\alpha)\nu(j, -\alpha-1) - \frac{1}{2j-1} (j^2-m^2)T_j(\alpha+1)\nu(j-1, -\alpha-1)\right\}B^*(\alpha+1). \end{aligned} \quad (7.41)$$

Now Eq. (7.41) holds for all values of m. In particular, it holds when m is replaced by $-m$. Let us subtract the equation which one would obtain by replacing m by $-m$ from (7.41). We see immediately that

$$D_j(m, \alpha) = -D_j(-m, \alpha). \quad (7.42)$$

Hence, the quantity $R(j, \alpha)$ in (7.39) is given by

$$R(j, \alpha) = 0. \quad (7.39a)$$

Again, since (7.41) holds for all m, we may equate the coefficients of m^2 on both sides of the equation. We obtain a relation between $E(\alpha)$ and $B(\alpha)$, namely,

$$B(\alpha) = -\frac{j-\alpha+1}{j+\alpha} \frac{\nu(j-1, \alpha-1)}{\nu(j, -\alpha)} \times \frac{T_{j+1}(\alpha-1)}{T_j(\alpha-1)} E^*(\alpha-1), \quad (7.43)$$

or using the expressions for T_j and ν

$$B(\alpha) = -\left[\frac{2|\alpha-1|+3}{2|\alpha|+3} \left(\frac{|\alpha-1|+1}{|\alpha|+1}\right)^3 \times \frac{|\alpha|-\alpha+1}{|\alpha-1|+\alpha} \frac{|\alpha|-\alpha+2}{|\alpha-1|+\alpha+1}\right]^{\frac{1}{2}} E^*(\alpha-1). \quad (7.44)$$

Having found $D_j(m, \alpha)$ we can obtain $A_j(m, \alpha)$ from (7.18).

Let us summarize our results:

$$\begin{aligned} A_j(m, \alpha) &= -\frac{m(j-\alpha+1)}{j(j+1)} \times \frac{\nu(j-1, \alpha-1)}{T_j(\alpha-1)} E^*(\alpha-1), \\ B_j(m, \alpha) &= \frac{(j-m)}{2j-1} \nu(j-1, -\alpha)B(\alpha), \\ C_j(m, \alpha) &= 2(j+1)(j+m+1) \times \nu(j, \alpha-1)E^*(\alpha-1), \\ D_j(m, \alpha) &= -\frac{m(j-\alpha)}{j(j+1)} \frac{\nu(j-1, \alpha)}{T_j(\alpha)} E(\alpha), \\ E_j(m, \alpha) &= \frac{j-m}{2j-1} \nu(j-1, \alpha)E(\alpha), \\ F_j(m, \alpha) &= 2(j+1)(j+m+1) \times \nu(j, -\alpha-1)B^*(\alpha+1). \end{aligned} \quad (7.45)$$

Since $B(\alpha)$ and $E(\alpha)$ are related by (7.44), we see that all our coefficients now depend on only one function of α, namely $E(\alpha)$. Therefore, it will be our objective to find this function. We shall use (7.5) toward this end.

We shall apply both sides of (7.5) to the ket $|E, j, m, \alpha\rangle$ and equate the coefficients of $|E, j, m, \alpha\rangle$. Thus, we must find how $\mathbf{T}^2$ acts on $|E, j, m, \alpha\rangle$.

We note that

$$2\mathbf{T}^2 = (T_2 - iT_1)(T_2 + iT_1) + (T_2 + iT_1)(T_2 - iT_1) + 2T_3^2. \quad (7.46)$$

Hence, to find $\mathbf{T}^2 |E, j, m, \alpha\rangle$ we must find $(T_2 - iT_1) |E, j, m, \alpha\rangle$ and $(T_2 + iT_1) |E, j, m, \alpha\rangle$. From the commutation rules (7.3) and from (3.46) and (7.15)

$$\begin{aligned}(T_2 - iT_1) |E, j, m, \alpha\rangle &= [T_3, J_2 - iJ_1] |E, j, m, \alpha\rangle, \\ &= [A_j(m+1, \alpha) - A_j(m, \alpha)] |E, j, m+1, \alpha-1\rangle \\ &\quad + [B_{j+1}(m+1, \alpha) - B_{j+1}(m, \alpha)] |E, j+1, m+1, \alpha-1\rangle \\ &\quad + [C_{j-1}(m+1, \alpha) - C_{j-1}(m, \alpha)] |E, j-1, m+1, \alpha-1\rangle \\ &\quad + [D_j(m+1, \alpha) - D_j(m, \alpha)] |E, j, m+1, \alpha+1\rangle \\ &\quad + [E_{j+1}(m+1, \alpha) - E_{j+1}(m, \alpha)] |E, j+1, m+1, \alpha+1\rangle \\ &\quad + [F_{j-1}(m+1, \alpha) - F_{j-1}(m, \alpha)] |E, j-1, m+1, \alpha+1\rangle. \end{aligned} \quad (7.47)$$

Likewise

$$\begin{aligned}(T_2 + iT_1) |E, j, m, \alpha\rangle &= [J_2 + iJ_1, T_3] |E, j, m, \alpha\rangle \\ &= (j+m)(j-m+1)[A_j(m, \alpha) - A_j(m-1, \alpha)] |E, j, m-1, \alpha-1\rangle \\ &\quad + [(j+m+1)(j-m+2)B_{j+1}(m, \alpha) - (j+m)(j-m+1)B_{j+1}(m-1, \alpha)] |E, j+1, m-1, \alpha-1\rangle \\ &\quad + [(j+m-1)(j-m)C_{j-1}(m, \alpha) - (j+m)(j-m+1)C_{j-1}(m-1, \alpha)] |E, j-1, m-1, \alpha-1\rangle \\ &\quad + [(j+m)(j-m+1)D_j(m, \alpha) - D_j(m-1, \alpha)] |E, j, m-1, \alpha+1\rangle \\ &\quad + [(j+m+1)(j-m+2)E_{j+1}(m, \alpha) - (j+m)(j-m+1)E_{j+1}(m-1, \alpha)] |E, j+1, m-1, \alpha+1\rangle \\ &\quad + [(j+m-1)(j-m)F_{j-1}(m, \alpha) - (j+m)(j-m+1)F_{j-1}(m-1, \alpha)] |E, j-1, m-1, \alpha+1\rangle. \end{aligned} \quad (7.48)$$

Then by equating the coefficients of

$$|E, j, m, \alpha\rangle \quad \text{in} \quad \mathbf{T}^2 |E, j, m, \alpha\rangle = s(s+1) |E, j, m, \alpha\rangle - \alpha^2 |E, j, m, \alpha\rangle,$$

we obtain an expression of the form (on using (7.45) for the coefficients A_j, B_j, etc.)

$$K(j, \alpha) + mL(j, \alpha) + m^2M(j, \alpha) = 0, \quad (7.49)$$

where L, K, and M are independent of m. Since the calculation is extremely tedious, we do not carry it out in detail. Equation (7.49) holds for every value of m. It follows that

$$K(j, \alpha) = L(j, \alpha) = M(j, \alpha) = 0. \quad (7.50)$$

On writing out $K(j, \alpha) = 0$ explicitly we have

$$\begin{aligned}&2j^2\left[\frac{2j+1}{j+1}\frac{j-\alpha+1}{j+\alpha-1} + 1\right] \\ &\times \nu^2(j-1, \alpha-1) |E(\alpha-1)|^2 \\ &+ 2(j+1)^2(2j+3)\left[\frac{1}{j}\frac{j-\alpha}{j+\alpha+2} + \frac{1}{2j+1}\right] \\ &\times \nu^2(j, \alpha) |E(\alpha)|^2 \\ &+ 2\frac{(j+1)^2(2j+3)}{2j+1}\nu^2(j, -\alpha) |B(\alpha)|^2 \\ &+ 2j^2\nu^2(j-1, -\alpha-1) |B(\alpha+1)|^2 \\ &= s(s+1) - \alpha^2. \end{aligned} \quad (7.51)$$

We can eliminate $|B(\alpha)|$ from (7.51) by using (7.44). We obtain

$$4c^2(\alpha) |E(\alpha)|^2 + 4c^2(\alpha-1) |E(\alpha-1)|^2 = \frac{s(s+1) - \alpha^2}{2}, \quad (7.52)$$

where

$$c(\alpha) = \left[\frac{2|\alpha| + 3}{2}\frac{|\alpha+1|^3}{(|\alpha| + \alpha + 1)(|\alpha| + \alpha + 2)}\right]^{\frac{1}{2}}. \quad (7.52a)$$

But from the definition of $E_j(m, \alpha)$ it follows that

$$E(-s-1) = 0.$$

One can then solve the recursion relation (7.52) for $|E(\alpha)|$:

$$|E(\alpha)| = \frac{1}{4c(\alpha)} [(s-\alpha)(s+\alpha+1)]^{\frac{1}{2}}. \tag{7.53}$$

Thus

$$E(\alpha) = \frac{\omega(\alpha)}{4c(\alpha)} [(s-\alpha)(s+\alpha+1)]^{\frac{1}{2}}, \tag{7.54}$$

where $\omega(\alpha)$ is a complex number of modulus 1:

$$|\omega(\alpha)| = 1. \tag{7.54a}$$

From (7.44)

$$B(\alpha) = -\frac{\omega^*(\alpha)}{4c(-\alpha)} [(s-\alpha+1)(s+\alpha)]^{\frac{1}{2}}. \tag{7.55}$$

We can now use (7.54) and (7.55) in (7.45) to obtain the coefficients $A_j(m, \alpha)$, etc. A useful identity is

$$\nu(j, \alpha) = c(\alpha)\left[\frac{2(j+\alpha+1)(j+\alpha+2)}{(2j+3)(j+1)^3}\right]^{\frac{1}{2}}. \tag{7.56}$$

We find that we obtain surprisingly simple results.

$$\begin{aligned}
A_j(m, \alpha) &= -\frac{m\omega^*(\alpha-1)}{2j(j+1)} \\
&\quad \times [(j-\alpha+1)(j+\alpha)(s-\alpha+1)(s+\alpha)]^{\frac{1}{2}}, \\
B_{j+1}(m, \alpha) &= -\frac{(j-m+1)\omega^*(\alpha-1)}{4(2j+1)} \\
&\quad \times \left[\frac{2(j-\alpha+1)(j-\alpha+2)(s-\alpha+1)(s+\alpha)}{(2j+3)(j+1)^3}\right]^{\frac{1}{2}}, \\
C_{j-1}(m, \alpha) &= (j+m)\omega^*(\alpha-1) \\
&\quad \times \left[\frac{(j+\alpha-1)(j+\alpha)(s-\alpha+1)(s+\alpha)}{2j(2j+1)}\right]^{\frac{1}{2}}, \\
D_j(m, \alpha) &= -\frac{m\omega(\alpha)}{2j(j+1)} \\
&\quad \times [(j-\alpha)(j+\alpha+1)(s-\alpha)(s+\alpha+1)]^{\frac{1}{2}}, \\
E_{j+1}(m, \alpha) &= \frac{(j-m+1)\omega(\alpha)}{4(2j+1)} \\
&\quad \times \left[\frac{2(j+\alpha+1)(j+\alpha+2)(s-\alpha)(s+\alpha+1)}{(2j+3)(j+1)^3}\right]^{\frac{1}{2}}, \\
F_{j-1}(m, \alpha) &= -(j+m)\omega(\alpha) \\
&\quad \times \left[\frac{(j-\alpha-1)(j-\alpha)(s-\alpha)(s+\alpha+1)}{2j(2j+1)}\right]^{\frac{1}{2}}.
\end{aligned} \tag{7.57}$$

It remains only to find the factors $\omega(\alpha)$. We shall now show that they can be chosen to be 1.

Let us introduce a new basis $|E, j, m, \alpha)$ which is related to the old one by

$$\begin{aligned}
|E, j, m, \alpha) &= \prod_{r=0}^{\alpha+s-1} \omega(-s+r)\, |E, j, m, \alpha\rangle \text{ for } \alpha > -s, \\
|E, j, m, -s) &= |E, j, m, -s\rangle.
\end{aligned} \tag{7.58}$$

It is easily seen that all our operators act in the new basis precisely as in the old one except that in the new basis $\omega(\alpha) = 1$.

Thus, we shall assume that the kets $|E, j, m, \alpha\rangle$ are chosen such that $\omega(\alpha) = 1$ and then the coefficients are given by (7.57) explicitly. From (7.2), (7.15), (7.47), (7.48), and (7.57) we know how the spin operators S_i act in our basis.

VIII. THE DERIVATION OF THE FORMS OF THE OPERATORS $\mathcal{J}_i$

Part III. Completion of the Nonzero Mass Case

We shall now find the operators Q_i and $K(E, j)$ which appear in Eqs. (4.46)–(4.48). Henceforth, however, we no longer supress the variables α in these equations. From the commutation rules for the infinitesimal generators,

$$[w_0, \mathcal{J}_3] = iw_3. \tag{8.1}$$

Let us apply both sides of (8.1) to the ket $|E, j, -j, \alpha\rangle$. In order to evaluate $[w_0, \mathcal{J}_3]\, |E, j, -j, \alpha\rangle$ we must use

$$w_0 \frac{\partial}{\partial E} |E, j, m, \alpha\rangle = \alpha\left(\frac{E}{p} + p\frac{\partial}{\partial E}\right) |E, j, m, \alpha\rangle. \tag{8.2}$$

Equation (8.2) follows from

$$\begin{aligned}
w_0 \frac{\partial}{\partial E} &|E, j, m, \alpha\rangle \\
&= w_0 \lim_{\Delta\to 0} \frac{1}{\Delta} [|E+\Delta, j, m, \alpha\rangle - |E, j, m, \alpha\rangle], \\
&= \alpha \lim_{\Delta\to 0} \frac{1}{\Delta} [p(E+\Delta)\, |E+\Delta, j, m, \alpha\rangle \\
&\qquad - p(E)\, |E, j, m, \alpha\rangle], \\
&= \alpha \frac{\partial}{\partial E} [p(E)\, |E, j, m, \alpha\rangle],
\end{aligned} \tag{8.3}$$

where $p(E)$ is used to indicate the functional dependence of p on E, i.e., $p(E) = [E^2 - \mu^2]^{\frac{1}{2}}$. Equation (8.2) then follows from (8.3) on using the product rule for the derivative.

On using (8.2), (3.46), and (4.46) we obtain

$$[w_0, \mathcal{J}_3]\,|E, j, -j, \alpha\rangle = (w_0 - \alpha p)\mathcal{J}_3\,|E, j, -j, \alpha\rangle$$

$$= -\frac{1}{(j+1)}\,[p \sum_{\alpha'} (\alpha' - \alpha)K_{\alpha,\alpha'}(E, j)$$

$$\times |E, j, -j, \alpha'\rangle - i\alpha^2 E\,|E, j, -j, \alpha\rangle]$$

$$+ [p \sum_{\alpha'} (\alpha' - \alpha)Q_{j+1:\alpha,\alpha'}(E)\,|E, j+1, -j, \alpha'\rangle$$

$$- i\alpha E T_{j+1}(\alpha)\,|E, j+1, -j, \alpha\rangle], \qquad (8.4)$$

where $Q_{j+1:\alpha,\alpha'}(E)$ is the quantity which we have previously denoted by $Q_{j+1}(E)$ when we indicate the matrix character in the variable α.

From (7.6), (7.15), and (7.57) we have

$$w_3\,|E, j, -j, \alpha\rangle = -E\alpha\left[-\frac{\alpha}{j+1}\,|E, j, -j, \alpha\rangle + T_{j+1}(\alpha)\,|E, j+1, -j, \alpha\rangle\right]$$

$$- \frac{\mu}{2(j+1)}\,[(j-\alpha+1)(j+\alpha)(s-\alpha+1)(s+\alpha)]^{\frac12}\,|E, j, -j, \alpha-1\rangle$$

$$+ \frac{\mu}{4}\left[\frac{2(j-\alpha+1)(j-\alpha+2)(s-\alpha+1)(s+\alpha)}{(2j+3)(j+1)^3}\right]^{\frac12}\,|E, j, -j, \alpha-1\rangle$$

$$- \frac{\mu}{2(j+1)}\,[(j-\alpha)(j+\alpha+1)(s-\alpha)(s+\alpha+1)]^{\frac12}\,|E, j, -j, \alpha+1\rangle$$

$$- \frac{\mu}{4}\left[\frac{2(j+\alpha+1)(j+\alpha+2)(s-\alpha)(s+\alpha+1)}{(2j+3)(j+1)^3}\right]^{\frac12}\,|E, j+1, -j, \alpha+1\rangle. \qquad (8.5)$$

By comparing the coefficients of like kets in the expression

$$[w_0, \mathcal{J}_3]\,|E, j, -j, \alpha\rangle = iw_3\,|E, j, -j, \alpha\rangle, \qquad (8.6)$$

we obtain the following results:

$$Q_{j+1:\alpha,\alpha'}(E) = K_{\alpha,\alpha'}(E, j) = 0$$

$$\text{if} \quad \alpha' \neq \alpha\, \alpha+1, \quad \text{or} \quad \alpha - 1; \qquad (8.7)$$

$$K_{\alpha,\alpha+1}(E, j)$$

$$= \frac{i\mu}{2p}\,[(j-\alpha)(j+\alpha+1)(s-\alpha)(s+\alpha+1)]^{\frac12};$$

$$K_{\alpha,\alpha-1}(E, j)$$

$$= \frac{-i\mu}{2p}\,[(j-\alpha+1)(j+\alpha)(s-\alpha+1)(s+\alpha)]^{\frac12};$$

$$Q_{j+1:\alpha,\alpha+1}(E)$$

$$= \frac{-i\mu}{4p}\left[\frac{2(j+\alpha+1)(j+\alpha+2)(s-\alpha)(s+\alpha+1)}{(2j+3)(j+1)^3}\right]^{\frac12};$$

$$Q_{j+1:\alpha,\alpha-1}(E)$$

$$= \frac{-i\mu}{4p}\left[\frac{2(j-\alpha+1)(j-\alpha+2)(s-\alpha+1)(s+\alpha)}{(2j+3)(j+1)^3}\right]^{\frac12}. \qquad (8.8)$$

The only unknown quantities are the diagonal elements $Q_{j+1:\alpha,\alpha}(E)$ and $K_{\alpha,\alpha}(E, j)$. We shall find them by a process which resembles the procedure used for the massless case where these are the only elements which exist. We compare $w_3\,|E, j, -j, \alpha\rangle$ as given by (5.8) and by (8.5). We obtain two equations for $K_{\alpha,\alpha}(E, j)$ and $Q_{j+1:\alpha,\alpha}(E)$ which are analogous to (6.2) and (6.3).

From these equations one finds that $K_{\alpha,\alpha}(E, j)$ is independent of j. To indicate this independence we shall introduce a real function of E and α which we shall denote by $k(E, \alpha)$ defined by

$$k(E, \alpha) = K_{\alpha,\alpha}(E, j). \qquad (8.9)$$

One also obtains

$$Q_{j+1:\alpha,\alpha}(E) = T_{j+1}(\alpha)\left[\frac{k(E, \alpha)}{\alpha} - i(j+1)\frac{E}{p}\right]. \qquad (8.10)$$

Thus we now know all quantities except $k(E, \alpha)$.

We now show that we can choose a basis such that

$$k(E, \alpha) = 0. \qquad (8.11)$$

Toward this end we apply both sides of the commutation relation

$$[\mathcal{J}_3, w_3] = -iw_0 \qquad (8.12)$$

to the ket $|E, j, -j, \alpha\rangle$. The coefficient of $|E, j, -j, \alpha+1\rangle$ in the expansion of $[\mathcal{J}_3, w_3]\,|E, j, -j, \alpha\rangle$ must be zero from (8.12). We use the Equation (7.6) for w_3 and (4.46) for $\mathcal{J}_3$, substituting (7.57), (8.7)–(8.10) for the various coefficients which appear in the expressions for $\mathcal{J}_3$ and T_3. On setting the coefficient of $|E, j, -j, \alpha+1\rangle$ equal to zero, we obtain after a considerable amount of reduction,

$$\frac{k(E, \alpha+1)}{\alpha+1} = \frac{k(E, \alpha)}{\alpha} \qquad (8.13)$$

Hence, $k(E, \alpha)$ must have the form

$$k(E, \alpha) = c(E)\alpha, \tag{8.14}$$

where $c(E)$ is generally a function of E. Since $k(E, \alpha)$ is real, so is $c(E)$.

We now show that we can choose a set of kets such that $c(E) = 0$. Let us define $|E, j, m, \alpha)$ by

$$|E, j, m, \alpha) = e^{i\beta(E)}\,|E, j, m, \alpha\rangle, \tag{8.15}$$

where

$$(d/dE)\beta(E) = c(E)/p. \tag{8.16}$$

In terms of this new basis all of the previous formulas hold with $c(E) = 0$.

We have thus completed our derivation of the form of the infinitesimal generators of inhomogeneous Lorentz group in an angular momentum basis.

It should be noted that the case of nonzero mass reduced to the case of zero mass when $\mu = 0$.

To summarize: we have given a set of kets $|E, j, m, \alpha\rangle$ which satisfy the completeness relation (3.47) with $W(p) = 1/p$. The operators H, J_i, P_i act on these kets in the manner shown by (3.46) where $T_i(\alpha)$ is given by (3.42). The operators ϑ_i act in the manner given by (3.46)–(3.48) where the matrices $K_{\alpha,\alpha'}(E, j)$ and $Q_{i:\alpha,\alpha'}(E)$ are given by (8.7) to (8.11).

To obtain the form of the operators as given in Part I we first choose a new basis in which the kets denoted by $|E, j, m, \alpha)$ are given by

$$|E, j, m, \alpha) = \left[\frac{(j-m)!}{(j+m)!}\frac{1}{(2j)!}\right]^{\frac{1}{2}} |E, j, m, \alpha\rangle. \tag{8.17}$$

The new kets satisfy the completeness relation

$$\sum_{j,m,\alpha} \int dE\, |E, j, m, \alpha)\, \frac{1}{p}\, (E, j, m, \alpha| = I. \tag{8.18}$$

It is an easy enough matter to find how the infinitesimal generators act in the new basis. Finally, we introduce functions $\varphi(E, j, m, \alpha)$ in Hilbert space defined by

$$\varphi(E, j, m, \alpha) = (E, j, m, \alpha \mid \Phi\rangle,$$

where $|\Phi\rangle$ is an abstract vector. The operators as given in Part I act on the functions $\varphi(E, j, m, \alpha)$ rather than the basis vectors. We refrain from describing the transcription of the operators to function space, since it is obvious.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 10 OCTOBER 1964

Fundamental Properties of Perturbation-Theoretical Integral Representations. III*

NOBORU NAKANISHI
Applied Mathematics Department, Brookhaven National Laboratory, Upton, New York
(Received 13 March 1964)

Asymptotic behavior and subtraction problem of the perturbation-theoretical integral representation (PTIR) are investigated in detail. Six theorems are rigorously proved in this connection. It is shown that a function represented by an unsubtracted PTIR may asymptotically increase in particular directions. The relation between the asymptotic behavior and the subtraction number is clarified for the subtracted PTIR. As a by-product one obtains a consistent definition of a finite part of the integral involving $x^{-1}\theta(x)$.

I. INTRODUCTION

ANALYTICITY and uniqueness properties of perturbation-theoretical integral representations (PTIR) were investigated in detail in our previous papers.[1,2] But asymptotic property of PTIR was studied only for some subseries of perturbation expansion,[2] and no general discussions are made on this subject as yet. Recently, PTIR has been applied to the investigation of the high-energy behavior of the scattering amplitude in the Bethe-Salpeter formalism,[3] and we have found that the *unsubtracted* PTIR can describe the Regge behavior and more general high-energy behaviors. This is an interesting feature of PTIR, which is not known in the conventional dispersion theory. It will, therefore, be desirable to explore the asymptotic behavior and subtraction problem of PTIR.

For simplicity, we consider the two-variable PTIR only:

$$f(s,\,t) = \int_0^1 dz \int_0^\infty d\alpha \,\frac{\rho(z,\,\alpha)}{\alpha - zs - (1-z)t}. \tag{1.1}$$

Here $\int_0^1 dz \int_0^\infty d\alpha$ should be understood as $\int_{-\infty}^\infty dz \int_{-\infty}^\infty d\alpha$, with an integrand having a support $\{0 \le z \le 1,\ \alpha \ge 0\}$. If (1.1) is not convergent at $\alpha = \infty$, it is necessary to make subtractions. From the knowledge of the conventional dispersion theory, one might make the following subtraction:

$$\frac{f(s,\,t) - f(s,\,t_0)}{t - t_0} = \int_0^1 dz \int_0^\infty d\alpha \,\frac{\tilde\rho(z,\,\alpha)}{\alpha - zs - (1-z)t}, \tag{1.2}$$

with $t_0 < 0$. But if one applies this subtraction procedure to a function for which (1.1) converges, then one finds that $\tilde\rho(z,\,\alpha)$ is generally completely different from $\rho(z,\,\alpha)$. Indeed, an elementary calculation shows

$$\tilde\rho(z,\,\alpha) = \int_0^z dx \int_0^\infty d\beta\, x\rho(x,\,\beta) \times \delta'(x(\alpha - t_0) - z(\beta - t_0)). \tag{1.3}$$

Since the invariance of the weight function is important in the subtraction procedure, (1.2) should not be accepted as a good one. Hence, we shall merely subtract a constant term $f(s_0,\,t_0)$ instead of a one-variable function. In general, we have the following subtracted PTIR[1]:

$$f(s,\,t) = P(s,\,t) + \int_0^1 dz \int_0^\infty d\alpha \left[\frac{z(s-s_0) + (1-z)(t-t_0)}{\alpha - zs_0 - (1-z)t_0}\right]^N \times \frac{\rho(z,\,\alpha)}{\alpha - zs - (1-z)t}, \tag{1.4}$$

where $P(s,\,t)$ is an $(N-1)$th-order polynomial of s and t, $(s_0,\,t_0)$ being a fixed point in the analyticity domain of PTIR, i.e., $(s_0,\,t_0) \in D_{st}$. Then the following questions will arise. Is (1.4) general enough to represent any function holomorphic in D_{st} and bounded by a polynomial of $|s|$ and $|t|$ at infinity? If the subtracted PTIR is possible, how can we determine the subtraction number N? To answer these questions was the motivation of the present work.

In the next section, we give six theorems which can be proved rigorously. The proof of each theorem is given in the Appendix of the same number. In Sec. 3, we discuss the singularity in z of the weight function in the unsubtracted PTIR. In Sec. 4, we consider the subtracted PTIR, and answer the questions of last paragraph with a reasonable degree of confidence. The final section is devoted to another

* This work was performed under the auspices of the U. S. Atomic Energy Commission.
[1] N. Nakanishi, Phys. Rev. **127**, 1380 (1962).
[2] N. Nakanishi, J. Math. Phys. **4**, 1385 (1963).
[3] N. Nakanishi, Phys. Rev. **133**, B214 (1964); **133**, B1224 (1964).

topic concerning a new consistent definition of the finite part of $x^{-1}\theta(x)$, which is a by-product of the consideration in Sec. 4. In Appendix VII, we derive various formulas given in concrete examples of the text.

Throughout this paper, n, m, N stand for zero or positive integers, and A, B, C, M, R denote big positive numbers, whereas ϵ, δ, γ, σ, η, etc., usually represent small positive numbers. While s, t, $\hat{s}$, $\hat{t}$, $\bar{s}$, $\bar{t}$ are complex variables, s', t', s'', t'' are used only as real variables.

II. THEOREMS

The first aim of this section is to investigate the asymptotic property of PTIR under certain assumptions. First, we consider (1.1). The convergence of the integral in the right-hand side is, of course, implicitly assumed in (1.1). Since it is too difficult to deal with the general case, we shall assume that $\rho(z, \alpha)$ decreases at least like $\alpha^{-\delta}(\delta > 0)$ at infinity, and similarly that $\rho(z, \alpha)$ has a finite number of singularities in z of order $z^{-1+\sigma}(\sigma > 0)$ at worst. For finite values of α, $\rho(z, \alpha)$ will be a distribution of α, which may be dependent on z as is suggested by the following simple example.

Example 1.

$$f(s, t) = (a - s)^{-1}(b - t)^{-1} \quad (a \geq 0,\, b \geq 0). \tag{2.1}$$

Its weight function is, of course, given by

$$\rho(z, \alpha) = \delta'(\alpha - za - (1 - z)b). \tag{2.2}$$

It will be natural to expect in a naïve sense that if the above conditions on $\rho(z, \alpha)$ are satisfied then the function $f(s, t)$ given by (1.1) vanishes at infinity in D_{st}. But if we want to *prove* this statement, we must express it in a mathematically well-defined manner. Care must be taken in the definition of the asymptotic region because, for example, no matter how large $|s|$ may be, the function (2.1) does not become small if t approaches b. That is, in general, we should avoid considering the asymptotic behavior in a neighborhood of an unbounded singularity. Now, we obtain the following theorem.

Theorem I. Let $f(s, t)$ be an analytic function which can be represented as (1.1), where the weight function $\rho(z, \alpha)$ has the following properties. There exists a function of z,

$$H(z) \equiv \prod_{i=1}^{m} |z - z_i|^{1-\sigma} \quad (0 < \sigma < 1,\, 0 \leq z_i \leq 1), \tag{2.3}$$

such that

$$\hat{\rho}(z, \alpha) \equiv H(z)\rho(z, \alpha) \tag{2.4}$$

satisfies the following conditions.

(i) There exists an integer n such that

$$\hat{\rho}(z, \alpha) = (\partial/\partial\alpha)^n \varphi(z, \alpha), \tag{2.5}$$

where $\varphi(z, \alpha)$ is a continuous function of z and α.

(ii) There exists a positive number R (z independent) such that when $\alpha > R$, $\hat{\rho}(z, \alpha)$ is a *function* of α and satisfies the following inequalities:

(a) $$|\hat{\rho}(z, \alpha)| < A\alpha^{-\delta}, \tag{2.6}$$

(b) $$|\hat{\rho}(z, \alpha + \Delta\alpha) - \hat{\rho}(z, \alpha)| < B\,|\Delta\alpha|^{\mu} \quad \text{for} \quad |\Delta\alpha| \leq \kappa, \tag{2.7}$$

where δ, μ, κ, A, B are positive constants.

Then, $f(s, t)$ has the following properties:

(A) For any *closed* subset K of D_{st}, we can always find positive numbers γ, C, M such that

$$|f(s, t)| < C(|s| + |t|)^{-\gamma}, \tag{2.8}$$

whenever $|s| + |t| > M$ and $(s, t) \in K$.

(B) We can always find a positive number M such that if for any z satisfying $0 \leq z \leq 1$ s and t satisfy the inequality

$$|zs + (1 - z)t| > M \tag{2.9}$$

and if $(s, t) \in D_{st}$, then (2.8) holds.

In the above theorem, the condition (2.7) is called the Hölder condition or the Lipschitz condition of order μ. This assumption is usually necessary if one wants to discuss bounds of a singular integral. The main result in the theorem is, of course, the property (A). The property (B) is a special consequence of the assumption (2.6). For instance, the following example does not have the property (B).

Example 2.

$$f(s, t) = (-t)^{-1} \sum_{n=0}^{\infty} (n^2 - s)^{-1}, \tag{2.10}$$

whose weight function is

$$\rho(z, \alpha) = \sum_{n=0}^{\infty} \delta'(\alpha - zn^2). \tag{2.11}$$

Theorem I can easily be generalized to the case of the subtracted PTIR.

Theorem II. If $f(s, t)$ is represented as (1.4) instead of (1.1), and if (2.6) is replaced by the condition (i) (a')

$$|\hat{\rho}(z, \alpha)| < A\alpha^{N-\delta}, \tag{2.12}$$

then one has

$$|f(s, t)| < C(|s| + |t|)^{N-\gamma} \tag{2.13}$$

instead of (2.8).

From Example 1, we can expect that the unboundedness of $f(s, t)$ at the boundary of D_{st} is generally caused by singularities in α of $\rho(z, \alpha)$. Indeed, this is true, namely, we have the following theorem.

Theorem III. If $f(s, t)$ is represented as (1.4), and if $\hat{\rho}(z, \alpha)$ defined by (2.4) is a continuous function satisfying (2.12) and (2.7) for any $\alpha \geq 0$, then we can always find positive numbers γ and C such that

$$|f(s, t)| < C(1 + |s| + |t|)^{N-\gamma} \tag{2.14}$$

in the *whole* D_{st}.

In the above theorem it should be remarked that the Hölder condition (2.7) is required also for $\alpha = 0$ and $\Delta\alpha < 0$ with the convention

$$\rho(z, \beta) = 0 \quad \text{for} \quad \beta < 0. \tag{2.15}$$

Now, our next task is to consider the inverse problem of Theorem I. Namely, we want to find PTIR for a given function $f(s, t)$ holomorphic in D_{st} and bounded at infinity. This problem was discussed already twice,[1,2] but the proof of the theorem was still incomplete.[4]

As was noticed previously,[2] it is not necessary to assume that $f(s, t)$ be holomorphic in the whole D_{st}. This is because we have the following theorem, which is essentially due to Glaser.[5]

Theorem IV. Let $f(s, t)$ be holomorphic in domains D_+ and D_- separately, and both boundary values on E coincide with each other, where

$$\begin{aligned} D_+ &\equiv \{s, t; \operatorname{Im} s > 0, \operatorname{Im} t > 0\}, \\ D_- &\equiv \{s, t; \operatorname{Im} s < 0, \operatorname{Im} t < 0\}, \\ E &\equiv \{s, t; \operatorname{Im} s = \operatorname{Im} t = 0, \\ &\qquad \operatorname{Re} s < 0, \operatorname{Re} t < 0\}. \end{aligned} \tag{2.16}$$

If $|f(s, t)|$ is bounded by a polynomial of $|s|$ and $|t|$ in any closed subset of D_+ and D_-, then $f(s, t)$ is holomorphic in D_{st}.

For completeness, we write here the explicit shape of the domain D_{st}.[1,2] Let

$$\begin{aligned} D^*[s, t] &= \{s, t; \operatorname{Im} s > 0, \operatorname{Im} t < 0, \operatorname{Im} st^* \geq 0\}, \\ D^*[s] &= \{s, t; \operatorname{Im} s = 0, \operatorname{Re} s \geq 0\}. \end{aligned} \tag{2.17}$$

Then D_{st} is the complement of

$$D^*[s, t] \cup D^*[t, s] \cup D^*[s] \cup D^*[t]. \tag{2.18}$$

Hence, of course, D_{st} includes D_+, D_-, and E. D_{st} is the envelope of holomorphy of $D_+ \cup D_- \cup E$.

Now, our main theorem is as follows.

Theorem V. Let $f(s, t)$ be holomorphic in domains D_+ and D_- separately, and both boundary values on E coincide with each other. Moreover, assume that $f(s, t)$ satisfies the following boundedness conditions.

(i) For any closed subset K of D_+ and D_-, there exist positive numbers δ, A, M such that

$$|f(s, t)| < A(|s| + |t|)^{-\delta}, \tag{2.19}$$

whenever $|s| + |t| > M$ and $(s, t) \in K$.

(ii) For the same K, one has

$$|(\partial/\partial t)f(s, t)| < B(|s| + |t|)^{-\gamma} \sum_{i=1}^{m} |z_i s + (1 - z_i)t|^{-1}, \tag{2.20}$$

whenever $|s| + |t| > M$ in K, where $\gamma > 0$, $B > 0$ and $0 \leq z_i < 1$ $(i = 1, 2, \cdots, m)$.

Then $f(s, t)$ can be represented as (1.1), where $\rho(z, \alpha)$ is defined for every z except for $z_1, z_2, \cdots, z_m$, and 1. For a fixed z, $\rho(z, \alpha)$ is a distribution of α, which is given by

$$\rho(z, \alpha) = (2\pi i)^{-1} \lim_{\epsilon \to 0+} [\psi(z, \alpha + i\epsilon) - \psi(z, \alpha - i\epsilon)]. \tag{2.21}$$

Here $\psi(z, w)$ is a holomorphic function of w except for $w \geq 0$, and we have for $w < 0$

$$\psi(z, w) = (1 - z)^{-1} \int_0^\infty ds' \frac{\partial}{\partial w} f_s\left(s', \frac{w - zs'}{1 - z}\right), \tag{2.22}$$

where $f_s(s', t)$ is the absorptive part of $f(s, t)$, i.e.,

$$f_s(s', t) \equiv (2\pi i)^{-1} \lim_{\epsilon \to 0+} [f(s' + i\epsilon, t) - f(s' - i\epsilon, t)]. \tag{2.23}$$

In the above theorem, the boundedness condition (i) is essentially equivalent to the result (A) of Theorem I. We may conjecture that the condition

---

[4] The proof given in Ref. 2 is incomplete in the following respects. (i) The integral (2.11) of Ref. 2 is not well defined because the contours necessarily pass through some zero points of the denominator of the integrand. (ii) $(\partial/\partial t)f(s, t)$ does not necessarily behave like $O(|t|^{-1-\delta})$ at infinity. (iii) The analytic continuation of $g(w, z)$ to the lower half-plane of w is not good because then some part of the s' contour belongs to the singularity region.

[5] J. Bros and V. Glaser (unpublished); A. Bottino, A. M. Longoni, and T. Regge, Nuovo Cimento **23**, 954 (1962). The proof given in the latter paper is incomplete because the ξ' contour crosses the singularity region of the integrand at $\xi' = 0$ and on the large semicircle. To avoid this difficulty, it is essential to use the edge-of-the-wedge theorem as is done in our proof (see Appendix IV). The equivalence of D_{st} and the envelope of holomorphy of $D_+ \cup D_- \cup E$ was first pointed out in Ref. 2.

(i) will be enough to give the essential results of Theorem V because we know no counterexample to this statement, but it seems to be technically extremely difficult to eliminate a condition on a partial derivative of $f(s, t)$.[6]

One might suppose that the right-hand side of (2.20) may be replaced by $B(|s| + |t|)^{-\gamma} |t|^{-1}$, but this bound is not general enough to admit a simple example $(-s)^{-\frac{1}{2}}(-s - t)^{-\frac{1}{2}}$. It should be remarked that z_i cannot be equal to unity in (2.20). We have, of course, some examples which satisfy (i) but not (ii).

Example 3.

$$f(s, t) = (-s)^{-\frac{1}{2}}(-t)^{-\frac{1}{2}} \times \exp[-(-t)^{\frac{1}{2}}], \qquad [\operatorname{Re}(-t)^{\frac{1}{2}} \geq 0]. \quad (2.24)$$

This function does not satisfy the condition (ii), but it still has the representation (1.1) with a weight function satisfying all the conditions of Theorem I.

Since the conditions of Theorem V are imposed only on (s, t) belonging to K, $\rho(z, \alpha)$ is not necessarily bounded by $\alpha^{-\sigma}$ $(\sigma > 0)$. For instance, Example 2 satisfies all the conditions of Theorem V. It seems to be very difficult to prove a fairly general theorem which gives the boundedness of $\rho(z, \alpha)$ at $\alpha = \infty$. The following theorem is too restrictive to be practical.

Theorem VI. If the condition (ii) of Theorem V is replaced by the stronger condition that

$$\text{(ii}') \quad |(\partial/\partial t)f(s, t)| < B(1 + |s| + |t|)^{-1-\gamma} \quad (2.25)$$

in the *whole* D_+ and D_-, then we have

$$|\rho(z, \alpha)| < C(1 + \alpha)^{-\sigma} \quad (2.26)$$

for $z \neq 1$, where $0 < \sigma < \gamma$ and C is a big positive number.

Both Theorems V and VI concern the unsubtracted PTIR. The extension to the subtracted PTIR is by no means trivial. The reason why it is difficult will be clarified in Sec. IV.

III. SINGULARITIES IN z

In the preceding section, we have assumed that singularities in z of $\rho(z, \alpha)$ are integrable ones in the usual sense. But this restriction is too stringent for practical applications, and we should admit for $\rho(z, \alpha)$ to include a distribution of z independent of α. For example, if $f(s, t)$ is independent of s, $\rho(z,\alpha)$ is necessarily proportional to $\delta(z)$.

When $\rho(z, \alpha)$ contains such a distribution of z, the asymptotic behavior (2.8) is no longer assured. The following examples will be instructive.

Example 4. If

$$\rho(z, \alpha) = \delta(\alpha - a)\delta^{(n)}(z - z_0), \quad (3.1)$$

with $a \geq 0$ and $0 \leq z_0 \leq 1$, then

$$f(s, t) = \frac{n!(t - s)^n}{[a - z_0 s - (1 - z_0)t]^{n+1}}. \quad (3.2)$$

Example 5. The weight function

$$\rho(z, \alpha) = \alpha^{-\sigma}\delta^{(n)}(z - z_0), \quad (3.3)$$

with $0 < \operatorname{Re} \sigma < 1$ and $0 \leq z_0 \leq 1$ corresponds to

$$f(s, t) = \Gamma(1 - \sigma)\Gamma(n + \sigma) \frac{(t - s)^n}{[-z_0 s - (1 - z_0)t]^{n+\sigma}}. \quad (3.4)$$

From the above examples we see that if $\rho(z, \alpha)$ contains an $(n + 1)$th order singularity at $z = z_0$, then $|f(s, t)|$ can increase as $(|s| + |t|)^n$ *only when* one goes to infinity in the direction in which $|z_0 s + (1 - z_0)t|$ remains small. The purpose of this section is to show that the above statement is generally true, but no claim of rigor is made for the reasoning in this section.

In what follows it is very important to consider the following distribution introduced by Schwartz[7]:

$$Y_\lambda(x) = [\Gamma(\lambda)]^{-1} \operatorname{Pf} x^{\lambda-1}\theta(x) \quad \text{for } \lambda \neq 0, -1, -2, \cdots,$$
$$= \delta^{(n)}(x) \quad \text{for } \lambda = -n. \quad (3.5)$$

If $\varphi(x)$ is a test function, $\int Y_\lambda(x)\varphi(x)\,dx$ is an entire function of a complex parameter λ. $Y_\lambda(x)$ can be understood as the discontinuity of an analytic function $\Gamma(1 - \lambda)(-x)^{\lambda-1}$. Thus a δ function and its derivatives can be regarded as special cases of a power of x.

For simplicity, we consider the singularity at $z=0$. In this case we expect a special asymptotic behavior of $f(s, t)$ when $s \to \infty$ but $t/s \to 0$. In a neighborhood of $z = 0$ the weight function may be written as

$$\rho(z, \alpha) \simeq F(z)\varphi(\alpha) \quad \text{at} \quad z \approx 0, \quad (3.6)$$

where the symbol $\simeq$ means that the ratio of both

[6] It is generally impossible to deduce $f'(x) = O(x^{-1-\delta})$, $(x > 0, \delta > 0)$, from $f(x) = O(x^{-\delta})$ even if $f(x)$ is holomorphic on the positive real axis, because, for example, $f(x) = x^{-\delta}\sin x$. The author is much indebted to Dr. Pincus and Dr. Marr for valuable discussions on the asymptotic behavior of a derivative.

[7] L. Schwartz, *Théorie des Distributions* (Hermann & Cie., Paris, 1950), Chap. II. The symbol Pf denotes Hadamard's finite part, which means to discard the divergent part of the integral in a consistent manner.

sides tends to unity. Then the behavior for $s \to \infty$ but $t/s \to 0$ is determined by

$$I \equiv \int_0^1 dz \frac{F(z)}{\beta - zs}, \tag{3.7}$$

where $\beta \equiv \alpha - t$. For practical applications, the following case is important and seems to be sufficiently general:

$$F(z) = c \operatorname{Pf} [z^{\lambda-1}(\log 1/z)^{\nu}(\log \log 1/z)^{\xi} \cdots]. \tag{3.8}$$

For simplicity, we consider the case

$$F(z) = \operatorname{Pf} [z^{\lambda-1}(\log 1/z)^{\nu}], \tag{3.9}$$

since the extension to the general case is straightforward. Putting $s = -re^{i\theta}$, $(r > 0, |\theta| < \pi)$ and $z = y/r$, we have

$$I = r^{-\lambda} \int_0^r dy \frac{\operatorname{Pf} [y^{\lambda-1}(\log r - \log y)^{\nu}]}{\beta + ye^{i\theta}}. \tag{3.10}$$

As is easily seen by a binomial expansion, the leading term of the numerator is $(\log r)^{\nu}$ because the singularity of $\log y$ at $y = 0$ does not lead to infinity for $r \to \infty$. Thus one gets

$$I \simeq r^{-\lambda}(\log r)^{\nu} J, \tag{3.11}$$

with

$$J \equiv \int_0^{\infty} dy \frac{\operatorname{Pf} [y^{\lambda-1}]}{\beta + ye^{i\theta}} = \frac{\pi}{\sin \pi\lambda} \beta^{\lambda-1} e^{-i\lambda\theta}, \tag{3.12}$$

When the integration of (3.12) is carried out, we first assume $0 < \operatorname{Re} \lambda < 1$, and then analytically continue the result with respect to λ. Hence, (3.12) is true even for $\operatorname{Re} \lambda \leq 0$. Thus we have

$$I \simeq (\pi/\sin \pi\lambda)\beta^{\lambda-1}(-s)^{-\lambda}[\log (-s)]^{\nu}. \tag{3.13}$$

Putting $\lambda = -\mu$, $(\operatorname{Re} \mu \geq 0)$, we obtain the following asymptotic behavior of $f(s, t)$:

$$f(s, t) \simeq \int_0^{\infty} d\alpha \frac{\pi}{\sin \pi(\mu + 1)} \cdot \frac{\varphi(\alpha)}{(\alpha - t)^{\mu+1}} (-s)^{\mu}[\log (-s)]^{\nu}. \tag{3.14}$$

Now, in the Bethe–Salpeter approach to the high-energy behavior of the scattering amplitude,[3] the momentum transfer in the crossed channel was treated as a parameter, hence μ and ν may be functions of the momentum transfer. Thus we were able to describe the Regge and Regge-cut behaviors by PTIR. However, if we consider the S-matrix theory and, hence, t is identified with the momentum transfer, μ and ν in (3.14) cannot be functions of the momentum transfer. In this case, μ and ν should be considered as functions of α. But, unfortunately, we cannot carry out the integration over α if μ and ν are dependent on α.

The special asymptotic behavior (3.14) is, of course, due to the choice of the special direction $t/s \to 0$. Our next task is to see that if the asymptotic behavior is considered in any *other* direction then $f(s, t)$ satisfies (2.8), provided that $\rho(z, \alpha)$ satisfies all the conditions of Theorem I except for $z = 0$ and

$$z^{n+1-\delta}\rho(z, \alpha) = 0 \quad \text{for} \quad z = 0, \quad (\delta > 0). \tag{3.15}$$

First, we shall show an inequality

$$\left| \int_0^1 dz \int_0^{\infty} d\alpha \frac{\rho(z, \alpha)}{\alpha - zs - (1 - z)t} \left(\frac{\alpha - t}{\alpha + a}\right)^n \right| < A(|s| + |t|)^{n-\gamma}, \tag{3.16}$$

for $|s| + |t| > M$ with $0 < \gamma < \delta$ and $a > 0$. Without loss of generality, a smaller value than unity can be taken as the upper limit of the z integral. Then the integrand can be rewritten as

$$\frac{1}{(1 - z)^n} \frac{\rho(z, \alpha)}{(\alpha + a)^n} \left[P(z(\alpha - s), (1 - z)(\alpha - t)) + \frac{(-z)^n(\alpha - s)^n}{\alpha - zs - (1 - z)t} \right], \tag{3.17}$$

where $P(x, y)$ is a certain $(n - 1)$th order polynomial of x and y. The integral coming out from this polynomial part is, of course, convergent in the sense of a distribution, and gives an $(n - 1)$th-order polynomial of s and t. As for the second term of (3.17), because of the assumption (3.15) the weight functions

$$(-1)^k(1 - z)^{-n} {}_nC_k\alpha^k(\alpha + a)^{-n}z^n\rho(z, \alpha), \quad (k = 0, 1, \cdots, n), \tag{3.18}$$

satisfy all the conditions of Theorem I. Thus we have established (3.16).

If we consider a special case

$$\rho(z, \alpha) = F(z)\delta(\alpha - \alpha_0), \qquad (\alpha_0 \geq 0), \tag{3.19}$$

(3.16) leads to

$$\left| \int_0^1 dz \int_0^{\infty} d\alpha \frac{\rho(z, \alpha)}{\alpha - zs - (1 - z)t} \right| < A' \frac{(|s| + |t|)^{n-\gamma}}{|t|^n}, \tag{3.20}$$

for $|s| + |t| > M$ and $|t| > R$ with $R \geq 2\alpha_0$. Next, we consider the case in which $\rho(z, \alpha)$ is an integrable function of α and

$$\rho(z, \alpha) = 0 \quad \text{for} \quad \alpha > R. \tag{3.21}$$

Such a function can be written as

$$\rho(z, \alpha) = \lim_{N\to\infty} N^{-1} \sum_{k=0}^{N} \rho(z, \alpha)\delta(\alpha - N^{-1}kR), \quad (3.22)$$

just as done in the definition of the Riemann integral. Using (3.20) with (3.19) and taking $M > 4R$, we obtain (3.20) for the present $\rho(z, \alpha)$. Thus the expected result has been established to a certain extent. As is indicated by Example 5, the contribution from $\alpha > R$ in the general case will not be important. Summarizing the above investigation, we have the following statement.

Conjecture I. Let $f(s, t)$ be a function holomorphic in D_{st}. If for any closed subset K of D_+ and D_-, whenever $|s| + |t| > M$, the inequality

$$|f(s, t)| < A(|s| + |t|)^{-\delta} \times \left[1 + \sum_{i=1}^{m} \left(\frac{|s| + |t|}{|z_i s + (1 - z_i)t|}\right)^{\lambda_i}\right], \quad (3.23)$$

with $\delta > 0$, $0 \leq z_i \leq 1$, and $\lambda_i \geq 0$ $(i = 1, 2, \cdots, m)$ holds, then $f(s, t)$ can be represented as (1.1), and in a neighborhood of $z = z_i$, $\rho(z, \alpha)$ is defined as a distribution of z and the order of the singularity does not exceed

$$|z - z_i|^{-\lambda_i - 1 + \delta} \quad (3.24)$$

apart from logarithmic factors. Conversely, if in (1.1) $\rho(z, \alpha)$ has singularities at $z = z_i$ $(i = 1, 2, \cdots, m)$ of order (3.24) and satisfies all the conditions of Theorem I in all other points, then $f(s, t)$ satisfies the inequality (3.23).

IV. SUBTRACTED PTIR

The purpose of this section is to extend the conjecture made in Sec. III to the case of the subtracted PTIR (1.4). Since the general mathematical consideration is prohibitively difficult, we shall deduce the general conclusion from the previous results for the unsubtracted PTIR and some concrete examples of the subtracted PTIR.

Example 6.

$$f(s, t) = \int_0^1 dz \int_0^\infty d\alpha \, \frac{[zs + (1 - z)t]^N \alpha^\lambda \delta^{(n)}(z)}{\alpha^N[\alpha - zs - (1 - z)t]}. \quad (4.1)$$

This example is a generalization of Example 5. The integral is convergent for

$$N > \operatorname{Re} \lambda > N - 1,$$

and

$$n \geq N > \operatorname{Re} \lambda > -1, \quad (4.2)$$

and then

$$f(s, t) = \Gamma(\lambda + 1)\Gamma(n - \lambda)(t - s)^n(-t)^{\lambda - n}. \quad (4.3)$$

Example 7.

$$f(s, t) = (-s)^\mu(-t)^\nu. \quad (4.4)$$

Its weight function is given by

$$\rho(z, \alpha) = Y_{-\mu}(z)Y_{-\nu}(1 - z)Y_{\mu+\nu+1}(\alpha), \quad (4.5)$$

where $Y_\lambda(x)$ is defined by (3.5).

It will be natural to inquire whether or not the existence of distributional singularities in z (in the sense of Sec. III) can be predicted by the asymptotic behavior of $f(s, t)$. In Example 6, the special asymptotic behavior of order $|s|^n$ in the direction $t/s \to 0$ is certainly owing to the $(n + 1)$th order singularity at $z = 0$. But consider Example 7 with $\mu > 0$ and $\nu > 0$. The $(\mu + 1)$th-order singularity at $z = 0$ gives the asymptotic behavior of order $|s|^\mu$ in the direction $t/s \to 0$. However, the asymptotic behavior in the general direction

$$[z_0 s + (1 - z_0)t]/s \to 0, \qquad (0 < z_0 < 1) \quad (4.6)$$

is of order $|s|^{\mu+\nu}$, which is *stronger* than $|s|^\mu$. Therefore, if we consider a function

$$f(s, t) = (-s)^\mu(-t)^\nu + (-s)^{\mu+\nu}, \quad (4.7)$$

it exhibits no special asymptotic behavior in the direction $t/s \to 0$. Thus the answer to the above question is negative. Namely, *we cannot say anything about the nonexistence of distributional singularities in z of $\rho(z, \alpha)$ from the asymptotic behavior of $f(s, t)$ in the subtracted* PTIR. This means that the introduction of distributions of z is quite natural and inevitable in the subtracted PTIR, and this is the reason why it is difficult to extend the proof of Theorem V to the case of the subtracted PTIR.

Thus, we arrive at the following conjecture.

Conjecture II. Let $f(s, t)$ be a function holomorphic in D_{st}. If for any closed subset K of D_+ and D_-, whenever $|s| + |t| > M$, the inequality

$$|f(s, t)| < A(|s| + |t|)^\lambda \times \left[1 + \sum_{i=1}^{m} \left(\frac{|s| + |t|}{|z_i s + (1 - z_i)t|}\right)^{\lambda_i}\right], \quad (4.8)$$

with $\lambda < N$, $0 \leq z_i \leq 1$, and $\lambda_i > 0$ $(i = 1, 2, \cdots, m)$ holds, then $f(s, t)$ can be represented as (1.4), and the singularities in z of $\rho(z, \alpha)$ of *more than* $(\lambda + 1)$*th order* can be located only at $z = z_i$ with the order of at most $|z - z_i|^{-\lambda - \lambda_i - 1}$. The second statement of Conjecture I is likewise extended.

We can now answer the questions in Sec. I. The answer to the first question is "yes." The subtracted PTIR will be general enough to represent any reasonable function holomorphic in D_{st} if $\rho(z, \alpha)$ is a distribution of not only α but also z. The answer to the second question will be as follows. If for any (nonzero and nonnegative) complex number k one always has

$$|f(s, ks)| < A\, |s|^{\lambda}, \tag{4.9}$$

when $|s| > M$, where (s, ks) belongs to a closed subset K of D_+ and D_-, then N is determined as the minimal nonnegative integer greater than λ. Therefore, the number of subtractions in PTIR is usually *less* than (sometimes equal to) that in the double dispersion representation.

The weight function $\rho(z, \alpha)$ sometimes contains a new distribution which is not well known so far. The following example will be such an interesting one.

Example 8. If we operate $\partial^2/\partial\mu\partial\nu$ on the function of Example 7, we see that a function

$$f(s, t) = (-s)^{\mu} \log(-s)(-t)^{\nu} \log(-t) \tag{4.10}$$

has a weight function

$$\begin{aligned}\rho(z, \alpha) = {} & Y_{-\mu}(z) Y_{-\nu}(1-z) Y_{\mu+\nu+1}(\alpha)\{[\psi(-\mu) - \log z \\ & + \log\alpha - \psi(\mu+\nu+1)][\psi(-\nu) - \log(1-z) \\ & + \log\alpha - \psi(\mu+\nu+1)] - \psi'(\mu+\nu+1)\}.\end{aligned} \tag{4.11}$$

Here $\psi(x)$ stands for the polygamma function, i.e.,

$$\begin{aligned}\psi(x) &\equiv \Gamma'(x)/\Gamma(x) \\ &= -\gamma + \sum_{n=0}^{\infty} [(n+1)^{-1} - (x+n)^{-1}],\end{aligned} \tag{4.12}$$

where γ is Euler's constant. We are interested in the limit $\mu \to 0$ and $\nu \to 0$, namely, we want to find the weight function of

$$f(s, t) = \log(-s)\log(-t). \tag{4.13}$$

From (4.11), noticing

$$\begin{aligned}&\lim_{\mu\to 0} Y_{-\mu}(z)\psi(-\mu) \\ &\quad = -z^{-1} + \delta(z) \lim_{\mu\to 0} [\Gamma(-\mu) + \psi(-\mu)] \\ &\quad = -z^{-1} - 2\gamma\cdot\delta(z),\end{aligned} \tag{4.14}$$

we have

$$\begin{aligned}\rho(z, \alpha) = {} & z^{-1} + \delta(z)(\log z + \gamma - \log\alpha) + (1-z)^{-1} \\ & + \delta(1-z)[\log(1-z) + \gamma - \log\alpha].\end{aligned} \tag{4.15}$$

Unfortunately, (4.15) is not well defined at $z = 0$ and $z = 1$. But (4.15) suggests that the sum $z^{-1} + \delta(z) \log z$ will give a meaningful result. Indeed, by a direct calculation we can show

$$\begin{aligned}&\log(-s)\log(-t) \\ &= \int_0^1 dz \int_0^{\infty} d\alpha\, \frac{[zs + (1-z)t + 1]\rho(z, \alpha)}{(\alpha+1)[\alpha - zs - (1-z)t]},\end{aligned} \tag{4.16}$$

with

$$\begin{aligned}\rho(z, \alpha) = {} & \mathrm{Pf}\,[z^{-1}\theta(z)] + \mathrm{Pf}\,[(1-z)^{-1}\theta(1-z)] \\ & - [\delta(z) + \delta(1-z)]\log\alpha.\end{aligned} \tag{4.17}$$

Here, the distribution $\mathrm{Pf}\,[x^{-1}\theta(x)]$ is defined by

$$\begin{aligned}&\int dx\, \mathrm{Pf}\,[x^{-1}\theta(x)]\varphi(x) \\ &\equiv \lim_{\epsilon\to 0+} \int dx\, [x^{-1}\theta(x-\epsilon) + \delta(x-\epsilon)\log x]\varphi(x),\end{aligned} \tag{4.18}$$

where $\varphi(x)$ is a test function. Comparison of (4.15) with (4.17) leads to the identification

$$\mathrm{Pf}\,[z^{-1}\theta(z)] = z^{-1} + \delta(z)(\log z + \gamma) \quad (z \geq 0). \tag{4.19}$$

V. FINITE PART OF $x^{-n}\theta(x)$

In (4.18) we have defined a distribution $\mathrm{Pf}\,[x^{-1}\theta(x)]$. A similar distribution was introduced by Schwartz in his famous book.[7] But his definition of the finite part of $x^{-1}\theta(x)$ is simply to discard the logarithmically divergent part. As was noticed by himself, his definition has a serious difficulty, namely, it is not invariant under the transformation of the integration variable. For instance, according to his prescription, one has

$$\int_0^1 dx\, \mathrm{Pf}\,[x^{-1}\theta(x)]_{\mathrm{Schwartz}} = 0. \tag{5.1}$$

But if one puts $x = 2y$, then one obtains

$$\int_0^{\frac{1}{2}} dy\, \mathrm{Pf}\,[y^{-1}\theta(y)]_{\mathrm{Schwartz}} = -\log 2. \tag{5.2}$$

Thus the value of the integral changes. This is quite unsatisfactory, and his distribution cannot be used for practical calculations. On the other hand, our definition of $\mathrm{Pf}\,[x^{-1}\theta(x)]$ is free from this difficulty as is easily checked. Therefore, it will be desirable to investigate the properties of our distribution.

We define

$$\mathrm{Pf}\,[x^{-n}\theta(x)] \equiv \frac{(-1)^{n-1}}{(n-1)!}\frac{d^n}{dx^n}[\theta(x-\epsilon)\log x], \tag{5.3}$$

for any positive integer n, where $\epsilon \to 0+$ should be taken after the integration over x is carried out.

It is easy to see that

$$\mathrm{Pf}\, x^{-n} = \mathrm{Pf}\, [x^{-n}\theta(x)] + (-1)^n \mathrm{Pf}\, [(-x)^{-n}\theta(-x)], \tag{5.4}$$

where Pf x^{-1} is identical with Cauchy's principal part and Pf x^{-n} can be derived from it by successive differentiations (with an appropriate coefficient). Thus (5.3) is a natural generalization of Pf x^{-n}.

Let $\varphi(x)$ be a test function, which is, of course, an infinitely differentiable function. We can easily calculate the integral

$$F[\varphi] \equiv \int_0^a dx\, \mathrm{Pf}\, [x^{-n}\theta(x)]\varphi(x), \qquad (a > 0) \tag{5.5}$$

according to the definition (5.3). We obtain

$$F[\varphi] = [(n-1)!]^{-1}\left\{-\sum_{j=0}^{n-2}(n-j-2)!a^{-n+j+1}\varphi^{(j)}(a) + \varphi^{(n-1)}(a)\log a - \int_0^a dx\, \varphi^{(n)}(x)\log x\right\}. \tag{5.6}$$

Especially, for $n = 1$ we have

$$\int_0^a dx\, \mathrm{Pf}\, [x^{-1}\theta(x)]\varphi(x) = \varphi(a)\log a - \int_0^a dx\, \varphi'(x)\log x. \tag{5.7}$$

According to (5.3), if one puts $x = ky$, $(k > 0)$, then one obtains

$$\mathrm{Pf}\, [x^{-n}\theta(x)] = k^{-n}\,\mathrm{Pf}\, [y^{-n}\theta(y)] + (-1)^{n-1}[(n-1)!]^{-1}k^{-n}\log k\,\delta^{(n-1)}(y). \tag{5.8}$$

The appearance of an additional term guarantees the invariance under the transformation of the integration variable.

Finally, carrying out the differentiation in (5.3), we have

$$\mathrm{Pf}\, [x^{-n}\theta(x)] = x^{-n}\theta(x-\epsilon) + \sum_{j=1}^{n-1}(-1)^j n(n-j)^{-1}(j!)^{-1}x^{-n+j}\delta^{(j-1)}(x-\epsilon) + (-1)^{n-1}[(n-1)!]^{-1}\delta^{(n-1)}(x-\epsilon)\log x. \tag{5.9}$$

Hence, we have the multiplication law

$$x^\mu\, \mathrm{Pf}\, [x^{-n}\theta(x)] = x^{-n+\mu}\theta(x) \tag{5.10}$$

only for $\mu > n - 1$, but

$$x^\mu\, \mathrm{Pf}\, [x^{-n}\theta(x)] \neq \mathrm{Pf}\, [x^{-n+\mu}\theta(x)], \tag{5.11}$$

for $0 < \mu \leq n - 1$. For instance,

$$\int_0^1 dx\, \mathrm{Pf}\, [x^{-1}\theta(x)] = 0, \tag{5.12}$$

but

$$\int_0^1 dx\cdot x\, \mathrm{Pf}\, [x^{-2}\theta(x)] = -1. \tag{5.13}$$

The above definition may be useful for practical calculations. We can now consistently calculate the finite part of a logarithmically divergent integral. It might be particularly useful for the calculation of an infrared-divergent transition amplitude.

ACKNOWLEDGMENT

The author is very grateful to Dr. Y. Shimamoto for his kind interest.

APPENDIX I. PROOF OF THEOREM I

We shall first prove the following lemma. The method is an extension of that of Källén[8] and Frye and Warnock.[9]

Lemma 1. Let

$$F(w) \equiv \int_0^\infty d\alpha\, \frac{\rho(\alpha)}{\alpha - w}, \tag{A 1.1}$$

where $\rho(\alpha)$ has the following properties.

(i) There exists a continuous function $\varphi(\alpha)$ such that

$$\rho(\alpha) = \varphi^{(n)}(\alpha). \tag{A 1.2}$$

(ii) There exists a positive number R such that for $\alpha > R$ $\rho(\alpha)$ is a function of α satisfying the following conditions

(a) $$|\rho(\alpha)| < A\alpha^{-\delta}, \tag{A 1.3}$$

(b) $$|\rho(\alpha + \Delta\alpha) - \rho(\alpha)| < B\,|\Delta\alpha|^\mu \quad \text{for} \quad |\Delta\alpha| \leq \kappa, \tag{A 1.4}$$

where δ, μ, κ, A, B are positive constants.

Then we can always find positive numbers δ', C, M such that

$$|F(w)| < C\,|w|^{-\delta'}, \tag{A 1.5}$$

whenever $|w| > M$ except for the positive real axis.

Proof: From (A1.3) we have for $\alpha > R$

$$|\rho(\alpha + \Delta\alpha) - \rho(\alpha)| < 2A\alpha^{-\delta}, \tag{A 1.6}$$

provided that $\alpha \gg |\Delta\alpha|$. Hence, (A1.4) and (A1.6) yield

$$|\rho(\alpha + \Delta\alpha) - \rho(\alpha)| < (2A\alpha^{-\delta})^\nu (B\,|\Delta\alpha|^\mu)^{1-\nu} \tag{A 1.7}$$

[8] G. Källén, Kgl. Danske Videnskab. Selskab, Mat.-Fys. Medd. 27, No. 12 (1953), Appendix.

[9] G. Frye and R. L. Warnock, Phys. Rev. **130**, 478 (1963), Appendix A.

with $0 < \nu < 1$. We can, therefore, take

(b′) $|\rho(\alpha+\Delta\alpha)-\rho(\alpha)| < B'\alpha^{-\delta'}|\Delta\alpha|^{\mu'}$ for $|\Delta\alpha| \leq \kappa$ (A 1.8)

with $0 < \delta' < \delta$ and $\mu' > 0$ instead of (A1.4) without loss of generality.

Let $r \equiv |w| > 2R$ and $\kappa < R$. We divide the integral (A1.1) into five parts:

$$\int_0^\infty = \int_0^{\frac{1}{2}r} + \int_{\frac{1}{2}r}^{r-\kappa} + \int_{r-\kappa}^{r+\kappa} + \int_{r+\kappa}^{2r} + \int_{2r}^\infty. \quad \text{(A 1.9)}$$

(1°) We may assume $n \geq 1$ without loss of generality.

$$\left|\int_0^{\frac{1}{2}r}\right| = \left|\int_0^{\frac{1}{2}r} d\alpha \frac{\varphi^{(n)}(\alpha)}{\alpha - w}\right| \leq \sum_{j=1}^n (j-1)! \left|\frac{\varphi^{(n-j)}(r/2)}{(r/2-w)^j}\right| + n!\left|\int_0^{\frac{1}{2}r} d\alpha \frac{\varphi(\alpha)}{(\alpha-w)^{n+1}}\right|. \quad \text{(A 1.10)}$$

From (A1.3), for $\alpha > R$ there is a jth-order polynomial $P_j(\alpha)$, $(n \geq j \geq 1)$, such that

$$|\varphi^{(n-j)}(\alpha)| < P_j(\alpha)\alpha^{-\delta}. \quad \text{(A1.11)}$$

Hence,

$$\left|\frac{\varphi^{(n-j)}(r/2)}{(r/2-w)^j}\right| < \frac{P_j(r/2)(r/2)^{-\delta}}{(r/2)^j} = O(r^{-\delta}), \quad \text{(A 1.12)}$$

$$\left|\int_0^{\frac{1}{2}r} d\alpha \frac{\varphi(\alpha)}{(\alpha-w)^{n+1}}\right| \leq \frac{\max\limits_{0\leq\alpha\leq r/2} \varphi(\alpha)}{(\frac{1}{2}r)^n} = O(r^{-\delta}). \quad \text{(A 1.13)}$$

(2°) $$\left|\int_{\frac{1}{2}r}^{r-\kappa}\right| \leq \int_{\frac{1}{2}r}^{r-\kappa} d\alpha \frac{|\rho(\alpha)|}{|\alpha-w|} \leq A\int_{\frac{1}{2}r}^{r-\kappa} d\alpha \frac{\alpha^{-\delta}}{r-\alpha} \leq A(\tfrac{1}{2}r)^{-\delta}\log(r/2\kappa) = o(r^{-\delta'}). \quad \text{(A 1.14)}$$

(3°) $$\int_{r-\kappa}^{r+\kappa} = \int_{r-\kappa}^{r+\kappa} d\alpha \frac{\rho(\alpha)-\rho(r)}{\alpha-w} + \rho(r)\int_{r-\kappa}^{r+\kappa}\frac{d\alpha}{\alpha-w}. \quad \text{(A 1.15)}$$

Because of (A1.8) we have

$$\left|\int_{r-\kappa}^{r+\kappa} d\alpha \frac{\rho(\alpha)-\rho(r)}{\alpha-w}\right| \leq B'\int_{r-\kappa}^{r+\kappa} d\alpha \frac{\alpha^{-\delta'}|\alpha-r|^{\mu'}}{|\alpha-r|} \leq 2B'(r-\kappa)^{-\delta'}\int_0^\kappa \beta^{-1+\mu'}d\beta = 2B'\mu'^{-1}\kappa^{\mu'}(r-\kappa)^{-\delta'} = O(r^{-\delta'}). \quad \text{(A 1.16)}$$

$$\left|\rho(r)\int_{r-\kappa}^{r+\kappa}\frac{d\alpha}{\alpha-w}\right| \leq Ar^{-\delta}\left|\log\frac{r+\kappa-w}{r-\kappa-w}\right| = o(r^{-\delta'}). \quad \text{(A 1.17)}$$

(4°) $$\left|\int_{r+\kappa}^{2r}\right| \leq A\int_{r+\kappa}^{2r} d\alpha \frac{\alpha^{-\delta}}{\alpha-r} \leq A(r+\kappa)^{-\delta}\log(r/\kappa) = o(r^{-\delta'}). \quad \text{(A 1.18)}$$

(5°) $$\left|\int_{2r}^\infty\right| \leq A\int_{2r}^\infty d\alpha\frac{\alpha^{-\delta}}{\alpha-r} \leq A\int_R^\infty d\alpha \frac{\alpha^{-\delta}}{(\frac{1}{2}\alpha)^{1-\delta'}r^{\delta'}} = O(r^{-\delta'}). \quad \text{(A 1.19)}$$

Thus we obtain (A1.5). Q.E.D.

Now, we apply Lemma 1 to

$$F(z, w) \equiv \int_0^\infty d\alpha \frac{\hat{\rho}(z, \alpha)}{\alpha - w}. \quad \text{(A 1.20)}$$

Then there exist positive numbers δ', C_1, M_1 such that

$$|F(z, w)| < C_1|w|^{-\delta'}, \quad \text{(A 1.21)}$$

whenever $|w| > M_1$ except for the positive real axis. Let

$$K_w \equiv \{w;\, w = zs + (1-z)t,\ 0 \leq z \leq 1,\ (s, t) \in K\}. \quad \text{(A 1.22)}$$

Then K_w is a closed set which does not intersect $\{w \geq 0\}$. Since the intersection of K_w and $\{|w| \leq M_1\}$ is compact and $F(z, w)$ is holomorphic there, $F(z, w)$ is bounded there, i.e.,

$$|F(z, w)| < C_2. \quad \text{(A 1.23)}$$

Thus,

$$|F(z, w)| < C_0(M_1 + |w|)^{-\delta'} \quad \text{in } K_w, \quad \text{(A 1.24)}$$

where $C_0 \equiv \max(2^{\delta'}C_1, M_1^{\delta'}C_2)$. By definition, we have

$$f(s, t) = \int_0^1 dz\,[H(z)]^{-1}F(z, zs+(1-z)t). \quad \text{(A 1.25)}$$

Hence,

$$|f(s, t)| < C_0\int_0^1 dz\,[H(z)]^{-1}\{M_1 + |zs+(1-z)t|\}^{-\delta'}, \quad \text{(A 1.26)}$$

for $(s, t) \in K$. We divide the integration range $[0, 1]$ into $I[s, t]$ in which the inequality

$$|zs + (1-z)t| < \tfrac{1}{2}M^{\frac{1}{2}}(|s| + |t|)^{\frac{1}{2}} \quad \text{(A 1.27)}$$

holds and the remaining part. Then

$$|f(s, t)| < C_0M_1^{-\delta'}\int_{I[s,t]} dz\,[H(z)]^{-1} + C_0\int_0^1 dz\,[H(z)]^{-1}\{M_1 + \tfrac{1}{2}M^{\frac{1}{2}}(|s|+|t|)^{\frac{1}{2}}\}^{-\delta'}. \quad \text{(A 1.28)}$$

Since the z integral is convergent, the second term evidently behaves like $O((|s| + |t|)^{-\delta'/2})$. Therefore, the problem is to estimate the first term. For this purpose, we use the following lemma.

Lemma 2. Let s and t be complex, $M > 0$, and

$$I[s, t] \equiv \{z; 0 \leq z \leq 1, |zs + (1-z)t| < \tfrac{1}{2}M^{\frac{1}{2}}(|s| + |t|)^{\frac{1}{2}}\}. \quad \text{(A 1.29)}$$

We denote the Lebesgue measure of a set S by $\mu(S)$. Then

$$\mu(I[s, t]) < 4M^{\frac{1}{2}}(|s| + |t|)^{-\frac{1}{2}} \quad \text{for} \quad |s| + |t| > 4M. \quad \text{(A1.30)}$$

Proof: Putting $\xi = 2z - 1$, $u = s + t$, and $v = s - t$, we have

$$zs + (1-z)t = \tfrac{1}{2}(u + \xi v),$$

$$4M \leq |s| + |t| \leq |u| + |v|. \quad \text{(A 1.31)}$$

Let

$$J[s, t] \equiv \{\xi; |\xi| \leq 1, |u + \xi v| < M^{\frac{1}{2}}(|u| + |v|)^{\frac{1}{2}}\}. \quad \text{(A 1.32)}$$

Then it is evident that

$$2\mu(I[s, t]) \leq \mu(J[s, t]). \quad \text{(A 1.33)}$$

The inequality

$$|u + \xi v| < M^{\frac{1}{2}}(|u| + |v|)^{\frac{1}{2}} \quad \text{(A 1.34)}$$

can be rewritten as

$$|v|^2 \xi^2 + 2 \operatorname{Re}(uv^*)\xi + |u|^2 - M(|u| + |v|) < 0. \quad \text{(A 1.35)}$$

The discriminant D of this quadratic form is

$$D = M(|u| + |v|)\,|v|^2 - (\operatorname{Im} uv^*)^2. \quad \text{(A 1.36)}$$

For $D \geq 0$

$$\mu(J[s, t]) \leq 2D^{\frac{1}{2}}\,|v|^{-2}, \quad \text{(A 1.37)}$$

and for $D < 0$

$$\mu(J[s, t]) = 0. \quad \text{(A 1.38)}$$

Therefore, in general, we have

$$\mu(J[s, t]) \leq 2M^{\frac{1}{2}}(|u| + |v|)^{\frac{1}{2}}\,|v|^{-1}. \quad \text{(A 1.39)}$$

On the other hand, (A1.34) leads to

$$|u| - |\xi|\cdot|v| < M^{\frac{1}{2}}(|u| + |v|)^{\frac{1}{2}}. \quad \text{(A 1.40)}$$

Since $|\xi| \leq 1$, we have

$$|v| \geq |\xi|\cdot|v| > |u| - M^{\frac{1}{2}}(|u| + |v|)^{\frac{1}{2}}. \quad \text{(A 1.41)}$$

Adding $|v|$ to (A1.41) and dividing it by two, we obtain

$$|v| > \tfrac{1}{2}(|u| + |v|)^{\frac{1}{2}}[(|u| + |v|)^{\frac{1}{2}} - M^{\frac{1}{2}}]. \quad \text{(A 1.42)}$$

Substitution of (A1.42) in (A1.39) yields

$$\mu(J[s, t]) < 4M^{\frac{1}{2}}[(|u| + |v|)^{\frac{1}{2}} - M^{\frac{1}{2}}]^{-1}. \quad \text{(A 1.43)}$$

Hence, (A1.33) together with (A1.31) leads to

$$\mu(I[s, t]) < 2M^{\frac{1}{2}}[(|s| + |t|)^{\frac{1}{2}} - M^{\frac{1}{2}}]^{-1} \leq 4M^{\frac{1}{2}}(|s| + |t|)^{-\frac{1}{2}}. \quad \text{(A 1.44)}$$

Thus, we have proved (A1.30). Q.E.D.

Now, it is obvious that

$$[H(z)]^{-1} < a\left(1 + \sum_{i=1}^{m} |z - z_i|^{-1+\sigma}\right) \quad \text{(A 1.45)}$$

if a is sufficiently large. Hence, putting $\eta \equiv \mu(I[s, t])$, we have

$$\int_{I[s,t]} dz\,[H(z)]^{-1} < a\eta + a\left[\sum_{i=1}^{m} \int_{z_i-\eta}^{z_i+\eta} dz\,|z - z_i|^{-1+\sigma} + \eta\cdot\eta^{-1+\sigma}\right] = O(\eta^\sigma). \quad \text{(A 1.46)}$$

Lemma 2 tells us that

$$\eta < 4M^{\frac{1}{2}}(|s| + |t|)^{-\frac{1}{2}}. \quad \text{(A 1.47)}$$

Thus the first term of (A1.28) behaves like, at most, $(|s| + |t|)^{-\sigma/2}$. Then putting

$$\gamma \equiv \min(\delta'/2, \sigma/2), \quad \text{(A 1.48)}$$

we obtain the statement (A) of Theorem I. The statement (B) is likewise obtained by applying (A1.21) to (A1.25) directly ($M_1 = M$, $C_0 = 2^{\delta'}C_1$).

APPENDIX II. PROOF OF THEOREM II

Lemma 1 of Appendix I together with

$$\rho(\alpha) = \frac{\hat{\rho}(z, \alpha)}{[\alpha - zs_0 - (1-z)t_0]^N} \quad \text{(A 2.1)}$$

leads to

$$\left|\int_0^\infty d\alpha\,\frac{\hat{\rho}(z, \alpha)}{[\alpha - zs_0 - (1-z)t_0]^N[\alpha - zs - (1-z)t]}\right| < C_0[M_1 + |zs + (1-z)t|]^{-\delta'}. \quad \text{(A 2.2)}$$

Hence,

$$|f(s, t)| \leq |P(s, t)| + C_0 \sum_{i=0}^{N} {}_NC_i\,|s - s_0|^i\,|t - t_0|^{N-i}\,\varphi_i(s, t) \quad \text{(A 2.3)}$$

with

$$\varphi_i(s,t) \equiv \int_0^1 dz\cdot z^i(1-z)^{N-i}[H(z)]^{-1} \times [M_1 + |zs+(1-z)t|]^{-\delta'} \quad \text{(A 2.4)}$$

Since $|z^i(1-z)^{N-i}| \le 1$, we see

$$\varphi_i(s,t) = O((|s|+|t|)^{-\gamma}), \quad \text{(A 2.5)}$$

according to the proof of Theorem I. Thus

$$|f(s,t)| = O((|s|+|t|)^{N-\gamma}). \quad \text{(A 2.6)}$$

APPENDIX III. PROOF OF THEOREM III

If we can prove the following lemma, then the rest of the proof is equivalent to that of Theorem II.

Lemma 3. If for any $\alpha \ge 0$

$$|\rho(\alpha)| < A\alpha^{-\delta}, \quad \text{(A 3.1)}$$

$$|\rho(\alpha+\Delta\alpha)-\rho(\alpha)| < B\,|\Delta\alpha|^{\mu} \text{ for } |\Delta\alpha| \le \lambda, \quad \text{(A 3.2)}$$

where $1 > \delta > 0$, $\mu > 0$, and $\lambda > 0$, then we can always find positive numbers δ' and C such that

$$\left|\int_0^\infty d\alpha \frac{\rho(\alpha)}{\alpha - w}\right| < C(1+|w|)^{-\delta'} \quad \text{(A 3.3)}$$

for any nonzero and nonpositive w.

Proof: Let $\kappa = \frac{1}{3}\lambda$. In the case $r \equiv |w| \ge 2\kappa$, we have[10]

$$\left|\int_0^\infty\right| < C'r^{-\delta'} \quad \text{(A 3.4)}$$

immediately from the proof of Lemma 1 of Appendix I. In the case $r < 2\kappa$, (A3.2) yields

$$|\rho(\alpha)| < B\alpha^{\mu} \quad \text{for} \quad 0 \le \alpha \le 3\kappa \quad \text{(A 3.5)}$$

because of the convention (2.15). Hence,

$$\left|\int_0^{3\kappa}\right| \le B\int_0^{3\kappa} d\alpha\, |\alpha - r|^{-1+\mu} + \left|\rho(r)\int_0^{3\kappa}\frac{d\alpha}{\alpha - w}\right|$$

$$< B\mu^{-1}[r^{\mu} + (3\kappa - r)^{\mu}] + Br^{\mu}\left|\log\frac{3\kappa - w}{-w}\right|$$

$$< \text{constant}, \quad \text{(A 3.6)}$$

$$\left|\int_{3\kappa}^\infty\right| < A\int_{3\kappa}^\infty d\alpha \frac{\alpha^{-\delta}}{\alpha - 2\kappa} = \text{constant}. \quad \text{(A 3.7)}$$

Thus we obtain

$$\left|\int_0^\infty\right| \le C(1+r)^{-\delta'} \quad \text{(A 3.8)}$$

for any w except for $w \ge 0$. Q.E.D.

[10] Instead of the part (1°) of the proof of Lemma 1, we now have simply

$$\left|\int_0^{r/2}\right| < A(1-\delta)^{-1}(\tfrac{1}{2}r)^{-\delta},$$

assuming $\delta < 1$.

APPENDIX IV. PROOF OF THEOREM IV

By multiplying $f(s,t)$ by $(st)^{-N-\delta}$, $(\delta > 0)$, we can assume

$$|f(s,t)| < A(|s|+|t|)^{-\delta} \text{ for } |s|+|t| > M \quad \text{(A 4.1)}$$

without loss of generality.

According to the edge-of-the-wedge theorem,[11] $f(s,t)$ is a single analytic function holomorphic in $D_+ \cup D_- \cup \mathfrak{N}(E)$, where $\mathfrak{N}(E)$ stands for a complex neighborhood of E. For $(\hat{s},\hat{t}) \in D_+$ we consider

$$F(\hat{s},\hat{t},\xi) \equiv (2\pi i)^{-1}\int_{-\infty}^{\infty} d\xi' \frac{f(\xi'\hat{s}-\epsilon, \xi'\hat{t}-\epsilon)}{\xi' - \xi}, \quad \text{(A 4.2)}$$

where $\operatorname{Im} \xi > 0$, and ϵ is an infinitesimal positive constant. The right-hand side of (A4.2) is well defined because of the analyticity and the boundedness (A4.1) of $f(s,t)$.

If we take particular points $\arg \hat{s} = \arg \hat{t}$, then we can close the ξ' contour of (A4.2) by adding a large semicircle because $(\xi'\hat{s}-\epsilon, \xi'\hat{t}-\epsilon) \in D_+ \cup D_- \cup \mathfrak{N}(E)$, and we obtain

$$F(\hat{s},\hat{t},\xi) = f(\xi\hat{s}-\epsilon, \xi\hat{t}-\epsilon). \quad \text{(A 4.3)}$$

Because of the uniqueness of analytic continuation, (A4.3) tells us that $F(\hat{s},\hat{t},\xi)$ is the analytic extension of $f(\xi\hat{s}-\epsilon, \xi\hat{t}-\epsilon)$ to the topological product of D_+ and $\{\operatorname{Im}\xi > 0\}$. Thus $f(s,t)$ is holomorphic in

$$D' \equiv \{s,t; s=\xi\hat{s}, t=\xi\hat{t}, (\hat{s},\hat{t}) \in D_+, \operatorname{Im}\xi > 0\}. \quad \text{(A 4.4)}$$

We will show $D' = D_{st}$ in the following.

For simplicity, we write

$$\theta \equiv \arg s, \quad \varphi \equiv \arg t,$$
$$\hat\theta \equiv \arg \hat{s}, \quad \hat\varphi \equiv \arg \hat{t}, \quad \psi \equiv \arg \xi, \quad \text{(A 4.5)}$$

then

$$\theta = \hat\theta + \psi, \quad \varphi = \hat\varphi + \psi, \quad \text{(A 4.6)}$$

and

$$0 < \hat\theta < \pi, \quad 0 < \hat\varphi < \pi, \quad 0 < \psi < \pi. \quad \text{(A 4.7)}$$

Since D_{st} is explicitly given as the complement of (2.18), we compare D' with it in the following.

(1°) $D' \supset D_+$ and $D' \supset D_-$ are evident (the choices of ψ are $\psi \approx 0$ and $\psi \approx \pi$, respectively).

[11] H. J. Bremermann, R. Oehme, and J. G. Taylor, Phys. Rev. **109**, 2178 (1958); F. J. Dyson, *ibid.* **110**, 579 (1958).

(2^0) When $\operatorname{Im} s > 0$ and $\operatorname{Im} t < 0$, the points belonging to D_{st} are characterized by $\operatorname{Im} st^* < 0$. This condition can be rewritten as

$$0 < \theta < \pi < \varphi < 2\pi, \quad 0 < \varphi - \theta < \pi. \qquad \text{(A 4.8)}$$

On the other hand, if we choose ξ as

$$\varphi - \pi < \psi < \theta, \qquad |\xi| = 1, \qquad \text{(A 4.9)}$$

then we have

$$\hat{s} = |s|\, e^{i(\theta-\psi)}, \qquad \hat{t} = |t|\, e^{i(\varphi-\psi)}, \qquad \text{(A 4.10)}$$

hence $(\hat{s}, \hat{t}) \in D_+$. Thus the points of D_{st} belong to D'. Conversely, if $\operatorname{Im} st^* \geq 0$, i.e., $\varphi - \theta \geq \pi$, which in turn implies $\hat{\varphi} - \hat{\theta} \geq \pi$. This contradicts (A4.7). Thus both domains in this portion coincide with each other.

(3^0) When $\operatorname{Im} s < 0$ and $\operatorname{Im} t > 0$, the problem is reduced to the above case by interchanging s and t.

(4^0) When $\operatorname{Im} s = 0$, the points of D_{st} is characterized by $\operatorname{Re} s < 0$ with $\arg t \neq 0$. As for D', (A4.6) and (A4.7) imply $\arg s \neq 0$ and $\arg t \neq 0$, hence $\operatorname{Im} s = 0$ gives $\operatorname{Re} s < 0$.

(5^0) The case $\operatorname{Im} t = 0$ is similar to the above.

APPENDIX V. PROOF OF THEOREM V

We consider a point $(s, t) \in D_{st}$ such that

$$\operatorname{Im} s > \epsilon, \qquad \operatorname{Im} t > \epsilon, \qquad \text{(A 5.1)}$$

where $\epsilon > 0$. Cauchy's Theorem leads to

$$f(s, t) = (2\pi i)^{-2} \int_C \frac{d\bar{s}}{\bar{s} - s} \int_C \frac{d\bar{t}}{\bar{t} - t} f(\bar{s}, \bar{t}), \qquad \text{(A 5.2)}$$

where the closed contour C is indicated in Fig. 1. Let R be the radius of the semicircle of C. As $R \to \infty$, the contribution from the semicircle vanishes because of the condition (i). Hence,

$$\begin{aligned} f(s, t) &= (2\pi i)^{-2} \int_{-\infty+i\epsilon}^{+\infty+i\epsilon} \frac{d\bar{s}}{\bar{s} - s} \int_{-\infty+i\epsilon}^{+\infty+i\epsilon} \frac{d\bar{t}}{\bar{t} - t} f(\bar{s}, \bar{t}) \\ &= (2\pi i)^{-2} \int_{-\infty}^{\infty} ds' \int_{-\infty}^{\infty} dt' \frac{f(s' + i\epsilon, t' + i\epsilon)}{(s' - s + i\epsilon)(t' - t + i\epsilon)} \\ &= (2\pi i)^{-2} \int_{-\infty}^{\infty} ds' \int_{-\infty}^{\infty} dt'\, f(s' + i\epsilon, t' + i\epsilon) \\ &\times \int_0^1 dz\, [zs' + (1 - z)t' - zs - (1 - z)t + i\epsilon]^{-2}. \end{aligned} \qquad \text{(A 5.3)}$$

We want to interchange the order of the s' and t' integrations and the z integration. But this is not trivial because the denominator of the integrand may not necessarily be large when $|s'|$ and $|t'|$ are large.

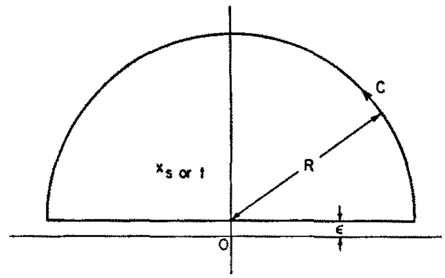

FIG. 1. The contour C on the $\bar{s}$ or $\bar{t}$ plane.

Lemma 4. Let

$$I[R, s', t'] \equiv \{z; 0 \leq z \leq 1, \max(|s'|, |t'|) > R > 1, \; |zs' + (1 - z)t'| < |s't'|^{\frac{1}{2}-\sigma}\}, \qquad \text{(A 5.4)}$$

where s' and t' are real and $0 < \sigma < \frac{1}{2}$. Then its Lebesgue measure $\mu(I[R, s', t'])$ *uniformly* tends to zero as $R \to \infty$.

Proof: We consider two cases $s't' \geq 0$ and $s't' < 0$ separately.

(1^0) The case $s't' \geq 0$. We may assume $s' \geq 0$ and $t' \geq 0$ without loss of generality. The main inequality in (A5.4) becomes

$$zs' + (1 - z)t' < (s't')^{\frac{1}{2}-\sigma}. \qquad \text{(A 5.5)}$$

When $s' = t'$, (A5.5) becomes $1 < s' < s'^{1-2\sigma}$, which is impossible. When $s' > t'$, the points z belonging to $I[R, s', t']$ satisfy

$$0 \leq z < \frac{(s't')^{\frac{1}{2}-\sigma} - t'}{s' - t'}, \qquad \text{(A 5.6)}$$

namely,

$$\mu(I[R, s', t']) \leq \frac{(s't')^{\frac{1}{2}-\sigma} - t'}{s' - t'}. \qquad \text{(A 5.7)}$$

If $t' \leq 1$, the right-hand side of (A5.7) is $O(R^{-\frac{1}{2}-\sigma})$. If $t' > 1$,

$$\mu(I[R, s', t']) < \frac{t'^{\frac{1}{2}-\sigma}(s'^{\frac{1}{2}} - t'^{\frac{1}{2}})}{s' - t'} = \frac{t'^{\frac{1}{2}-\sigma}}{s'^{\frac{1}{2}} + t'^{\frac{1}{2}}} < s'^{-\sigma} < R^{-\sigma}. \qquad \text{(A 5.8)}$$

When $s' < t'$, by interchanging (s', z) and $(t', 1 - z)$ the problem is reduced to the above.

(2^0) The case $s't' < 0$. We may assume $s' > 0$ and $t' < 0$ without loss of generality. Let $t'' \equiv -t' > 0$. The main inequality in (A5.4) becomes

$$\pm[zs' - (1 - z)t''] < (s't'')^{\frac{1}{2}-\sigma}. \qquad \text{(A 5.9)}$$

If $zs' - (1 - z)t'' \geq 0$, the points z belonging to

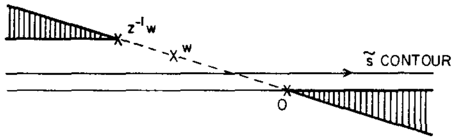

FIG. 2. The singularity regions of $f(\tilde{s}, (w - z\tilde{s})/(1 - z))$ on the $\tilde{s}$ plane when w lies in the second quadrant.

$I[R, s', t']$ satisfy

$$\frac{t''}{s' + t''} \leq z < \frac{(s't'')^{\frac{1}{2}-\sigma} + t''}{s' + t''}. \qquad (A\ 5.10)$$

If $zs' - (1 - z)t'' < 0$, we have only to interchange (s', z) and $(t', 1 - z)$. Hence,

$$\mu(I[R, s', t']) \leq 2 \cdot \frac{(s't'')^{\frac{1}{2}-\sigma}}{s' + t''} \leq 2R^{-2\sigma}. \qquad (A\ 5.11)$$

Thus we have

$$\mu(I[R, s', t']) = O(R^{-\sigma}) \qquad (A\ 5.12)$$

for any case. Q.E.D.

Now, denoting the interval $[0, 1]$ by I, we can rewrite (A5.3) as

$$\int_{-\infty}^{\infty} ds' \int_{-\infty}^{\infty} dt' \int_0^1 dz$$
$$= \int_{-\infty}^{\infty} ds' \int_{-\infty}^{\infty} dt' \left[\int_{I - I[R,s',t']} dz + \int_{I[R,s',t']} dz \right]. \qquad (A\ 5.13)$$

In the first term of the right-hand side, s' and t' satisfy either $\{|s'| \leq R, |t'| \leq R\}$ or

$$|zs' + (1 - z)t'| \geq |s't'|^{\frac{1}{2}-\sigma}. \qquad (A\ 5.14)$$

Therefore, if we choose σ so as to satisfy $\frac{1}{4}\delta > \sigma > 0$, the order of the s' and t' integrations and the z integration can be interchanged on account of the condition (i). The second term tends to zero as $R \to \infty$ because of Lemma 4. Thus,

$$f(s, t) = \int_0^1 dz\, \psi_+(z, zs + (1 - z)t), \qquad (A\ 5.15)$$

with

$$\psi_+(z, w) \equiv (2\pi i)^{-2}$$
$$\times \int_{-\infty}^{\infty} ds' \int_{-\infty}^{\infty} dt' \frac{f(s' + i\epsilon, t' + i\epsilon)}{[zs' + (1 - z)t' - w + i\epsilon]^2},$$
$$(\mathrm{Im}\, w > \epsilon). \qquad (A\ 5.16)$$

For w fixed, $\psi_+(z, w)$ is a function of z defined almost everywhere in $0 \leq z \leq 1$. Since the contribution from $z = 1$ is infinitesimal, we always assume $z \neq 1$ hereafter.

We can carry out one of integrations in (A5.16) as follows:

$$\psi_+(z, w) = (2\pi i)^{-2} \int_{-\infty+i\epsilon}^{+\infty+i\epsilon} d\tilde{s}$$
$$\times \int_{-\infty+i\epsilon}^{+\infty+i\epsilon} d\tilde{t} \frac{f(\tilde{s}, \tilde{t})}{[z\tilde{s} + (1 - z)\tilde{t} - w]^2},$$
$$= (2\pi i)^{-2} \int_{-\infty+i\epsilon}^{+\infty+i\epsilon} d\tilde{s}$$
$$\times \int_C d\tilde{t} \frac{f(\tilde{s}, \tilde{t})}{[z\tilde{s} + (1 - z)\tilde{t} - w]^2}. \qquad (A\ 5.17)$$

Cauchy's theorem leads to

$$\psi_+(z, w) = (2\pi i)^{-1}(1 - z)^{-1}$$
$$\times \int_{-\infty+i\epsilon}^{+\infty+i\epsilon} d\tilde{s} \frac{\partial}{\partial w} f\left(\tilde{s}, \frac{w - z\tilde{s}}{1 - z}\right). \qquad (A\ 5.18)$$

Because of the condition (ii), the integral (A5.18) is convergent if $z \neq z_i$. Thus $\psi_+(z, w)$ is well defined except for $z = z_1, z_2, \cdots, z_m, 1$.

Our next task is to investigate the analyticity of $\psi_+(z, w)$ in w for z fixed. It is evident from (A5.16) that $\psi_+(z, w)$ is holomorphic in $\mathrm{Im}\, w > \epsilon$. Next, we consider the analytic continuation to

$$\{w; \epsilon \geq \mathrm{Im}\, w \geq 0, \mathrm{Re}\, w < 0\}. \qquad (A\ 5.19)$$

For this purpose, we investigate the analyticity of $f(\tilde{s}, (w - z\tilde{s})/(1 - z))$ in $\tilde{s}$ when w is fixed in the second quadrant. This can be easily done by using (2.18). The result is illustrated in Fig. 2 in case of $z \neq 0$. The shaded areas stand for singularity regions, which are defined by

$$z^{-1}\, \mathrm{Im}\, w \leq \mathrm{Im}\, \tilde{s} \leq (\mathrm{Im}\, w/\mathrm{Re}\, w)\, \mathrm{Re}\, \tilde{s}, \qquad (A\ 5.20)$$

and

$$0 \geq \mathrm{Im}\, \tilde{s} \geq (\mathrm{Im}\, w/\mathrm{Re}\, w)\, \mathrm{Re}\, \tilde{s}. \qquad (A\ 5.21)$$

In case of $z = 0$, there is no singularity in the upper half-plane. Thus we can analytically continue $\psi_+(z, w)$ to (A5.19) by deforming the $\tilde{s}$ contour of (A5.18). For $\mathrm{Im}\, w = 0$, the $\tilde{s}$ contour becomes like Fig. 3. The singularity regions now become two cuts shown in Fig. 3. [The change of the limit $-\infty + i\epsilon$ into $-\infty - i\epsilon$ causes no trouble because of the condition (ii) and continuity.]

In the next step, we fix w on the negative real axis. Then we can further deform the $\tilde{s}$ contour

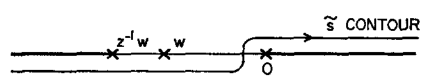

FIG. 3. The deformed $\tilde{s}$ contour when w lies on the negative real axis.

into the lower half-plane. Since the contribution from a large semicircle vanishes because of the condition (ii),[12] we finally obtain

$$\psi_+(z, w) = (2\pi i)^{-1}(1 - z)^{-1} \int_{C'} d\bar{s} \frac{\partial}{\partial w} f\left(\bar{s}, \frac{w - z\bar{s}}{1 - z}\right), \tag{A 5.22}$$

where the contour C' is shown in Fig. 4.

All the above procedure can be done quite analogously for a point $(s, t) \in D_-$ such that

$$\operatorname{Im} s < -\epsilon, \qquad \operatorname{Im} t < -\epsilon. \tag{A 5.23}$$

Then we obtain

$$\psi_-(z, w) \equiv (2\pi i)^{-1}(1 - z)^{-1} \times \int_{+\infty - i\epsilon}^{-\infty - i\epsilon} d\bar{s} \frac{\partial}{\partial w} f\left(\bar{s}, \frac{w - z\bar{s}}{1 - z}\right). \tag{A 5.24}$$

For w on the negative real axis, we have

$$\psi_-(z, w) = (2\pi i)^{-1}(1 - z)^{-1} \int_{C'} d\bar{s} \frac{\partial}{\partial w} f\left(\bar{s}, \frac{w - z\bar{s}}{1 - z}\right). \tag{A 5.25}$$

Therefore, we see

$$\psi_+(z, w) = \psi_-(z, w) \tag{A 5.26}$$

on the negative real axis. This means that $\psi_+(z, w)$ and $\psi_-(z, w)$ define an analytic function $\psi(z, w)$ which is holomorphic except for the ϵ neighborhood of the positive real axis.

Finally, we investigate the asymptotic behavior of $\psi(z, w)$. For this purpose, we again apply the condition (ii) to (A5.18). In the present case, since $|w|$ is large, it is necessary to investigate the behavior of the integrand much more closely. Since the intersection of K and the disc $|s| + |t| \leq M$ is compact, we have

$$|(\partial/\partial t) f(s, t)| < B_0. \tag{A 5.27}$$

Therefore, the condition (ii) can be rewritten as

$$|(\partial/\partial t) f(s, t)| < B'(M + |s| + |t|)^{-\gamma} \times \sum_{i=1}^{m} (\epsilon_i + |z_i s + (1 - z_i)t|)^{-1} \tag{A 5.28}$$

[12] The inequality (2.20) holds also for an arbitrary closed subset of D_{st}. This can be proved as follows. From (A4.2) together with (A4.3) we have

$$|(\partial/\partial t) f(s, t)| \leq (2\pi)^{-1} \int_{-\infty}^{\infty} d\xi' \left| \frac{(\partial/\partial \hat{t}) f(\xi'\hat{s} - \epsilon, \xi'\hat{t} - \epsilon)}{\xi' - \xi} \right|,$$

where $s = \xi\hat{s} - \epsilon$, $t = \xi\hat{t} - \epsilon$, and $|\xi| = 1$. Substituting (2.20) in the integrand and taking $\epsilon \to 0+$, we easily obtain the desired result.

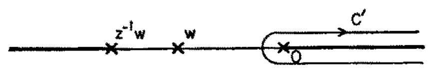

FIG. 4. The contour C' on the $\bar{s}$ plane.

in the whole K, where

$$B' \equiv \max(2^{1+\gamma}B, M^\gamma \epsilon_1 B_0), \quad \epsilon_i \equiv \min_{(s,t)\in K} |z_i s + (1 - z_i)t| > 0. \tag{A 5.29}$$

For $\operatorname{Im} w > \epsilon$, we use (A5.28).

$$\left|\int_{-\infty+i\epsilon}^{+\infty+i\epsilon} d\bar{s} \frac{\partial}{\partial w} f\left(\bar{s}, \frac{w - z\bar{s}}{1 - z}\right)\right| \leq \int_{-\infty}^{\infty} ds' \left|\frac{\partial}{\partial w} f\left(s' + i\epsilon, \frac{w - zs' - iz\epsilon}{1 - z}\right)\right| < (1 - z)^{-1} B' \sum_{i=1}^{m} I_i, \tag{A 5.30}$$

where

$$I_i \equiv \int_{-\infty}^{\infty} ds' (M + |s'| + |v - ks'|)^{-\gamma} \times (\epsilon_i + |(1 - z_i)v - k's'|)^{-1}, \tag{A 5.31}$$

with

$$v \equiv (1 - z)^{-1}(w - iz\epsilon), \qquad (\operatorname{Im} v > 0),$$
$$k \equiv z(1 - z)^{-1} \geq 0,$$
$$k' \equiv (z - z_i)(1 - z)^{-1} \leq (1 - z_i)k. \tag{A 5.32}$$

The assumption $z \neq z_i$ implies $k' \neq 0$. Writing $\operatorname{Re} v \equiv v'$ and $\operatorname{Im} v \equiv v'' > 0$, we have

$$I_i < 2^\gamma \int_{-\infty}^{\infty} ds' (M + |s'| + |v' - ks'| + |v''|)^{-\gamma} \times (\epsilon_i + |(1 - z_i)v' - k's'|)^{-1}. \tag{A 5.33}$$

The transformation $u = k's' - (1 - z_i)v'$ leads to

$$I_i < \frac{2^\gamma}{|k'|} \int_{-\infty}^{\infty} du \left(M + \frac{|u + v_i'|}{|k'|} + \frac{|ku + k_i''v_i'|}{|k'|} + |v''|\right)^{-\gamma} (\epsilon_i + |u|)^{-1}, \tag{A 5.34}$$

with $v_i' \equiv (1 - z_i)v'$ and $k_i'' \equiv k - k'(1 - z_i)^{-1} \geq 0$ ($z_i \neq 1$ by assumption). We make use of the following inequality, which can be easily proved: If $a \geq b \geq 0$, one has

$$|X + Y| + |aX + bY| \geq c(|X| + |Y|) \tag{A 5.35}$$

for any real values of X and Y, where

$$c \equiv \min\left(1, \frac{a - b}{2}, \frac{a - b}{2a}\right). \tag{A 5.36}$$

Applying (A5.35) to the first factor of the integrand of (A5.34), we obtain

$$I_i < 2^{\gamma}\ |k'|^{-1}$$
$$\times \int_{-\infty}^{\infty} du\ [M + h(|u| + |v_i'|) + |v''|]^{-\gamma}(\epsilon_i + |u|)^{-1}, \tag{A 5.37}$$

where $h > 0$ because of (A5.36) with $k' \neq 0$. Let $h' \equiv \min\ (1,\ h(1 - z_i))$. Then choosing σ such that $0 < \sigma < \gamma$, we have

$$\begin{aligned} I_i &< 2^{\gamma}\ |k'|^{-1} \int_{-\infty}^{\infty} du\ (\epsilon_i + |u|)^{-1} \\ &\quad \times (M + h\ |u|)^{-\gamma+\sigma}(M + h'\ |v|)^{-\sigma} \\ &= O(|v|^{-\sigma}) \\ &= O(|w|^{-\sigma}). \end{aligned} \tag{A 5.38}$$

Thus it has been established that

$$\psi(z, w) = O(|w|^{-\sigma}), \tag{A 5.39}$$

for $\operatorname{Im} w > \epsilon$. The same is true also for $\operatorname{Im} w < -\epsilon$. Hence, Lindelöf's asymptotic theorem tells us that (A5.39) is true also in

$$\{w;\ |\operatorname{Im} w| \leq \epsilon,\ \operatorname{Re} w < 0\}. \tag{A 5.40}$$

Now, the analyticity and the boundedness (A5.39) of $\psi(z, w)$ yield

$$\psi(z, w) = (2\pi i)^{-1} \int_{C'} d\bar{w}\, \frac{\psi(z, \bar{w})}{\bar{w} - w}. \tag{A 5.41}$$

Taking the *improper* limit $\epsilon \to 0+$ in (A5.22) and (A5.41), we obtain

$$\psi(z, w) = (1 - z)^{-1} \int_0^{\infty} ds'\, \frac{\partial}{\partial w}\, f_{\epsilon}\!\left(s',\ \frac{w - zs'}{1 - z}\right), \tag{A 5.42}$$

for $w < 0$, and

$$\psi(z, w) = \int_0^{\infty} d\alpha\, \frac{\rho(z, \alpha)}{\alpha - w}, \tag{A 5.43}$$

for $w \ngeq 0$. To interchange the order of $\epsilon \to 0+$ and and an integration is not made in the usual sense, but it *defines* a distribution. Therefore, the asymptotic behavior (A5.39) is not necessarily inherited by $\rho(z, \alpha)$. Theorem V has now been established by (A5.15) with $\psi_+ = \psi$, (A5.43), and (A5.42).

APPENDIX VI. PROOF OF THEOREM VI

The proof is the same with that of Theorem V except for the asymptotic behavior of $\psi(z, w)$. In the present case, instead of (A5.30), we have

$$\begin{aligned} &\left|\int_{-\infty+i\epsilon}^{+\infty+i\epsilon} d\bar{s}\, \frac{\partial}{\partial w}\, f\!\left(\bar{s},\ \frac{w - z\bar{s}}{1 - z}\right)\right| \\ &\leq B \int_{-\infty}^{\infty} ds' \left(1 + |s'| + \frac{|w - zs' - iz\epsilon|}{1 - z}\right)^{-1-\gamma} \\ &\leq 2^{1+\gamma} B \int_{-\infty}^{\infty} ds'\ (1 + |s'| + |v' - ks'| + |v''|)^{-1-\gamma}, \end{aligned} \tag{A 6.1}$$

in the whole D_+, where $v = v' + iv''$ and k are given in (A5.32). Since the last expression of (A6.1) is nothing but a special case of the right-hand side of (A5.33), we obtain

$$|\psi_+(z, w)| \leq C'(1 + |w|)^{-\sigma}. \tag{A 6.2}$$

The same is true also for $\psi_-(z, w)$ with $w \in D_-$. In the present case, (A6.2) holds regardless to ϵ, hence we can interchange the order of $\epsilon \to 0+$ and the integration in the ordinary sense. Thus we obtain (2.26).

APPENDIX VII. DERIVATION OF FORMULAS IN EXAMPLES

Examples 1 and 2. Trivial.

Example 3. The weight function can be easily calculated by using the representation of $(-s)^{-\frac{1}{2}}(-t)^{-\frac{1}{2}}$ (see Example 7) and

$$\exp\left[-(-t)^{\frac{1}{2}}\right] = \pi^{-1} \int_0^{\infty} d\alpha\, \frac{\sin \alpha^{\frac{1}{2}}}{\alpha - t}. \tag{A 7.1}$$

The result is

$$\begin{aligned} \rho(z, \alpha) &= \tfrac{1}{2}\pi^{-\frac{1}{2}}[\Gamma(\tfrac{1}{4})]^{-2}z^{-\frac{1}{4}} \\ &\times \int_z^1 dx \cdot x^{\frac{1}{4}}(1 - x)^{-\frac{3}{4}}(x - z)^{-\frac{3}{4}} J_0([x\alpha/(x - z)]^{\frac{1}{2}}), \end{aligned} \tag{A 7.2}$$

whose singularities are located at $z = 0$ (order $z^{-\frac{1}{4}}$) and at $\alpha = 0$ (order $\alpha^{-\frac{1}{4}}$) only, and (A7.2) behaves like $O(\alpha^{-\frac{1}{4}})$ as $\alpha \to \infty$.

Example 4. Trivial.

Example 5. See the next.

Example 6. When $N > \operatorname{Re} \lambda > N - 1$, we have

$$\begin{aligned} &\int_0^1 dz \int_0^{\infty} d\alpha\, \frac{[zs + (1 - z)t]^N \alpha^{\lambda} \delta(z - z_0)}{\alpha^N[\alpha - zs - (1 - z)t]} \\ &\qquad = \Gamma(\lambda + 1)\Gamma(-\lambda)[-z_0 s - (1 - z_0)t]^{\lambda}. \end{aligned} \tag{A 7.3}$$

Differentiating (A7.3) by z_0 n times, we obtain (4.3). If $n \geq N$, one has an identity

$$\begin{aligned} &\int dz\, \frac{[zs + (1 - z)t]^N}{\alpha - zs - (1 - z)t}\, \delta^{(n)}(z) \\ &\qquad = \int dz\, \frac{\alpha^N \delta^{(n)}(z)}{\alpha - zs - (1 - z)t}. \end{aligned} \tag{A 7.4}$$

Hence, for $n \geq N > \operatorname{Re}\lambda > -1$,

$$f(s, t) = n!(t - s)^n \int_0^\infty d\alpha \frac{\alpha^\lambda}{(\alpha - t)^{n+1}}$$

$$= \Gamma(\lambda + 1)\Gamma(n - \lambda)(t - s)^n(-t)^{\lambda-n}. \quad \text{(A 7.5)}$$

Example 7. First, we assume $0 > \operatorname{Re}\mu > -\frac{1}{2}$ and $0 > \operatorname{Re}\nu > -\frac{1}{2}$. Then we can use (2.21) with (2.22).

$$\psi(z, w) = (1 - z)^{-1} \int_0^\infty ds' \frac{\partial}{\partial w} \times \left[\frac{s'^\mu}{\Gamma(\mu + 1)\Gamma(-\mu)} \left(-\frac{w - zs'}{1 - z}\right)^\nu\right]$$

$$= [\Gamma(\mu + 1)\Gamma(-\mu)]^{-1}(-\nu)z^{-\mu-1}(1 - z)^{-\nu-1} \times \int_0^\infty dx \cdot x^\mu(x - w)^{\nu-1}$$

$$= [\Gamma(-\mu)\Gamma(-\nu)]^{-1}\Gamma(-\mu - \nu)z^{-\mu-1} \times (1 - z)^{-\nu-1}(-w)^{\mu+\nu}. \quad \text{(A 7.6)}$$

Then we obtain (4.5). For the general case, we analytically continue (4.5) with respect to μ and ν after making subtractions. In this way we get the correct result because of the invariance property of the weight function in the subtraction procedure (1.4).

Example 8. We want to prove (4.16). We denote the right-hand side of (4.16) by $f(s, t)$. Then

$$\frac{\partial}{\partial s}\frac{\partial}{\partial t} f(s, t) = \int_0^1 dz \int_0^\infty d\alpha \frac{2}{[\alpha - zs - (1 - z)t]^3} = \frac{1}{st}. \quad \text{(A 7.7)}$$

Since $f(s, t)$ is a symmetric function of s and t, (A7.7) leads to

$$f(s, t) = \log(-s) \log(-t) + \varphi(s) + \varphi(t), \quad \text{(A 7.8)}$$

where φ is an unknown function. On the other hand, because of (5.12), we obtain

$$f(s, s) = -2 \int_0^\infty d\alpha \frac{(s + 1) \log \alpha}{(1 + \alpha)(\alpha - s)} = [\log(-s)]^2, \quad \text{(A 7.9)}$$

which implies $\varphi(s) \equiv 0$.

N-Dimensional Total Orbital Angular-Momentum Operator. II. Explicit Representations*

KENNETH D. GRANZOW†
Sandia Laboratory, Albuquerque, New Mexico
(Received 14 April 1964)

A method of generating orthogonal polar coordinate systems in N-dimensional space is given. Commuting angular-momentum operators are easily found in the coordinate systems generated; these operators are of a single form which depends on two parameters. A short table of coordinate systems and the resulting structure of quantum numbers and eigenfunctions is given.

INTRODUCTION

IN a previous paper[1] an invariance property of the N-dimensional total angular-momentum operator (L^2) as defined by Louck was pointed out. It was shown that other polar coordinate systems exist beside that used by Louck in which the eigenvalue problem of L^2 can be solved. In this paper a systematic method of constructing polar coordinate systems is given. The eigenfunctions, eigenvalues, and quantum number structure of L^2 are found in the polar coordinate systems constructed. Thus, a systematic method is given for generating bases of the Hilbert space of functions on the $(N-1)$-dimensional unit sphere. The bases so generated can be used in the analysis of the N-dimensional harmonic oscillator and hence in molecular vibration analysis. They should also prove of use in N-dimensional theories of elementary particles (see, for instance, the recent paper of de Broglie *et al.*[2]). If elementary particles are describable as rotators in a hyperspace, then a comparison of the structure of elementary particle quantum numbers and the quantum number structures produced by the methods given in this paper could provide a semi-empirical method of finding the proper coordinate system for their description.

A PROPERTY OF POLAR COORDINATES

In this paper the words "polar coordinate system" shall refer to any generalized polar coordinates as defined in I.

Definition. In N-dimensional Euclidean space, *polar coordinates* consist of a radial coordinate R and $(N-1)$ angle coordinates θ^α such that $R = [\sum_{n=1}^N (x^n)^2]^{\frac{1}{2}}$ and $\theta^\alpha = h^\alpha(x^i)$, where the x^i are Cartesian coordinates and the functions $h^\alpha(x^i)$ are restricted such as to transform the metric to the form $ds^2 = \sum_{n=1}^N (dx^n)^2 = dR^2 + R^2(ds^*)^2$, where $(ds^*)^2 = g^*_{\alpha\beta}\, d\theta^\alpha\, d\theta^\beta$ (summation assumed over α and β), and where the functions $g^*_{\alpha\beta}$ are independent of R.

The transformation from polar coordinates to Cartesians has a unique form.

Theorem. The transformation from any polar coordinate system to Cartesian coordinates is of the form $x^n = Rf^n(\theta)$, $n = 1, 2, \cdots, N$; $\theta = (\theta^1, \theta^2, \cdots, \theta^{N-1})$.

Proof. It is assumed that $x^n = Rf^n(R, \theta)$ and it is then shown that $\partial f^n/\partial R \equiv 0$, i.e., f^n is independent of R for all n.

$$R^2 = \sum_{n=1}^N (x^n)^2 \Rightarrow \sum_{n=1}^N [f^n(R, \theta)]^2 = 1,$$

$$ds^2 = dR^2 + R^2(ds^*)^2 \Rightarrow \sum_{n=1}^N \left(R\frac{\partial f^n}{\partial R} + f^n\right)^2 = 1,$$

(By evaluating the coefficient of dR^2 in the metric.)

$$\Rightarrow \sum_{n=1}^N \left[R^2\left(\frac{\partial f^n}{\partial R}\right)^2 + R\frac{\partial (f^n)^2}{\partial R} + (f^n)^2\right] = 1,$$

$$\Rightarrow R^2 \sum_{n=1}^N \left(\frac{\partial f^n}{\partial R}\right)^2 + R\frac{\partial[\sum_{n=1}^N (f^n)^2]}{\partial R} + \sum_{n=1}^N (f^n)^2 = 1,$$

$$\Rightarrow R^2 \sum_{n=1}^N \left(\frac{\partial f^n}{\partial R}\right)^2 = 0 \Rightarrow \frac{\partial f^n}{\partial R} \equiv 0 \quad \text{for all } n, \quad R \neq 0.$$

In terms of the functions $f^n(\theta)$ it is easily seen that $g^*_{\alpha\beta} = \sum_{n=1}^N (\partial f^n/\partial\theta^\alpha)(\partial f^n/\partial\theta^\beta)$. Note also that the functions $f^n(\theta)$ have the property $\sum_{n=1}^N [f^n(\theta)]^2 = 1$.

GENERATION OF NEW COORDINATE SYSTEMS

If a polar coordinate system of N dimensions is known, then a polar coordinate system of $(N+1)$

---

* This work performed under the auspices of the U. S. Atomic Energy Commission.
† Present address: The Dikewood Corporation, Albuquerque, New Mexico.
[1] K. D. Granzow, J. Math. Phys. 4, 897 (1963). This paper will be referred to as I in the sequel.
[2] L. de Broglie, D. Bohm, P. Hillion, F. Halbwachs, T. Takabayasi, and J.-P. Vigier, Phys. Rev. 129, 438 (1963).

dimensions can be constructed. This construction will be referred to later as method (A). Suppose the transformation to Cartesians of the known polar coordinate system is

$$x^n = Rf^n(\theta), \qquad n = 1, \cdots, N. \tag{1}$$

The transformation $x'^n = R' \sin\theta'^N f^n(\theta')$, $n = 1, \cdots, N$; $x'^{N+1} = R' \cos\theta'^N$, $0 \le \theta'^N \le \pi$ defines a polar coordinate system of $(N+1)$ dimensions.[3] The coefficients of the metric for the new coordinate system are easily computed.

$$g'^*_{\alpha\beta} = \sin^2\theta'^N g^*_{\alpha\beta}(\theta'), \qquad \alpha, \beta < N+1,$$

$$g'^*_{(N+1)\alpha} = \begin{cases} 1, & \alpha = N+1, \\ 0, & \alpha < N+1. \end{cases}$$

Thus, in the new coordinate system the metric has the form $ds'^2 = (dR')^2 + (R')^2[(d\theta'^N)^2 + \sin^2\theta'^N(ds^*)^2]$, where the coordinates of $(ds^*)^2$ are given primes to indicate they are of the new coordinate system. This process generates the ordinary spherical polar coordinates from 2-dimensional polar coordinates. Iteration of this procedure generates the polar coordinates used by Louck.[4]

Since the coefficients of the metric are now known, the total angular momentum operator L'^2 can be evaluated. $L'^2 = \hbar^2\nabla'^{*2}$, where

$$\nabla'^{*2} = \frac{1}{(g'^*)^{\frac{1}{2}}}\frac{\partial}{\partial\theta'^\alpha}\left[(g'^*)^{\frac{1}{2}} g'^{*\alpha\beta}\frac{\partial}{\partial\theta'^\beta}\right],$$

$$\nabla'^{*2} = \frac{1}{\sin^{N-1}\theta'^N}\frac{\partial}{\partial\theta'^N}\left(\sin^{N-1}\theta'^N\frac{\partial}{\partial\theta'^N}\right) + \frac{\nabla^{*2}}{\sin^2\theta'^N},$$

and where the coordinates of the operator ∇^{*2} are given primes to indicate they are of the new coordinate system. Since ∇'^{*2} and ∇^{*2} commute, ∇^{*2} can be replaced by its eigenvalue (which is known from Louck's work and from I). Hence,

$$L'^2 = -\hbar^2\left\{\frac{1}{\sin^{N-1}\theta'^N}\frac{\partial}{\partial\theta'^N}\left(\sin^{N-1}\theta'^N\frac{\partial}{\partial\theta'^N}\right) - \frac{l(l+N-2)}{\sin^2\theta'^N}\right\}, \tag{3}$$

where l takes on nonnegative integer values.[5] The equation for the part of the eigenfunction dependent on θ'^N is

$$\left[\frac{1}{\sin^{N-1}\theta'^N}\frac{d}{d\theta'^N}\left(\sin^{N-1}\theta'^N\frac{d}{d\theta'^N}\right) - \frac{l(l+N-2)}{\sin^2\theta'^N}\right]\Theta_N(\theta'^N) = -\lambda'\Theta_N(\theta'^N). \tag{4}$$

The total eigenfunction is a product of $\Theta_N(\theta'^N)$ and the eigenfunction of L^2 with quantum number, l, and the variables primed. From Louck's work it is known that $\lambda' = l'(l'+N-1)$, where $l' = l, l+1, \cdots$.

To solve Eq. (4) a change of variables is convenient. Let $z = \cos\theta'^N$ and $\Theta_N(\theta'^N) = (1-z^2)^{l/2}T(z)$. Equation (4) becomes

$$\left\{(z^2-1)\frac{d^2}{dz^2} + (2l+N)z\frac{d}{dz} + [l(l+N-1) - l'(l'+N-1)]\right\}T(z) = 0, \tag{5}$$

which is Gegenbauer's equation. The solutions of interest are the Gegenbauer polynomials $T^\beta_\alpha(z)$, where $\alpha = l' - l$ and $\beta = l - 1 + N/2$ (in the notation of Morse and Feshback).[6]

If two polar coordinate systems are known of N_1 and N_2 dimensions, a third polar coordinate system can be constructed of $(N_1 + N_2)$ dimensions.[7] This construction will be referred to as method (B). The transformation

$$x^n = R\sin\theta^{N_1+N_2-1}f_1^n(\theta^1, \cdots, \theta^{N_1-1}), \quad n = 1, \cdots, N_1;$$

$$x^{n+N_1} = R\cos\theta^{N_1+N_2-1}f_2^n(\theta^{N_1}, \theta^{N_1+1}, \cdots, \theta^{N_1+N_2-2}),$$

$$n = 1, \cdots, N_2; \quad 0 \le \theta^{N_1+N_2-1} \le \tfrac{1}{2}\pi$$

defines the new coordinate system. The coefficients of the metric are given by

$$g^*_{\alpha\beta} = \sin^2\theta^{N_1+N_2-1}g^*_{1\alpha\beta}(\theta^1, \cdots, \theta^{N_1-1}), \qquad \alpha, \beta \le N_1 - 1,$$

$$g^*_{\alpha\beta} = \cos^2\theta^{N_1+N_2-1}g^*_{2\alpha\beta}(\theta^{N_1}, \theta^{N_1+1}, \cdots, \theta^{N_1+N_2-2}), \qquad N_1 \le \alpha, \beta \le N_1 + N_2 - 2,$$

$$g^*_{\alpha\beta} = 1, \qquad \alpha = \beta = N_1 + N_2 - 1,$$

$$g^*_{\alpha\beta} = 0, \qquad \alpha = N_1 + N_2 - 1 \text{ and } \beta < \alpha$$

or

$$N_1 \le \alpha \le N_1 + N_2 - 2 \text{ and } \beta < N_1. \tag{6}$$

[3] Primed quantities are associated with the new coordinate system; unprimed quantities are associated with the known coordinate system.

[4] James D. Louck, J. Mol. Spectry. 4, 298 (1960).

[5] An exception to the assumption of nonnegative l occurs for $N = 2$, $l = m = 0, \pm1, \pm2, \cdots$. To avoid treating this as a special case l is set equal to $|m|$ in which case l is again nonnegative and Eq. (3) is left unchanged.

[6] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), pp. 547–549, 600–604, 782–784.

[7] In the following, subscripts 1 and 2 refer to the two known coordinate systems. Unsubscripted variables refer to the new coordinate system.

Thus, in the new coordinate system, the metric has the form

$$ds^2 = dR^2 + R^2[(d\theta^{N_1+N_2-1})^2 + \sin^2\theta^{N_1+N_2-1}(ds_1^*)^2 + \cos^2\theta^{N_1+N_2-1}(ds_2^*)^2],$$

where the coordinates of $(ds_1^*)^2$ and (ds_2^*) are those appearing as arguments in f_1^n and f_2^n, respectively.

Evaluating the operator ∇^{*2} using the coefficients of the metric (6), one obtains

$$\nabla^{*2} = \left[\frac{1}{\sin^{N_1-1}\theta^k\cos^{N_2-1}\theta^k}\frac{\partial}{\partial\theta^k}\left(\sin^{N_1-1}\theta^k\cos^{N_2-1}\theta^k\frac{\partial}{\partial\theta^k}\right)\right] + \frac{\nabla_1^{*2}}{\sin^2\theta^k} + \frac{\nabla_2^{*2}}{\cos^2\theta^k}, \quad (7)$$

where $k = N_1 + N_2 - 1$ and the coordinates of ∇_1^{*2} and ∇_2^{*2} are those appearing as arguments in f_1^n and f_2^n, respectively. Since ∇^{*2}, ∇_1^{*2}, and ∇_2^{*2} mutually commute they can each take on eigenvalues. Since the eigenvalues are known from previous work one can immediately write

$$L^2 = -\hbar^2\left\{\left[\frac{1}{\sin^{N_1-1}\theta^k\cos^{N_2-1}\theta^k}\frac{\partial}{\partial\theta^k}\left(\sin^{N_1-1}\theta^k\cos^{N_2-1}\theta^k\frac{\partial}{\partial\theta^k}\right)\right] - \frac{l_1(l_1+N_1-2)}{\sin^2\theta^k} - \frac{l_2(l_2+N_2-2)}{\cos^2\theta^k}\right\}, \quad k = N_1+N_2-1, \quad (8)$$

which yields the equation

$$\left\{\frac{1}{\sin^{N_1-1}\theta^k\cos^{N_2-1}\theta^k}\frac{d}{d\theta^k}\left(\sin^{N_1-1}\theta^k\cos^{N_2-1}\theta^k\frac{d}{d\theta^k}\right) - \frac{l_1(l_1+N_1-2)}{\sin^2\theta^k} - \frac{l_2(l_2+N_2-2)}{\cos^2\theta^k}\right\}\Theta_k(\theta^k) = -\lambda\Theta_k(\theta^k), \quad (9)$$

where λ is known to equal $l(l + N_1 + N_2 - 2)$. Equations (3) and (4) can be viewed as degenerate cases of Eqs. (8) and (9) in which $N_2 = 1$ and l_2 is set equal to zero. If N_1 is also equal to one and l_1 is set equal to zero, Eq. (8) gives $L^2 = -\hbar^2\,\partial^2/\partial\theta^2 = L_z^2$. Thus, in all the polar coordinate systems that can be generated by iterations and combinations of the two methods (A) and (B), the eigenvalue problem reduces to finding solutions of Eq. (9) and/or its degenerate forms.

To solve Eq. (9) for $N_1 > 1$ and $N_2 > 1$ it is convenient to define a new independent variable by the formula $z = \cos^2\theta^k$. Equation (9) becomes

$$\left\{\frac{d^2}{dz^2} + \left[\frac{N_2/2}{z} + \frac{N_1/2}{z-1}\right]\frac{d}{dz} - \left[\frac{l_1(l_1+N_1-2)/4}{z-1} - \frac{l_2(l_2+N_2-2)/4}{z} + \frac{l(l+N_1+N_2-2)}{4}\right]\frac{1}{z(z-1)}\right\}\Theta_k = 0. \quad (10)$$

The solution of this equation is given in I for $N_1 = 3n$, n an integer and $N_2 = 3$. It is just as easy to solve Eq. (10), in general. Using the general solution given by Morse and Feshback,[8] it is found that the solution which is analytic at $z = 0$ is

$$\Theta_k = z^{l_2/2}(z-1)^{l_1/2}F(a, b\,|c|\,z), \quad (11)$$

where

$$a = \tfrac{1}{2}(l_1 + l_2 - l),$$
$$b = \tfrac{1}{2}(l_1 + l_2 + l + N_1 + N_2 - 2),$$
$$c = l_2 + \tfrac{1}{2}N_2,$$

and $F(a, b\,|c|\,z)$ is the hypergeometric function.[9] To make Θ_k a polynomial in z and, hence, analytic at $z = 1$, a is set equal to a nonpositive integer.

$$\tfrac{1}{2}(l_1 + l_2 - l) = -\kappa, \qquad \kappa = 0, 1, 2, \cdots,$$
$$l = l_1 + l_2 + 2\kappa. \quad (12)$$

Thus, l is even (or odd) if and only if $(l_1 + l_2)$ is even (or odd).

COORDINATE SYSTEMS, EIGENFUNCTIONS, AND STRUCTURE OF QUANTUM NUMBERS

By starting with ordinary two-dimensional polar coordinates and then generating coordinate systems of more dimensions by the repeated application of methods (A) and/or (B), one can construct many polar coordinate systems. Whenever method (A) is applied, (1) the new angle coordinate assumes values in the interval $[0, \pi]$, (2) the eigenfunctions of L^2 are products of $\Theta_N(\theta^N)$ [where Θ_N is a solution of Eq. (4)] and eigenfunctions of L^2 found in the preceding coordinate system,[10] and (3) the new quantum number takes on integer values equal to or greater than the quantum number of the eigenfunction of L^2 in the preceding coordinate system. Whenever method (B) is applied (1) the the new angle coordinate assumes values in the interval $[0, \pi/2]$, (2) the eigenfunctions of L^2 are products of $\Theta_{N_1+N_2-1}(\theta^{N_1+N_2-1})$ [where $\Theta_{N_1+N_2-1}$ is a solution of Eq. (10)] and the product of the eigen-

[8] Reference 6, pp. 539, 542.

[9] It has been assumed in this treatment that l_1, and l_2 are nonnegative. Hence, if $N_1 = 2$ let l_1 equal $|m_1|$; if $N_2 = 2$ let l_2 equal $|m_2|$. This does not change Eqs. (8), (9), or (10).

[10] The terminology *preceding coordinate system* or *systems* is used to refer to the coordinate system or systems to which methods (A) or (B) are applied.

TABLE I. Coordinate systems of up to and including six dimensions which can be generated by application of methods (A) and (B) are given. The resulting eigenfunctions of L^2 and the structure of orbital angular-momentum quantum numbers are included in the table. The table can be extended *ad inf.* by repeated application of methods (A) and (B). See the text for definitions of the eigenfunctions, quantum numbers, and methods of construction.

No.	Dimensions	Constructed from No.	Method of construction	Structure of quantum numbers	Eigenfunction
1[a]	2			m	$(1, m)$
2[a,b]	3	1	A	$l' \geqq \lvert m\rvert$	$(2, l' \mid \lvert m\rvert)(1, m)$
3[a]	4	2	A	$l_2' \geqq l_1' \geqq \lvert m\rvert$	$(3, l_2' \mid l_1')(2, l_1' \mid \lvert m\rvert)(1, m)$
4	4	1, 1	B	$l \geqq (\lvert m_1\rvert + \lvert m_2\rvert)$	$(2, 2, l \mid \lvert m_1\rvert, \lvert m_2\rvert)(1, m_1)(1, m_2)$
5[a]	5	3	A	$l_3' \geqq l_2' \geqq l_1' \geqq \lvert m\rvert$	$(4, l_3' \mid l_2')(3, l_2' \mid l_1')(2, l_1' \mid \lvert m\rvert)(1, m)$
6	5	4	A	$l' \geqq l \geqq (\lvert m_1\rvert + \lvert m_2\rvert)$	$(4, l' \mid l)(2, 2, l \mid \lvert m_1\rvert, \lvert m_2\rvert)(1, m_1)(1, m_2)$
7	5	1, 2	B	$l \geqq (\lvert m_1\rvert + l'); l' \geqq \lvert m_2\rvert$	$(2, 3, l \mid \lvert m_1\rvert, l')(1, m_1)(2, l' \mid \lvert m_2\rvert)(1, m_2)$
8[a]	6	5	A	$l_4' \geqq l_3' \geqq l_2' \geqq l_1' \geqq \lvert m\rvert$	$(5, l_4' \mid l_3')(4, l_3' \mid l_2')(3, l_2' \mid l_1')(2, l_1' \mid \lvert m\rvert)(1, m)$
9	6	6	A	$l_2' \geqq l_1' \geqq l \geqq (\lvert m_1\rvert + \lvert m_2\rvert)$	$(5, l_2' \mid l_1')(4, l_1' \mid l)(2, 2, l \mid \lvert m_1\rvert, \lvert m_2\rvert)(1, m_1)(1, m_2)$
10	6	7	A	$l_2' \geqq l \geqq (\lvert m_1\rvert + l_1'); l_1' \geqq \lvert m_2\rvert$	$(5, l_2' \mid l)(2, 3, l \mid \lvert m_1\rvert, l_1')(1, m_1)(2, l_1' \mid \lvert m_2\rvert)(1, m_2)$
11	6	1, 3	B	$l \geqq (\lvert m_1\rvert + l_2'); l_2' \geqq l_1' \geqq \lvert m_2\rvert$	$(2, 4, l \mid \lvert m_1\rvert, l_2')(1, m_1)(3, l_2' \mid l_1')(2, l_1' \mid \lvert m_2\rvert)(1, m_2)$
12	6	1, 4	B	$l_2 \geqq (\lvert m_3\rvert + l_1); l_1 \geqq (\lvert m_1\rvert + \lvert m_2\rvert)$	$(2, 4, l_2 \mid \lvert m_3\rvert, l_1)(1, m_3)(2, 2, l_1 \mid \lvert m_1\rvert, \lvert m_2\rvert)(1, m_1)(1, m_2)$
13[b]	6	2, 2	B	$l \geqq (l_1' + l_2'); l_1' \geqq \lvert m_1\rvert; l_2' \geqq \lvert m_2\rvert$	$(3, 3, l \mid l_1', l_2')(2, l_1' \mid \lvert m_1\rvert)(1, m_1)(2, l_2' \mid \lvert m_2\rvert)(1, m_2)$

[a] Louck's coordinates.
[b] Included in $3M$-dimensional coordinates treated in I.

functions of L^2 found in the preceding coordinate systems, and (3) the new quantum number takes on even or odd values equal to or greater than the sum of the quantum numbers of the eigenfunctions of L^2 found in the preceding coordinate systems according to whether this sum is even or odd, respectively. Table I lists thirteen coordinate systems with the resulting eigenfunctions and quantum number structure which can be constructed by combinations of the two methods. The exponential is denoted $(1, m)$; solutions of Eq. (4) are denoted $(N, l' \mid l)$; solutions of Eq. (10) are denoted $(N_1, N_2, l \mid l_1, l_2)$; m denotes a quantum number of L^2 in two-dimensional polar coordinates where the eigenfunction is $\exp(im\theta)$; l' denotes a quantum number which arose by application of method (A); l denotes a quantum number which arose by application of method (B). If more than one of a given type quantum number occurs, the quantum numbers of that type are subscripted. The table can, of course, be extended *ad inf.* to higher and higher dimensionalities. The left-most quantum number and factor in the eigenfunction are those resulting from the last application of methods (A) or (B).[11]

[11] This left-most quantum number which results from the last application of methods (A) or (B) is the quantum number referred to by Rakavy as the *seniority* in his study of the harmonic oscillator. [G. Rakavy, Nucl. Phys. 4, 289 (1957)].

For a given N one can represent the coordinate systems (and hence, the corresponding commuting operators and eigenfunctions) which can be generated by the methods (A) and (B) by writing N as various partitioned sums of ones and twos. For instance, the two possible four-dimensional coordinate systems can be represented by the sums $(2 + 1) + 1$ and $2 + 2$. In this scheme each parenthesis must contain exactly two terms. Each plus sign represents an application of either method (A) or (B); if one of the two terms that a particular plus sign operates on is a *1*, then that plus sign represents an application of method (A), otherwise it represents an application of method (B). Since there is only one three-dimensional coordinate system derivable by this scheme, the number *3* can unambiguously be substituted for $(2 + 1)$. The numbers *2* and *3* and the arithmetic values of quantities in parenthesis always have the meaning of the dimensionality of the coordinate systems on which methods (A) or (B) are being applied. Using this representation, the four-dimensional coordinate systems are represented by $3 + 1$ and $2 + 2$; the five-dimensional by $(3 + 1) + 1$, $(2 + 2) + 1$, and $3 + 2$; the six-dimensional by $[(3 + 1) + 1] + 1$, $[(2 + 2) + 1] + 1$, $(3 + 2) + 1$, $(3 + 1) + 2$, $(2 + 2) + 2$, and $3 + 3$. Thus, for any N, each such partitioning corresponds to a given coordinate system and hence to a unique set of commuting operators and eigenfunctions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 10 OCTOBER 1964

Inequalities Relating the Nearest-Neighbor Spin Correlation and the Magnetization for the Heisenberg Hamiltonian

HAROLD FALK*
U. S. Army Electronics Research and Development Laboratories, Fort Monmouth, New Jersey
(Received 28 April 1964)

Convexity properties of the free energy are used to obtain inequalities relating the nearest-neighbor spin correlation and the magnetization for the Heisenberg Hamiltonian.

I. INTRODUCTION

OF some value are exact relationships in the form of inequalities which connect the properties of a system described by a soluble model Hamiltonian H_0 with the properties of a system described by $H = H_0 + (H - H_0)$. Inequalities, like symmetry considerations, may often greatly reduce the amount of calculation necessary to obtain a desired result, and will sometimes indicate when an approximation method has violated some intrinsic property of the system. Variational methods such as those based on the weak form of Peierls' variational theorem[1] reflect general convexity[2] characteristics of the partition function and the free energy. Although the weak form of Peierls' theorem is frequently used[3] as the basis for minimizing a trial free energy, the following discussion will provide an example which utilizes the theorem to obtain inequalities relating quantities other than free energies.

In particular, for a system described by the Heisenberg Hamiltonian with either nearest-neighbor ferromagnetic or antiferromagnetic interaction, inequalities are obtained which relate the nearest-neighbor spin correlation (short-range order), here defined as $\rho_1 = (Nz)^{-1} \sum_{(f,g)} \langle \mathbf{S}_f \cdot \mathbf{S}_g \rangle$, to the magnetization per spin. The sum is over all nearest-neighbor pairs (f, g) of N spins each spin having z nearest neighbors, and the bracket $\langle\ \rangle$ denotes the canonical ensemble average of the quantity enclosed.

II. INEQUALITIES IMPLIED BY CONVEXITY

First recall that for the canonical ensemble the free energy $F(\lambda)$ and average $\langle O \rangle_\lambda$ of any operator O are, respectively,

$$F(\lambda) = (-\beta)^{-1} \ln \operatorname{tr} \{\exp [-\beta(H_0 + \lambda H_1)]\}, \qquad (\beta^{-1} = kT), \tag{1}$$

and

$$\langle O \rangle_\lambda = \frac{\operatorname{tr} \{O \exp [-\beta(H_0 + \lambda H_1)]\}}{\operatorname{tr} \{\exp [-\beta(H_0 + \lambda H_1)]\}}, \tag{2}$$

where the β dependence is suppressed on the left side of (1) and (2). The Hamiltonian of the system is $H = H_0 + \lambda H_1$.

Use will be made of the general property

$$\langle H_1 \rangle_\lambda \leq \langle H_1 \rangle_{\lambda_0}, \qquad (0 \leq \lambda_0 \leq \lambda), \tag{3}$$

which may be obtained from the weak form of the Peierls variational theorem often written[1]

$$F(\lambda) \leq F(0) + \lambda F'(0).$$

It is only necessary to use the arbitrariness[2] in defining H_0 and H_1 to write

$$F(\lambda) \leq F(\lambda_0) + (\lambda - \lambda_0) F'(\lambda_0), \tag{4}$$

where λ_0 is arbitrary. The inequality (4) implies that $F''(\lambda) \leq 0$; consequently,

$$F'(\lambda) = F'(\lambda_0) + \int_{\lambda_0}^{\lambda} d\lambda_1 F''(\lambda_1) \leq F'(\lambda_0). \tag{5}$$

Since $F'(\lambda) = \langle H_1 \rangle_\lambda$, it is clear that (3) is simply a consequence of the λ-convexity of $F(\lambda)$.

Another property to be used is the β-convexity of $\beta F(\lambda)$ [alternatively, the T-convexity of $F(\lambda)$]; i.e.,

$$k^{-2}\beta^{-3}\partial^2 F(\lambda)/\partial T^2 = \partial^2[\beta F(\lambda)]/\partial\beta^2 = \partial\langle H \rangle_\lambda/\partial\beta \leq 0, \tag{6}$$

since

$$\partial\langle H \rangle_\lambda/\partial\beta = -\langle (H - \langle H \rangle_\lambda)^2 \rangle_\lambda. \tag{7}$$

The statement that the specific heat is nonnegative or that

$$(\langle H \rangle_\lambda)_{T_0} \leq (\langle H \rangle_\lambda)_T, \qquad (0 \leq T_0 \leq T), \tag{8}$$

is equivalent to (6).

* Present address: Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania.

1 R. E. Peierls, Phys. Rev. **54**, 918 (1938); T. D. Schultz, Nuovo Cimento **8**, 943 (1958); J. Czerwonko, Bull. Acad. Polon. Sci. Cl. III **7**, 639, 699 (1959); H. Falk, Physica **29**, 1114 (1963); the weak form has been attributed to N. N. Bogoliubov, by J. Kvasnikov, Doklady Akad. Nauk SSSR. **110**, 755 (1956).

2 A concise summary of convexity properties of the free energy will be found in: R. B. Griffiths, J. Math. Phys. **5**, 1215 (1964).

3 H. Falk, Phys. Rev. **133**, A 1382 (1964); references to other work are contained therein.

The following examples will indicate how the inequalities (3) and (8) may be used to obtain certain bounds relating perturbed averages $\langle O\rangle_\lambda$ to unperturbed averages $\langle O\rangle_0$; the latter being associated with a truncated, soluble Hamiltonian.

III. FERROMAGNET

Consider the Heisenberg Hamiltonian for a system with nearest-neighbor ferromagnetic interaction of the spins (each of magnitude S). The Zeeman energy due to an external magnetic field H_z is added so that

$$H = -2J \sum_{(f,g)} \mathbf{S}_f\cdot\mathbf{S}_g - g\mu H_z \sum_f S_f^z. \tag{9}$$

By adding and subtracting the term

$$2J \sum_{(f,g)} \mathbf{S}_f\cdot(\sigma S)\hat{k},$$

where $\hat{k}$ is a unit vector along the z axis, (9) may be written

$$H = H_0 + \lambda H_1, \tag{10}$$

with

$$H_0 = -(g\mu H_z + Jz\sigma S) \sum_f S_f^z, \tag{11}$$

and

$$H_1 = -J \sum_f \sum_\delta (\mathbf{S}_{f+\delta} - \sigma S\hat{k})\cdot\mathbf{S}_f. \tag{12}$$

In (10) $\boldsymbol{\delta}$ is the vector from site f to one its z nearest neighbors, and σ is arbitrary. Perturbed quantities are labeled with the parameter λ, which is implicitly equal to unity.

Now it is well known and easily demonstrated that H_0, which is proportional to the z projection of the total magnetic moment, commutes with H_1. For convenience the trace will be taken over the complete set of eigenstates of H_0 for which

$$\langle S_f^xS_g^x + S_f^yS_g^y\rangle_0 = 0 \qquad (f \neq g). \tag{13}$$

Equations (12) and (13) may be combined with (3) to provide

$$\sum_f \sum_\delta [\langle \mathbf{S}_{f+\delta}\cdot\mathbf{S}_f\rangle_\lambda - \sigma S\langle S_f^z\rangle_\lambda] \geq \sum_f \sum_\delta [\langle S_{f+\delta}^zS_f^z\rangle_0 - \sigma S\langle S_f^z\rangle_0]. \tag{14}$$

Explicit calculation shows that for the free-spin Hamiltonian H_0,

$$\langle S_{f+\delta}^zS_f^z\rangle_0 = \langle S_{f+\delta}^z\rangle_0\langle S_f^z\rangle_0. \tag{15}$$

Upon substituting (15) into (14) and selecting

$$\sigma S \equiv \langle S_{f+\delta}^z\rangle_0, \tag{16}$$

the right side of (14) vanishes and

$$\sum_f \sum_\delta \langle \mathbf{S}_{f+\delta}\cdot\mathbf{S}_f\rangle_\lambda \geq \sum_f \sum_\delta \langle S_{f+\delta}^z\rangle_0\langle S_f^z\rangle_\lambda. \tag{17}$$

With H_0 translationally invariant, $\langle S_{f+\delta}^z\rangle_0$ is independent of site and (17) may be written

$$(Nz)^{-1} \sum_f \sum_\delta \langle \mathbf{S}_{f+\delta}\cdot\mathbf{S}_f\rangle_\lambda \geq \langle S^z\rangle_0 N^{-1} \sum_f \langle S_f^z\rangle_\lambda. \tag{18}$$

Now H_0 is the Hamiltonian which yields the usual molecular-field theory[4] of ferromagnetism; therefore, the left side of (18), or ρ_1 for the Heisenberg ferromagnet, is greater than or equal to the product of the known magnetization per site given by molecular-field theory and the magnetization per site given by the exact Heisenberg Hamiltonian including the Zeeman term.

For the trivial case of $H_z = 0$ the result that ρ_1 must be nonnegative follows from (18) since $\langle S_f^z\rangle_\lambda$ vanishes by symmetry. The further result that ρ_1 must be a nonincreasing function of T for $H_z = 0$ follows from (8), since ρ_1 is then proportional to $\langle -H\rangle_\lambda$.

IV. ANTIFERROMAGNET

Now consider the Heisenberg Hamiltonian for a system with nearest-neighbor antiferromagnetic interaction of the spins (each of the magnitude S). Let f and g refer to two mutually exclusive, identical sublattices. The Hamiltonian

$$H = 2J \sum_{(f,g)} \mathbf{S}_f\cdot\mathbf{S}_g + g\mu H_z\Big(\sum_f S_f^z + \sum_g S_g^z\Big)$$

may be written

$$H = H_0 + \lambda H_1,$$

with

$$H_0 = (g\mu H_z + Jz\sigma_2 S) \sum_f S_f^z + (g\mu H_z + Jz\sigma_1 S) \sum_g S_g^z,$$

and

$$H_1 = 2J \sum_{(f,g)} \tfrac{1}{2}[(\mathbf{S}_f - \sigma_1 S\hat{k})\cdot\mathbf{S}_g + \mathbf{S}_f\cdot(\mathbf{S}_g - \sigma_2 S\hat{k})].$$

The notation follows the ferromagnetic case, except that there are now two arbitrary quantities σ_1 and σ_2 which were introduced by adding and subtracting equal terms. The trace is conveniently taken over the same set of eigenstates as for the ferromagnetic case.

[4] J. H. Van Vleck, Rev. Mod. Phys. 17, 27 (1945).

From (3), (13), and (15),

$$\sum_{(f,g)} [\langle \mathbf{S}_f\cdot\mathbf{S}_g\rangle_\lambda - \sigma_1 S\langle S_g^z\rangle_\lambda + \langle \mathbf{S}_f\cdot\mathbf{S}_g\rangle_\lambda - \sigma_2 S\langle S_f^z\rangle_\lambda] \leq \sum_{(f,g)} [\langle S_f^z\rangle_0\langle S_g^z\rangle_0 - \sigma_1 S\langle S_g^z\rangle_0 + \langle S_f^z\rangle_0\langle S_g^z\rangle_0 - \sigma_2 S\langle S_f^z\rangle_0]. \tag{19}$$

By selecting

$$\sigma_1 S \equiv \langle S_f^z\rangle_0; \qquad \sigma_2 S \equiv \langle S_g^z\rangle_0, \tag{20}$$

the right side of (19) vanishes and

$$\sum_{(f,g)} \langle \mathbf{S}_f\cdot\mathbf{S}_g\rangle_\lambda \leq \sum_{(f,g)} \tfrac{1}{2}[\langle S_f^z\rangle_0\langle S_g^z\rangle_\lambda + \langle S_g^z\rangle_0\langle S_f^z\rangle_\lambda]. \tag{21}$$

In contrast with the ferromagnetic case[5] the antiferromagnetic short-range order must, for $H_z = 0$, be a nonpositive, nondecreasing function of temperature.

ACKNOWLEDGMENT

The author is grateful to Dr. R. B. Griffiths for kindly reading the preprint of this article, and for offering useful suggestions.

[5] T. Nagamiya, K. Yosida, and R. Kubo, Advan. Phys. 4, 2 (1955).